Studies on Entrepreneurship, Structural Change and Industrial Dynamics

Tessaleno Devezas
João Leitão
Askar Sarygulov *Editors*

Industry 4.0

Entrepreneurship and Structural Change in the New Digital Landscape

Springer

Studies on Entrepreneurship, Structural Change and Industrial Dynamics

The 'Studies on Entrepreneurship, Structural Change and Industrial Dynamics' series showcases exceptional scholarly work being developed on the still unexplored complex relationship between entrepreneurship, structural change and industrial dynamics, by addressing structural and technological determinants of the evolutionary pathway of innovative and entrepreneurial activity.

The series invites proposals based on sound research methodologies and approaches to the above topics. Volumes in the series may include research monographs and edited/contributed works. Please send the completed proposal form downloadable from this page (see "For Authors and Editors") to the series editors.

More information about this series at http://www.springer.com/series/15330

Tessaleno Devezas • João Leitão •
Askar Sarygulov
Editors

Industry 4.0

Entrepreneurship and Structural Change in the New Digital Landscape

Springer

Editors
Tessaleno Devezas
University of Beira Interior
Covilhã, Portugal

João Leitão
University of Beira Interior
Covilhã, Portugal

Askar Sarygulov
Saint-Petersburg State University of Architecture and Civil Engineering
Saint Petersburg, Russia

ISSN 2511-2023 ISSN 2511-2031 (electronic)
Studies on Entrepreneurship, Structural Change and Industrial Dynamics
ISBN 978-3-319-49603-0 ISBN 978-3-319-49604-7 (eBook)
DOI 10.1007/978-3-319-49604-7

Library of Congress Control Number: 2017934476

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Acknowledgment

This book was prepared on the basis of contributions presented at the Third International Conference "Breakthrough technologies of the 21st century and their transformative power on the existing industrial structure and socioeconomic realm," held at Peter the Great Saint-Petersburg Polytechnic University, on May 18–20, 2016, in the framework of the Russian Science Foundation Project No. 14-28-00065.

Contents

Introduction

Tessaleno Devezas, João Leitão, and Askar Sarygulov

1 The Arguments

At the dawn of the nineteenth century, the human civilization inaugurated a new era in its way of wealth production: the industrial production, or in other words, the so-called Industrial Revolution. Mainstream economics tell us that since the inception of this radical transformation, human technology has enabled and promoted a series of disruptive transformations in the form of industrial production. Following this line of thought, the first industrial revolution, located at the transition from the eighteenth to nineteenth centuries, was characterized by the introduction of water- and steam-powered mechanical manufacturing facilities. The second industrial revolution, which happened at the transition from the nineteenth to twentieth centuries, was based on the introduction of electrically powered mass production and the intensive division of labor. Then followed the third industrial revolution, which entrenched from the 1960s until the 1990s, whose main driving force was the usage of electronics and information technologies (digital revolution) to achieve further automation of manufacture. These three successive revolutions can be coined as Industry 1.0, Industry 2.0, and Industry 3.0, respectively.

Most economists now agree that we are entering the fourth industrial revolution, whose main characteristic is the usage of cyber-physical systems or, in other words, the linkage of real objects and people with information-processing/virtual objects via

J. Leitão acknowledges the financial support from Fundação para a Ciência e a Tecnologia, IP (FCT) (Grant: SFRH/BSAB/127908/2016).

T. Devezas (✉) • J. Leitão
University of Beira Interior, Covilhã, Portugal
e-mail: tessalen@ubi.pt; jleitao@ubi.pt

A. Sarygulov
St.-Petersburg State University of Architecture and Civil Engineering, St. Petersburg, Russia
e-mail: asarygulov@lan.spbgasu.ru

T. Devezas et al. (eds.), *Industry 4.0*, Studies on Entrepreneurship, Structural Change and Industrial Dynamics, DOI 10.1007/978-3-319-49604-7_1

information networks (Internet-of-things, 3D printing, artificial intelligence, bioengineering, cloud computing, etc.) but also using nanotechnologies and new efficient and intelligent materials. The term appeared for the first time in 2011 during the famous Hannover Fair, as a kind of project in high-tech strategy of the German Industry, and in the following year was created the (German) Working Group on Industry 4.0, which delivered its final report in April 2013 again at the Hannover Fair. This report defined the Industry 4.0 environment, which includes the strong customization of products under the conditions of high flexibility of mass production (improved automation technology), requiring the introduction of methods of self-organized systems (self-optimization, self-configuration, self-diagnosis, etc.) to get the suitable linkage between the real (machines, workers) and the virtual worlds.

The term was definitively adopted at the 2015 World Economic Forum (WEF) Annual Meeting held in January 2015 in Davos, Switzerland, soon followed by the publication of the book "The Fourth Industrial Revolution," signed by Klaus Schwab, founder and president of WEF. More recently (January 2016), WEF delivered a very important report entitled "The Future of Jobs: Employment, Skills, and Workforce Strategy for the Fourth Industrial Revolution."

In a similar vein, the evolutionary economics current wants to see not just a succession of industrial revolutions but approaches the succession of industrial transformations under the perspective of long-term fluctuations of the world economy, the well-known socioeconomic long waves or Kondratieff Waves (K-waves for short). This perspective, also known as the Neo-Schumpeterian economic school, is based on the evolutionary concept of the global economy as a far from equilibrium open complex system with an endogenous self-regulatory mechanism, characteristic of the self-organized complex systems. Such mechanism consists in the periodic swarming of basic (or disruptive) innovations that concentrate during the phase of economic recession, with a period of about a half-century. According to the Schumpeterian concept of creative destruction, these downwave swarming of innovations triggers the following upwave phase of economic expansion, during which the bunch of new radical technologies interact synergistically giving birth to a completely new "technosphere."

Thus, this current of economic thought advocates the occurrence of at least four K-waves (some adepts of this school speak about previous five waves) since the inception of the Industrial Revolution at the turn of the eighteenth to nineteenth centuries and that we are now entering the expansion phase of the fifth K-wave (or sixth according to other). Each of these K-waves defines a clear "technosphere" designed by the leading basic innovations of that period, a not very different view from that of the successive industrial revolutions defended by mainstream economists.

But whatever school of economic thought we follow, the fact is that we are witnessing now the birth of a completely new socioeconomic paradigm, probably the most important global change ever experienced by the human civilization. Things are now very different; the concept of innovation itself has changed a lot; means of wealth production are no longer what they used to be. During the previous industrial revolutions, or the equivalent K-waves, we have had the emergence of huge materialized industrial sectors based on the introduction of radically new artifacts, generating a myriad of new jobs and professions. Today's means of wealth production are not based more on the mass production of artifacts, but instead are increasingly decoupled from material consumption and significantly based on intangible innovations and assets. Millions of jobs are being vaporized in a rhythm never seen, while others are emerging, toward the creation of billion-dollar companies, the so-called unicorn companies, which are conducted by a reduced number of highly skilled professionals.

In summary, we have now a very different reality, strongly based on the virtual cyber-physical systems world, which carries with itself the necessity of a profound structural change of production means, trade, education, and social organization. Governments, entrepreneurs, businesses, and the ordinary people need to adapt to this "brave new world" in the scope of a new digital landscape.

2 Setting the Stage

Some authors of the chapters of this book are carrying out research activities within the project *"Structural and Cyclical Paradigm of Economic and Technological Renewal of Macro-systems (World and Russia in the first half of the XXI century),"* supported by the Russian Science Foundation (Project 14-28-00065), lead by Professor Askar A. Akaev from the Moscow State University. Some of the conclusions reached until now by the research team can be summarized as:

– Classic long Kondratieff waves tend to shorten, and the Kuznets waves are becoming increasingly determinant in the development of innovation.
– These long economic waves are asymmetrical and occur as asynchronous processes.
– Economic development is in fact a non-equilibrium process.
– Structure of the economy plays an important role in the formation of economic cycles and consequently is paramount to promote economic growth.
– In the same vein, economic structure affects the incentives to innovate and exerts influence on the formation of competitive industrial sectors.
– The loss of economic systems due to structural imbalances is not only commensurable with the losses from the cyclical crises but greatly exceeds them due to the inertia of the process of structural reform.
– The creation of a monitoring system is required in order to consistently find the trend of structural changes, evaluating their intensity and calculating the imbalance in the structural development.
– Such monitoring system should also take into account the effect of technological progress and its uneven nature, when structural readiness (or non-readiness) can speed up or slow down the diffusion of new technologies in the industry and thus predetermine the pace of economic development in general.
– New breakthrough technologies and basic innovations have not yet become the object of investors' attention, and national economies are in different phases of

structural readiness, which determines the asynchronous nature of their economic and technological renewal.
– A new technological platform is being formed, the result of which will be the creation of a fundamentally new manufacturing industry (Industry 4.0).

An important goal of the project was to consider the role of the foreseen breakthrough technologies of the first half of the twenty-first century in the formation of the new industrial structures and the consequently changing environmental, social, and economic landscape. With this in mind, the research team organized the *Third International Conference "Breakthrough technologies of the 21st century and their transformative power on the existing industrial structure and socioeconomic realm"* held at Peter the Great Saint-Petersburg Polytechnic University, on May 18–20, 2016, where the main topics for discussion were identified as:

– Emerging disruptive technologies, materials, and energy
– Industrial and structural change
– Impact on the environment
– Technological clusters
– Entrepreneurial activities
– Digital economy
– Social consequences of new technologies and the future of jobs

This book includes some of the contributions presented at this conference, as well as others especially conceived for this edited volume, which is the first to appear in the scope of the new Springer series titled "Entrepreneurship, Structural Change and Industrial Dynamics." This series aims to disseminate exceptional scholarly work being developed on the still unexplored complex relationship between entrepreneurship, structural change, and industrial dynamics by addressing structural and technological determinants of the evolutionary pathway of innovative and entrepreneurial activity.

3 The Book

The book is divided into three parts: Part I—Structural Change & Cycles; Part II—Technological Change; and Part III—Entrepreneurship Development.

In Part I, various approaches are presented by established scholars, funded in the technological and economic cycles and waves tradition and in the innovative change theory, following a Neo-Schumpeterian approach. They point for the need of deepening the knowledge on long-term tendencies and structural changes connected not only to the macroeconomic conditions, but rather with the structural characteristics of different industries that are experiencing fast changes, under the new paradigm based on the convergence between the connectivity and interoperability dimensions.

There follows a summarized review of the contributions brought together in the three main parts of the book. Concerning Part I, in chapter "Economic Potential of Breakthrough Technologies and Its Social Consequences", Askar Akaev and Andrey Rudskoi devote special attention to the role played by new breakthrough technologies, namely, the so-called NBIC-technologies (N—nano-, B—bio-, and I—information- and C—cognitive-), which are primarily due to the powerful synergies generated through their convergence. The authors present an application to the US economy, which uses about 10 % of the NBIC technologies' capacity. Through the use of an original mathematical model, it is revealed that the use of such type of technologies can ensure economic growth by 1 % or more. An analysis of the social impact of the use of NBIC technologies is also provided, addressing the social consequences of the replacement process of highly skilled and paid jobs by efficient new information technologies.

In chapter "Structural and Technological Stalemate in Eurozone: If This Is the Reality, What We Can Expect?", Askar Akaev, Yuri Ichkitidze, and Valentin Sokolov analyze structural and cyclical imbalances in the context of the European economy. The authors reveal that, since 2009, both the Eurozone and the USA are facing an unstable equilibrium that is prone to buckling under the influence of small internal or external price shocks. For better characterizing a hypothetical bifurcation process, several nonlinear dynamics models are developed that aim, on the one hand, to describe five possible stages of the economic system and, on the other hand, to show that the Eurozone economy is entering into a very important stage of bifurcation. In addition, the consequences of the bifurcation process are identified, for determining the nature of the future economic development not only in Europe but also in global economy.

In chapter "On the Asymmetry of Economic Cycles", Valentin Sokolov, Tessaleno Devezas, and Svetlana Rumyantseva offer a brief review on economic cycle theories, outlining the key parameters that describe their dynamic behavior, and propose a model of asymmetric economic cycles. The model is based on the assumption that economic cycles are nonequilibrium processes which are characteristic of complex open-ended systems. It uses two different seesaw-like sine functions in its construction: an additive version and a multiplicative version. An empirical verification of the model is presented for the GDP cyclic dynamics for several countries and the asymmetries observed on a global scale. For some countries, the asymmetric behavior of energy production and unemployment is also analyzed.

In chapter "Modern Trends in Evaluation of Macroeconomic Structural Change", Igor Yevsikov, Konstantin Korovin, and Askar Sarygulov outlined the pronounced structural allometry with disproportionally high share of the financial and banking sectors, in the context of a global financial crisis. The authors propose new approaches for evaluating structural shifts, based on the attractor theory, namely, the use of proportionality factor and World-Mensch's nonlinear model. The empirical findings provide the basis for constructing a structural changes monitoring system, which can also be used in the scope of structural shift management.

In chapter "Distribution and Clusters of Basic Innovations", Michael Bolotin and Tessaleno Devezas present an innovative approach about the temporal distribution of basic innovations, which are here considered as the main driver of economic growth. A retrospective analysis of statistical evidence and a proof of the Poisson distribution of basic innovations based on the theory of random processes are presented. In the innovative empirical approach, the authors use a unified supersample time series rather than smaller different datasets as previously done by other authors. The empirical evidence points to a strong evidence favoring the Poisson distribution of innovations. The main findings reveal that long-term changes of appearance of basic innovations occur less frequently than expected by the theory of long waves. Therefore, long-term changes in occurrence of basic innovations are not strictly periodic. In what concerns the wavelike dynamics of global economies, for future studies, the authors suggest exploring its correlation with innovation diffusion, rather than with their clustering.

In chapter "Financial Instability Under Innovation Development: Reasons and Regulation Within the Model of Evolutionary Processes", Yuri Ichkitidze uses evolutionary process modeling for studying the still unexplored topic of financial instability. The author presents some empirical facts confirming that the stock's price dynamics is better described by a Markov switching type model. Furthermore, using the equilibrium model of price formation, it is demonstrated that the temporary price trends on stock markets are evolutionary processes that occur in the conditions of a duality of the equilibrium between the market price and the fair value. The main results reveal that the causes of financial instability are the capital concentration and the lack of investment opportunities, as compared with the available financial resources, whereas the symptoms are frequently recurring financial bubbles and crises.

In Part II, various contributions address the broad topic of technological change, which here covers the complex and nonlinear processes from idea generation to introduction of innovation into the market. Nevertheless, the set of contributions addresses a caveat found in the literature, that is, the lack of geostrategic and macro perspectives on supranational interconnections and long-term tendencies that are able to guide nonlinear process of innovation and entrepreneurship, with a diversified focus on renewable energy, new materials, additive manufacturing technologies, private astronautics, eco-efficiency, and sustainability.

In chapter "The Challenges in the Transition from Fossil Fuel to Renewable Energy", Unurjargal Nyambuu and Willi Semmler advocate that the implementation of climate stability requires a change in energy technologies. Moreover, the authors defend that renewable energy exhibits no externality when produced and is of infinite supply, although it implies some costs. These are the triggers for studying the cost and price trends of fossil versus renewable energy and the effects of those industries' performance on stock markets. A growth model is presented that explains the dynamics of how fossil fuel is phased out and renewables are phased in. Considering the transition from fossil to renewable energy, an interesting discussion on employment effects is also provided, as well as guidelines for future industrial policies.

In chapter "Racing to a Renewable Transition?", William Thompson and Leila Zakhirova advocate that over the past 500 years, historical patterns of energy transitions suggest that systemic leadership transitions have become tied to transitions in the primary source of energy. In this line of reasoning, the authors raise a critical question: what might be the next big source of energy that will power the world economy? The authors argue that the conversion to non-fossil fuel, from the part of the USA and China, is critical for ensuring the welfare of their own economies and the movement to address threats emanating from climate change. Nevertheless, neither state, albeit for different reasons, is embracing renewable or non-fossil fuel energy to the extent necessary to feel very confident that an energy transition will occur.

In chapter "Metal Matrix/Nanocarbons Composites Based on Copper and Aluminum", Oleg Tolochko, Vesselin Michailov, and Andrey Rudskoi present recently discovered carbon nanostructures such as carbon nanofibers (CNFs), nanotubes (CNTs), and graphene, which are promising components for next-generation high-performance structural and multifunctional composite materials. The authors develop a new approach to fabricate composite materials on the basis of aluminum and copper matrix. This study provides a solution to one of the biggest challenges for creating strong, electrically or thermally conductive CNTs/CNFs or graphene composites, that is, the difficulty of achieving a good dispersion of the carbon nanomaterials in a metal matrix. It is also revealed that these discontinuously reinforced new materials can be successfully developed for industrial applications.

In chapter "Additive Technologies: The Basis of Digital Custom Manufacturing", Anatoliy Popovich, Vadim Sufiiarov, and Alexey Grigoriev address the hot topic of additive manufacturing (AM) technologies, also known as 3D printing, devoting special attention to the development of first polymer machines for manufacturing functional metal parts with advanced characteristics and 3D bioprinting. The authors present and describe an approach for designing and producing a hip prosthesis from titanium alloy, using additive manufacturing technologies. The study focuses on titanium alloy powders containing 6 % aluminum and 4 % vanadium, which are raw materials for selective laser melting based on gas and plasma atomization technologies. Empirical findings about the influence of selective laser melting modes on the quality of samples are presented. In addition, research results on the microstructure and phase composition of compact material before and after the heat treatment are revealed. The set of results open a new research avenue for creating and manufacturing customized products, in accordance with the conception of a fully digital production.

In chapter "Private Astronautics and Its Role in Space Exploration", Alexander Zheleznykov and Vadim Korablev describe the current status of private astronautics activities, which are associated with non-state financing of the development of space capabilities, such as launch vehicles, spacecraft, and satellites, for either applied or scientific purposes. The authors present distinct scenarios, outlining that if in 2015 the contribution of private astronautics into world space activities is estimated at some percent, then by 2020 we can expect about 15–20 % of the world market, and by 2030—up to 40 %. This only applies only if there is no significant deterioration in the international political scenario. Any crisis in relations between the leading space powers will certainly affect the development of “private astronautics.” Following the authors’ vision, the favorable development of the situation will make the bold entrepreneurs major players in the field of space activities.

In chapter “The MANBRIC-Technologies in the Forthcoming Technological Revolution”, Leonid Grinin, Anton Grinin, and Andrey Korotayev analyze the relationship between Kondratieff waves and major technological revolutions on the basis of the theory of production principles and production revolutions. The authors present some forecasts about the features of the Sixth Kondratieff Wave/the Fourth Industrial Revolution. In this line of thought, they show that the technological breakthrough of the Sixth Kondratieff Wave may be interpreted as both the Fourth Industrial Revolution and the final phase of the Cybernetic Revolution. It is assumed that the sixth K-wave in the 2030s and 2040s will merge with the final phase of the Cybernetic Revolution, that is, a phase of self-regulating systems. According to the authors, this same period will be characterized by the breakthrough in medical technologies (e.g., MANBRIC technologies—i.e., medical, additive, nano-, bio-, robo-, info-, and cogno-technologies) that will be capable of combining a number of other technologies into a single system of new and innovative technologies.

In chapter “Global Pattern in Materials Consumption: An Empirical Study”, Tessaleno Devezas, António Vaz, and Christopher Magee survey the worldwide production of a basket of 98 important industrial materials between 1960 and 2013, divided into nine main groups. The authors raise the following question: if it is possible to keep the global economy growing while simultaneously reducing the amount of materials resources necessary to maintain today’s growth rate? In the context of the still unexplored framework of dematerialization theory and decoupling concept, the authors reveal an increasing consumption of some key materials in the first decade of the present century, but which seems to reveal a declining trend in most recent years. The growth trend evidenced in the first decade of the present century is interpreted as a consequence of China’s modernization in the past three decades. The empirical findings do not support the conclusion that human society is experiencing dematerialization of the global economy, but some patterns found show that there are reasons for optimism.

In Part III, entrepreneurship development is addressed under an eclectic basis, trying to conjoin different although complementary perspectives on technology entrepreneurship, corporate entrepreneurship, family business, innovation and entrepreneurial networks, entry, exit, and mobility determinants and coopetition and co-innovation.

In chapter “Challenges in Technology Entrepreneurship: Managing Interprofessional Conflicts in Biopharmaceutical Enterprises”, Călin Gurău investigates biopharmaceutical enterprises, which are classical examples of knowledge organizations with multiple professional cultures. The two main professional groups that should cooperate for the success of the new biopharmaceutical venture—business people and scientists—have, however, different sets of values, which can determine intra-organizational tensions and conflicts. The author makes an attempt to identify the existing conflicts between the professional subcultures: business people and scientists, which coexist in a biopharmaceutical SME but denote different sets of values, which can determine intra-organizational tensions and conflicts. Thus, the conflict management procedures employed by the manager-entrepreneur to reconcile the interests, requirements, and visions of various groups that work in the organization are analyzed. The research findings show that most organizations present hybrid profiles. However, despite its theoretical limitations, the Conflict Management Matrix can be successfully used by practitioners to diagnose their type of organization and to identify the best procedures to reconcile the existing internal conflicts.

In chapter "Perception Gaps in International Corporate Entrepreneurship: The Role of Knowledge Transfer's Mechanisms", Davide Gamba, Mara Brumana, Tommaso Minola, and Lucio Cassia address the core issue of this edited volume, that is, Industry 4.0, defining it as a wave that has changed the economic environment through the massive introduction of new technological tools, which in turn permit to approach knowledge transfer process in a very different way from the past. Having in mind this scenario of change, the authors analyze the perception gaps associated with the knowledge sharing process within a multinational corporation and their effect on performance. The main findings outline the role played by different typologies of knowledge transfer mechanisms, making available a contextualization of the current state of the art of multinational companies' knowledge transfer practices.

In chapter "Family Business and Entrepreneurship: Competencies and Organizational Behavior", Serena Cubico, Giuseppe Favretto, Piermatteo Ardolino, Stefano Noventa, Diego Bellini, Giovanna Gianesini, and João Leitão, selected as unit of analysis the family-owned enterprise, which faces various stages of growth and development over time once the second and subsequent generations enter into the business. The authors develop a qualitative study that aims at in-depth knowledge on increasing the understanding of which competencies are relevant to family business success and how they significantly influence both organizational behavior and family dynamics. Empirical findings highlight the importance of family business as a unique context to study organizational behavior and confirmed how family enterprises are influenced by family culture and kinship ties.

In chapter "Metrics for Innovation and Entrepreneurial Networks", Luís Farinha and João J. Ferreira made a serious attempt to present a model for measuring the regional dynamics of the Triple and Quadruple Helixes (Academia-Firms-Local Governments and Civil Society), based on innovativeness pillars. The authors provide a rich overview on the state of the art of research and is duly assessed in order to develop and propose a framework of analysis that enables an integrated approach in the context of collaborative dynamics. They present the Regional Helix Scoreboard (RHS) for the study of regional collaborative dynamics, which integrates and associates different backgrounds and identifies a number of key performance indicators for future research challenges.

In chapter "Determinants and Interdependence of Firm Entry, Exit and Mobility", Rui Baptista and Murat Karaöz develop an innovative approach where they test the phenomenon of entry and exit activities symmetrically, simultaneously, and with delay response to each other considering mobility. The authors use a longitudinal dataset from Portugal covering the 1986–1993 period and 320 six-digit industries. The industries are classified as growing, mature, and declining parts. The main results reveal that growing industries experience significantly more entry, exit, and changes in market shares among incumbents. Although the argument that declining industries experience less entry has support, the authors advise that its positive significance on exit equation has some weak support. Overall, the results support the entry, exit, and incumbent mobility response to unobserved common determinants. Moreover, additional estimation results suggest that also there is interdependence among entry, exit, and incumbent mobility.

Finally, in chapter "Coopetition and Co-innovation: Do Manufacturing and Service Providers Behave Differently?", Dina Pereira, João Leitão, and Tessaleno Devezas analyze the firms' behavior in generating co-innovative products and services to reveal their innovative performance and the dynamics of coopetition targeted at open innovation. For this purpose, they use a dataset of 3682 manufacturing firms and 1221 service firms that participated in the European Community Innovation Survey (CIS) 2008. A "probit" analysis is conducted separately for manufacturing and service firms and, within each sector, according to firms' category of technological intensity. The results reveal the significant influence of manufacturing and service firms' capacity to generate co-innovative products and services, such as coopetition arrangements between competing firms and other R&D stakeholders, and also firms' capacity to introduce innovations into the market. The empirical evidence also reveals that for service firms the effects of introducing process innovations inside the firm and the existence of internal R&D activities are of major significance for creating a capacity to generate co-innovations.

This book is made available under a conjecture of high risk and uncertainty, as well as global interconnection of markets and industries. That is why it is so critical to understand the four main characteristics of Industry 4.0, namely: (1) vertical networking (e.g., smart production systems, smart factories, smart products, and smart logistics); (2) horizontal integration (that is, new generation of global value-creation networks); (3) throughout engineering (embracing the entire product life cycle); and (4) acceleration through exponential technologies (aiming at mass market application, as their cost and size came down). Increasing its level of understanding, it will act as a key element for better addressing structural changes, managing economic cycles, increasing competitiveness, ensuring technological change, reducing risk, and fostering the pathway into the fourth industrial revolution funded in new exponentially growing technologies but necessarily balanced at the resources level.

A final message is directed to policymakers; it is urgent to (re)design a Smart Industrial Policy, targeted at sustainability, eco-efficiency, equity, responsible production, and fair commerce, if the ultimate goal is to implement a sustainable growth strategy aiming to reduce structural imbalances and social inequality. Recovering a timeless masterpiece performed by Louis Armstrong… "what a wonderful world this would be!"

Part I
Structural Change and Cycles

Economic Potential of Breakthrough Technologies and Its Social Consequences

Askar A. Akaev and Andrey I. Rudskoi

Abstract Since 2008 the world economy is undergoing a systematic structural and technological crisis. The economic potential of the technologies based on silicon semiconductor microelectronics is exhausted. Experts are now trying to identify new breakthrough technologies that can play the same role in the development of society. In this paper we consider the NBIC-technologies (N-nano-, B-bio-, I-information- and C-cognitive-) as such technologies, which are characterized by the their powerful synergies. Exemplifying the US economy, which uses about 10 % of the NBIC-technologies' capacity, and based on an original mathematical model, we estimate that the use of such technologies can ensure economic growth by one percentage point or more. Using data from the medium-term forecast of the US NNI, we provide a long-term forecast of the dynamics of the volume of product innovation and the number of new jobs created in the US and around the world, describing the economic potential of NBIC-technologies. An analysis of the social impact of the use of NBIC-technologies is also provided. In particular, the issue of social consequences of the process of replacement of the highly skilled and paid jobs by more profitable new information technologies, is outlined.

Keywords NBIC-technologies • Basic innovations • Social development • Nonlinear models

1 The Global Economy Slowdown During the 5th KEC (1982–2018)

It is well known that economic development today largely depends on scientific and technological progress. The Nobel prizewinner in Economics Robert Solow proved that technological innovation fosters economic growth (Solow 2000), and

A.A. Akaev (✉)
Research Institute of Complex Mathematical Systems on Behalf of I.R. Prigogine, Moscow State University, Leninskie Gory 1, Moscow, Russia
e-mail: askarakaev@mail.ru

A.I. Rudskoi
Saint Petersburg Peter The Great Polytechnic University, Saint-Petersburg, Russia
e-mail: rector@spbstu.ru

T. Devezas et al. (eds.), *Industry 4.0*, Studies on Entrepreneurship, Structural Change and Industrial Dynamics, DOI 10.1007/978-3-319-49604-7_2

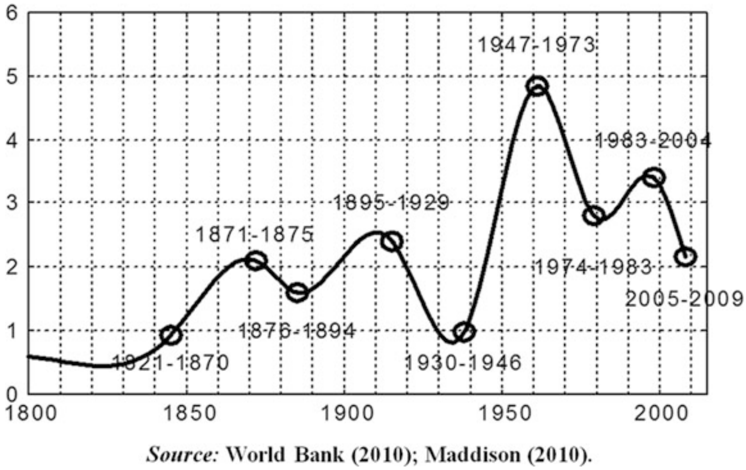


Fig. 1 Change in relative annual average growth rate of the world GDP, 1800–2009 (%)

demonstrated that today more than ¾ of economic growth rate is due to technological progress. Indeed, in the twentieth century it was observed an unprecedented global economic growth mainly due to the scientific and technological revolution powered by breakthrough innovations, as shown in Fig. 1. The average annual GDP growth rate worldwide amounted to 4.9 % in 1948–1973. The world economic crisis in the 1970s was followed by a major recession with the reduction of the average annual GDP growth rate to 3.1 % in 1973–2001. After the recovery from the crisis in 2001, the world economy gained some momentum growing by 3.6 % per year on average in 2003–2007. However, it should be noted that this was mostly generated from fast growing BRIC economies. The crisis of the world financial and economic system in 2007–2008 caused another global economic slowdown. As a result, in 2013, global GDP growth was only 3.1 %, being the lowest growth rate since 2009, when the acute phase of the global financial crisis ended. In 2014, global GDP increased up to 3.3 %, whereas in 2015 its expected growth was about 3.7 %. Thus, after the past record levels, the global economy has entered the period of slowing growth rates, recurring and deepening crises and increasing uncertainty.

The global economic crisis of 2000–2001 which affected the field of information technology and "knowledge economy", sparked off the debate on the prospects of technological progress in the twenty-first century and its impact on the economic development. There were two main propositions. Some scientists stated that the acceleration of scientific and technological progress is still ongoing, therefore there is a high probability of new breakthrough technologies that will again trigger an unprecedented economic growth and provide the solution of current problems in power industry, ecological and social spheres, thereby raising substantially living standards worldwide (Yakovets 2004). Others are skeptical about the future of modern technological progress holding that there was a shift from the revolutionary development of science and technology to the evolutionary type of development in the 1970s, which will result in a gradual technological slowdown in the twenty-first century, at least in its first half (Hirooka 2006; Chernov 2006). They point out that there were no science and technology breakthroughs in the last 30 years and the technological progress has slowed down.

In a recent paper developed by Gordon (2013), it is argued that the economic leap over the last 250 years is a once-off experience in the history of mankind and there are six obstacles for further innovative development of the U.S. economy. These obstacles are the following one: (1) demographic problems due to the 'baby boom' generation retirement; (2) lower level of education; (3) increased income inequalities; (4) expansion of outsourcing, whereby American workers have to compete with cheap labor from developing countries; (5) the deterioration of environment and energy problems; and (6) a huge households' debt and state deficit.

According to Heinberg (2011), this new economic reality has been caused by the impending shortage of the cheap accessible oil. We hold a different view, since the world's proven oil reserves will last for most of the twenty-first century, and petroleum production engineering is steadily developing. It should be noted, though, that high oil prices are among major factors impeding the recovery of the world economy after the crisis of 2008–2009.

So, at the end of the twentieth century mankind entered into the information age, which will reach its peak in 2050, and the 300-year era of industrial civilization will be completed. We expect that breakthrough information technologies, coupled with the NBIC-technologies to be put into practice on a large scale in the next decade (2015–2025), will be able to trigger the economic growth, that is why we are considering here this possibility.

2 On Kondratiev-Schumpeter's Theory of Innovative Cyclical Economic Development

Scientific and technological progress as well as the process of innovation are developing unevenly, in cycles (Schumpeter 1982; Hirooka 2006), which results in the cyclical nature of economic development. The theory of long economic cycles, or waves, introduced by the great Russian economist Nikolai Kondratiev, was one of the core research focus during the twentieth century (Kondratiev 2002). As early as in the 1920s Kondratiev looked into the patterns of the world economy development and discovered long cycles of economic growth with a 50 years duration, which are called "Kondratiev long economic cycles" (KEC). He outlined the relationship between the rising stages of these cycles and waves of technical inventions and their practical implementation in the form of innovations (innovative products and technologies). This type of relationship is revealed in Fig. 2.

The basic innovations of 4th KEC were the breakthrough achievements of the twentieth century technological revolution as shown in Fig. 3: nuclear energy,

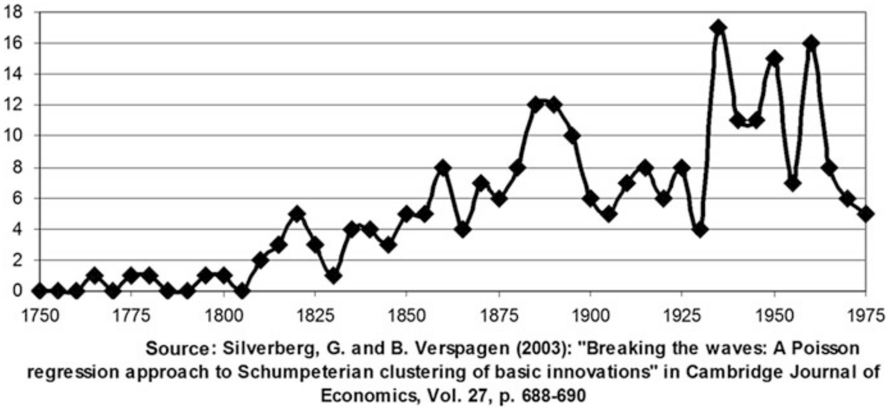


Fig. 2 Technological progress (number of innovations per 5-year period)

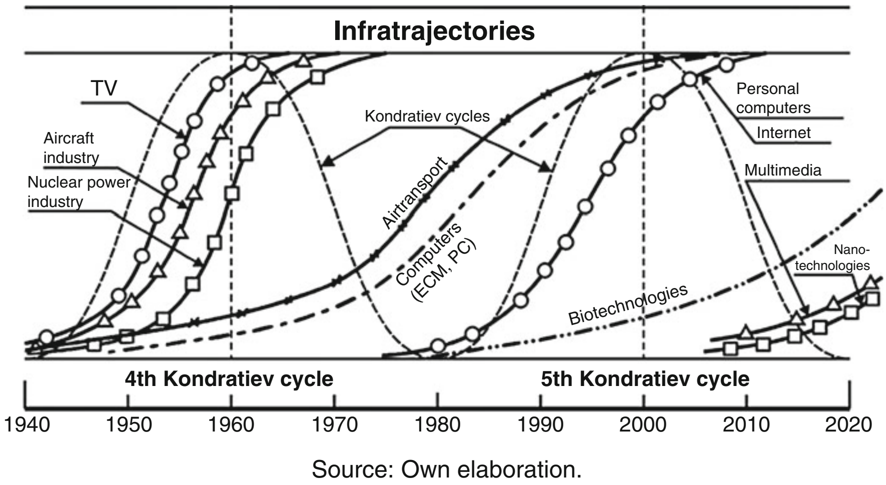


Fig. 3 Diffusion of innovations during the rising periods of Kondratiev cycles of economic activity

quantum electronics and laser technologies; computers and automation of production; jet and rocket engines; Earth covering satellite communication and television were the main innovations of the 4th KEC. It should be mentioned that all these technologies were developed and widely applied for the first time in the history of mankind. It is quite natural that the 4th technological mode (TM) induced the global economic growth of 4.9 % in the period from 1948 to 1973, which was unprecedented in the history of mankind (Fig. 1). The core of the 5th (TM) was composed of microelectronics, personal computers, information technology and biotechnology, which were derivative from the major innovations of the 4th TM. Therefore,

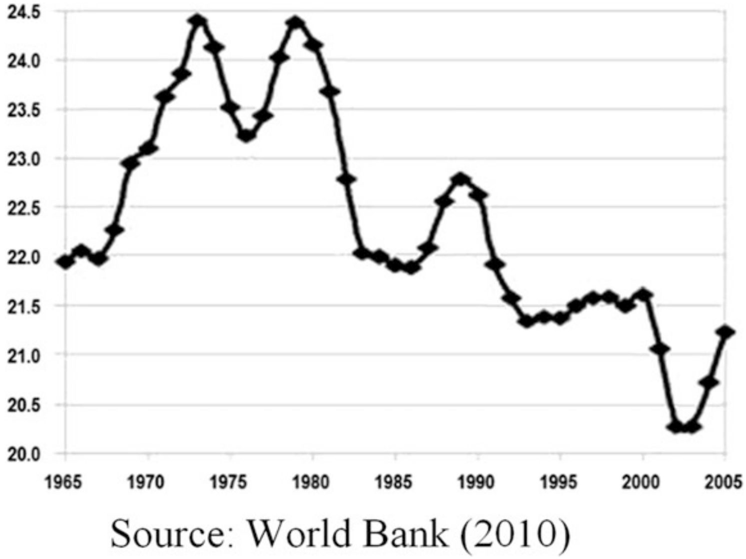


Fig. 4 The dynamics of gross investment into fixed capital (% of world GDP)

the economic efficiency of the 5th TM was much lower: the global economic growth rate decreased to 3.1 % (Fig. 1).

The fact that this trend coincides with the decrease in the number and effectiveness of innovative technologies is confirmed by the reduced share of gross investment in the world GDP over the same period, as well as the lower effectiveness of macroeconomic investment measured in U.S. dollar GDP growth to U.S. dollar investment, as shown in Figs. 4 and 5 respectively. At the beginning of the 5th Kondratiev cycle (1980s) the world entered into the information age, which will reach its peak by the end of the 6th Kondratiev cycle (about 2050). Perhaps then there will be a new technological revolution with breakthrough innovations (Yakovets 2004), which may start a new era of long-term technological and economic growth likewise the one seen in the middle of the twentieth century.

Joseph Schumpeter enhanced Kondratiev's theory of long economic cycles with an innovative theory of long waves, by integrating the former into the innovation theory of economic development (Schumpeter 1939). Hirooka (2006) proved that there is a close correlation between innovations and Kondratiev long cycles. Funded on the analysis of a large array of empirical data, the same author revealed that the diffusion of innovations is strictly synchronized with a rising wave of Kondratiev cycle and reaches its peak simultaneously with the highest peak of the cycle as previously shown in Fig. 3. Various major innovations, due to the self-organization mechanism, form a cluster at the stage of depression. This phenomenon was found by Mensch (1979), who called it a "trigger effect of depression". In other words, depression forces enterprises to seek opportunities for survival, thus powering the process of innovation, i.e. depression triggers the process of

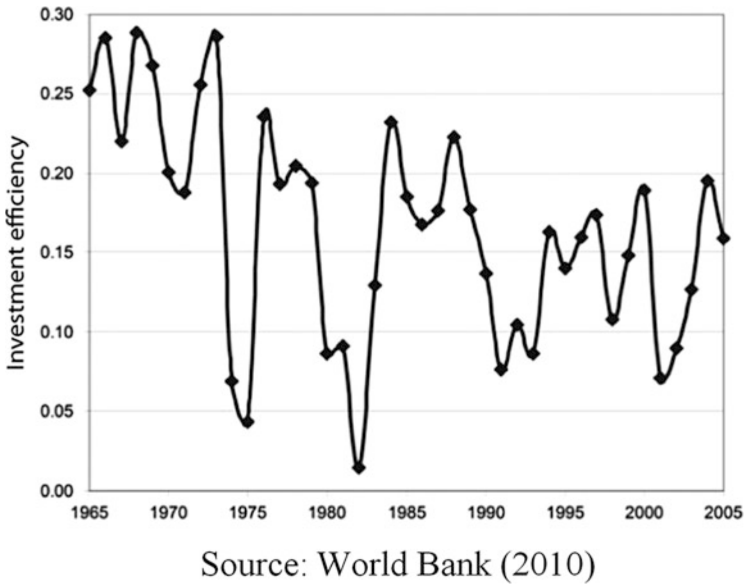


Fig. 5 The dynamics of investment efficiency (USD)

innovation! Clusters of major technologies give rise to new industries, thus triggering another Kondratiev long cycle. The synergetic effect of innovations and their interaction within the cluster can cause the intensive cumulative economic growth that is the main driving force for economic development.

3 Hirooka's Innovative Paradigm

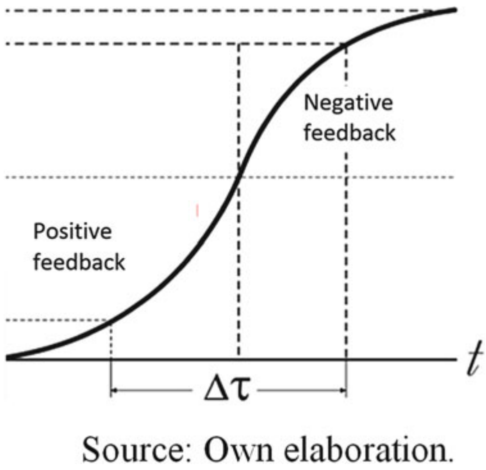
The process of diffusion of innovations was thoroughly studied by Hirooka (2006), who stated that it can be described by a Verhulst logistic function:

$$\frac{dy}{dt} = ay(y_0 - y), \tag{1}$$

where y is the demand for the product at the moment t; y_0 is a limit value of the total market; and a is constant. The solution of this equation is a nonlinear logistic curve described by the equation (see Fig. 6):

$$y = \frac{y_0}{1 + c\exp(-ay_0 t)}, c = const. \tag{2}$$

Fig. 6 Logistic curve of innovation diffusion



In practice, it is usual to express the duration of the innovative product diffusion on the market by the length of time Δt between $\frac{y_{\min}}{y_0} = 0.1$ and $\frac{y_{\max}}{y_0} = 0.9$ that quite accurately reflects the real-time of diffusion (Hirooka 2006).

Nonlinear character of the diffusion of innovations (2) implies that every innovation follows the development trajectory that reaches its peak within a specific time period that signals the completion of the life cycle of innovation. This enables us to identify any innovation and determine the temporal segment of its development. Hirooka (2006) founded that the life cycle of innovations was 90 years during the first industrial revolution (eighteenth century) and it gradually declined to 25–30 years today.

Some innovations extend beyond one Kondratiev cycle to another (Fig. 3), contributing to the emergence of new infrastructures and networks, thus forming a longer trajectory of development, which Hirooka (2006) called infratrajectory (e.g., computers, aviation, biotechnology, etc). These innovations are referred to as mainstream innovations; first, they create new markets, and then its wider application contributes to a new infrastructure in the economy. Infratrajectories form particular clusters, each cluster characterized with its core innovation. For example, in the current 5th Kondratiev cycle, an important cluster has been formed by computer technologies. Mainstream innovations also bring various new applications and institutional changes that contribute to significant market expansion in the subsequent Kondratiev cycle. Mainstream innovations of the forthcoming 6th KEC will be based on information and communication technologies that will form powerful infratrajectories by convergence with nano-, bio-, and cognitive technologies.

Hirooka (2006) also found that the innovative paradigm consists of three logistic trajectories as shown in Fig. 7: technology, development and diffusion. Technological trajectory is a set of core technologies relevant to the innovation that emerged as a result of some significant technological invention or scientific

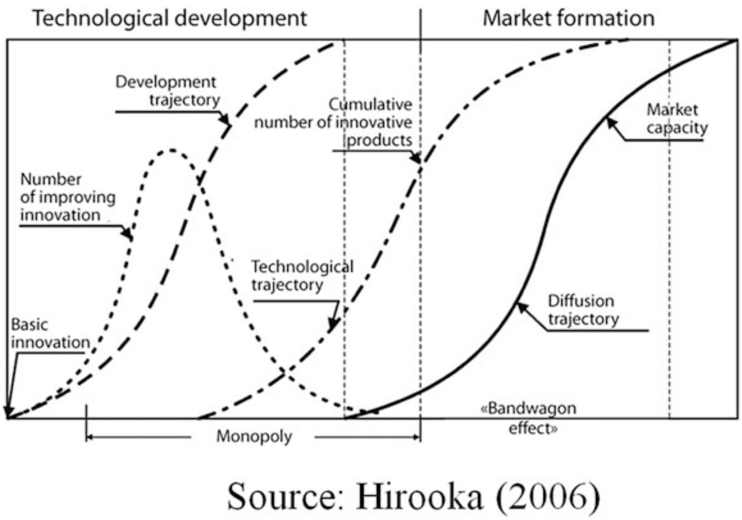


Fig. 7 The structure of innovative paradigm with three trajectories

discovery. The trajectory of development (development of innovation) is a range of innovative products put out through the application of core technologies. The trajectory of development plays a very important role in the innovation paradigm, because it transmits technological knowledge from academic institutions to industries, which results in the establishment of venture enterprises aimed at the industrial development of an innovative product and its marketing. There are more opportunities for venture enterprises in the first 10–15 years of the first half of development trajectory, since the intensive marketing of the innovative product starts soon after the completion of the technological trajectory and lasts about 25–30 years before the market is saturated.

Hirooka (2006) was the first to study the technology's trajectory, by showing that the later is described by the logistic curve and lasts for about 30 years, starting with some significant discoveries or technological inventions. Hence, the innovative paradigm has a cascade structure with three logistic trajectories spaced from each other at a certain distance that is determined in empirical terms. Thus, the innovation paradigm enables us to accurately predict the trajectory of the innovative products diffusion on the market, based on a predetermined trajectory of technology (see Fig. 8), which is exemplified by electronics (Hirooka 2006). Since the latter is 25–30 years ahead of the former, it can easily be built before new products are launched into the market. Thus, Hirooka (2006) analyzed the development trajectories of the most prospective technologies of the future, which are likely to form a technological cluster of the upcoming 6th Kondratiev cycle: multimedia; nanotechnology; biotechnology; genetic engineering and regeneration of human organs; superconductors; and quantum computers. He built development trajectories for all these technologies and found that all of them will reach the peak in the

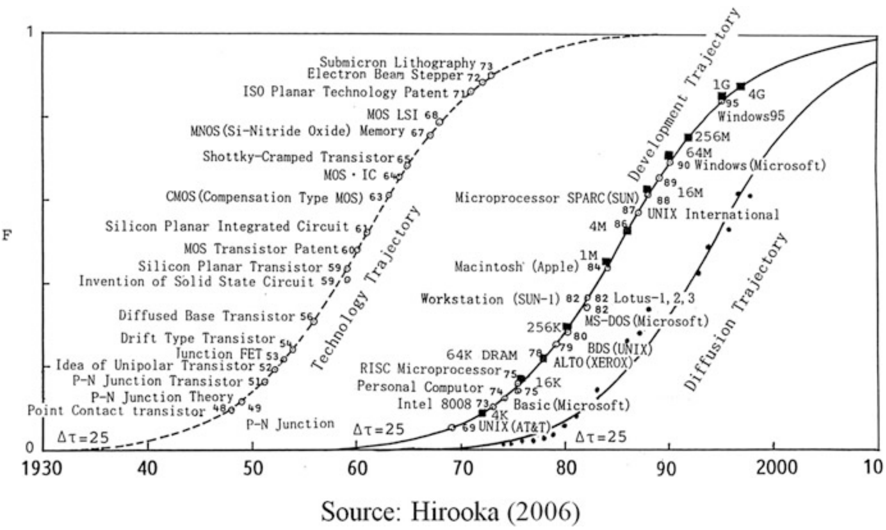


Fig. 8 Innovation paradigm of electronics

2010–2015 period, which implies their intensive diffusion in the market and the start of the upcoming 6th Kondratiev cycle with the rising wave in 2020–2030.

Nevertheless, all innovation technologies of the 6th Kondratiev cycle are evolving from the previous cycles, none of them being breakthrough, which was the case of the 4th technological mode that generated the most powerful Kondratiev economic wave. Therefore, many economists suggest that during the forthcoming 6th KEC the growth rates are unlikely to exceed those in the previous periods of the 4th and 5th KEC. However, most of them neglect the potentially powerful synergetic effect produced by the convergence of NBIC technologies, due to which the 6th TM is able to trigger a greater world economic growth than the 5th TM. Nonetheless, it will not be as intensive as in the 4th TM.

4 Forecasting a Rising Phase of the 6th Long Wave According to Kondratiev's Economic Theory

Using Hirooka (2006)'s innovation approach (see Fig. 7), we will do an attempt to forecast the beginning of the rising phase of the Kondratiev's 6th Economic Cycle (KEC). We definitely know the beginning of the development trajectory of nanotechnologies: 1985 marked the discovery and synthesis of fullerenes, which are carbon compounds C_{60} composed of 60 carbon atoms in the form of a hollow sphere, and it is considered to be the first nanostructure. The atomic force microscope (AFM), which was invented in 1986, enabled imaging and manipulating single atoms. AFM is the main tool for creating new nanostructures and their measurements (Williams and Adams 2007). Leading countries did their best in the above-mentioned areas and showed enormous achievements. In the area of new nanostructures creation, the carbon-related achievements were: (1) in 1991, the discovery of carbon nanotubes with a great potential for a number of industries; and (2) in 2004, the discovery of graphene, a one-atom-thick carbon film, likely to be a mostly prospective material for nanoelectronics. In general, nanomaterials, or smart materials, with their unique properties can be used in almost every sphere of human activity, bringing about innovations and breakthroughs (Rudskoi 2007).

Furthermore, nano tools continuously evolved and progressed resulting into the invention of such essential nano tools as computerized scanning probe microscopes (SPM) for high precision on-line imaging and manipulating nanoparticles; optical tweezers for three-dimensional picking up and manipulating nanostructures. Nanomanipulators with piezoelectric motors enable smooth controllable manipulations in any directions. Summing up, we can argue that nano tools have reached perfection and revealed great opportunities for researchers and scientists to create new nanostructures, measure their properties and find new practical applications. The industrial application of nano tools contributes to the production of nanoparticles and nanomaterials in amounts, which could satisfy the market demand (Williams and Adams 2007).

Thus, it is obvious that the development trajectory of nanotechnologies (nanomaterials and nano tools) is successfully evolving. According to Hirooka's (2006) approach (Fig. 7), the trajectory will reach saturation approximately in 2016 (30 years after AFM invention in 1986) and shortly afterwards the technological trajectory will result into the large-scale diffusion of innovation nanoproducts onto markets. This will lead to the economic upturn in the developed countries and later the global economy. Akayev et al. (2009) confirm that on the basis of the Schumpeter-Kondratiev theory of innovative cyclical economic development, the current slowdown will be prolonged and last till 2016–2018, when it starts the rising phase of the 6th KEC.

Thus, we expect that in 2014–2015 developed countries after a turbulent phase, will start recovering, and in 2017–2018 these same countries will see the upturn of the 6th long wave, based on the powerful impact of the 6th technological mode (TM) where the core effect is associated with NBIC converging technologies. Therefore, the governments of these economies need to concentrate all resources and efforts on capturing the NBIC cluster, which forms the 6th technological mode, that is, the new structure of the global economy. The period from 2014 till 2018–2020, is considered as the most favorable to embrace a new wave of major innovations based on NBIC technologies (Akayev and Rudskoi 2013).

Presently, the nanotechnological industry is growing at 24 %, per year. According to the U.S. National Science Foundation, the potential world market of nanoproducts and nanotechnologies will account for $ 1.1 trillion in 2015. The most contributing areas will be nanomaterials (31 %), nanoelectronics (28 %) and pharmaceuticals (17 %). Two million qualified jobs will emerge as a result of establishing a new high tech industry (Roco 2011). Thus, the supply of nanoproducts will be intensified. Similarly, new industries, based on NBIC technologies, are expected to get off the ground and fuelled the 6th long wave of the world economic development.

5 Forecasting the Economic Potential of NBIC Technologies

Akayev and Rudskoi (2013) present the properties of NBIC technologies that resulted from intensive mutual penetration and influence of nano-, bio-, information and cognitive sciences and technologies. Being the focus of scientists' attention for a long time, this phenomenon has been called NBIC convergence technologies Roco and Bainbridge (2003). Due to convergence, NBIC technologies produce a considerable synergetic effect. In respect of the expected global socio-economic changes, NBIC convergence is considered to be a breakthrough (Bainbridge and Roco 2006). The synergetic effect caused by NBIC convergence, or by mutual cooperation, can be so strong that its contribution to the increase of the total factor productivity will reach a significant impact (about 4.9 %) on the global economy's growth of the 4th KEC (1948–1973).

How to assess this synergetic effect? Let us consider a neoclassical model of the economic growth with correctly identified physical and human capital offered by Mankiw et al. (1992):

$$Y(t) = K^{\alpha}(t)H^{\beta}(t)[A(t)L(t)]^{1-\alpha-\beta}, \tag{3}$$

where $Y(t)$ is the current national income (GDP); $K(t)$ is the level of reproducible physical capital; $H(t)$ is the level of reproducible human capital; $L(t)$ is the level of raw labor involved in the economy; $A(t)$ is the Harrod-neutral technical progress; α and β are parameters of the production function.The empirical analysis revealed that developed countries of the Organization for Economic Cooperation and Development (OECD) have $\alpha = 0.14$ and $\beta = 0.37$, whereas non oil-producing developing countries have $\alpha = 0.31$ and $\beta = 0.28$ (Mankiw et al. 1992, p. 240).Equation (3) in the exchange form is as follows:

$$q_Y = \alpha q_K + \beta q_H + (1 - \alpha - \beta)(q_A + q_L), \tag{4}$$

where $q_Y = \frac{\dot{Y}}{Y}$; $q_K = \frac{\dot{K}}{K}$; $q_H = \frac{\dot{H}}{H}$; $q_L = \frac{\dot{L}}{L}$; $q_A = \frac{\dot{A}}{A}$.

Since the "total factor productivity" can be interpreted as an indicator of collaborative synergetic effects into the economic growth of labor and capital factors. Equation (4) results into the evaluation formula for:

$$q_Y^{syn} = (1 - \alpha - \beta)q_A. \tag{5}$$

Hereby the evaluation formula for OECD countries is $q_{Y_{HD}}^{syn} \cong 0.49q_A$, but for non oil-producing developing countries the value obtained is slightly lower, that is, $q_{Y_{LD}}^{syn} \cong 0.41q_A$.

Thus, the assessment of the synergetic effect, or its contribution to the economic growth, means calculating the growth rates of the technological progress q_A (5). Following Akayev and Rudskoi (2013), we consider a model to calculate growth rates of the average technological level $A(t)$ for the whole economy referring to the relative economic efficiency of newly introduced major innovation technologies, using Dubovsky (1989)'s equation:

$$q_A = \frac{\dot{A}}{A} = s(t)\xi(t)\left(\frac{a}{A} - 1\right), \tag{6}$$

where $s(t)$—savings ratio ($s = \frac{I}{Y}$, I—gross investments, Y—GDP); $\xi(t)$—return on the invested capital ($\xi = \frac{Y}{K}$, K—physical capital). Savings ratio $s(t)$ usually changes cyclically around a slowly changing trend. Therefore, in the first approximation it can be accepted as a constant $s(t) = s_0$.

We have pointed out before that diffusion of further major innovations will be based on the logistic law (2), which is expressed by the following:

$$a = \frac{a_0(1 + c)}{1 + c\exp[-d(t - T_0)]} \tag{7}$$

where T_0 is a year of emergence for a cluster of basic technologies of a new generation; a_0, c, d are constant parameters. Return on the invested capital $\xi(t)$ in Eq. (6) in the long run is also considered as a constant within one KEC. However, inside KEC it significantly diverges from its trend. Synchronizing with KEC, it goes up with the KEC rise and falls in the periods of slowdown and recession. The function $\xi(t)$ for the U.S. economy can be described by the following sinusoid (Akayev and Rudskoi 2015):

$$\xi(t) = \xi_0 + \xi_1 \sin\omega(t - 1987), \tag{8}$$

where $\xi_0 = 0.34$; $\xi_1 = 0.03$; $\omega = \frac{\pi}{15}$.

Differential equation (6) is solved with a numerical method at different values for the coefficient of innovation technology efficiency ρ ($\rho_1 = 0.9$; $\rho_2 = 1.05$; $\rho_3 = 1.2$). Estimation and calculation techniques for all constant parameters in equation (6) are originally developed in Akayev and Rudskoi (2015).The calculations for the growth rates of the technological breakthrough are presented in Fig. 9. Previously, we assessed the efficiency coefficient for NBIC technologies: $\rho = 1.2$. For comparison purposes, we calculated the technological breakthrough with low efficiency of NBIC technologies ($\rho = 0.9$). However, from the analysis of the forecast curves (Fig. 9), it results that growth rates of the technological

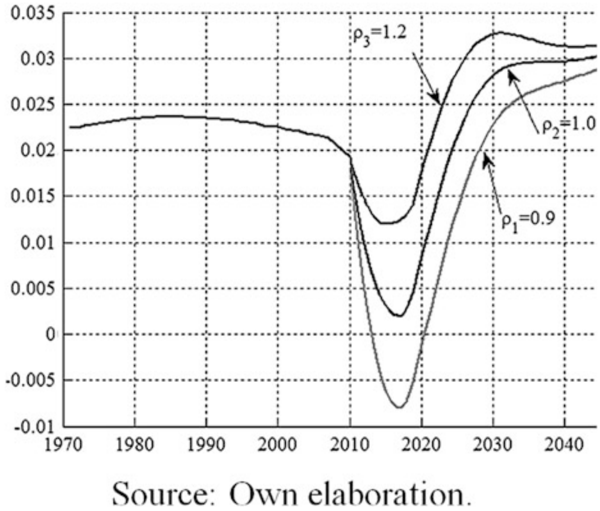


Fig. 9 Growth rate dynamics for the technological breakthrough (q_A) in the 5th and 6th KEC (forecast)

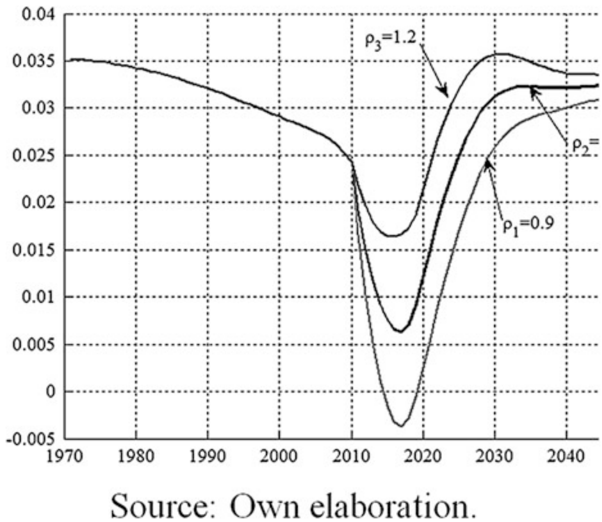


Fig. 10 Growth rate dynamics (q_Y) for the economic development in the 5th and 6th KEC (forecast)

breakthrough will be higher than the ones of the economic progress in the 5th KEC. The main reason for this is a significant rise of synergetic effects.

Thus, a powerful cooperative influence of NBIC technologies does produce the great synergetic effect, accelerating the technological progress to 3.3 % by 2030 (Fig. 9) on the rising wave of the 5th KEC, which is considerably higher than in the 1980–2009 period. Figure 10 shows the graphs of the expected growth rates in the US economy till 2050. From observing Fig. 10, we retain that for the expected efficiency coefficient for NBIC technologies $\rho = 1\,, 2$, the US economic growth will reach the level of the prosperous 1990s by 2020s, and then will show a steady 3.4 %

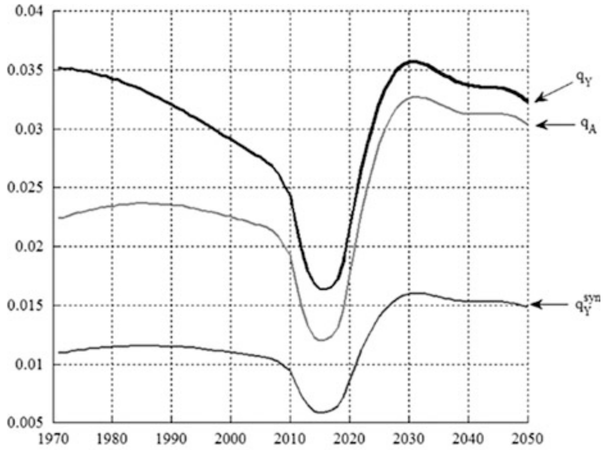


Fig. 11 Expected growth rates of the U.S. economy (q_Y) till 2050, technological progress (q_A) and synergetic effects (q_Y^{syn}) for $\rho = 1,2$

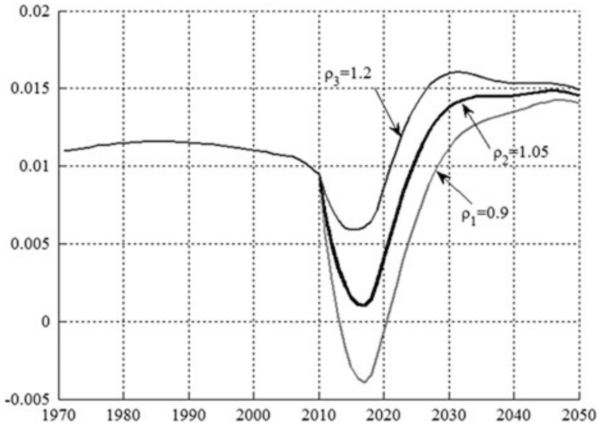
Source: Own elaboration.

annual growth till 2050. The U.S. economy had a similar growth in the 1980s, on the rising wave of the 5th KEC. However, at that time it was caused by the labor force growth ($q_Y = q_A + q_N$). The effect q_N in the 6th KEC will be lower, which is illustrated by comparing the graphs q_Y and q_A in Fig. 11. This figure presents dynamics of the three growth rates q_A; q_Y; q_Y^{syn}.

In Fig. 12 graphs of synergetic effects are presented, according to different possible values of the NBIC technology efficiency coefficient, as compared to the major technologies of the 5th TM. The analysis of the graphs confirms that synergetic effects in the 6th KEC are 1.4 times higher than the ones in the 5th KEC, considering an expected efficiency coefficient of major technologies in the 6th TM equal to $\rho = 1,2$. It is worth noting that even with lower efficiency of NBIC technologies ($\rho = 0,9$), as compared to the major technologies of the 5th TM, the synergetic effect of NBIC technologies in 2030s will greatly exceed the effect of major technologies in the 5th TM (Fig. 12).

According to the Schumpeter-Kondratiev theory of innovative cyclical economic development and Hirooka's innovation paradigm, we expect that economies of developed countries after a turbulent phase, will show a marked growth in 2014–2015, which will be followed by the rising wave of the Kondratiev's 6th Economic Cycle in 2017–2018 (the 6th KEC, 2018–2050). Most significant innovations of the 6th KEC will be NBIC technologies (nano-, bio-, information and cognitive technologies) forming the core of the 6th technological mode (TM). The next upturn in the world economy we are expecting in 2020s on the rising wave of the 6th KEC. NBIC technologies have a powerful synergetic effect produced by the convergence of nano-, bio-, info- and cognitive technologies. These technologies will enable accelerating the technological progress, which will exceed the growth achieved during the rising wave of the 5th KEC (1982–2006). Hereby this will lead to the shift of the slowdown trends in the world economy all along the 5th KEC

Fig. 12 Expected curves of NBIC technology synergetic effects on the growth rates of the U.S. economy in the 6th KEC (2018–2050)



Source: Own elaboration.

(1982–2013) into the upward trends. It has been shown on the example of the U.S. economy that the technological progress will grow from 2.2 % in the 1980s to 3.3 % in the 2030s.

Above we calculated the expected growth rates of the technological progress and U.S. economy triggered by a wide-spread launch of NBIC technologies. Now we will consider a long-term forecast for other important properties of the NBIC technologies economic potential: jobs in new innovation industries and economic sectors, output of the innovative products. Roco (2011) analyzed indicators of NBIC-technology development at the initial stage (2000–2010) and the vision for their development to 2020. The author regards the National Nanotechnology Initiative (NNI) as a driving force for the development and application of nanoscale science and technologies in the U.S.A. and throughout the world. Obviously, the NNI made a decisive impact on the establishment and development of nanotechnologies in the U.S.A. which has become an unassailable leader in this prospective economic area. It helped the U.S. economy recover after the recession of 2008–2009 and go up to the rising wave of the 6th KEC. This growth was fuelled by NBIC technologies.

The NNI is also widely recognized as a new effective coordination mechanism for government agencies that support scientific and engineering research. Motivated by the NNI, coherent and sustained R&D programs in the field were soon announced by other nations, such as, Japan (2001); Korea (2001); the European Community (2002); Germany (2002); and China (2002). From 2001 till 2004, over 60 countries established programs at a national level. However, the first and largest program was the NNI itself. In the 2000–2010 period, its cumulative funding was $13.5 billion, including $1.8 billion in 2010. The NNI became second only to the space program in the US civilian science and technology investments. In the first decade of the twenty-first century the total amount of R&D funding in various areas of NBIC technologies in the U.S.A. exceeded significantly the funding in other countries: Germany ($1.55 billion); Japan ($874.7 million); France ($726.9 million); and Russia ($711.4 million). The leader of NBIC technologies in the developing countries is China (about $600 million). Mihail Roco assumes that R&D funding in the area of NBIC technologies will show an annual steady rise of 25 % within the second decade.

According to Roco (2011) by 2020 nanotechnologies will have been widely used in industry, medicine, information and communication technologies, sustainable use of natural resources. Furthermore, in this same period, there will be a transition from simple nanoscale components to complex active nanosystems, and also from passive nanostructures and nanocomponents to active nanostructures and nanosystems.

Roco (2011) forecasted the dynamics in the number of researchers and workers involved in one domain or another of NBIC technologies all over the world, and predicts the growth rates of innovative products:

i. In 2008 the number of researchers and workers was estimated at about 400,000, of which 150,000 were located in the U.S.A.; In 2015, it is expected to have 2 million nanotechnology workers worldwide, including 800,000 in the U.S.A.; By 2020 the number of researchers and workers is expected to be 6 million, of which 2 million will be in the U.S.A.; and
ii. Also, in 2008, the value of products incorporating nanotechnology as the key component reached about $200 billion, of which near $80 billion was in the U.S. A.; By 2015 a nanoproduct value is expected to be about $1 trillion worldwide, of which $0.4 trillion will be in the U.S.A.; By 2020 a nanoproduct value is estimated at the rate of $3 trillion worldwide, of which $1 trillion will be in the U.S.A.

Thus, the market is doubling every 3 years as a result of successive introduction of new products. The growth rates in new industries account for 24 % per year. Experts forecast the same growth rates in the nearest 5–10 years. Extrapolating the available data in three points ($T_1 = 2008$, $T_2 = 2015$ and $T_3 = 2020$) and based on the logistic curve, we obtained trajectories of the dynamics in the number of new jobs (Fig. 13) and in the output of innovative products in areas of NBIC technologies (Fig. 14). Building the forecast logistic curves, we took into account the fact that the inflection point was located at 2025 according to formula (7). Figure 13 shows that by 2040–2045 the industry of NBIC technologies will have provided about 18 million new jobs worldwide, including 7 million jobs in the U.S.A. However, this is not a great number.

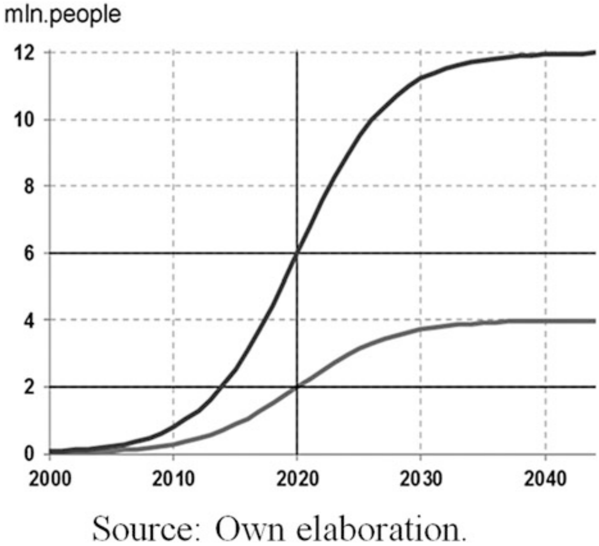


Fig. 13 Dynamics in the number of new jobs in areas of NBIC technologies

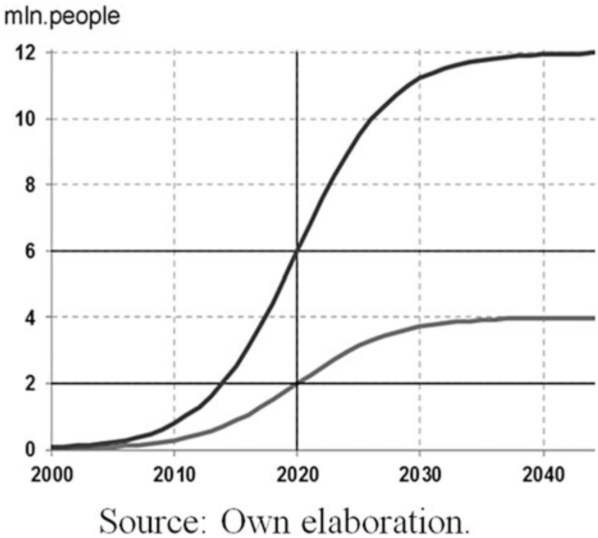


Fig. 14 Rise in the output of innovative products in areas of NBIC technologies

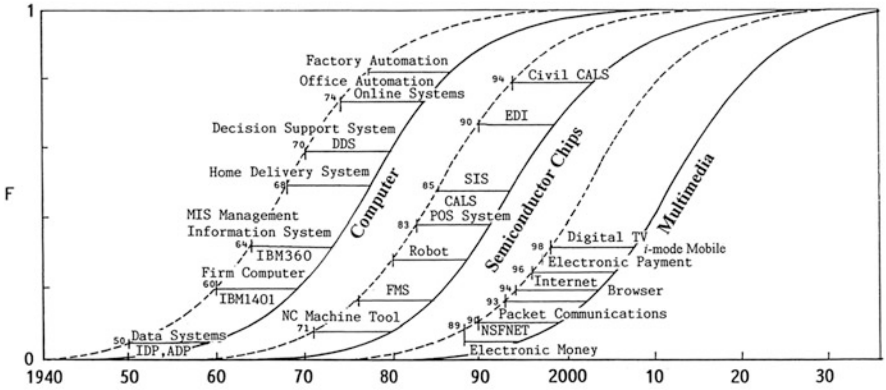
6 Mainstream Innovations of the 6th KEC and Technological Convergence

A new innovation paradigm contributes for developing innovative products and establishing new industries. However, it is extremely important to provide a stream of innovations into existing industries, that will expand the value added in these industries and ensure higher productivity throughout the economy. The transition of technologies from new industries to conventional ones results into the convergence of technologies and evolution of innovation paradigms. Hence, we can identify two directions in the development of major innovations. The first is the establishment of a new industry to produce innovative products. The second is the penetration into the conventional industries that triggers the productivity and the invention of new products through the technological convergence. The innovations are able to promote economic growth, if they have a universal nature and permeate a lot of economic areas. During the 4th and the 5th Kondratiev cycles such a universal innovation were computers and electronics (microprocessors). The most outstanding example is the convergence of electronics and metal-cutting machines, which has led to the invention of high-precision efficient CNC metal-cutting machines.

It should be stressed that the automobile industry was significantly improved by the adoption of electronics. Electronic components are used in today's cars to check engines, control driving etc. Generally, microprocessors have revolutionized the production technology in all industries ranging from metal-cutting machines to cars and planes. The first microprocessor Intel 8088, launched in 1973, was a simple modified chip that can be regarded as a "founding father" of the Information Era based on extremely powerful and cheap microelectronics. Intel 8088 was the foundation for the first commercial IBM PC in 1981. The technological convergence of computers with the industries of steel, cement, chemicals resulted into a quality leap in these industries. Thus, innovations will embrace the whole economy and its institutions via mechanisms of converging technologies and institutional changes. It is important that institutions should meet the tasks at every stage. Accordingly, the usage of major technologies of the 6th KEC in the conventional industries need to be the focus of attention and the stimulation of this process shall be a priority. Today the usage of nanotechnologies is limited to household, medical, agricultural and power engineering areas, i.e. conventional industries.

The fruitful convergence of technologies in the 6th KEC can be exemplified by the convergence of NBIC technologies with the technologies of 3D printing. The latter, known as additive manufacturing, has revolutionized manufacturing, technology, design, medicine, due to the fact that this process is cheaper, user-friendly, environmentally friendly and waste-free. 3D printing is widely used in the production of supporting elements for planes, major parts for spacecraft engines. NASA's experts predict that in the near future it will be possible to apply 3D printing in weightlessness, using robotic spiders for building structures of dozens of meters in diameter and hundreds of meters in length. Raw materials for 3D printing can include a variety of powders, fibers, minerals and other materials. In the short term, it will be possible to observe amazing uses of 3D printing in medicine: printing human organs; intervertebral cartilage printing; remodeling injured bones, etc. Using unique properties of NBIC technologies will allow establishing 3D industry that is able to manufacture diverse customized and personalized products. Moreover, diversification of products is a prerequisite for the sustainable economic development.

It was noted earlier, computers became the mainstream innovation as early as the 4th KEC, bringing to life the digital world, software, microelectronics, Internet,



Source: Hirooka (2006).

Fig. 15 Development of ICT

multimedia etc., which interact, strengthen and enrich each other. All the above-listed is the foundation for the area of information and communication technologies (ICT). Hirooka (2006) calculated and built the ICT innovation paradigm that is shown in Fig. 15. The ICT—the greatest achievement of the twentieth century—are one of the major resources for the economic development in the twenty-first century. At the beginning of the twenty-first century the global ICT market exceeded $1 trillion with the steady growth within the boundaries 10 % per year. These growth rates for this market are expected to be maintained at least till 2020. As Fig. 15 shows their life cycle will continue till 2030s.

The ICT are getting mainstream technologies, infratrajectories that connect the 5th and 6th technological modes, 5th and 6th KEC in the global economy. ICT transform traditional technologies not only by changing their properties, but also completely modifying them. This is related to unlimited opportunities that ICT provide to process large arrays of information using most powerful supercomputers and relevant software. These are the computer technologies that have contributed to the success in genetic engineering which has revolutionized agriculture.

The ICT's revolution has occurred. Computer technologies have transformed telecommunications. The Internet and networks have become the backbone of today's global society. The Internet is not only a current technological innovation, this is a key technology of the Information Age. The Internet has become a fundamental lever for increasing productivity and competitiveness, which for its turn help to set up new network businesses, hereby establishing a new order for the economy. Using the Internet by private companies has become common in OECD countries, e.g. in 25 countries over 89 % of companies have access to the Internet and more than a half have their own web sites. Highly developed countries have recently charted their course toward the Intelligent Internet of Things, based on Wi-Fi protocol and promotion of the mobile internet. The new economy emerged in the U.S.A. in middle 1990s, in late 90s it was disseminated across dynamic sectors of other world economies. For instance, in the 1985–1995 period, labor productivity in the U.S.A. had an average increase of 1.4 % per year, whereas from 1996 to 2000 this increase doubled and accounted for 2.8 %.

Due to the innovation breakthrough in 1980s–1990s, information economic and technological indicators became dominant, and the developed world moved into the post industrial era. According to Jorgenson and Motohashi (2005), ICT made a contribution to the 20–40 % growth of GDP in the developed countries in middle 2000s, with ICT amounting to 70–80 % in the positive dynamics of the aggregate factor productivity. Annual ICT costs rose by 5–6 % throughout the world. ICT costs reached 9–20 % of companies' revenues and account for 5 % of their capital. ICT costs in the GDP amount to 7–10 % for the developed countries (Sweden: 9 %, Great Britain: 8 %, U.S.A.: 8 %, Japan: 7 %), whereas in developing countries had a smaller weight (India: 3 %, Brazil: 5.5 %, Russia: 3 %).

Investments in ICT products and services lead to capital deepening with higher labor productivity. However, this is not typical for every country. The comprehensive empirical research, conducted by the Economist Intelligence Unit, showed that ICT could contribute to the economic growth only in the case of reaching a certain vantage point (EIU 2003). The ICT had to reach a critical amount before making a significant impact on the country's economy. Therefore, countries with the highly developed ICT infrastructure tend to experience faster economic growth. Among these countries are Scandinavian countries—Denmark, Norway, Finland and Sweden. They are the countries which showed the greatest contribution to the rise in labor productivity from 1996 to 2006. Countries, especially developing ones, with lower ICT development indices have either zero ICT effect or a negative one.

Today it is obvious that ICT have a dominant role in a new Kondratiev economic wave and in determining higher economic growth rates, followed by achievements in nanotechnologies, biotechnologies, genetic engineering, creation of new materials, adoption of alternative energy sources and aerospace engineering. ICT provide the growth in creating innovative products and services based on NBIC technologies, in their manufacturing and selling. Using ICT, the developed countries aspire to ensure high automation and optimization of manufacturing processes, to reduce power and material consumption. Most outstanding achievements in this sphere belong to U.S.A. Compared to 1997, the consumption of power in mechanical engineering became halved in 2003. Within the same period in the processing industry this indicator reduced by a third, whereas the U.S. GDP reduced by 15 %. The consumption of materials in the same period in mechanical engineering decreased by 25 %, in the processing industry—by 20 %, GDP went down by 10 % (Smith and Lum 2005).

The ICT speed up the economic development, as countries are increasingly integrated in the global economy, which contributes also for increasing the living standards and economic activity of population. In addition, the ICT can help developing countries to improve accessibility and quality of education, by fostering the use of scientific, technological and organizational achievements accumulated in developed countries. Nevertheless, the developing countries need to develop firstly their ICT infrastructure, for reaching the advantage point that leads to economic growth, technology borrowing and usage of NBIC technologies and innovative products.

7 Social Consequences of NBIC-Technological Revolution

The development of capitalism has gone through an extensive and deep economic crisis in 2008–2009 that is referred to as the first global crisis of capitalism in the twenty-first century and, due to the number of losses incurred, it was called "the Great Recession". The crisis caused a discussion on how long it would take capitalist countries to bottom out of the stagnation, and whether a long-term sustainable growth and economic prosperity would be resumed. The most profound researchers were concerned with the question: "Is there a future for capitalism?" This is also the title of the recently published book written by the prominent historians and sociologists Wallerstein, R. Collins, M. Mann, G. and Derlugyan K. and Calhoun. Collins (2015) highlights the main mechanism leading to the collapse of modern capitalism: it is the ongoing replacement of highly skilled and highly paid jobs by the new information technologies that are becoming increasingly cost-effective.

In this regard Collins (2015) formulated the following question of extreme importance: what are the political and social consequences of the future situation, when two-thirds of the educated middle class in Western countries and around the world, will be unemployed? Millions of people around the world still remember the events of " the Arab spring 2011" when young people who graduated from European universities, but yet unemployed, an instant swept away the authoritarian regimes in a number of Middle East and North African countries. R. Collins, who along with Wallerstein, predicted the demise of the Soviet communism in 1970, is now concerned with the question: how long the capitalism will function as a system and how the middle class being the basis of this system would behave when faced up with the inevitably lower standards of living? However, he does not give a clear answer to the question: Should we expect another great bourgeois revolution in the twenty-first century? Even a detailed analysis of all possible constructive ways to cope with these problems does not provide any alternative to the future course of events when high technologies of the twenty-first century will inevitably result in a high unemployment rate of the middle class.

But what is the situation with the middle class in reality? Let's look at the problem in terms of the coming NBIC-technological revolution. It is widely known that capitalism in nineteenth–twentieth centuries led to mechanization and automation of the manual labor, but, at the same time, caused an increase in the number of managerial positions for the educated middle class on various levels of management. For its turn, this has resulted in the growth of productivity and of the average household income, although the number of employees at enterprises was steadily declining. For example, in the U.S. manufacturing sector the production volume increased by 1.46 times, and productivity rose by 2.15 times in the period from

1987 to 2010, while the employment rate declined by 1.5 times (Sherk 2010). However, the situation began to change dramatically in the 2000s. Computer technology helped to automate many routine tasks, which caused the reduction of traditional white color jobs. Starting from the 1980s, computers were increasingly used to solve the tasks such as accounting and office paper work, as well as full-scale production work at plants—types of work that secured employment and comfortable income for middle class. Thus, unemployment and income inequality will only increase with the growing capacity of computers.

As concerns the innovative sectors of the economy, there will be fewer jobs too. During the upward phase of the forthcoming 6th KEC (2018–2042) new sectors of the economy can use the following formula related to the dynamics of the aggregate labor productivity (a), physical capital ($K_{i\text{в}}$), and employment

$$(L_{i\text{в}}) : a(t) = \nu \left(\frac{K_{i\text{в}}}{L_{i\text{в}}} \right)^{\theta}, \tag{9}$$

where ν—coefficient of proportionality; θ—elasticities stock of workers knowledge in the capital. The previous Eq. (9) describes the model of technological progress, proposed by Arrow (1962). In this model, technological progress (total productivity) is determined by the amount of fixed capital involved in the production, depending of the knowledge volume and employees' skills acquired at work. Arrow (1962) assumed that the value θ for the aviation industry equals approximately to 0.7. For many traditional industries $\theta = 0.32$. Thus, Eq. (9) leads to the following:

$$K_{i\text{в}} = \left(\frac{a}{\nu} \right)^{\frac{1}{\theta}} L_{i\text{в}}, \tag{10}$$

then the marginal rate of labor replacement by capital can be written as:

$$\frac{\partial K_{i\text{в}}}{\partial L_{i\text{в}}} = \left(\frac{a}{\nu} \right)^{\frac{1}{\theta}}. \tag{11}$$

Hence, we conclude that the higher the technical progress is, the more labor will be replaced by the capital.

Taking the logarithmic derivative of Eq. (10) we obtain in growth rates the following equation:

$$\frac{\dot{K_{i\text{в}}}}{K_{i\text{в}}} = \frac{\dot{L_{i\text{в}}}}{L_{i\text{в}}} + \frac{1}{\theta} \frac{\dot{a}}{a}. \tag{12}$$

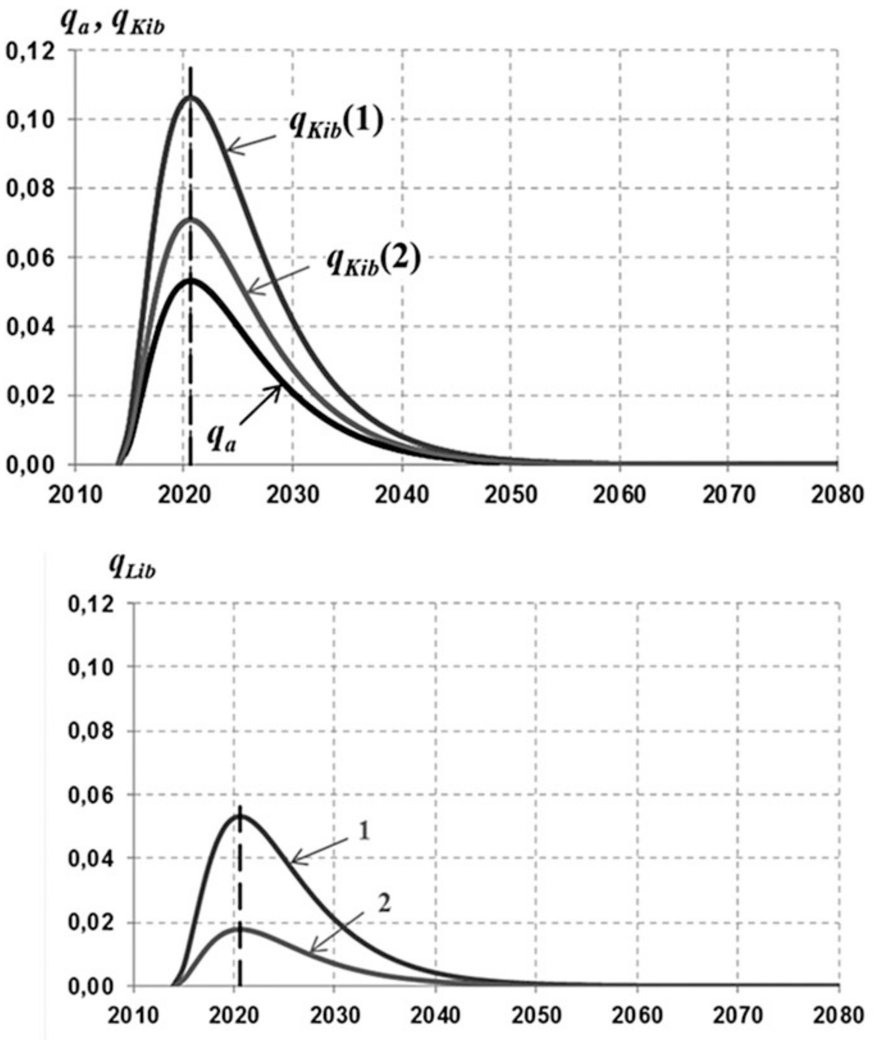
During the next few decades, the pace of technological progress $\left(q_a = \frac{\dot{a}}{a} \right)$, as mentioned in paragraph 5, will show a significant growth. As for the rate of physical capital accumulation $\left(q_{K_{i\text{в}}} = \dot{K_{i\text{в}}} / K_{i\text{в}} \right)$, there is an assumption that it will decline in the period of the 6th KEC as NBIC technologies are less capital intensive than the previous platform technologies of the 5th technological mode.

This conclusion is born out of practice. Today, the slow growth of investment in the world is the concern of many experts. Earlier, when the economies came out of the doldrums of investment into new equipment, new technologies and businesses started growing at a faster pace. But now this is not the case. The experts believe that this is due to the fact that the nature of production itself is changing. Powerful software and the ability to process large volumes of data reduce the need for conventional equipment. The development of cloud computing technology reduces the need for high-performance computers and servers. In addition, as noted above, NBIC technologies are of low capital intensity. Therefore, companies are currently increasing the investment into intangible assets, which means lower rates of accumulation of physical capital.Combining both trends, that is, the increase in the growth rate of technological progress (q_a) and the reduction in the accumulation rate of physical capital $\left(q_{K_{i6}}\right)$, (see Eq. 12), we can conclude that there will be an imminent slowdown in the number of new jobs $\left(q_{L_{i6}} = {}^{\dot{L}_{i6}}/_{L_{i6}}\right)$ in new industries. This process is illustrated in Fig. 16. Today, this trend can be seen in Germany. The main reason lies in the fact that information technology has been introduced into the production process too extensively.

Impressive innovations in computer technology–from advanced industrial robots to automatic translation services, are largely to blame for sluggish employment growth over the past 10–15 years (Brynjolfsson and McAfee 2013). The fact that powerful innovative technologies are getting wider application not only in manufacturing, office work and sales, where they have become commonplace, but also in areas such as legal and financial services, education and medicine, becoming a major concern of middle class today. Moreover, the growing technological progress eliminates jobs faster than creates them, contributing to the stagnation of average available income and an increasing social inequality in the U.S.A.

Brynjolfsson and McAfee (2013) also argue that technological progress facilitates productivity increase, but at the same time eliminates the need for many types of work, sweeping away millions of jobs. The productivity growth is followed by the increase of GDP, but the average income stopped growing, as can be observed in Fig. 17. There is also a growing income inequality. This process began in the 1980s, but moved faster in the 2000s. The authors called this process "Great Separation" and concluded: "People are not keeping pace with the technologies that are improving so quickly that our skills and organizational structures do not correspond to their level".

Thus, the world is going through the next stage of the labor market evolution caused by the transition to the high-tech and knowledge-based economy in which the main workforce is concentrated in the SIEM-sectors (science, information, engineering and mathematics). There will be an increase in the demand for highly skilled jobs requiring the use of computer technology, software development, computer systems, networks and databases analysis, as well as the application of robots, but jobs of the average qualification will become obsolete. Experts from



Source: Own elaboration.

Fig. 16 The influence of the slowing trend of capital accumulation $\left(q_{K_{iв}}\right)$ on the number of new jobs $\left(q_{L_{iв}}\right)$ in innovative sectors of the economy

McKinsey Global Institute predict that by 2020 companies from SIEM-branches all over the world will need 40 million of highly qualified employees (McKinsey Global Institute Analysis 2013). As for the low-skilled labor, the number of people involved in it will exceed the number of working places by 100 million, which will contribute to unemployment growth and stratification in advanced economies. Presently, the income of highly skilled personnel is about three times as high as of the average ones.

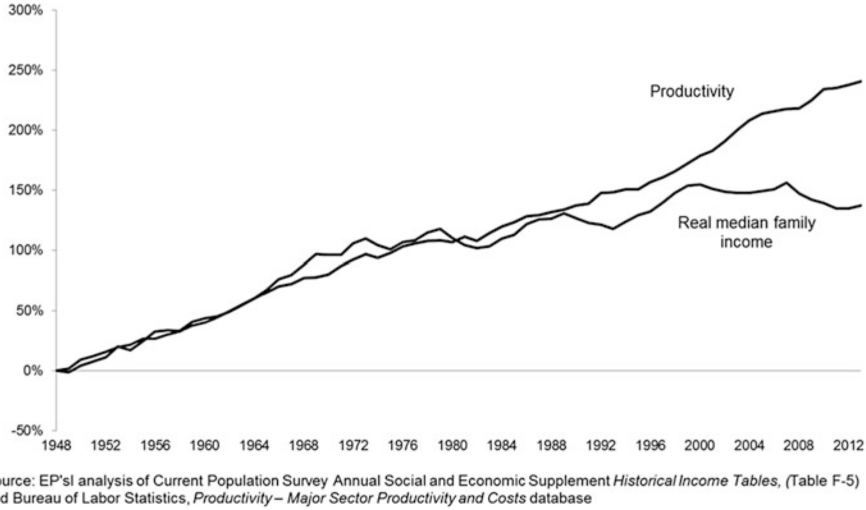


Fig. 17 Dynamics of growth of labor productivity and real average household income in the U.S.A

The usual working environment in the near future will radically change under the influence of ICT, artificial intelligence, intelligent robots, new technologies data processing and cloud computing. The rapid development of these technologies will reduce the huge number of middle class jobs. Recently, John Gaunder, the Forrester analyst, published the results of his research into the issue with the analysis of both well-established firms and start-ups. He expects that by 2025, in the U.S.A., the robots will displace 22.7 million jobs and 16 % of employees from the current number of employees. Due to the convergent NBIC technologies, we are approaching the creation of intelligent robots and intelligent computers. The wide-spread use of household humanoid robots (androids) will begin in 2017–2018. In Japan, for instance, robots are generally regarded as the only means to help a growing number of the elderly: it is planned that they will cook, do domestic chores and brighten their days. Robots were widely used by the Japanese to cope with the aftermath of the accident at the nuclear power plant "Fukushima" caused by tsunami.

Presently, the high-tech sector of robotics is one of the most dynamic and fast growing industrial sectors of developed countries and global production. Robots with primitive forms of intelligence are now widely used in the nuclear power, oil and gas industries, in selecting, sorting and transporting of goods, especially hazardous ones, in large warehouses. Today, an intelligent vehicle is able to drive independently even in city streets. Use of robots in the medical and pharmaceutical field is also impressive. For example, according to U.S. data for 2011, 85 % of all operations on patients with prostate cancer were performed using the surgeon robot Da Vinci. More than two thousand surgical robots are currently under use in the U.S.A. and more than 600—in Europe. Thus, patients themselves are in favor of robotic surgery, because it provides them a new standard of quality of life.

In the coming decades, we are likely to see how people and robots, humans and computers will work together, complementing each other. But then, the time will come when computers and robots will replace humans in many workplaces, which can become commonplace in the 2020s and 2030s. Indeed, in 2025 there will be brand new smart robots that could be threatening many human workplaces. In 2020–2025 there will also be computers with the artificial intelligence, capable of using advanced methods of data processing and the analysis of human language. Such computers with the access to large volumes of data related to the area of application, will be capable of much more applicable recommendations than the “advice” from highly qualified experts. As soon as in 2030s there will appear first cognitive computers, designed to imitate the human brain’s work and then the technological displacement will be ever more intensive. All this may lead to a double reduction of the middle class in the period from 2025 till 2045. Thus, in the 2030s, the extensive use of NBIC technological achievements will create an immense social problem for most developed capitalist countries.

In order to assess the dynamics and scale of jobs loss for the middle class as a result of the coming technological displacement in the first half of the twenty-first century, we have considered similar dynamics of technological substitution of low-skilled workers in various fields of human activity. At the same time, we relied on the results of the research contained in the McKinsey Global Institute Analysis (2013), which were presented in Fig. 18.

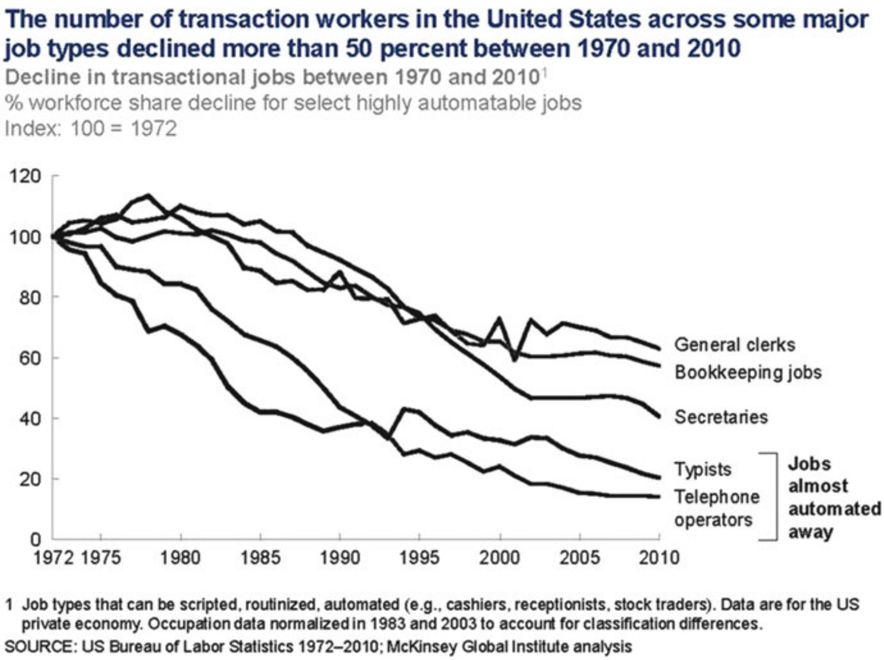


Fig. 18 Dynamics of reduction of certain professions by technical substitution of jobs

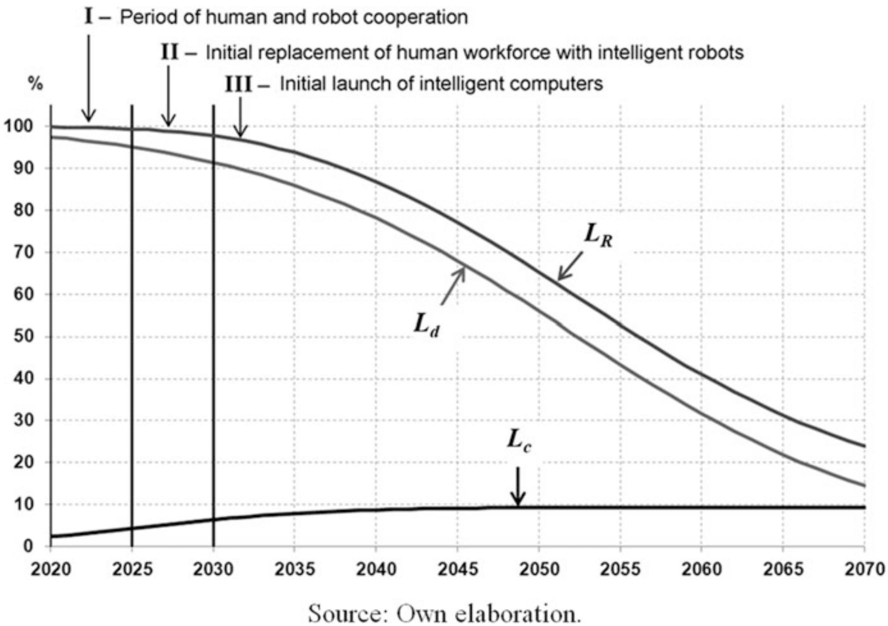


Fig. 19 The dynamics of technological substitution of middle class jobs and the creation of new jobs in the NBIC-sectors

As it can be observed in Fig. 19, the curves illustrate the technological displacement of different professions: they are all very similar to the logistic curve downward. Consequently, technological displacement of human labor follows the same logistic law, which was originally introduced by Fisher and Pry (1971), for replacing another technology.

When we calculated the technical substitution parameter that characterizes the substitution rate in the logistic model for all the curves previously presented in Fig. 19, we found that they were very close to one another. After computing the average value, we calculated and plotted the downward logistic curve (Fig. 19), which is characterized by the technological substitution of jobs of lower middle class on the assumption that it will be active after 2025. As can be seen, as early as 2040–2045, the middle class can be reduced by half. Of course, in parallel, new jobs will be created in NBIC sectors, which can be seen in Fig. 19.

References

Akayev, A. A., & Rudskoi, A. I. (2013). Analiz i prognoz vliianiia shestogo tekhnologicheskogo uklada na dinamiku mirovogo ekonomicheskogo razvitiia [Analysis and forecast of the impact of the sixth technological mode on the global economic development dynamics]. In *Mirovaia* *dinamika: zakonomernosti, tendentsii, perspektivy* [World dynamics: Patterns, trends and prospects] (pp. 142–166). Moscow: Librokom.

Akayev, A. A., & Rudskoi, A. I. (2015). Ob odnoj matematicheskoj modeli dlya prognoznyx raschetov sinergeticheskogo effekta nbic texnologij i ocenki ego vliyaniya na ekonomicheskij rost v pervoj polovine xxi veka [About a mathematical model for predictive calculations of synergies of NBIC technologies and assessment of their impact on economic growth in the first half of the 21st century]. *DAN [The proceedings of the Russian Academy of Sciences], 461*(4), 383–386.

Akayev, A. A., Pantin, V. I., & Ayvazov, A. E. (2009, December 10). *Analiz dinamiki dvizheniya mirovogo ekonomicheskogo krizisa na osnove teorii tsiklov. Doklad na pervom rossijskom ekonomicheskom kongresse* [Analysis of the global economic crisis dynamics on the basis of the theory of cycles. Report on the First Russian Economic Congress]. Lomonosov Moscow State University.

Arrow, K. J. (1962). The economic implications of learning by doing. *Review of Economic Studies, 24*, 155–173.

Bainbridge, W. S., & Roco, M. (2006). *Managing nano-bio-info-cogno innovations/converging technologies in society*. Dordrecht: Springer.

Brynjolfsson, E., & McAfee, A. (2013). How technology is destroying jobs. MIT Technology Review, July/August 2013. See also: http://www.technology-review.com/featuredstroy/515926/how-technology-is-destroying-jobs/

Chernov, A. I. (2006). Chto proiskhodit s sovremennym NTP? [What is happening to modern NTP?]. *EKO*, No. 4, 38–52.

Collins, R. (2015). *Srednij klass bez raboty: vykhody zakryvayutsya/Est li buduschchee u kapitalizma?* [The unemployed middle-class: The ways out are closed/Does capitalism have a future?] (pp. 61–112). Moscow: Publishing House of the Gaidar Institute.

Dubovsky, S. V. (1989). *Nestantsionarnii NTP v globalnom modelirovanii/Sistemnoe modelirovanie: modeli I metodi* [Unsteady STP global modeling/system modeling: Models and methods] (Vol. 20, pp. 20–50). Moscow: VNIISI.

EIU. (2003). http://www.microsoft.com/Rus/Government/analytics/EUI/part1.mspx

Fischer, J. S., & Pry, R. H. (1971). A simple substitution model of technological change. *Technological Forecasting and Social Change,* No. 3, 75–88.

Gordon, R. J. (2013). Zakonchen li ekonomicheskii rost? Shest' prepiatstvii dlia innovatsionnogo razvitiia (na primere SShA) [Is U.S. economic growth over? Faltering innovation confronts the six headwinds]. *Voprosy ekonomiki,* No. 4, 49–67.

Heinberg, R. (2011). *The end of growth: Adapting to our new economic reality*. Gabriola Island: New Society.

Hirooka, M. (2006). *Innovation dynamism and economic growth. A nonlinear perspective*. Cheltenham: Edward Elgar.

Jorgenson, D., & Motohashi, K. (2005). Information technology and the Japanese economy. *Journal of the Japanese and International Economies, 19*(4), 460–481.

Kondratiev, N. D. (2002). *Bol'shie tsikly kon'iunktury i teoriia predvideniia* [The great cycles of conjuncture and theory of prediction]. Moscow: Ekonomika.

Mankiw, G., Romer, D., & Weil, D. A. (1992). Contribution to the empirics of economic growth. *Quarterly Journal of Economics, 107*(2), 407–437.

McKinsey Global Institute Analysis. (2013). Disruptive technologies: Advances that will transform life, business, and global economy, May 2013.

Mensch, G. (1979). *Stalemate in technology—Innovations overcame the depression*. New York: Ballinger.

Roco, M. C. (2011). The long view of nanotechnology development: The National Nanotechnology Institute at 10 years. *Journal of Nanoparticle Research*, No. 12, 427–445.

Roco, M. C., & Bainbridge, W. S. (2003). Converging technologies for improving human perfomance, WTEC.

Rudskoi, A. I. (2007). *Nanotekhnologii v metallurgii* [Nanotechnology in metallurgy]. St. Petersburg: Nauka.
Schumpeter, J. A. (1939). *Business cycles*. New York: McGraw-Hill.
Schumpeter, I. (1982). *Teoriia ekonomicheskogo razvitiia* [The theory of economic development]. Moscow: Progress.
Sherk, J. (2010). Technology explains drop in manufacturing jobs. http://www.heritage.org/research/reports/2010/10/technology-explains-drop-in-manufacturing-jobs
Smith, G. M., & Lum, S. K. S. (2005). Annual industry accounts. Revised estimates for 2002–2004. In: *Survey of current business*, December 2005 (Vol. 85, No. 12). BEA.
Solow, R. (2000). *Growth theory*. Oxford: Oxford University Press.
Williams, L., & Adams, W. (2007). *Nanotechnology*. McGraw-Hill, New York.
Yakovets, I. V. (2004). *Epokhal'nye innovatsii XXI veka* [Epochal innovations of the 21st century]. Moscow: Ekonomika.

Structural and Technological Stalemate in Eurozone: If This Is the Reality, What We Can Expect?

Askar Akaev, Yuri Ichkitidze, and Valentin Sokolov

Abstract Even 6 years after the acute phase of the Great Recession 2007–2009, Euro area economy does not show strong growth, which signals severe structural and cyclical imbalances in the European economy. Empirical data reveals that both the Eurozone and the U.S.A. are facing since 2009 an unstable equilibrium that is prone to buckling under the influence of small internal or external price shocks. For detecting the bifurcation process, i.e. transition to a state of metamorphosis, we have specially developed models of nonlinear dynamics, which describe five possible stages of the economic system and, in particular, show that the Eurozone economy is entering into a very important stage of bifurcation and the consequences of which are fundamental to determine the nature of the future economic development of both European and global economy.

Keywords Bifurcation • Macroeconomic evolution • Economic dynamics • Structural shifts and cyclical changes

1 Introduction

The overcoming of the global economic crisis is currently the main focus of attention from politicians, economists, and scholars from around the world. In everyday awareness, the crisis is considered a sudden disease, which affected

A. Akaev
Research Institute of Complex Mathematical Systems on Behalf of I.R. Prigogine, Moscow State University, Leninskie Gory 1, Moscow, Russia
e-mail: askarakaev@mail.ru

Y. Ichkitidze (✉)
Department of Finance, National Research University Higher School of Economics, Kantemirovskaya Street 3, St-Petersburg, Russia
e-mail: ichkitidze@reflexivity.ru

V. Sokolov
Center for Fundamental Research of Economic Development in Russia, St-Petersburg State University of Economics, Marata str. 27, St-Petersburg, Russia
e-mail: svn@engec.ru

T. Devezas et al. (eds.), *Industry 4.0*, Studies on Entrepreneurship, Structural Change and Industrial Dynamics, DOI 10.1007/978-3-319-49604-7_3

world economy. But the current crisis is a phenomenon that is not a strange event, but rather expectable for the market economy. During the last 200 years there were five deep cyclic crises, with the Great Depression of the 1930s and today's Great Recession standing out. This property of the market economy was for the first time observed in a consistent way in the 1920s of the last century by Kondratiev (1926), who also described its basic properties. Kondratiev's theory had experienced further development in the works of Schumpeter (1939) and Mensch (1979). More recently, the work of Hirooka (2006) deserves also to be outlined.

It is well known that the structure of Kondratiev cycles is very simple. Each cycle comprises two phases or waves: upswing and downswing. The upswing is the period of long domination of good market conditions in world economy, when it develops dynamically, overcoming easily short-term small crises. Downswing phase is the period of bad market conditions when, despite short-time temporary upswings, dominate the depression and low business activity, as a result of which world economy experience unstable development with decline and deep crises. For interleaving cycles, the upswings are preceded by the periods of crisis and depression.

In the development course of each Kondratiev cycle, there were always strong structural disparities, both on the scale of national economies and on the scale of world capitalism, which in turn caused structural crises, and in doing so, made substantial impact on further economic development (Menshikov and Klimenko 1989). The structural crisis and lengthy depression that impacted the world economy since the 1970s, have stimulated the studies on the evolution of structural instability and its impact on economic development (Entov 1987). The experiences of the short world economic crises of 1974–1976 and 1980–1982 have revealed that structural crisis materially increased the destructive power of cyclic crises, facilitating accelerated growth of expensiveness and growing instability in world economic relations (Entov 1987).

In the German context, Mensch (2006) demonstrated that the structural instability, taking place in 1971–1974, was the cause of the deep crisis in the production and occupation in 1975–1976. The structural analysis developed by Mensch (2006), revealed that structural instability was also present in other developed countries in the 1971–1974 period. Thus, the structural crisis of the 1970s was coincident with the depression phase of the fourth Kondratiev cycle, causing deep economic crisis and paved the way for the following fifth Kondratiev cycle on the world scale. Thus, structural instability in separate parts of an economy goes hand-in-hand with structural susceptibility to the breakthrough of basis innovations, a phenomenon to which Mensch (2006) phrased a rule, according to which «innovations overcome depression». Hence, it follows, that the best time for the start of basis innovations is depression. That is why the same author coined the depression period as the 'time of structural transformation'.

In order to explain the structural crisis of the 1970s Mensch developed the 'theory of metamorphoses'. He took as the basis of the metamorphosis model a biquadratic function applied to the time series for the US and Germany (Wold and Mensch 1985). After a qualitative analysis, the same authors concluded that the start of structural crisis in world economy has implied serious changes in the economies of the European countries, which were much more profound than during usual cyclic downturns. Indeed, the introduction within the European Community of a single currency system in 1979 was an important step towards a unified economic space with the free movement of goods, services, capital, technology and persons; today the Euro area includes 19 states out of 27 EU member states. Since the start of the Great Recession in 2009 the euro area economy has not revealed strong growth, which signals structural and cyclical imbalances across the European economy. In this regard, we can confidently state that the European and American economy in 2009 entered the phase of bifurcation—a very unstable economic development. In addition, we have developed models of nonlinear dynamics, for detecting bifurcation process, that is, a transition to the state of metamorphosis.

2 Some Evidences on Bifurcation and Structural Disparities

For determining the bifurcation periods in dynamic series we used trend-stationary model, based on the summation of logistic curves. This model fully conforms to Mensch's theory and, as well as established models of Markov switching processes (Hamilton 1989, 1996, 2008; Timmermann 2000; Taylor 2008), fitting for the description of non-linear system dynamics. It allows us to single out long-term cycles in dynamic rows, which, within the framework of difference-stationary models (Box and Jenkins 1976) are taken as simple casual concurrence of circumstances, and to consider, in the description of structural changes, historical stages of the system, i.e. sequential variable time intervals, describing its evolution. The use of this model, from the viewpoint of econometrics, is correct, because hypothesis of a unit root with significance level 0.01 is rejected, with Fisher's criterion proving statistic validity of the model for a given number of parameters. The model, however, is not so simple for forecasting purposes: at bifurcation intervals, even insignificant noise can result in substantial shift of the trend parameters estimation. Nevertheless, in despite of not widely used in econometrics, this model has been used by some authors (Kapitsa 2000; Kwasnicki 2013; Modis 2013; Taagepera 2014), as well as us, considering the fact that its application in the analysis of dynamic series to be expedient.

Starting from 1960, the dynamics of logarithms of the USA and Eurozone GDPs in constant (2005) prices, is presented in Fig. 1. By making use of the model:

$$Y_t = \sum_{i=1}^{k} a_i \cdot L_{i,t} + d + \varepsilon_t$$

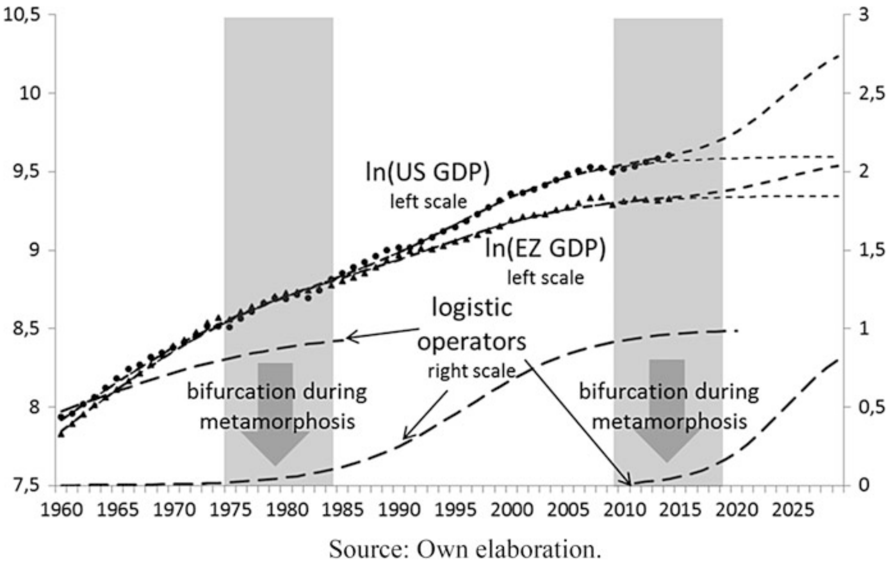


Fig. 1 The USA and Eurozone GDP dynamics, long-term cycles and bifurcation periods

where $L_{i,t} = \frac{1}{1+b_i \cdot e^{-\rho_i(t-t_0)}}$ is logistic operator, both indicators, i.e. logarithm of GDP of U.S.A. (Y_t^{US}) and logarithm of GDP of Eurozone (Y_t^{EU}), become trend-stationary at $k = 2$,[1] and could be described as:

$$\begin{cases} \widehat{Y}_t^{us} = 1.854 \cdot L_{1,t} + 0.7069 \cdot L_{2,t} + 7.038 \\ \widehat{Y}_t^{eu} = 1.9688 \cdot L_{1,t} + 0.4119 \cdot L_{2,t} + 7.16 \end{cases}$$

Corresponding logistic operators are equal to: $L_{1,t} = \frac{1}{1+1.2364 \cdot e^{-0.1052t}}$; and $L_{2,t} = \frac{1}{1+10.5846 \cdot e^{-0.1818(t-24)}}$; considering all model parameters determined by the least squares technique. Autocorrelation of the residuals after trend elimination is equal to 0.65 for the GDP of the U.S.A. and 0.75 for the GDP of the Eurozone. Performing Dickey-Fuller stationarity test allows us to reject hypothesis of a unit root for a 0.01 significance level. The crossing of logistic trends in this model (similar to Mensch's metamorphosis) corresponds to the period of system bifurcation, since the further trajectory of the movement is connected either with the slowing-down inertial development, or with addition of a new logistic trend. By observing Fig. 1, we can retain that the first crossing falls on the 1970–1980 period, while 2008 crisis made necessary conditions for the second crossing. In case the

[1]For $k = 1$ the observed autocorrelation coefficients are equal to 0.83 and 0.88, for the GDP of the U.S.A. and Eurozone, which together with the results obtained with the Dickey-Fuller test indicates the non-stationary type of the residual series.

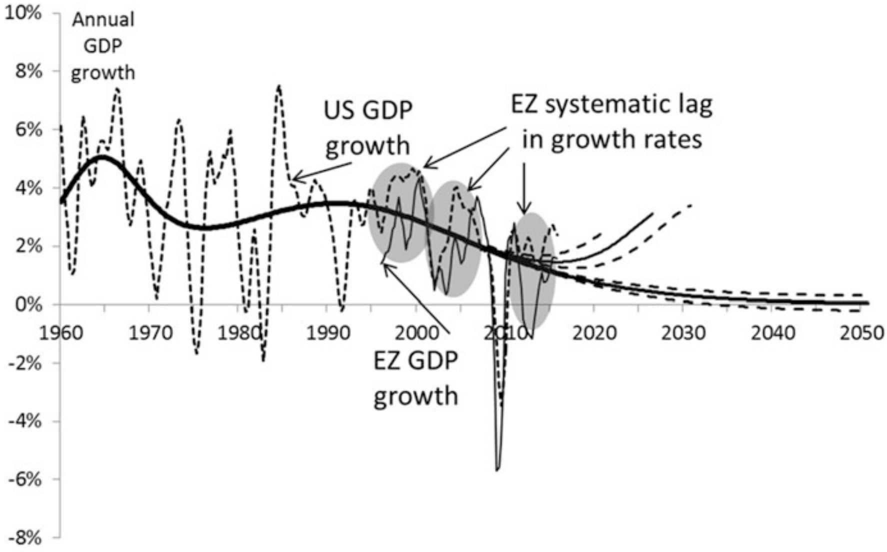
inertial development continues, annual average growth rate in 2020s will comprise from 0.5 to 1.5 %, while if new logistic trend appears, they increase to 2–3 %. The bifurcation period will last until further prospects clear up: either new trend becomes obvious, or the unforeseen delay on inertial trajectory will lead to considerable reduction of development potential.

This model is not just a quantitative description of dynamic series, since it also reveals evolutionary pathways of economic development. Obviously the defining impact on the long-term growth of economy is exerted by technological modes (Glazyev 1993; Perez 2004; Hirooka 2006). New breakthrough NBIC technologies (for more detail, see Hirooka 2006; Akaev and Rudskoy 2014) are going to become the source of economic growth in 2020–2040. The process of innovation, however, not only defines the state of economy, but it is also an endogenous variable. The result depends on the volume of investment and condition of institutional environment; it creates multiple equilibriums, which make given GDP trajectories not only scenarios, but dynamic attractors. When system loses inertia, considerably higher efforts have to be made to resume its development. Since even now, the growth rate of investments in the Eurozone lags behind the growth rate of GDP (the share of investments in GDP of the Eurozone during 2010–2015 has dropped from 24 to 20 %, while in the USA it grew from 13 to 17 % over the same period), the inertial attractor for the Eurozone looks quite realistic prospective. Undoubtedly, considerable impact on reduction of investments was exerted by the financial crisis of 2011. Recent studies (De Grauwe and Ji 2013; Panizza and Presbitero 2014) have shown that the economy of the Eurozone has passed into a condition of "bad" equilibrium as a result of financial system instability.

Annual growth rates of economies of the U.S.A. and the Eurozone in comparison with a long-term trajectory, and the specified system attractors are presented in Fig. 2. In general terms, we are able to observe that growth rate (2–3 % a year) gained by the U.S. economy after crisis of 2008 fits the goal of creation of a new long-term trajectory, it is only necessary to keep these rates of development as long as possible. The same does not apply to the Eurozone; present-day 1.5–2 % of annual growth are insufficient to change the existing inertial trend.

Besides, what stands out is the fact, that at the beginning of business cycle, the growth rate of Eurozone's economy lag considerably behind that of the U.S.A. (see Fig. 2), the correspondent increase occurs right before the recession stage. This is due to the role played by external demand in determining the economic development in the Eurozone; internal incentives (consumer or investment demand) are rather weak, therefore the growth export is the most important growth factor (Anaraki 2014; Fojtíková 2014).

Let us consider short-term trends in dynamics of real GDP of the U.S.A. and the Eurozone in more detail. Unlike long-term trends which are defined by technologies, investments into technologies and institutional environment, short-term trends are more dependent on specific economic incentives. The main sources of short-term impulses which impact macroeconomic dynamics over the period of the next 3–5 years are the fiscal and monetary policies, and domestic and external demand. These sources denote a weak influence on the long-term trend; more likely, they



Source: Own elaboration.

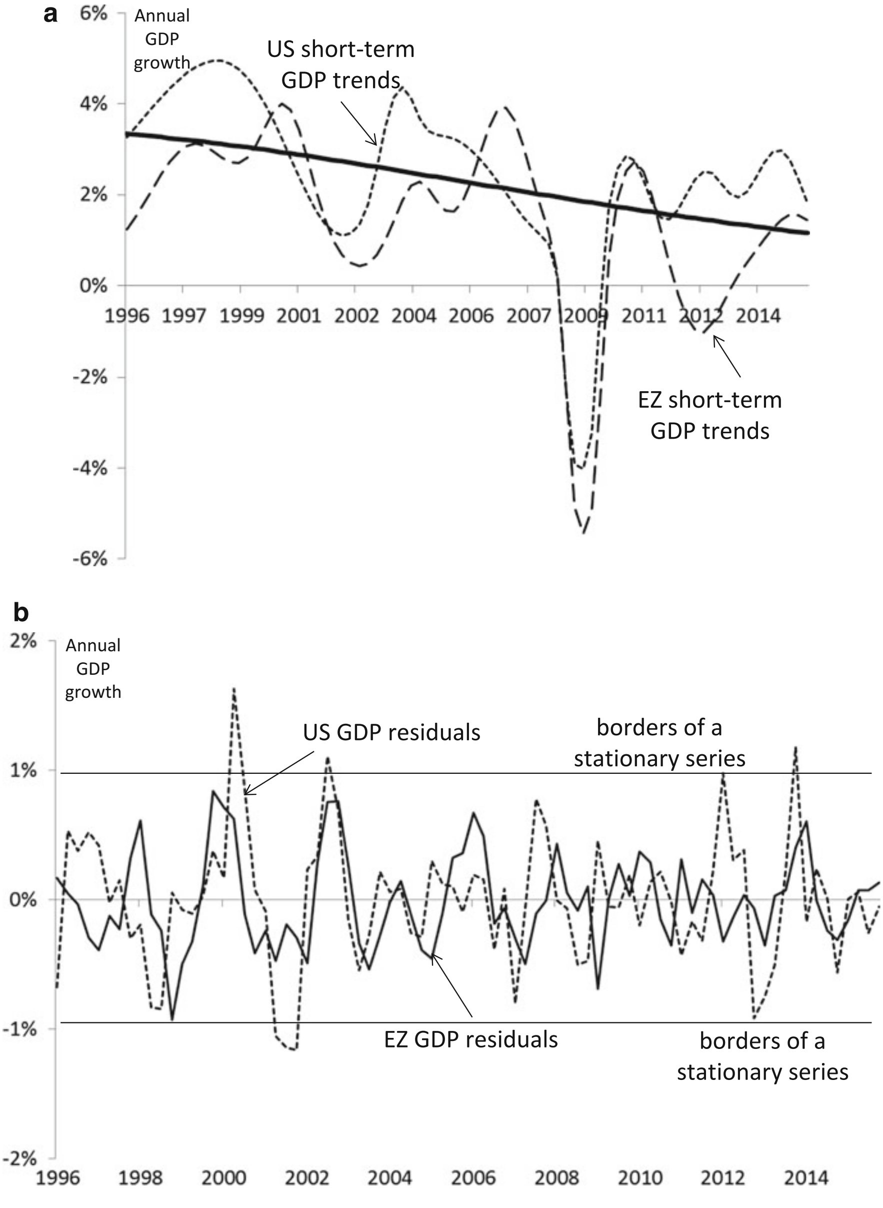
Fig. 2 Annual GDP growth rates, long-term cycles and bifurcation periods

choose (implement) implicit potential, which was created as a result of scientific and technical discoveries, but don't influence it.

However, active short-term trends are necessary for maintaining the long-term innovative potential in shape (for example, the long absence of demand for results of scientific studies from business, even with sufficient funding and the efficient institutional environment, causes a slowdown in scientific work intensity) therefore between short and long-term trends of economy there is a cross (mutual) impact. It is this coupling that also defines the particular mechanism of origin of new long-term trends during the bifurcation periods, excepting for innovative potential, intensive persistent short-term trends are necessary, otherwise, it would be very difficult to overcome the effect associated with the inertial attractor.

Short-term trends in GDP dynamics of the U.S.A. and the Eurozone, as well as the long-term trends, obtained by series summing of logistic curves are presented in Fig. 3a. For Eurozone's economy their average duration is almost twice as small as that for economy of the USA (3–5 years vs. 6–8 years), which signals the impulses' weakness.

The last observed trend (continuing now), began at the end of 2012 and reached the maximum 2 % annual growth rate (as compared with 3.5 % in the U.S.A.) and gradually began to decline. This is not enough for changing a long-term tendency, since the minimum required impulse as follows from Fig. 3b, is 1–1.5 % of additional growth during 1–2 years which corresponds to annual growth of 3–3.5 % of GDP. Through multipliers, it will impact not only the creation of a new short-term trajectory, but long-term trends as well. In general, from the analysis of Figs. 2 and 3 follows that now only the U.S.A. creates the active



Source: Own elaboration.

Fig. 3 (**a**) Short-term trends in the dynamics of GDP of the USA and Eurozone. (**b**) Residuals after de-trending in the dynamics of GDP of the USA and Eurozone

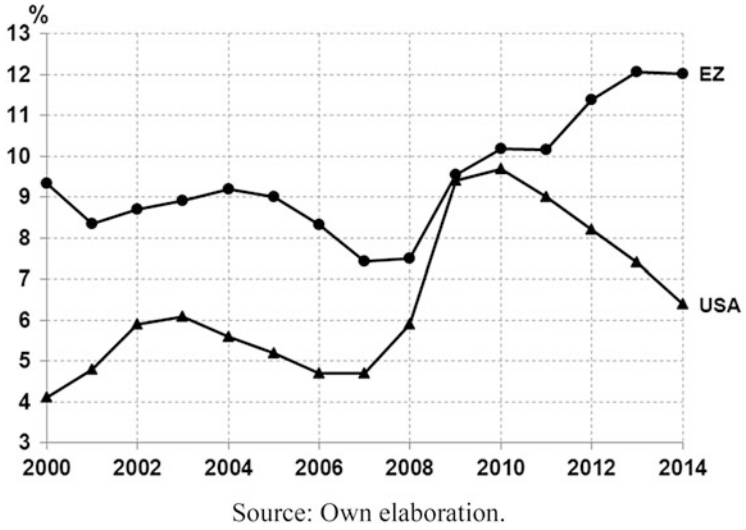


Fig. 4 Unemployment rate dynamics for the Eurozone and USA since 2000 (http://data.worldbank.org/indicator/SL.UEM.TOTL.ZS)

short-term trajectory of growth, required for transition to a new long-term trajectory (see also Akaev et al. 2014), while the Eurozone is the passive witness of formation of new technical mode. It explains the Eurozone's stalemate in Mensch's terms; under the conditions, the Eurozone isn't able to take actions for overcoming an inertial trajectory during the bifurcation period. The dynamics of unemployment rate confirms this finding (see Fig. 4); unlike the U.S.A., the Eurozone doesn't completely use the short-term potential of economic growth.

The above stated trends cannot be analyzed in a separate way from structural changes. The main justification for this is connected with the structural imbalances of the Eurozone, which are mainly observable in the development period, from 1995 till 1998. Afterwards, the period of growth in 2000–2008 has been accompanied by financial incentives, including monetary integration of the European Union and public procurement in certain countries, what allowed poorer countries of the Eurozone to obtain a considerable capital inflow (see Hein et al. 2011; Perez-Caldentey and Vernengo 2012). Under these conditions, the convergence of living standards of the Eurozone's countries was facilitated and prerequisites for equalization of business cycles were created.

Nevertheless, this development model wasn't able to ignore global imbalances in industry. After 2011, with the disruption of financial incentives, unemployment grew above 15 % in several EU countries, such as Greece, Portugal, Ireland, Spain and Italy (comprising 30 % of the population of the Eurozone), which apart from seriously restraining consumer demand, it has also contributed for restraining long-term innovative capacity. According to Industrial Structure Report (Pashev et al. 2015) such countries as Greece and Portugal are the most closed Eurozone economies; development of their industry depends to the highest extent on resources and work, their technological structure provides the smallest potential for innovative growth in the Eurozone. In Italy and Spain, low innovative sectors also command an essential part of the industry; in Ireland the limited high-tech cluster doesn't exert considerable impact on the development of remaining economic activities, besides, only a small part of exports is endogenously generated (Foster et al. 2013).

Under the conditions when the service sector is no longer capable to satisfy inquiries from labor market, it appears that without distinct industrial policy the Eurozone will not manage to eliminate disproportions among industries and to overcome divergence in the rates of economic development. Key system limitations are the decrease in competitiveness and investment flows (Zemanek et al. 2009). In countries such as Cyprus, Greece, Portugal and Italy, a share of investments into GDP after 2011 crisis, ended up at a critical low level of 10–16 %, contrasting with the situations of Germany, France, Belgium and Austria where it comprises 19–23 % of GDP. Against the background of reduction industry's share (from 1995 for 2014 it has decreased from 29 to 24 % of GDP), there is an essential reduction of high-tech export share (see Table 1) in Finland, Greece, Ireland of Italy, the Netherlands and Portugal. Altogether, these factors contribute for strengthening imbalances; and for overcoming it, the innovative capacity of the peripheral countries need to be reinforced, through industrial policy oriented to investment attraction (Botta 2014).

The link between structural transformations and cyclic processes previously displayed, characterizes the state of system metamorphosis; only evolutionary approach is capable of explaining how a form, which is a natural result of the previous development, prevents system from continuous development. In case of Eurozone economy, we faced the structural imbalances caused by process of integration of 1999–2009 which prevent system from creating a uniform business cycle. Now, under the conditions of uniform monetary policy, synchronization only occurs in central countries while in the peripheral countries the contingency of business cycles decreases (Lehwald 2013; Bekiros et al. 2015). It limits synergetic effect and creates clusters of stagnation; towards this mechanism it occurs a loss in the long term development potential.

3 Nonlinear Model of Bifurcation Process

For recognizing a bifurcation process, i.e. the transition to the state of metamorphism, Akaev (2007) has developed a nonlinear dynamic model, in which was derived a general differential equation describing the joint interaction of long-term economic growth and cyclic oscillations of business activity. The equation contains built-in accelerator of investments, which maintains non-damping oscillations in the economy. The same author proposed an algorithm of separation of the fast-oscillating business cycles and slowly-varying trend trajectory of long-term growth

Table 1 Change in high-tech exports (as % of industrial export)

Year	Austria	Belgium	Cyprus	Estonia	Euroarea	Finland	France
2000	14.60	10.73	1.67	29.93	20.02	27.36	24.60
2005	13.74	8.86	16.97	14.66	17.14	25.06	20.27
2013	13.72	11.36	7.15	10.50	15.85	7.21	25.84
Year	Germany	Greece	Ireland	Italy	Latvia	Luxembourg	Lithuania
2000	18.63	13.75	13.75	9.48	3.96	16.56	4.46
2005	17.42	10.58	10.58	7.98	5.31	11.86	6.15
2013	16.08	7.54	7.54	7.25	13.01	8.10	10.33
Year	Malta	Netherlands	Portugal	Slovenia	Slovak Republic	Spain	
2000	71.74	35.81	6.23	4.91	3.63	7.99	
2005	52.00	30.89	8.88	4.93	7.44	7.26	
2013	38.55	20.41	4.26	6.22	10.31	7.67	

Source: Worldbank

using the averaging method of Krylov-Bogolyubov-Mitropol'skii (Bogolyubov and Mitropol'skii 1962).

The nonlinear differential equation, describing cyclic oscillations of business activity around the trend growth curve can be written as follows (Akaev 2008):

$$\frac{d^2y}{dt^2} - \left[\sigma_0 - \frac{4}{3}æ\lambda v^3\left(\frac{dy}{dt}\right)^2\right]\frac{dy}{dt} + \omega_0^2\left[1 - \frac{s(1-s)}{æ}i\right]y = \phi(t), \qquad (1)$$

where $\sigma_0 = -[\lambda + æ - æ\lambda\nu - \lambda(1 - s)\beta/\gamma]$; $\omega_0{}^2 = \lambdaæ$; $y = Y - \bar{Y}$; $Y(t)$ is current production volume (current GDP level); $\bar{Y}(t)$ is the GDP level, corresponding to the trajectory of long-term growth; $\varphi(t)$ is quasi-periodic function, expressing the action of external driving force; v is accelerator's power; λ is the rate of the reaction lag of supply from demand; s is the saving ratio; æ is the rate of the reaction lag of actual capital investments from the moment of making decision on induced investments; β expresses the labor elasticity of output; γ is Okun's parameter; and i is interest rate.

The differential equation that describes the trajectory of economic growth can be written as follows (Akaev 2008):

$$\frac{d^2\bar{Y}}{dt^2} + \overline{\sigma}_0\frac{d\bar{Y}}{dt} + \overline{\omega}_0^2\bar{Y} = \lambda\left(\frac{d\bar{A}}{dt} + æ\bar{A}\right), \qquad (2)$$

where $\overline{\sigma}_0 = \lambda + æ - æ\lambda v$; $\overline{\omega}_0^2 = \lambda sæ$, $\bar{A}$ is the trend component of independent investments ($A(t) = \bar{A}(t) + \varphi(t)$).

It should be noted that, in deriving Eq. (1) concerning cyclic oscillations, it is taken into account cyclic unemployment, appearing in the periods of decline, which allows us to consider real economic system with underemployment, while most of existing models consider only full employment. It is common knowledge that the variations of unemployment rate are related to the oscillations in real output by Okun's law:

$$\frac{Y_F - Y}{Y_F} = \gamma(u - u^*),$$

where Y_F is the potential output at full employment; Y is the real output in the presence of cyclic unemployment; u^* is natural unemployment rate, corresponding to full employment; u is real unemployment rate; and γ is Okun's parameter.

The solution of Eq. (1) defines not only oscillations of output but also, at the same time, oscillations in employment. Since cyclic unemployment has negative economic and social consequences, it is obvious that the economic policy should be targeted at the limitation of the rate of cyclic unemployment, and, therefore, oscillations in output. Thus, decreasing the oscillation amplitude of total output we simultaneously decrease the rate of cyclic unemployment.

Since the level of output has impact on the volume of capital required for its production, the fluctuations of output influence the volume of required capital and, therefore attract net investments. Thus, if the output flow changes, available required capital changes, and positive or negative induced investment takes place, being then derivatives of the total available capital (basic and operating):

$$\dot{K} = I - \mu K$$

where μ is retirement rate; I is actual capital investment; K is available capital.

The above equation was also taken into account in derivation of Eqs. (1) and (2).

Let us now consider the system dynamics expressed by Eq. (1). If we introduce the variable $x = \frac{dy}{dt}$, Eq. (1) is reduced to a system of two first order differential equations:

$$\dot{y} = x; \quad \dot{x} = \left(\sigma_0 - \frac{4}{3}æ\lambda v^3 x^2\right)x - \omega_0^2\left[1 - \frac{s(1-s)}{æ}i\right]y + \phi(t) \tag{3}$$

For our purpose, it is sufficient to study the phase portrait of a given dynamic system by qualitative analysis, therefore we assume $\varphi(t)=0$. Concerning the effect of external forces, this can be deeply studied through qualitative methods, for example, averaging after Krylov-Bogolyubov-Mitropolsky (Bogolyubov and Mitropol'skii 1962). Since $s(1 - s)i/æ << 1$, then this can be neglected in additional analyses. Thus, the following autonomous dynamic system is obtained, which can be described by standard differential equations as follows:

$$\dot{y} = x; \quad \dot{x} = \left(\sigma_0 - \frac{4}{3}æ\lambda v^3 x^2\right)x - \omega_0^2 y \text{ or } \dot{x} = P(x,y); \dot{y} = Q(x,y). \tag{4}$$

The stationary state of a dynamic system as expressed in previous Eq. (4) ($\dot{x} = \dot{y} = 0$) takes place in the origin of phase plane coordinates: $x = 0, y = 0$. Thus, zero point ($x = 0, y = 0$) is a special point or the equilibrium point of the dynamic system.

For studying the behavior and stability of equilibrium point, the method of stability analysis by the first iteration (Kuznetsov et al. 2005) is used. Thus, the Jacobian matrix corresponding to equations of system (4), is linearized in the vicinity of zero point and is given by the following:

$$J_0 = \begin{vmatrix} P'_x(0,0) & P'_y(0,0) \\ Q'_x(0,0) & Q'_y(0,0) \end{vmatrix} = \begin{vmatrix} \sigma_0 & -\omega_0^2 \\ 1 & 0 \end{vmatrix} \tag{5}$$

From here it follows: $J_0 = {\omega_0}^2$; $S_0 = \sigma_0$

where S_0 is the trace of Jacobian matrix. The characteristic equation of the linearized system is written as follows: $p^2 - S_0p + J_0$. The roots of the characteristic equation are as follows:

$$p_{1,2} = \frac{1}{2}\left(S_0 \pm \sqrt{S_0^2 - 4J_0}\right) = \frac{1}{2}\left(\sigma_0 \pm \sqrt{\sigma_0^2 - 4\omega_0^2}\right) \tag{6}$$

Thus, the behavior of phase trajectories and, therefore, pattern of phase portrait in the vicinity of the critical point, is defined by the roots of characteristic equation $p_{1,2}$ that are expressed via trace S_0 and determinant of Jacobian matrix J_0. Taking into account the expressions for σ_0 and ω_0^2 (1), we obtain from (6), the following expression:

$$p_{1,2} = -\frac{1}{2}\left[\lambda + æ - æ\lambda\nu - \lambda(1-s)\frac{\beta}{\gamma}\right] \pm \frac{1}{2}\sqrt{\left[\lambda + æ - æ\lambda\nu - \lambda(1-s)\frac{\beta}{\gamma}\right]^2 - 4\lambda æ}. \tag{7}$$

The classification of the roots of the characteristic equation (7) and the type of corresponding equilibrium state are presented below in Table 2.

From observing Table 2, we are able to retain that for low values of accelerator power ν (to be exact $\nu \leq 1.05$), the equilibrium state of the system is quite stable, with asymptotic stability. This means that if the system is brought out of the equilibrium state by external shocks, it will, in the course of time, asymptotically converge to an equilibrium state, i.e. no development. Hence, the system will keep the original equilibrium state. Naturally, the rate of economy growth is unchanged.

At the value $\nu = 1.05$, i.e. in point $\sigma_0 = 0$, the roots of characteristic equation will become purely imaginary ($p_{1,2} = \pm\sqrt{\lambda æ}i$) and the critical point will be of central type. In this case the system becomes coarse and, consequently, structurally unstable and very sensitive to small fluctuations. Since at $\sigma_0 > 0$ ($\nu > 1.05$) equilibrium state loses stability, there is a clear indication that a new stable limiting cycle is created, i.e. $\sigma_0 = 0$ is a bifurcation point. To prove this suggestion, we will use the theorem enunciated below (Arrowsmith and Place 1986).

Theorem: If point O *(0,0)* is a special point of the dynamic system (4), at all values of parameter σ_0, roots $p_1(\sigma_0)$ and $p_2(\sigma_0)$ of the characteristic equation of corresponding linearized system are pure imaginary at $\sigma_0 = \sigma_0^*$, Re $p_1(\sigma_0)$= Re $p_2(\sigma_0)$, $\frac{d}{d\sigma_0}(\mathrm{Re}\,p_1(\sigma_0))_{|\sigma_0=\sigma_0^*} > 0$, and point O *(0,0)* is asymptotically stable at $\sigma_0 = \sigma_0^*$. Thus the value $\sigma_0 = \sigma_0^*$ is the point of bifurcation for the system (4). One can easily see that all conditions of the theorem hold true for the system (4) at point $\sigma_0 = 0$.

Thus, by the transition of parameter σ_0 over zero critical value from left to right occurs a bifurcation, called Andronov-Hopf bifurcation (Kuznetsov et al. 2005): the stable focus becomes unstable and generates a limiting cycle of small amplitude ~

Table 2 Roots classification and equilibrium state

Variation interval of accelerator's power	Roots of characteristic equation	Equilibrium state
$0 < v \leq \frac{1}{\lambda æ}\left[\lambda + æ - 2\sqrt{\lambda æ} - \lambda(1-s)\frac{\beta}{\gamma}\right]$ $[0 < v \leq 0.05]$	Real and negative ($\sigma_0 < 0$)	Stable node
$\frac{1}{\lambda æ}\left[\lambda + æ - 2\sqrt{\lambda æ} - \lambda(1-s)\frac{\beta}{\gamma}\right] < v < \frac{1}{\lambda æ}\left[\lambda + æ - \lambda(1-s)\frac{\beta}{\gamma}\right]$ $[0.05 < v < 1.05]$	Complex Negative real part ($\sigma_0 < 0$)	Stable focus
$v = \frac{1}{\lambda æ}\left[\lambda + æ - \lambda(1-s)\frac{\beta}{\gamma}\right]$ $[v = 1.05]$	Pure imaginary ($\sigma_0 = 0$)	Bifurcation point
$\frac{1}{\lambda æ}\left[\lambda + æ - \lambda(1-s)\frac{\beta}{\gamma}\right] < v < \frac{1}{\lambda æ}\left[\lambda + æ + 2\sqrt{\lambda æ} - \lambda(1-s)\frac{\beta}{\gamma}\right]$ $[1.05 < v < 2.05]$	Complex Positive real part ($\sigma_0 > 0$)	Unstable focus
$v \geq \frac{1}{\lambda æ}\left[\lambda + æ + 2\sqrt{\lambda æ} - \lambda(1-s)\frac{\beta}{\gamma}\right]$ $[v \geq 2.05]$	Real and positive ($\sigma_0 > 0$)	Unstable node

Source: Own elaboration

$\sqrt{\sigma_0}$. Such loss of stability, it is called the 'soft loss of stability'. Hence, the lost of stability turns into a cycle and the equilibrium becomes unstable. Structural stability is lost in a similar way.

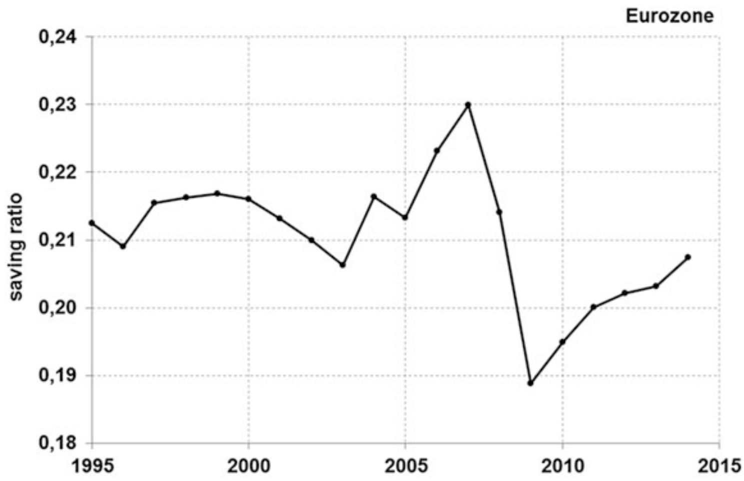
Then, at power $\nu > 1.05$ any small fluctuations in the system cease to damp and become growing, until they stabilize at a certain level ~ $\sqrt{\sigma_0}$, which is due to nonlinear effects (Kuznetsov et al. 2005). The change to limiting cycle is preceded by a transition process. At σ_0<<0 the oscillations are quasi-harmonic. At σ_0 ~ 1 (ν ~ 1.3) the oscillations will be seriously nonlinear, i.e. non-harmonic. Thus, an economic system with nonlinear accelerator (4) is a classical self-oscillating system, where positive feedback plays the part of nonlinear accelerator, and accelerator power ν is gain. If the gain is sufficiently large ($\nu > 1.05$), self-supporting oscillation process occurs in the system, whose main parameters are determined by the system parameters and are independent of start conditions.

We estimated numerical values of three key parameters (λ, æ, ν) of nonlinear model, describing cyclic oscillations of business activity for Eurozone countries:

$\lambda = 5$ is the rate of the reaction lag of supply from demand;
æ $= 0.9$ is the rate of the reaction lag of actual capital investments from the moment of making decision on induced investments;
$\nu = 1.1$ is the accelerator's power.

Figure 5 shows the results for the parameter s, the saving ratio calculated for the Eurozone in our nonlinear model dynamics as the ratio of gross national saving to GDP (Gross domestic product at current market prices).[2]

[2]Data source—http://ec.europa.eu/economy_finance/db_indicators/ameco/zipped_en.htm



Source: Own elaboration.

Fig. 5 The dynamics of saving ratio for the Eurozone in the nonlinear model of bifurcation process

The value of parameter β (output elasticity in labor) was determined via estimation of production function parameters

$$\widetilde{Y} = \chi \cdot K^{\alpha} \cdot L^{1-\alpha}, \quad \beta = 1 - \alpha$$

where Y is the gross domestic product at 2005 market prices[3];

K—net capital stock at 2005 prices[4];
L—employment[5];
$\chi = 0.14$; $\alpha = 0.51$; $\beta = 0.49$.

The behavior of the GDP dynamics for the Eurozone is shown in Fig. 6 below.

For the estimation of parameter γ (Okun's parameter) we used an approximate formula, valid under the assumption that the natural level of unemployment remains stable in time:

$$\frac{\Delta Y}{Y} \approx k - \gamma \cdot \Delta u$$

where Y corresponds to the gross domestic product at 2005 market prices[6];

[3]Data source—http://ec.europa.eu/economy_finance/db_indicators/ameco/zipped_en.htm
[4]Data source—http://ec.europa.eu/economy_finance/db_indicators/ameco/zipped_en.htm
[5]Data source—http://www.imf.org/external/pubs/ft/weo/2014/02/weodata/weoselgr.aspx
[6]Data source—http://ec.europa.eu/economy_finance/db_indicators/ameco/zipped_en.htm

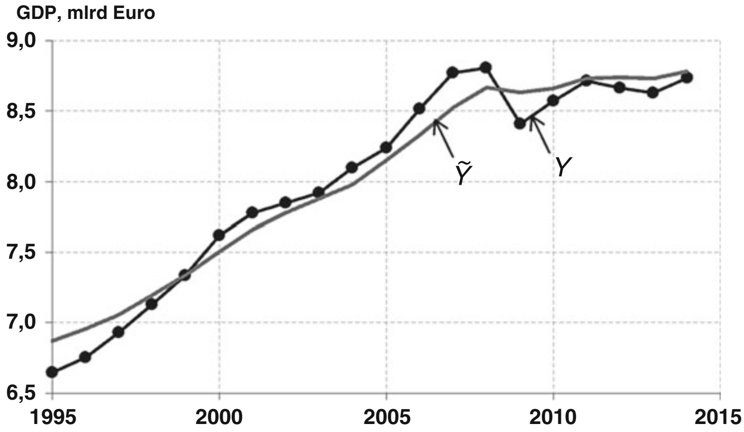


Fig. 6 GDP dynamics of the Eurozone in the nonlinear model of bifurcation process

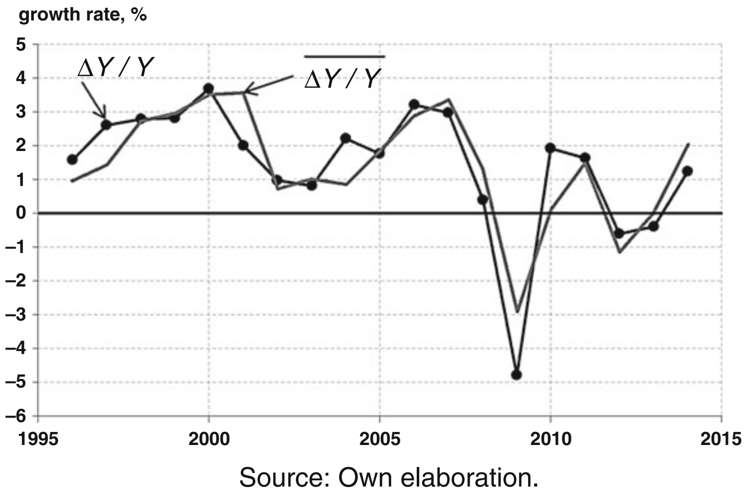


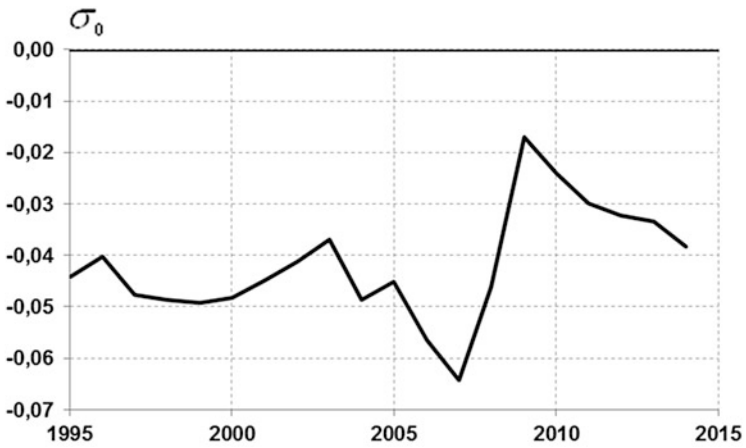
Fig. 7 GDP growth rate in the Eurozone in the nonlinear model of bifurcation process

ΔY is the variation of GDP;
Δu is the change of real unemployment level[7];
$k = 1.46$; $\gamma = 2.13$.

Figure 7 presented below exhibits the result for the GDP growth rate obtained with the nonlinear model under these assumptions.

Thus, for final calculation of σ_0 we can assume the following values of the parameters:

[7]Data source—http://data.worldbank.org/indicator/SL.UEM.TOTL.ZS



Source: Own elaboration.

Fig. 8 Dynamics of indicator σ_0

$\lambda = 5$; $æ = 0.9$; $\nu = 1.1$; $s = 0.21$; $\beta = 0.49$; $\gamma = 2.13$

The calculation after the expression (1) yields the result $\sigma_0 = -0.041$, i.e. virtually equal zero.

Considering parameter s in the proposed dynamics, one can trace the variation of σ_0 over time as shown in Fig. 8.

As it can be observed in Fig. 8, the maximum nearness to zero level occurred in 2009, which means that it was over this year that the bifurcation occurred; in the following years the value of σ_0 diminished slightly. This means that until now the economy of the Eurozone is still in recession.

4 Eurozone Is in the State of Metamorphism: Some Statistics Evidence

Today the economy of Eurozone is very vulnerable and is prone to stability losses under small demand shocks. Cardinal structural reforms were not carried out and most of the Eurozone's countries are burdened with excessive debts. Moreover, debt structure of the European economy remained unchanged: debts are still being repaid at the expense of new loans. Unemployment still stays at high level: the average for the European zone is 11.4 %, and notably, among young people it reaches 24 %. Internal demand in the Eurozone recovers extremely slowly due to the policy of austerities implemented by the European Union. Economic growth in Eurozone in 2014 was very weak, about only 0.8 %. The comparison of primary economic indicators (basis GDP growth rate, unemployment rate, rate of

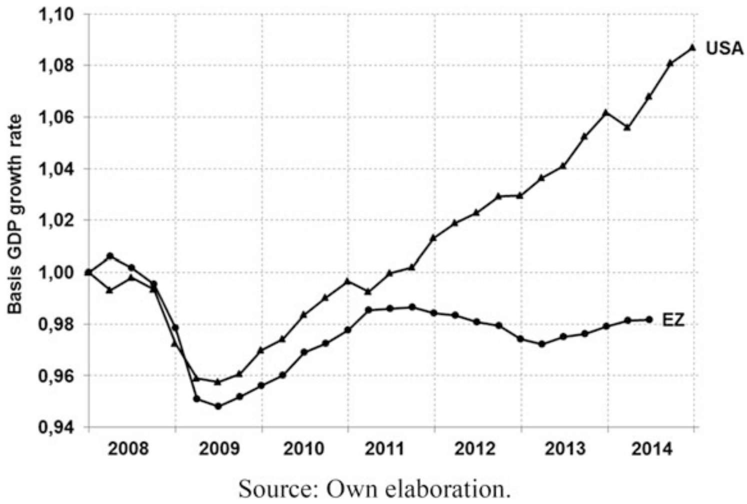


Fig. 9 Basis GDP growth rate for the Eurozone and USA since 2008 (In comparable prices. Calculated after http://www.economagic.com/)

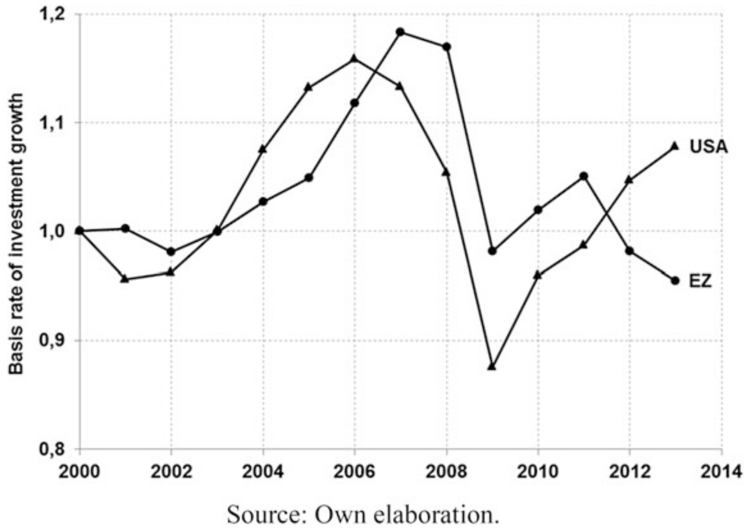


Fig. 10 Basis rate of investment growth for the Eurozone and USA since 2000 (In comparable prices. Computed from: http://data.worldbank.org/indicator/SL.UEM.TOTL.ZS)

investment growth and multifactor productivity dynamics) between the Eurozone and U.S.A. starting from 2008 can be observed in Figs. 9, 10 and 11.

Unfolding the information provided in the previous Figs. 9, 10 and 11 suggests that the Eurozone's economy is facing an unstable equilibrium. This means that the Eurozone's economy may plunge into repeated recession. As one knows from

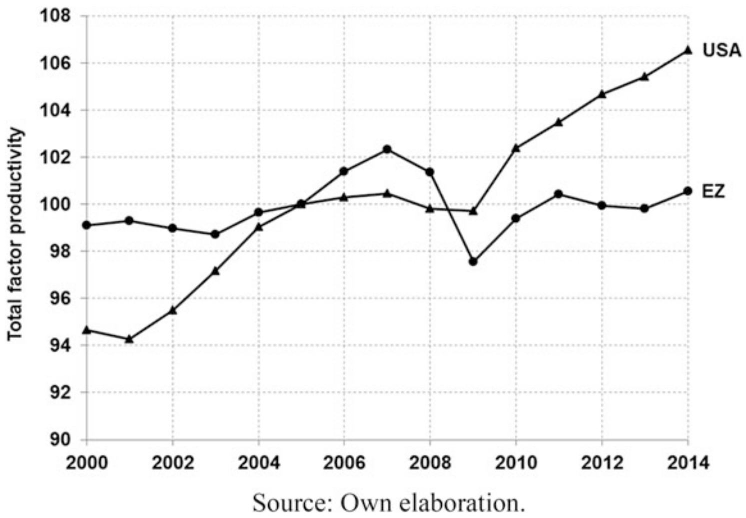


Fig. 11 Total factor productivity for the Eurozone and USA since 2000 (2005 = 100) (http://ec.europa.eu/economy_finance/db_indicators/ameco/zipped_en.htm)

nonlinear dynamics theory, small fluctuations of external action (for example, price shocks) in the neighborhood of a bifurcation point are strongly amplified. For this reason, a small shock impact will be sufficient for facing positive and negative development scenarios, in subsequent terms.

5 Conclusion

We are witnessing the stage of latent metamorphosis in the Eurozone economy, which is characterized by the appearing of signs of significant system deformation and their growth in time. Some Eurozone's countries have experienced extreme situations, creating social conflicts in the European Union. Some countries, such as Greece, are trying to turn the tide in their favor, creating conditions for the transition to the following stages of metamorphism. All this may result in radical changes in existing structure and morphology of the Eurozone's economy.

Starting from 2009, the year of bifurcation, and until now the economy of the Eurozone is still in recession. On the basis of the nonlinear model of bifurcation process, we can consider the following for the Eurozone's economy:

– the economy of the Eurozone is experiencing a system structural crisis similar to that of the 1970s. However, it is characterized by an extremely high level of unemployment, while in the 1970s it was relatively low;
– the European Union was rescued by a powerful social security system, which has considerably smoothed the negative effect of large-scale unemployment;

– the crisis revealed serious conflict of interests between different groups of the Eurozone countries, which can turn weaker the European Union;
– the Andronov-Hopf bifurcation has already occurred in 2009, which is the soft form of stability loss. This fact means that the economy of the Eurozone is now highly susceptible to structural changes, and consequently highly sensitive to small fluctuations of the internal and external environments;
– if correct economic decisions are made, the European Union has an exceptional chance to form a well-balanced and harmonious economic structure for the near future, in order to face the challenges of the twenty-first century. For addressing all the structural changes, the different interests of all countries in the core of the Eurozone should be taken into account;
– in this kind of situation, when structural crisis and chaos occur, it is impossible to return to the former equilibrium state of the economy, and any even insignificant shift will push the system in a new direction of development. The entrenching new stage of metamorphosis inevitably entails radical changes of the economic system. It is worth to note that a metamorphosis transformation should go in the positive way, since the changes occurring in the system during this phase are irreversible. Therefore, nations can't take bad decisions like 'let matters drift'. The transformation of any big economic system undergoing a metamorphosis phase is decisively related to the social-economic dynamics of the leading countries (these are Germany and France, in the Eurozone's case).

Thus, in the course of the actual fast metamorphosis changes the nations of the European Union need to maintain the structural change processes under their permanent control. The European Union has now the chance to create the most attractive model of capitalism that may design the contours of the next sixth Kondratiev cycle of economic development (2018–2050). Pragmatic steps in decreasing unemployment rate, meeting the daily people's needs and decreasing their sufferings, should become a priority, and only then, the transition to long-term goals in political and economic transformation should follow. At this point one should act uncompromisingly. The EU will need expansionary fiscal policy, at a cost of 300 billion Euros annually to ensure economic growth at 1.5–2 % per year. The solution of this problem will be simplified if, for instance, the trade with Russia is reestablished and geopolitical tensions relieved, which conforms to the vital interests of a new Europe. If this does not happen, it is fairly reasonable to expect that other developed countries will carry out planned policies for implementation of innovations—technological breakthroughs, based on clusters of converging technologies involving nanotechnology, biotechnology, information technology, and cognitive science (NBIC-technologies), creating the new technological platform and new structure of world economy. In simultaneous terms, vanguard developing countries (such as China and India) will start to accelerate industrialization of their economies through the full-scale borrowing of general purpose technologies (microelectronics, computer and information-communication technologies, energy—and resource-saving technologies, etc.).

Acknowledgments The research has been supported by the Grant of the Russian Science Foundation. Research project No. 14-28-00065: "Structural and cyclical paradigm of economic and technological renovation of macro-systems (World and Russia in the first half of the twenty-first century)".

References

Akaev, A. A. (2007). Derivation of the general macroeconomics dynamics equation describing the joint interaction of long-term growth and business cycles. *Doklady Mathematics, 76*(3), 879–881.

Akaev, A. A. (2008). Influence of business cycles on long-term economic growth. *Doklady Mathematics, 78*(1), 1–5.

Akaev, A. A., & Rudskoy, A. I. (2014). Synergetic effect of NBIC—Technologies and economic growth in the first half of the XXI century [Sinergeticheskij effekt NBIC—tekhnologij i mirovoj economicheskij rost v pervoj polovine XXI veka (in Russian)]. *Ekonomicheskaja politika, 2*, 25–46.

Anaraki, N. K. (2014). Effects of Euro devaluation on Eurozone exports. *International Journal of Economics and Finance, 6*(2), 19–24.

Arrowsmith, D., & Place, C. (1986). *Obyknovennie differenzial'nie uravnenija. Kachestvennja teorija s prilozhenijami* (in Russian) [Ordinary differential equations. A qualitative approach with applications]. Moscow: MIR.

Bekiros, S., Nguyen, D. K., Uddin, G. S., & Sjö, B. (2015). Business cycle (de)synchronization in the aftermath of the global financial crisis: Implications for the Euro area. *Studies in Nonlinear Dynamics & Econometrics, 19*(5), 609–624.

Bogolyubov, N. N., & Mitropol'skii, Y. A. (1962). *Asymptotic methods in the theory on nonlinear oscillations*. New York: Gordon & Breach.

Botta, A. (2014). Structural asymmetries at the roots of the eurozone crisis: What's new for industrial policy in the EU? *PSL Quarterly Review, 67*(269), 169–216.

Box, G. E. P., & Jenkins, G. M. (1976). *Time series analysis: Forecasting and control*. San Francisco: Holdenday.

De Grauwe, P., & Ji, Y. (2013). Self-fulfilling crises in the Eurozone: An empirical test. *Journal of International Money and Finance, 34*, 15–36.

Entov, R. (1987). Main way of interaction of the structural and cyclic crisis [Osnovnie formi vzaimodejstvija strukturnikh i ziklicheskikh krizisov (in Russian)]. *Mirovaja economika i mezhdunarodnije otnoshenija, 11*, 109–114.

Fojtíková, L. (2014). Performance and growth of the Eurozone export. *Procedia Economics and Finance, 12*, 154–163.

Foster, N., Stehrer, R., & Timmer, M. (2013). *International fragmentation of production, trade and growth: Impacts and prospects for EU member states*. (European Economy—Economic Papers, No. 484). Brussels: European Commission.

Glazyev, S. Y. (1993). *The theory of long-term technical and economic development* [Teoriya dolgosrochnogo thechniko-economicheskogo razvitiya (in Russian)]. Moscow: VlaDar.

Hamilton, J. B. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica, 57*(2), 357–384.

Hamilton, J. B. (1996). Specification testing in Markov-switching time-series models. *Journal of Econometrics, 70*, 127–157.

Hamilton, J. B. (2008). Regime-switching models. In *New Palgrave dictionary of economics* (2nd ed.). Basingstoke: Palgrave McMillan.

Hein, E., Truger, A., & van Treeck, T. (2011). *The European financial and economic crisis: Alternative solutions from a (Post-) Keynesian perspective* (IMK Working Paper, No. 9). Düsseldorf: Hans Böckler Stiftung.

Hirooka, M. (2006). Innovation dynamism and economic growth. In *A nonlinear perspective*. Cheltenham: Edward Elgar.

Kapitsa, S. P. (2000). The model of growth and economic development of the Earth's human population [Model rosta naselenia Zemli i ekonomicheskogo razvitiya chelovechestva (in - Russian)]. *Voprosy Economiki*, No. 12, 85–106.

Kondratiev, N. D. (1926). Die Langen Wellen der Konjuntur, Arch. Sozialwiss. *Sozialpolitik, 56* (3), 573–609.

Kuznetsov, A. P., Kuznetsov, S. P., & Ryiskin, N. M. (2005). *Nonlinear oscillations* [Nelinej'nie kolebanija (in Russian)]. Moscow: Fizmatlit.

Kwasnicki, W. (2013). Logistic growth of the global economy and competitiveness of nations. *Technological Forecasting and Social Change, 80*(1), 50–76.

Lehwald, S. (2013). Has the Euro changed business cycle synchronization? Evidence from the core and the periphery. *Empirica, 40*(4), 655–684.

Mensch, G. (1979). *Stalemate in technology—Innovations overcome the depression*. New York: Ballinger.

Mensch, G. (2006). If this long wave steeps-up and breaks: What then? In T. C. Devezas (Ed.), *Kondratieff waves. Warfare and world security* (pp. 80–90). Amsterdam: IOS Press.

Menshikov, S. M., & Klimenko, L. A. (1989). *Long waves in economics. When society casts the skin* [Dlinnie volni v ekonomike. Kogda obshestvo menyaet kozhu (in Russian)]. Moscow: Mezhdunarodnie otnoshenija.

Modis, T. (2013). Long-term GDP forecasts and the prospects for growth. *Technological Forecasting and Social Change, 80*(8), 1557–1562.

Panizza, U., & Presbitero, A. F. (2014). Public debt and economic growth: Is there a causal effect? *Journal of Macroeconomics, 41*, 21–44.

Pashev, K., Casini, P., Kay, N., & Pantea, S. (2015). EU structural change 2015. In *European Commission*. Luxembourg: Publications Office of the European Union.

Perez, C. (2004). Finance and technical change: A long-term view. In H. Hanusch & A. Pyka (Eds.), *The Elgar companion to neo-Schumpeterian economics*. Cheltenham: Edward Elgar.

Perez-Caldentey, E., & Vernengo, M. (2012). T*he Euro imbalances and financial deregulation: A post-Keynesian interpretation of the European debt crisis* (Working Paper, No. 702). Annandale-on-Hudson, NY: Levy Economics Institute of Bard College.

Schumpeter, J. (1939). *Business cycles: A theoretical, historical, and statistical analysis of the capitalist process*. London: McGraw-Hill.

Taagepera, R. (2014). A world population growth model: Interaction with Earth's carrying capacity and technology in limited space. *Technological Forecasting and Social Change, 82*, 34–41.

Taylor, S. J. (2008). *Modelling financial time series* (2nd ed.). Hackensack, NJ: World Scientific.

Timmermann, A. (2000). Moments of Markov switching models. *Journal of Econometrics, 96*, 75–111.

Wold, H., & Mensch, G. O. (1985). *The BIEQ-PLS model of a Schumpeterian wave* (Working Paper WSOM WP 85–017). Cleveland, OH: Weatherhead School of Management, Case Western Reserve University.

Zemanek, H., Belke, A., & Schnabl, G. (2009). Current account balances and structural adjustment in the euro area. *International Economics and Economic, 7*(1), 83–127.

On the Asymmetry of Economic Cycles

Valentin Sokolov, Tessaleno Devezas, and Svetlana Rumyantseva

Abstract This paper offers a brief review of economic cycle theories, pointing out the key parameters that describe their dynamic behavior, and proposes a model of asymmetric economic cycles. The model is based on the assumption that economic cycles are non-equilibrium processes characteristic of complex open-ended systems and uses two different see-saw-like sine functions in its construction: an additive version and a multiplicative version. An empirical verification of the model is presented for the GDP cyclic dynamics for several countries and the asymmetries observed on a global scale. For some countries we have also examined the asymmetric behavior for other economic variables, like energy production and unemployment. Conclusions are drawn about the suitability of the model to forecast economic cyclic dynamics.

Keywords Economic cycles • Non-equilibrium processes • Asymmetry of economic cycles • Amplitude of economic cycles • Periodicity of economic cycles • Complex systems

1 Review of Economic Cyclic Dynamics Theories

Throughout the long history of studying economic cycles, two key questions attracted researchers: firstly, what are the causes of cyclic oscillations? And, secondly, which are the regularities in their formation?

Fruitful attempts to determine the answer for the first question led to development of fundamental theories describing the cyclic mechanism of various factors

V. Sokolov (✉)
Center for Fundamental Research of Economic Development in Russia, St-Petersburg State University of Economics, Marata str. 27, St-Petersburg, Russia
e-mail: svn@engec.ru

T. Devezas
Department of Eletromechanics, University of Beira Interior, Ladeira do Lamego, Covilhã, Portugal
e-mail: tessalen@ubi.pt

S. Rumyantseva
Department of Economic Theory, St. Petersburg State University, Sankt-Peterburg, Russia
e-mail: svetlana-rumyantseva@mail.ru

T. Devezas et al. (eds.), *Industry 4.0*, Studies on Entrepreneurship, Structural Change and Industrial Dynamics, DOI 10.1007/978-3-319-49604-7_4

affecting the economic dynamics. Among the basic theories we can find: the innovation and technology theory (Mensch 1979; Kleinknecht 1987; Van Duijn 1983; Dosi 1991, and others), finance and credit theory (Delbeke 1985; Korpinen 1985), explanation in terms of war cycles (Goldstein 1988), explanation in terms of cycles in energy consumption (Nakicenovic 1987) and other considerations.

On the other hand, there were discovered cycles with different periodicity: 7–11 years – Juglar cycle (Juglar 1862), 3–5 years—Kitchin cycle (Kitchin 1923), 18–30 years—Kuznets cycle (Kuznets 1926), 45–60 years—Kondratieff cycle (Kondratieff 1935), and 90–120 years—leading power cycle (Modelski 1987). Thus, the cyclic picture of the global economy has been thoroughly studied in terms of their causes and their types. However, two main issues remained unclear: detailed representation of the cycle as an equilibrium or non-equilibrium process as well as accurate analysis and forecast of economic cycles. The issue on the equilibrium nature of economic cycles is important in terms of management of economic dynamics: 1—if the equilibrium concept is accepted, then management is unnecessary and the economic policy comes down to the laissez-faire principle; 2—if the cyclic nature is considered to be a non-equilibrium process, then development of management principles with regard to economic dynamics (for example, conduct of proactive industrial policy and development of national innovation systems) becomes vital.

The issue of accurate analysis and forecast is necessary because, in the majority of approaches, cycles are characterized as strictly synchronized processes, while break points of trends in real statistical data have asynchronous nature. This cycle peculiarity is considered in this paper in detail.

In terms of representation of cycles as equilibrium or non-equilibrium processes, the following should be noted. In the 1980s, under influence of the ideas of N.D. Kondratieff (Kondratieff 1922, 1925, 1928, 1935) and J. Schumpeter (1912, 1939) on the role of radical innovations in the formation of economic cycles, there was initiated an ample debate related to the nature of economic cycles, and, in particular, long waves and their consequences (appearance of new branches of production, new structures of the finance sector, alteration in the usage of primary energy resources, burgeoning of new professional activities, resuming, a radical change in the industrial structure of the economy) (Vasko 1985). Both Kondratieff and Schumpeter proceeded from the assumption that the economic system is essentially in a non-equilibrium state or oscillating around several levels of equilibrium. Later W.C. Mitchell (1959) studied cycles as trend deviations, i.e. non-equilibrium processes. The Keynesian theory of economic cycles uses also essentially a non-equilibrium approach (Keynes 2007). Keynes' famous recommendations on necessary government intervention in the economy were actually based on that fact. Neo-Keynesians, like R. Harrod (Harrod 1959) and R. Hansen (Hansen, 1951), based on studies of Mitchell, Keynes, Kondratieff, and Tugan-Baranovsky (Tugan-Baranovsky 1997), considered also economic cycles as non-equilibrium processes.

But on the other hand monetarists have presented a different view of the phenomenon. M. Friedman (Friedman and Friedman 1962; Friedman and Schwartz 1982), for instance, assumed that the economy is an equilibrium structure and actions of financial authorities cause price shocks leading to occurrence of cyclic oscillations which are, actually, initiated by the government.

Economic cycles were studied in detail by representatives of the theory of rational expectations, separated from the pure monetarism. Monetarists assumed that the instability of the monetary unit is a result of changes in the nominal money supply and in this way is the cause of economic cycles. They proposed to manage economic dynamics through changes in money supply. Representatives of the theory of rational expectations also consider government policy as the cause of crises and cycles. They think that this unpredictable policy leads to oscillations of money supply, causing changes in Consumer Price Index (CPI) and cyclic behavior of the economy. These ideas form the basis of a proposition of R. Lukas (Lucas 1977, 1987), which states that the government cannot forecast rational expectations of economic agents and acts blindfold, responding to mistakes of individuals. As a result of actions of individuals (companies), actions of the government become unpredictable as well, therefore initiating a cycle. In his works Lukas analyzed such issues as neutrality of money, natural rate of unemployment and efficiency of government intervention in the economy. In the works of T. Sargent (Sargent 1986) the analysis is centered on rationality of business units and the consideration of the economic system as essentially at equilibrium.

Representatives of the modern theory of the real business cycles, like E. Preskott and F. Kudland (1982, 1990), associate oscillations of the GDP level with external shocks considered as a stochastic process. They accept oscillations of manufacturing and employment as a consequence of technological factors (which are considered in their theory as external). In their works they assume that equilibrium is observed during economic growth. Oscillations of the GDP level, even those below zero, are a characteristic of an in-equilibrium state (Zamulin 2005; Ostapenko 2013). By means of combination of the stochastic general equilibrium model of a closed economy under conditions of perfect competition, and long-term analysis of economic growth, Kudland and Preskott showed that technological progress is the most important factor of economic development that causes short-term cyclic oscillations in the manufacturing and employment rate, and therefore shocks from aggregate supply are the cause of oscillations. Thus, they consider technical progress as an external factor of the economic system, a source of shocks. As a result, as V.A. Tsvetkov points out (Tsvetkov 2013), “in modern macroeconomic models supply shocks are considered along with demand shocks.”

S. Verick and I. Islam (Verick and Islam 2010) from the German Institute for the Study of Labor confirm the periodicity of economic crises occurring approximately every 10 years. They associate the appearance of crises in the economy mostly with price shocks, as well as with such factors as currency instability, systemic problems of the bank sector, and national debt.

The theory of the real business cycles states that the economy is at equilibrium, and it can be drawn away from this state by external shocks only. In general, development factors are the following macroeconomic parameters: prices, salary, and interest rate. Changes in labor efficiency, oil prices and other factors lead to changes in the equilibrium level. Founders of this theory assume that the economy can reach the equilibrium level by itself and measures coming from economy management rather hinder economic progress instead of facilitating it. This is the key difference of the theory of the real business cycles from the earlier Keynesianism and modern post-Keynesianism.

A similar concept of the mechanism of economic cycles, based on the idea that the economic system is essentially at equilibrium, belongs to the theory of the neo-Austrian school. This theory assumes that banks conducting expansionary monetary policy are responsible for cycle occurrence. During economic recovery they believe that the economy should be flood with new credit as means of payment, causing inflation and overheating. From the point of view of representatives of neo-Austrian school (Soto 2008), the assumptions of representatives of Chicago school, stating that economic crises are exogenous, are wrong. In their opinion, representatives of Keynesian school are wrong too as they classify economic crises as a characteristic of the market economy. Economic cycles are result of problems of the institutional system, i.e., problems related to the preferred fractional reserve banking system. Neo-Austrian school proposes to rely on spontaneous actions of market forces leading the system to equilibrium, and reintroduction of the gold standard as an optimal policy of government control.

The equilibrium approach where the cycle is, however, represented as combination of asymmetric oscillations, was introduced by M. Aoki (1996, 1998, 2001). His models are based on the concept of multiple equilibrium of the evolving closed economy. Therefore, his approach based on application of Markov processes is close to the approach of Kondratieff (1922, 1928, 1935), who assumed existence of three levels of equilibrium in systems where oscillations with different frequencies are observed (Kondratieff 1928).

In Russian economic thought the nature of cycles is mostly associated with non-equilibrium processes (Glaziev 1990, 2009; Yakovets 1999) and is studied in terms of the concept of technological paradigms occurring during some given periods when the economy is in non-equilibrium, i.e. during bifurcation periods. Non-equilibrium of economic dynamics is characterized by the fact that under conditions of chaos the system chooses one of the available competing lines of development. Such line becomes an attractor to which economic dynamics tends all the way to the saturation point of the technological paradigm and the corresponding institutional structure (Osipov and Shurgalina 1994). A similar concept on the mechanisms of economic cycles is based on the synergistic theory applied in studies of complex open-ended systems (Haken 1980; Prigogine and Stengers 1986) including the economy (Knyazeva and Kurdyumov 2002).

Devezas and Corredine (Devezas and Corredine 2002) proposed a cybernetic framework model that explains successfully many specific aspects of the timely unfolding of recurrent socioeconomic phenomena like economic longwaves. In this model the time-evolution of a techoeconomic system is described discretely as a logistically growing numbers of 'interactors' adopting an emerging set of basic technological innovations and exchanging and processing information in a finite niche of available information. The authors demonstrate that the rate of information entropy change exhibits a wave-like aspect evidenced by a four-phased behavior denoting the complete unfolding of a complete long-wave. Moreover, they speculate that social systems mimic living systems as efficient negentropic machines, and making use of Prigogine's entropy balance equation for open systems it is suggested that its cyclical behavior is probably the best way to follow nature's efficiency strategy.

An essentially non-equilibrium synergistic model of economic dynamics was proposed by Metcalfe (2007). It is based on the concept that the knowledge available to economic agents constantly shift the trajectory of economic development maintaining it essentially in a non-equilibrium state. In other words, in this model economics is an evolutionary process characteristic of open-ended complex systems. If technological evolution is considered as an attractor, new knowledge coming into the system leads to shifts in relation to the attractor. Such system is essentially at non-equilibrium and difficult to forecast. Constant system shift in relation to the attractor implies essentially in the non-equilibrium state of the system.

In cases when the economic cycle is considered as a phenomenon caused by the monetary sphere of the economy, it can be seen that researchers tend to a rough description of its causes in the form of optimistic/pessimistic mistakes and inappropriate actions of money authorities. In cases when the study area includes processes happening at the level of economy structure, i.e. structural shifts and alteration of technical and economic paradigms, economic development is immediately considered as a process which is mainly at non-equilibrium. This can be seen in the work of Buhm and Punzo (2001), who consider the economy as a non-equilibrium process which is mainly related to the alteration of structural elements of the economy.

It should be noted that modern economic literature still exhibits the confrontation of neoclassical and Keynesian schools, in the same vein as in decades ago, essentially the dispute between non-equilibrium post-Keynesian process (Arestis 1988, 2009) and neoclassical equilibrium process. Within the framework of new economic synthesis, these approaches are, however, moving closer to each other (Woodford 2010). In his study, Solomou (2001) analyzed institutional conditions affecting cycle duration and amplitude and demonstrated that during the last 200 years cycles changed their nature.

Therefore, modern economic literature dedicated to the nature of economic cycles presents an ever-lasting confrontation of equilibrium and non-equilibrium approaches to study this phenomenon. Several theories, which do not analyze structural changes in the economy, consider cycles as equilibrium processes (for example, neo-Austrian school and the theory of real business cycles). Those researchers, who take into account micro- and meso-level processes occurring at the level of sectors and branches of the economy, and provide synergistic explanation for processes of economic evolution, consider the cyclic behavior as a non-equilibrium process interrelated to economic growth (Akaev et al. 2011).

2 On the Key Parameters of Cyclic Dynamics

In the preceding section we described various approaches characterizing cyclic processes in terms of their nature, i.e., if they are equilibrium or non-equilibrium processes.

In our work we assume that economic cycles are manifestations of non-equilibrium systems, and that random factors affecting cycle amplitude and duration, as well as the characteristic lengths of decline and growth phases, play a major role in their unfolding. If we consider statistical data characterizing cycles we will see that from cycle to cycle their duration and amplitude may change, and the lengths of decline and growth phases vary as well. Such fact has led some researches, as for instance S. Solomou (1990) to deny regularities in the unfolding of economic oscillations, not considering them as cycles in the full sense of the word.

In our framework we consider that the unfolding of economic cycles has its own regularities, but, however, the influence of a variety of random factors leads to some form of asymmetry in their behavior and shifts the system away from the attractor. Therefore, economic cycles (of almost any duration) are asymmetrical in their up and down unfolding, and can present also different amplitudes. One of the influence-leading external factors is information entering the economic system from other systems (technological, ecological, social). However, the distribution of information in the economic system is asymmetrical as proposed by Stigler (1961) several decades ago. Therefore, information asymmetry feeding the economic system can be one of the reasons of the asymmetrical nature of economic cycles, as was also suggested by Devezas-Corredine (2002) in their cybernetic framework.

On the basis of this asymmetrical nature of cycles it is necessary to provide a quantitative evaluation of their parameters. Based in the analysis of vast amount of empirical data on economic cycles, we can define the following key parameters: 1—Periodicity, 2—Amplitude, 3—Asymmetry. To provide their quantitative evaluation we performed a series of calculations in relation to GDP dynamics and some other economic variables for several countries. The corresponding results are presented in the next sections.

2.1 Periodicity

The period of an economic cycle is understood to be the duration of a local dynamic process, in which the alternation of crisis, depression, recovery and growth phases is observed, and periodicity is the measure of the time interval between successive periods The mechanism of occurrence of economic cycles under the influence of technical progress is based on the periodic appearance and expansion of basic innovations. As it is well known, the diffusion of innovations in a socioeconomic system follow the pattern of simple logistic curves. Figure 1a below exhibits the

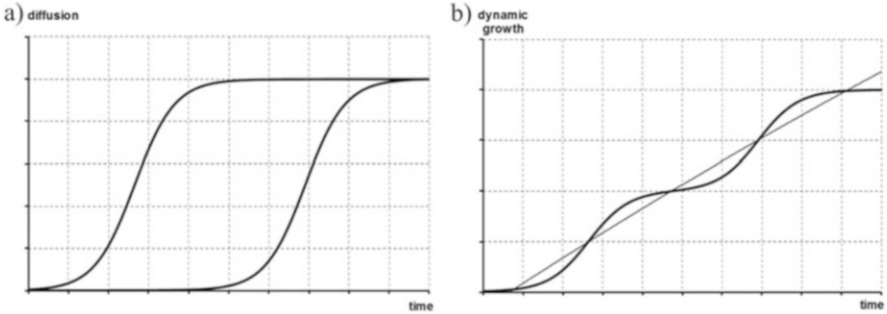


Fig. 1 Pictorial representation of the development of cyclic economic dynamics: (**a**) two successive basic innovations; (**b**) cyclical deviation of the general growth trend of an imaginary economic variable

pictorial representation of the diffusion of two successive basic innovations, which constitutes the basic mechanism of the development of a cyclic process in economic dynamics. Figure 1b presents the pictorial representation of the periodical deviation of the growth trend of a given economic measure as result of the impact of successive innovations.

Many authors studied the issues of appearance frequency and diffusion time of basic innovations. As examples we can mention the works of Gerhard Mensch (1973, 1979) and Masaaki Hirooka (2006), proving that conjuncture cycles are determined by innovation dynamics.

Recent published results of Akaev (2011) demonstrate that economic dynamics has fundamental cyclicality of at least three types, with periodicity from 115 to 180 years, from 45 to 60 years, and from 20 to 30 years. These types correspond to Modelski (leading powers), Kondratieff and Kuznets cycles respectively, already commented on in the previous section. The same study has found yet oscillations with periodicity from 75 to 100 years, named as cycles of infra-trajectories, as well as the phenomenon of contraction of Kondratieff cycles down to a periodicity of 30–40 years.

Regarding this last one, Kodratieff cycles (or Kondratief waves, or K-waves for short), there has been an intense debate about their exact duration and periodicity, and researchers are currently divided about the position where we are today in relation to the succession of K-waves: if we are at the end of the fourth wave and transition to the fifth, or perhaps already in the full ascending phase of the fifth K-wave, or if we have already reached the peak of the fifth and traveling the downward path that will define the innovation context for the start of the sixth wave. Devezas (2010) has interpreted the crisis of 2007–2008 as a strong anomaly of the expansion phase of the fifth K-wave that will reach the peak around 2020–2025, while Korotayev and Tsirel (2010), using the same set of data (Maddison's GDP series) and the same methodology (Fourier analysis), treated the 2007–2008 crisis as a temporary dip during the peak of the fifth K-wave, signaling the peak turning point. On the other hand Gladkikh (2014) identified

Table 1 Distribution of countries according to periodicity in GDP dynamics

Continent	Share of countries with periodicity length, %		
	30–40 years	40–50 years	50 and more years
America*	30	40	30
Europe	50	29	21
Asia	–	75	25
Oceania	–	–	100

*North and Latin

the downward phase of the fourth K-wave as it were located between 1977 and 2013, the year signaling the start of the upward phase of the fifth wave, which will reach the ceiling around 2028. Such differences of opinion highlight the need for further development of the methodology for estimation of the cyclic dynamics.

In order to evaluate the periodicity of economic cycles we performed studies on the GDP dynamics for 45 countries, using data from Maddison (Maddison 2010) and from The Conference Board (The Conference Board Total Economy Database 2014) (total GDP, in millions 2012 US$, converted to 2013 price level with updated EKP PPP). Table 1 below presents the results for countries grouped by continents, indicating the share distribution of the found periodicities—30/40 years, 40/50 years, and 50/60 years. The detailed results of the found periodicities for all countries investigated are presented in Appendix 1.

Analysis of the results shows that the process of contraction of Kondratieff waves is well defined in Europe where half of the countries presents periodicity of 30–40 years. It is less observed in Asian countries and is not observed at all in Australia and New Zealand.

2.2 *Amplitude*

Along with periodicity, the height (amplitude) of the decline and growth phases in relation to the basic development trend plays an important role in the description of the cyclic process. The range of these oscillations determines the significance of the cyclic component in the whole process of economic growth. If oscillations are insignificant, we can ignore them and consider only the general trend; otherwise, forecasts, which do not take into account the wave-like process, may seem to be incorrect.

From the economic point of view the intensity of trend fluctuations characterizes the degree of growth stability and depends on various factors which determine specific features of operation of a given economic system. Unfortunately, the issue of amplitude of cyclic oscillations has been poorly studied. To fill this gap, we performed evaluation of trend fluctuation characteristics in GDP dynamics for several countries, which were classified according to their level of development, to which we add separately the group of oil producing nations:

Appendix 1 Parameters of GDP cycles for 30 countries

Region	Country	Periodicity, years	Asymmetry coefficient
		p	K_{as}
North America	USA	43.4	0.83
	Canada	53.1	0.34
Latin America	Mexico	34.1	0.40
	Brazil	35.5	0.56
	Argentina	44.4	0.30
	Venezuela	52.6	0.36
	Chile	39.0	0.61
	Colombia	42.2	0.79
	Peru	48.7	0.54
	Uruguay	74.3	0.59
Western Europe	Great Britain	30.0	0.67
	France	46.2	0.24
	Germany	38.7	0.54
	Italy	39.9	0.67
	Spain	41.8	0.36
	Netherlands	41.1	0.29
	Belgium	44.0	0.52
	Portugal	30.4	0.19
	Austria	51.8	0.24
	Denmark	38.0	0.30
	Sweden	51.9	0.15
	Norway	37.8	0.40
	Finland	37.8	0.62
	Switzerland	51.1	0.29
Asia	Japan	41.7	0.51
	India	45.0	−0.42
	Indonesia	56.1	0.55
	Sri Lanka	48.5	0.50
Oceania	Australia	50.9	0.25
	New Zealand	54.5	0.32

Developed Countries USA, Canada, Great Britain, France, Germany, Italy, Netherlands, Sweden, and Japan.

Developing Countries Mexico, Brazil, Argentina, Venezuela, Russia, China, India, Indonesia, Malaysia, South Korea, Portugal, Thailand, Turkey, Algeria, Egypt, Nigeria, RSA.

Emerging Countries Poland, Hungary, Bulgaria, Romania, Ukraine, Belarus, Moldova, Armenia, Georgia, Latvia, Lithuania, Estonia, Azerbaijan.

Oil Countries Norway, Iran, Iraq, Saudi Arabia, Kuwait.

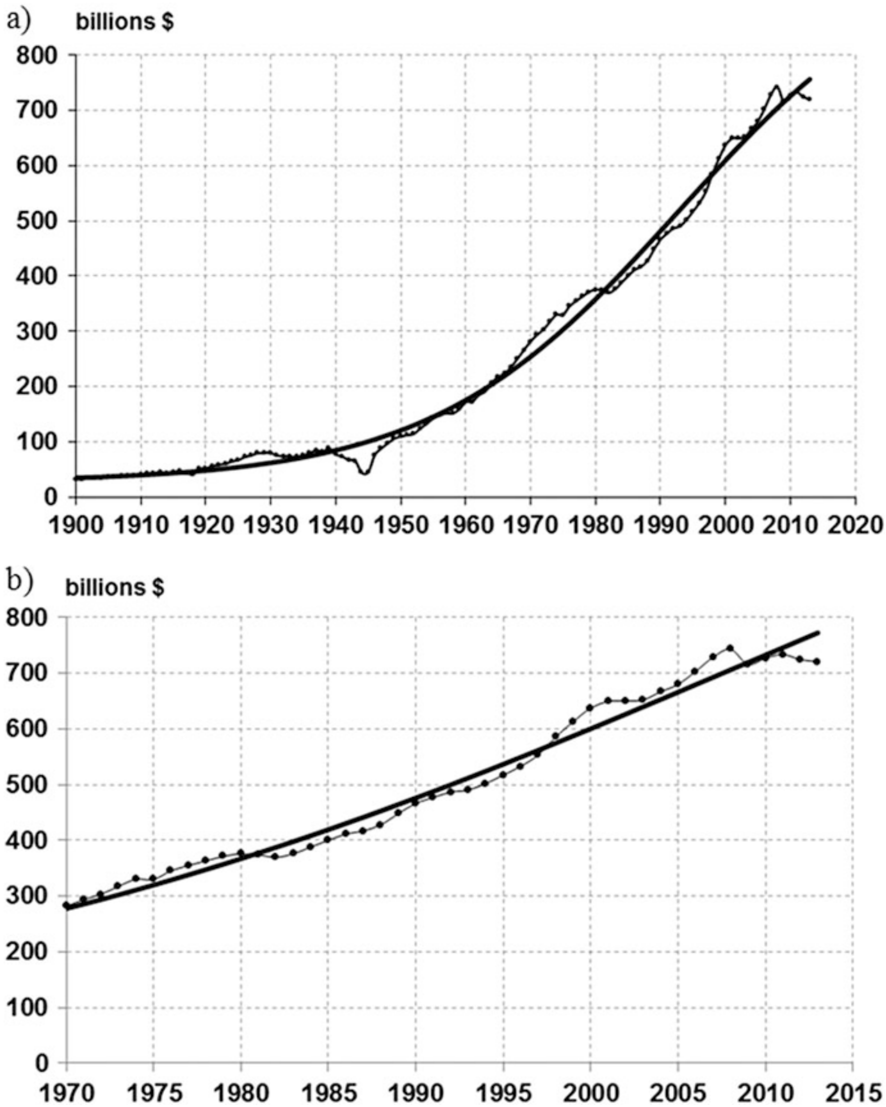


Fig. 2 Logistic fit of the GDP growth in the Netherlands—(a) period 1900–2013 (b) period 1970–2013 used for calculating the amplitude of deviations of the main logistic trend. Fitting parameters; a = 1462 (ceiling), b = 4.4, c = 0.037 (growth parameter)

For the most developed and developing countries analyzed it was possible to fit a simple logistic curve, as expressed by Eq. (1), to the GDP growth since the start of the twentieth century. As example it is shown in Fig. 2 the fitting in the case of the Netherlands; in Fig. 2a we see the result for the whole period (1900–2013), and in Fig. 2b the specific period 1970–2013 we used for calculating the amplitude of

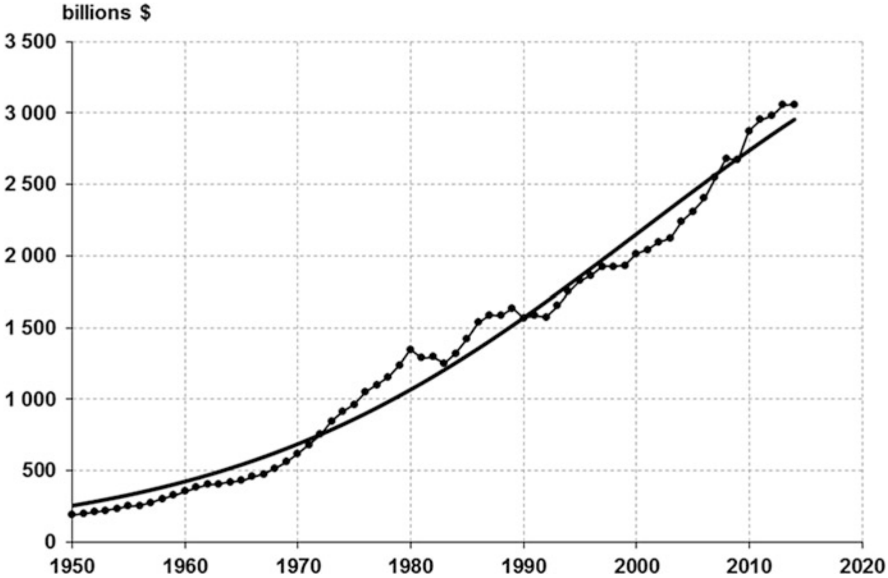


Fig. 3 Logistic fit of the GDP growth in Brazil. Fitting parameters; a = 4345 (ceiling), b = 17, c = 0.055 (growth parameter)

oscillations and asymmetry related to the main trend. Figure 3 shows the fitting for Brazil in the period 1950–2013.

$$y = \frac{a}{1 + b.e^{-c(t-t_0)}} \tag{1}$$

However, in the case of the emerging countries it was not possible to fit a simple logistic curve, and in these cases we performed the fitting with a modified logistic function in the form presented in Eq. (2) below:

$$y = \left[a + b.e^{-c(t-t_0)}\right].(t - t_0) + d \tag{2}$$

In Fig. 4 we can see the fitting of Eq. (2) to real GDP data of Azerbaijan:

Table 2 presents by groups of countries the results obtained for the trend fluctuations—amplitudes—expressed as percentiles. The complete set of results for individual countries can be seen in Appendix 2.

Analyzing the numbers presented in Table 2, we may conclude that economic growth in developed countries is more stable than in developing countries. In the group of emerging countries, the range of cyclic oscillations is formed at the level of developed countries, but is slightly wider. The range calculated for the group of oil countries (Iran, Iraq, Saudi Arabia, and Kuwait) differs vastly, and the trend fluctuations reach significant values. Such lack of stability in economic growth is related to violent oscillations of conjuncture in global energy markets, which affects the economy of the oil countries significantly. In general, we can state that countries

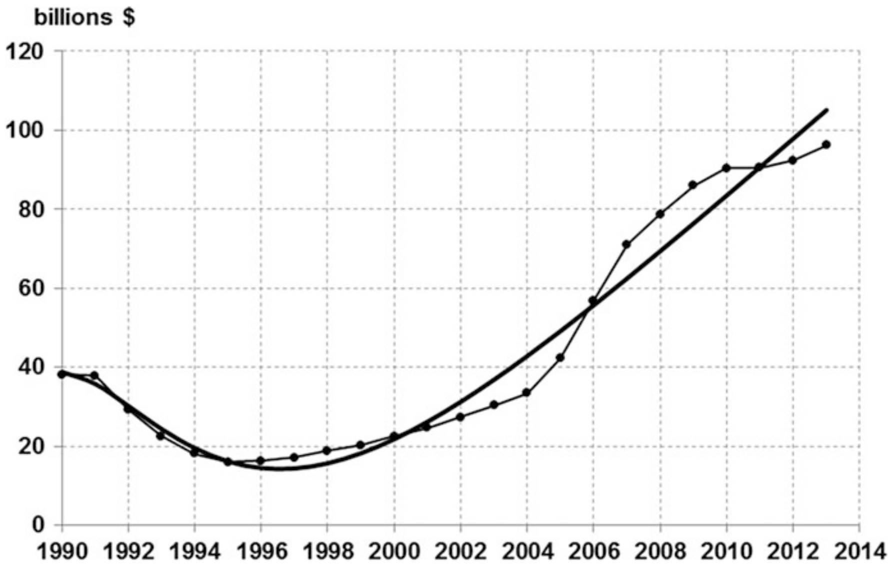


Fig. 4 Logistic fit of the GDP growth for Azerbaijan using the modified logistic function expressed by Eq. (2). Fitting parameters: a = 7.35, b = 98.7, c = 0.34, d = −92.

Table 2 Range of Cyclic oscillations, broken down by groups of countries

Groups of countries	Trend fluctuations, %	
	Average	Maximum
Developed countries	2–5	4–11
Developing countries	3–9	6–30
Emerging countries	2–10	5–25
Oil countries	3–33	7–72

geared to raw materials cannot rely on the stability of economic development. An exception to this series is Norway; it managed to create a more balanced structure in the sector of oil and gas production, using R&D results in the whole process flow: from engineering, production, and transportation to the maintenance system.[1]

[1]In Norway there is complete independence of the internal electricity market from oil and gas production: about 97 % of all electricity generation in Norway comes from hydropower. (Source: US Energy Information Administration. Norway, April 28, 2014 p.10). Moreover there is technological and institutional diversification of the sector, allowing to Norwegian companies to be involved in projects of various levels of technological complexity (from Sakhalin and the North Sea to Angola) and use small and medium-sized companies with an annual turn-over below10 million dollars. Many of them are supplying a wide range of cost effective and flexible quality products and services to the global market. (Sources: *Norwegian Petroleum Technology*: *A success story*. Editor H. Keilen. Norwegian Academy of Technological Sciences (NTVA) in co-operation with Offshore Media Group and INTSOK. 2005; *The Norwegian oil field service analysis* 2013. EYGM Limited, 2014).

Appendix 2 Range of cyclic oscillations for 45 countries

Region	Country	Trend fluctuations, %	
		Average	Maximum
North America	USA	2.3	6.0
	Canada	3.0	6.2
Latin America	Mexico	5.0	16.1
	Brazil	6.2	19.0
	Argentina	5.3	15.9
	Venezuela	5.7	21.9
Western Europe	Great Britain	4.3	11.0
	France	1.9	4.1
	Germany	1.8	4.2
	Italy	2.1	5.2
	Netherlands	3.6	6.7
	Sweden	3.4	9.8
	Norway	2.7	7.3
Eastern Europe	Poland	2.1	4.9
	Hungary	4.0	10.2
	Bulgaria	4.0	11.5
	Romania	4.9	14.2
Asia	Japan	2.7	8.7
	China	5.8	14.3
	India	2.3	5.8
	Indonesia	6.1	20.1
	Malaysia	3.8	12.9
	Korea	3.1	8.3
	Thailand	5.4	15.8
Eurasia	Russia	3.0	10.0
	Ukraine	6.7	16.5
	Belarus	3.0	8.0
	Moldova	4.2	12.6
	Armenia	8.7	20.2
	Georgia	4.5	11.2
	Azerbaijan	9.5	21.9
	Latvia	6.7	25.1
	Lithuania	4.8	16.2
	Estonia	6.4	18.0
Near East	Turkey	3.9	10.2
	Iran	12.7	55.6
	Iraq	32.6	66.7
	Saudi Arabia	10.6	38.0
	Kuwait	24.8	72.5

(continued)

Appendix 2 (continued)

Region	Country	Trend fluctuations, %	
		Average	Maximum
Africa	Algeria	8.9	29.9
	Egypt	7.0	22.7
	Nigeria	6.2	27.5
	RSA	5.1	12.2
Oceania	Australia	1.6	6.1
	New Zealand	3.3	9.8

2.3 Asymmetry

By asymmetry of an economic cycle we understand the difference in duration of the separate phases—crisis, depression, recovery and growth. According to empirical observations, crisis and growth phases quickly go by, while depression and recovery cover significant periods of time.

Applying the innovation theory of development of cyclic processes, we should take into account that any innovation has a life cycle. In correspondence with its life cycle, it does not remain at the maximum level after reaching its limits. The market potential decreases gradually, and it becomes history, being substituted by new innovations.

The examples of innovations-substituting processes abound. See for instance the case of radio development. In the beginning of the twentieth century valve receivers were invented, which gave way to transistor receivers in 1950s, and at the turn of the twenty-first century digital broadcasting settled definitely. Nowadays, there is almost no demand for valve and transistor receivers.

Upon considering that the market for a given innovation decrease after its height, the picture represented in Fig. 1 changes significantly, for the superposition of the life cycles leads to a cyclic process of asymmetric nature as shown in Fig. 5.

Asymmetry can be considered as a key concept in various spheres of science. Representatives of various sciences (biology, anthropology, cryptography, sociology, etc.) pay attention to asymmetry both in nature and society. We can state that perfect symmetry is a subjective theoretical construction like the category of the "perfect market". In reality (incl. economic reality) all phenomena and processes are more or less asymmetrical.

For example, with regard to the economy, the issue of asymmetry is studied in detail in the corresponding section of the theory of economic information proposed by George Stigler (1961). In 1996 the Nobel Prize in Economics was awarded jointly with James Mirrlees and William Vickrey "for their fundamental contributions to the economic theory of incentives under asymmetric information".

To model the asymmetric cyclic process, we propose two mathematical see-saw-like sine functions: an additive version and a multiplicative version.

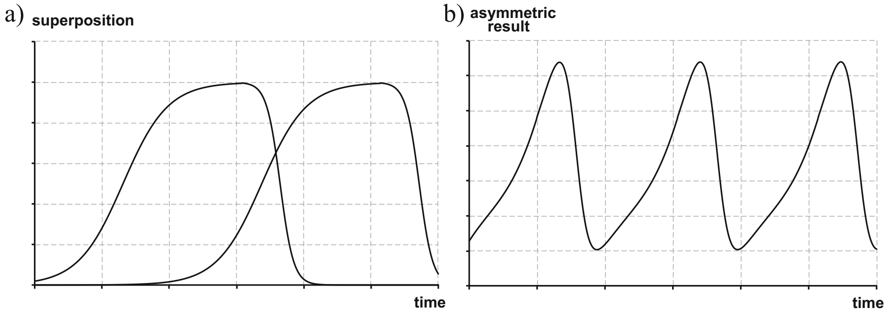


Fig. 5 Development of an asymmetric cyclic process. (a) basic innovations life cycle (b) asymmetric result

Additive version

$$\hat{z}=d\cdot\sin\left\langle\frac{\pi}{2\cdot arctg\left(\frac{\Delta}{2}\right)}\cdot\left\{arctg\left[\frac{1}{tg\left(\frac{\pi\cdot t}{p}-f\right)}-\Delta\right]-arctg\left[\frac{1}{tg\left(\frac{\pi\cdot t}{p}-f\right)}\right]\right\}+\frac{\pi}{2}\right\rangle \quad (3)$$

Multiplicative version

$$\hat{z}=d\sin\left\langle\frac{\pi}{\left[arctg\left(\frac{\Delta}{2}\right)\right]^2+\frac{\pi^2}{4}}\cdot\left\{arctg\left[\frac{1}{tg\left(\frac{\pi\cdot t}{p}-f\right)}-\Delta\right]\cdot arctg\left[\frac{1}{tg\left(\frac{\pi\cdot t}{p}-f\right)}\right]-\frac{\pi^2}{4}\right\}+\frac{\pi}{2}\right\rangle \quad (4)$$

where d—amplitude; p—cycle periodicity; f—shift along the X-axis; Δ—asymmetry; t—argument of time.

To measure the asymmetry level, we propose a special indicator:

$$K_{as}=\begin{cases}0 & \Delta=0\\ 1-2\cdot\dfrac{arctg(2/\Delta)}{\pi} & \Delta>0\\ -1-2\cdot\dfrac{arctg(2/\Delta)}{\pi} & \Delta<0\end{cases} \quad (5)$$

Values of this indicator range from -1 to +1. The zero value characterizes the case of perfect symmetry and the range limits correspond to extreme asymmetries.

These functions allow the description of asymmetric cycles in the dynamics of different economic processes. They differ in the pattern formation of upward, peak, and downward phases of the cycle. If the dynamic process is characterized by the formation of similar peaks and downturns, the model that gives a better

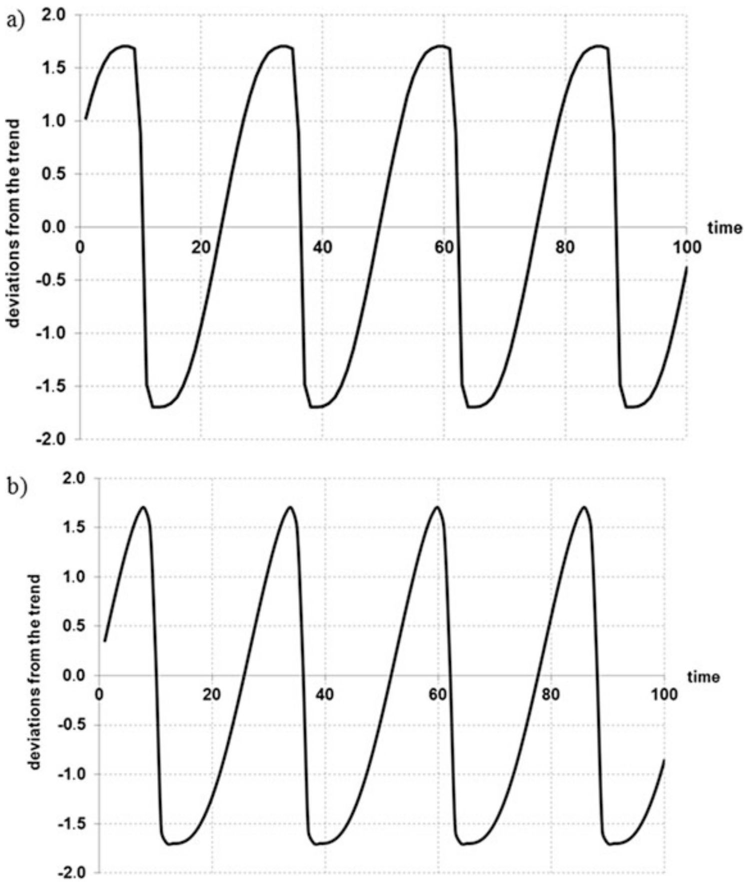


Fig. 6 Graphical representation of the proposed asymmetric sine functions: (a) additive form given by Eq. (3), (b) multiplicative form given by Eq. (4)

representation of the cycle function is the additive form, whose graphical representation is shown in Fig. 6a. In the case where the analyzed process has sharp peaks and faster cycle downturns the data are most successfully fitted using the function in the multiplicative form, whose graphical representation appears in Fig. 6b,

Let's give a closer look at the case of Azerbaijan for instance, whose general behavior of the GDP growth dynamics was already commented on earlier, when it was shown that the modified logistic function presents the best fit to the data (Fig. 4). In this case it was found that the additive sine function is the one that better describes the GDP trend deviation as shown in Fig. 7 below.

By combining in additive form both components—growth dynamics and trend deviation—we can show a consolidated model description of the dynamics of the analyzed indicator (GDP) for the case of Azerbaijan, as shown in Fig. 8.

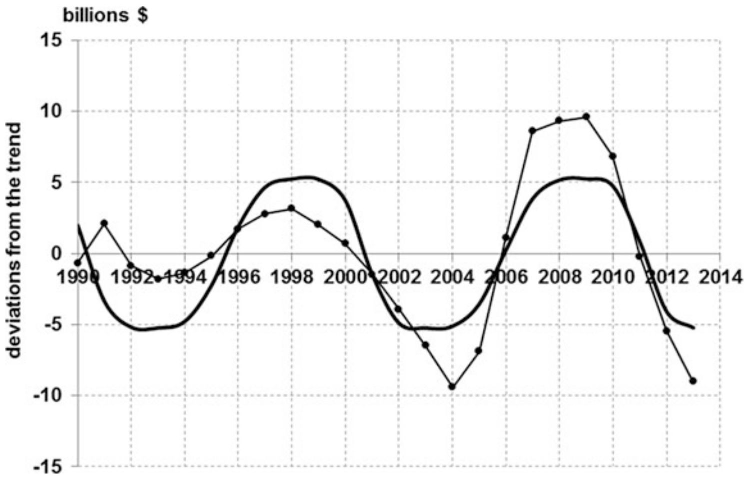


Fig. 7 GDP trend deviation for Azerbaijan fitted with the additive sine function. Fitting parameters: $p = 10.4, f = 2.5, \Delta = 0.31, d = 5.2$

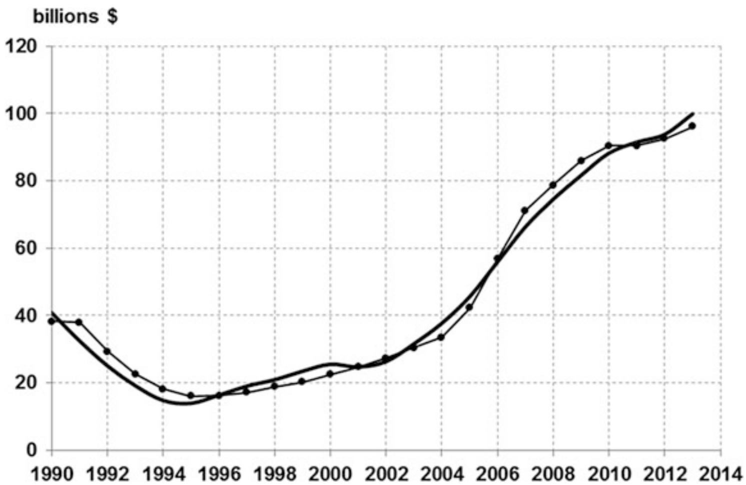


Fig. 8 Consolidated form of the GDP dynamics (trend plus cyclical component) for Azerbaijan ($r^2 = 0.986$)

Also for developing countries it was found that the additive sine function presents the best fitting. Figure 9 shows the cyclic dynamics of the GDP trend deviation for Brazil described with Eq. (3) and Fig. 10 presents the resulting consolidated model description of the Brazilian GDP dynamics.

For the asynchronous cyclic GDP growth dynamics of the developed countries we found that the multiplicative formulation is the one that better describes the oscillatory behavior along the last 100 years. As example Fig. 11 exhibits the GDP

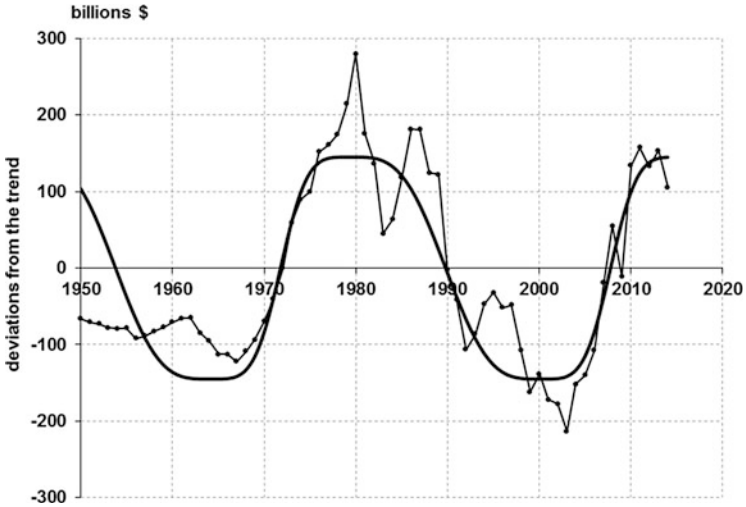


Fig. 9 Cyclic dynamics of the GDP trend deviation for Brazil described with the additive sine function. Fitting parameters: $p = 36, f = 2.65, \Delta = -0.5, d = 145$

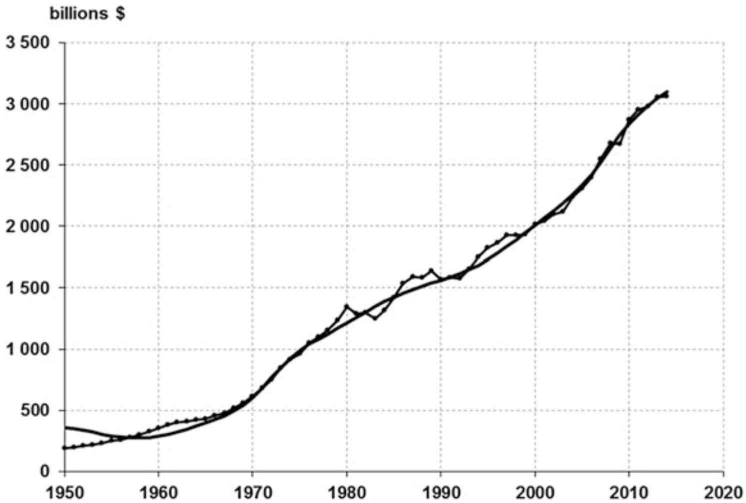


Fig. 10 Consolidated form of the GDP dynamics (trend plus cyclical component) for Brazil ($r^2 = 0.988$)

dynamics of Great Britain fitted with the multiplicative version of the asymmetric function.

Asymmetric cyclic processes are typical for many economic variables. In Fig. 12 it is shown the cyclic trend deviation of the global primary energy production, fitted with the multiplicative sine function. It is worth to observe in this graph the clear pattern of economic longwaves (K-waves), with a period of about 40 years, from

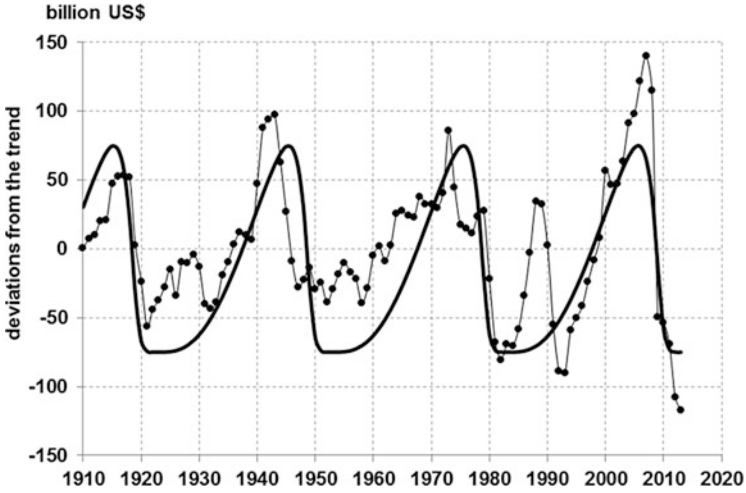


Fig. 11 Cyclic dynamics of the GDP trend deviation for the United Kingdom fitted with the multiplicative sine function

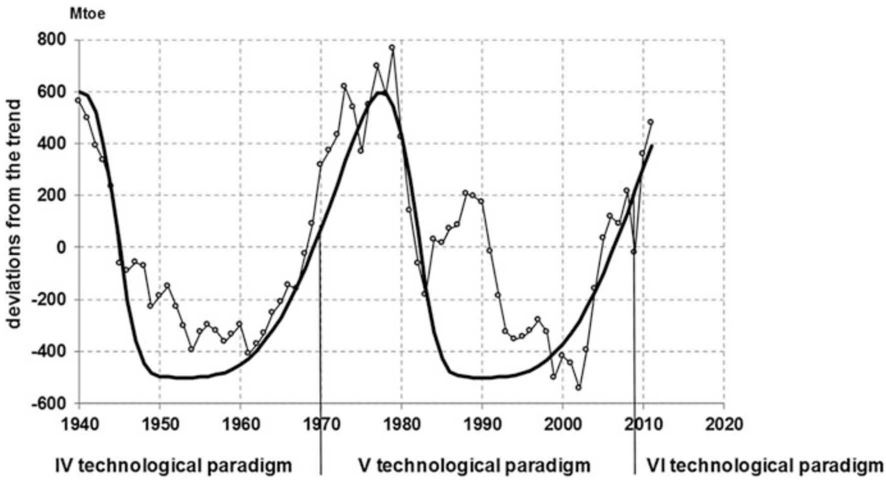


Fig. 12 Cyclic trend deviation of global energy production fitted with the multiplicative sine function.

which we can infer the typical transitions between technological paradigms. It is shown in the graph the points of transition from the fourth to the fifth technological paradigm (1970) and from the fifth to the sixth technological paradigm (2009). As observed by Rumyantseva (2003), in the beginning of each paradigm happens an explosive growth in energy production, alternated with steep decline, subsequent gradual recovery and preparation to the transition to the next paradigm. An

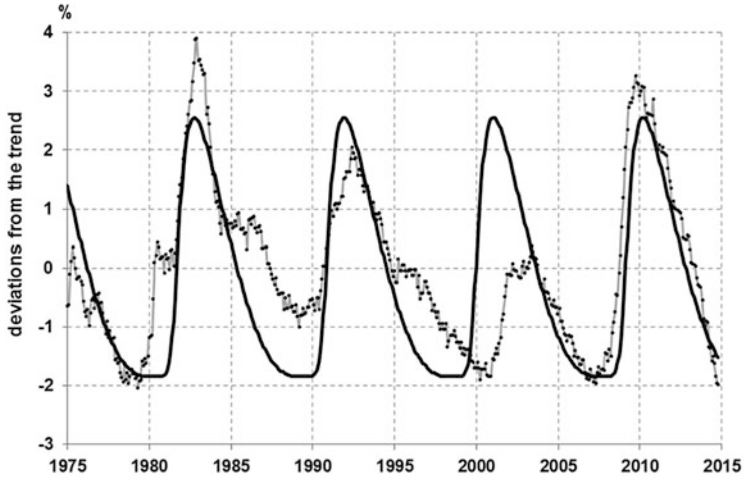


Fig. 13 Reverse asymmetry of cyclic oscillations of the civilian unemployment rate in the USA

Table 3 Asymmetry of GDP cycles per continent

Continent	Average value of asymmetry coefficient K_{as}
America*	0.53
Europe	0.39
Asia	0.50
Oceania	0.29

*North and Latin

interesting point worth to observe in this graph is the sharp and sudden decline in energy production observed in 2008–2009, caused probably by the great recession observed at this time, but that was quickly recovered in the years that followed, clearly resuming the growth trend.

In some cases, reverse asymmetry can be observed, whereby the duration of the declining phase significantly exceeds the period of growth of the economic attribute. Here we can cite as example the cyclic dynamics of the unemployment rate in the USA, whose graph is shown in Fig. 13.

The performed experimental calculations of the GDP dynamics for the 30 countries listed in Appendix 1 show that asymmetric cyclic processes are the common rule for all types of economies, independently of the development level. The maximum asymmetry level (K_{as}) was registered for the largest economy, i.e. the US economy. One of the possible reasons for this may lie in the very large share of the financial sector in GDP industry structure, which, by its economic nature, quickly and violently responds to changes in market conditions.

It can be observed that in different continents the asymmetry coefficient of the GDP dynamics varies significantly as shown in Table 3.

In most countries of North and Latin America the asymmetry coefficient is relatively high, while in European continent countries a smoother nature of the cyclic process prevails. The lowest asymmetry level is observed in distant countries of Oceania.

3 Cyclic Deviations of Linear Trends

In order to demonstrate the good performance of our proposed approach we have applied it to some economic variables that do not present necessarily a logistic unfolding during the observed time span, but instead can be fitted optimally as oscillatory variations around a linear (increasing or decreasing) general trend.

As first example Figs. 14, 15 and 16 exhibit our results of the analysis of the investment processes dynamics for the region of North-America in the last 50 years (data from the World bank). Figure 14 shows the linear decreasing trend (Y = a − bt) of gross capital formation expressed as percent of GDP; Fig. 15 shows the trend deviation fitted with the additive sine function, and Fig. 16 exhibits the final result of the trend plus cyclical component.

As a second example we present our results about a completely different economic variable related with the trade volume of a small European country. Figures 17, 18, and 19 present the results of our research on the foreign trade (exports of goods and services expressed as % of GDP) for Portugal during the last 55 years (data from the World bank). Following the same sequence as in the previous case, Fig. 17 shows the linear increasing trend (Y = a + bt), Fig. 18 presents the cyclical trend deviation, this time fitted with the multiplicative sine version, and Fig. 19 shows the final result for the linear trend plus cyclical component.

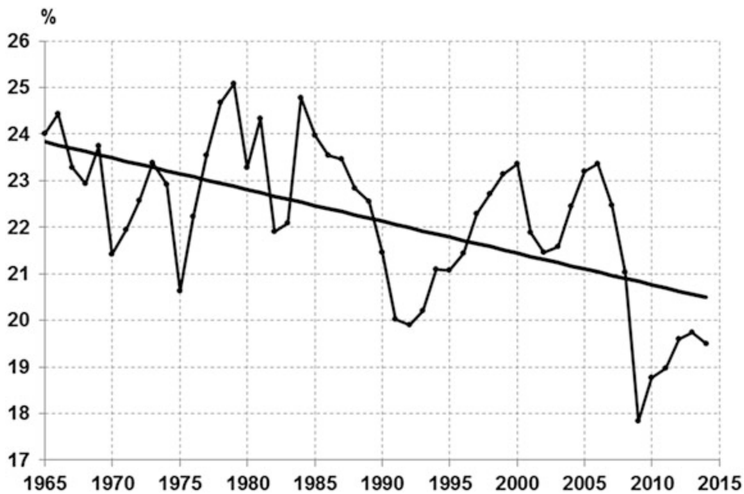


Fig. 14 Linear decreasing trend of gross capital formation in North America in the last 50 years. Fitting parameters: a = 24.24, b = 0.07

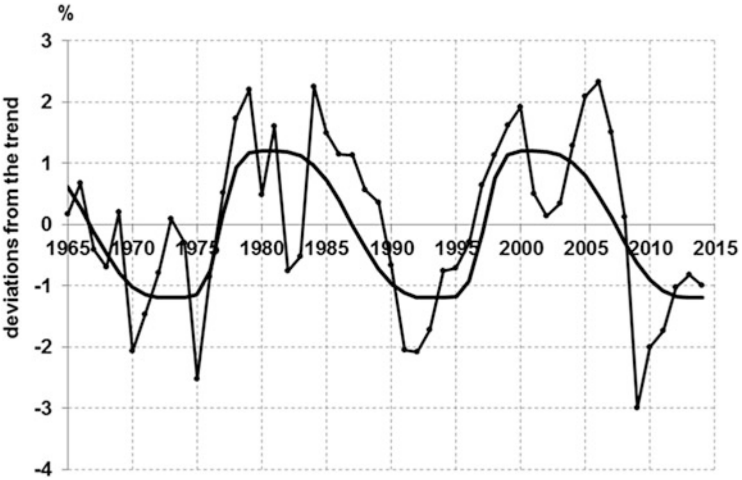


Fig. 15 Trend deviation of gross capital formation in North-America fitted with the additive sine function. Fitting parameters: $p = 20$, $f = 3.3$, $d = 1.2$, $\Delta = -1$

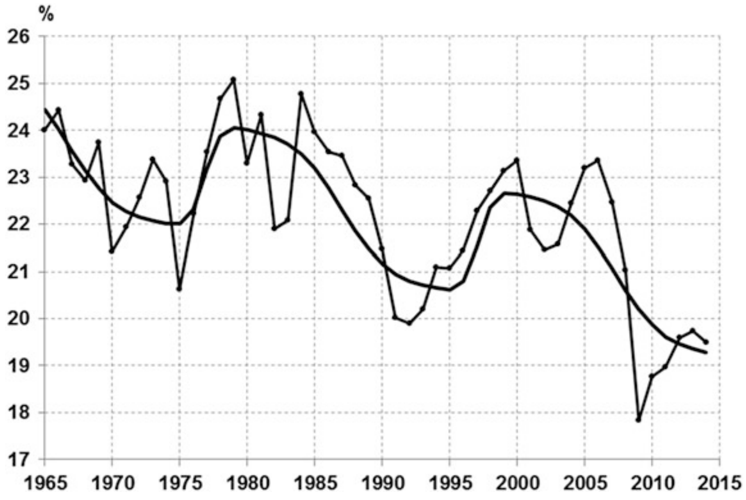


Fig. 16 Consolidated form (trend plus cyclical component) of the gross capital formation in North America ($r^2 = 0.781$)

As can be seen for both cases, our proposed model of asymmetric economic cycles works perfectly in describing very different economic variables such as capital formation or intensity of foreign trend. It is worth to point out a common denominator found in the analysis of these two cases: both reveal the clear unfolding of Kuznets-like cycles with a periodicity of about 25 years. Note that this fact demonstrates the versatility of the model to make predictions about the future behavior of economic variables.

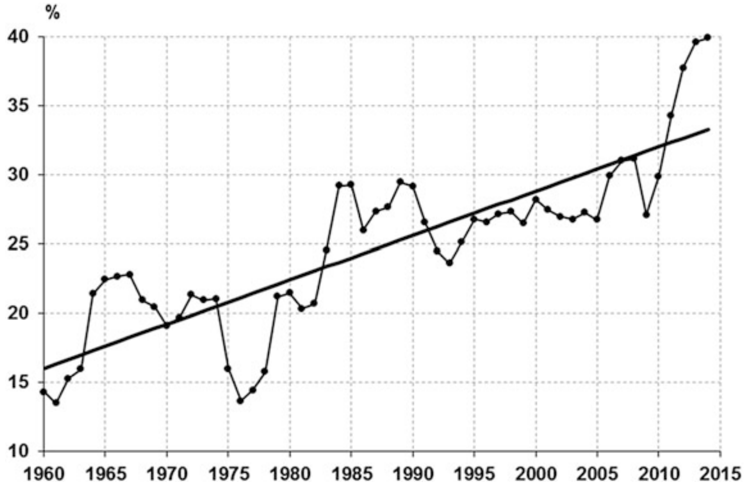


Fig. 17 Linear increasing trend of foreign trade (exports of goods and services) for Portugal in the last 55 years. Fitting parameters: a = 15.7, b = 0.032

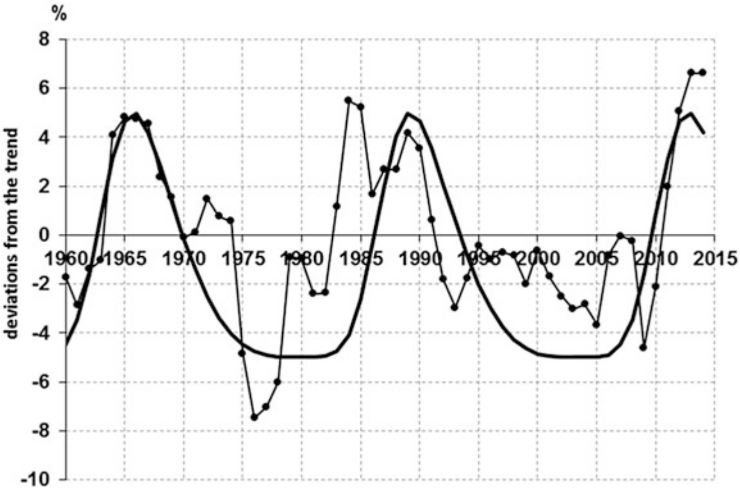


Fig. 18 Trend deviation of the Portuguese foreign trend fitted with the multiplicative sine function. Fitting parameters: p = 24, f = 0.9, d = 5, Δ = −0.7

4 Conclusion

Theories of economic cycles abound in the scholarly literature and the available models consider such cycles as either equilibrium or essentially non-equilibrium processes. Equilibrium in the dynamics of the economic cycles is considered when it is assumed that the cycle is caused by exogenous random factors of the economic

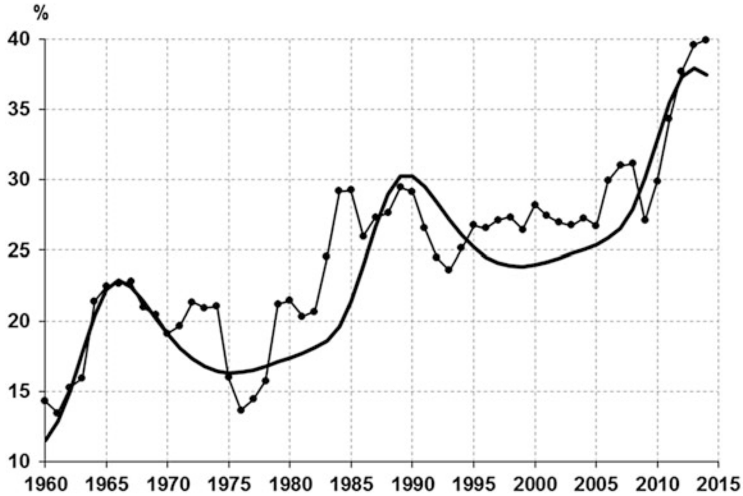


Fig. 19 Consolidated form (trend plus cyclical component) of the Portuguese foreign trade ($r^2 = 0.693$)

system, e.g. government intervention, intervention of monetary authorities, implementation of technologies, etc. In these cases, the cycle is considered as a process of fluctuations between demand and offer, i.e. a temporary phenomenon occurring in the process of "wandering search" of market equilibrium. Such approach is mainly typical for neoclassical and monetary explanations of economic cycles.

But in the cases when the detailed mechanism of the economic cycle is analyzed, and it is related to shifts in the industrial structure of the economy, the implementation of innovations is considered as an endogenous regularity of the economic cycle. Then the cycle is interpreted as an essentially non-equilibrium process. In this paper we analyze economic cycles as non-equilibrium processes characteristic of complex open-ended systems where technologies and industry shifts are the key development factor.

Keeping in mind a global perspective we have performed an extensive statistical analysis of the three key parameters of the economic cyclic dynamics: periodicity, amplitude, and asymmetry. Our results show that the average cycle periodicity varies significantly among different countries. Moreover, it was found that the most developed countries present contraction of the duration of long economic cycles, while in developing countries the duration of these cycles is higher. This implies in a larger influence of scientific and technical progress on the duration of economic cycles in developed countries, and significant asynchrony in the development of global cyclic process. On the other hand, our results show that the amplitude of economic cycles is higher in developing countries than in the developed ones. At the same time, the maximum cycle asymmetry is observed in the USA as the leading global economy. Considering its role in global economic development, this fact needs to be further studied to understand the true nature of cycles and develop a suitable economic policy.

In general, upon preparation of this paper the issue of asymmetry of economic cycles was subject to a deeper analysis. We associate cycle asymmetry with the fact that the economy is a complex open-ended system that presents a marked synergistic behavior. The economy as a complex open-ended system interacts with other systems of the socioeconomic realm, which provide it with resources and information, and this interaction impel it to move in the direction of the attractor in each economic cycle. This movement is non-linear and accompanied by decline and growth phases during the unfolding of the whole cycle. Information asymmetry is the cause of the asymmetric behavior, that is, differences in depth and height of amplitudes of the economic cycle, and in the growth and declining rates as well, a phenomenon that has not yet been well studied in economics. The conventional technical literature uses simple sine functions to describe and modeling economic cycles, but such representation is rough and inaccurate. To make decisions and forecasting future scenarios it is necessary for the model to correspond to the actual dynamics of economic cycles.

We propose a solution to this problem applying a model of asymmetric economic cycles, which allow to fit real economic cycles with various durations of growth and declining phases, various ratios of the growth period to the declining period, and various amplitudes. Our empirical results demonstrated that the proposed approach is able to describe very closely the real dynamics of economic cycles and, therefore, contributes to the issue of their forecasting, and to the analysis of the state of the economic system as well.

These first results confirm the possibility of analysis of the state of economic systems taking in account the fact that cycles in the global economy, and in individual countries as well, are asynchronous and vary in duration and amplitude (Akaev et al. 2012). We suppose that the proposed approach forms a foundation for a more precise tool of analysis and forecast, and for an adequate representation of the cyclic dynamics of the global economy.

References

Akaev, A. A., Rumyantseva, S. Y., Sarygulov, A. I., & Sokolov, V. N. (2011). *Ekonomicheskiye tsykly i ekonomicheskiy rost* [Economic cycles and economic growth] (456 p). Saint Petersburg: Publishing house of the Polytechnical University.

Akaev, A. A., Rumyantseva S. Y., Sarygulov, A. I., & Sokolov, V. N. (2012). Prichiny dlinnykh voln i problema neravnomernogo ekonomicheskogo razvitiya mirovoy ekonomiki, Kondratievskiye volny: aspekty i perspektivy [Causes of long waves and issue of uneven economic development of global economy. In A. A. Akaev, R. S. Grinberg, L. Y. Grinin, A. V. Korotaev, S. Y. Malkov (Eds.), Kondratieff waves: Dimensions and prospects (384 p.—pp. 110–136). Volgograd: Uchitel Publishing House (In Russian).

Aoki, M. (1996). *New approaches to macroeconomic modeling: Evolutionary stochastic dynamics, multiple equilibria, and externalities as field effects*. New York: Cambridge University Press.

Aoki, M. (1998). Simple model of asymmetrical business cycles: Interactive dynamics of a large number of agents with discrete choices. *Macroeconomic Dynamics, 2*, 427–442.

Aoki, M. (2001). Asymmetrical cycles and equilibrium selection in finitary evolutionary economic models. In L. F. Punzo (Ed.), *Cycles, growth and structural change* (pp. 203–218). London, NY: Routledge.

Arestis, P. (1988). Post-Keynsian monetary economics. Aldershot: Brookfield.

Arestis, P. (2009, May). New consensus macroeconomics: A critical appraisal. Working Paper No. 564. Annandale-on-Hudson: The Levi Economics Institute of Bard College.

Buhm, B., & Punzo, L. F.. (2001). Productivity–investment fluctuations and structural change. In L. F. Punzo (Ed.), *Cycles, growth and structural change* (pp. 47–92). London, NY: Routledge.

Delbeke, J. (1985). Long-term trends in Belgian money supply, 1877–1984. In T.Vasko (Ed.), *The long wave debate*. Berlin: Springer.

Devezas, T. C. (2010). Crisis, depressions, expansions—Global trends and secular analysis. *Technological Forecasting and Social Change, 77*, 739–761.

Devezas, T. C., & Corredine, J. (2002). The nonlinear dynamics of technoeconomic systems: An informational interpretation. *Technological Forecasting and Social Change, 69*, 317–357.

Dosi, G. (1991). *Diffusion of technologies and social behaviour* (N. Nakicenovic, A. Grubler, eds.). Berlin: IIASA, Springer.

Friedman, M., & Friedman, R. (1962). *Capitalism and freedom*. Chicago: University of Chicago Press.

Friedman, M., & Schwartz, A. J. (1982). Monetary trends in the United States and the United Kingdom: Their relations to income, prices and interest rates, 1869–1975.Chicago: University of Chicago Press.

Gladkikh, I. P. (2014). Institutsional'nye podkhody i osobennosti dlinnykh voln v postindustrial'noi ekonomike [Institutional approaches and specific features of long waves in postindustrial economy]. In L. Y. Grinin, A. V. Korotaev (Eds.), Kondrat'evskie volny: dlinnye i srednesrochnye tsikly [Kondratieff waves: Long and medium length cycles] (pp. 276–319). Volgograd: Uchitel (in Russian).

Glaziev, S. Y. (1990). Ekonomicheskaya teoriya tekhnichekogo razvitiya [Economic theory of technological change] (p. 230). Moscow: Nauka (in Russian).

Glaziev, S. Y. (2009). Mirovoy ekonomicheskiy krizis kak protsess smeny tekhnologicheskikh ukladov [The global economic crisis as a process of changing technological structures]. *Voprosy economiki, 3*, 26–38 (In Russian).

Goldstein, J. S. (1988). *Long cycles: Prosperity and war in the modern age*. New Haven: Yale University Press.

Haken, H. (1980). *Sinergetika* [Synergetics] (460 p). Moscow: MIR.

Hansen, A. (1959). *Business cycles and national income, 1951*. Moscow: Izdatelstvo inostrannoy literatury.

Harrod, R. (1959). *Towards a dynamic economics. Some recent developments of economic theory and their application to policy, 1956*. Moscow: Izdatelstvo inostrannoy literatury.

Hirooka, M. (2006). *Innovation dynamism and economic growth. A nonlinear perspective*. Cheltenham: Edward Elgar.

Juglar, C. (1862). *Des Crises commerciales et leur retour periodique en France, en Angleterre, et aux Etats-Unis*. Paris: Guillaumin.

Keynes, J. M. (2007) *General theory of employment, interest and money, 1936* (960 p). Moscow: Exmo.

Kitchin, J. (1923). Cycles and trends in economic factors. *Review of Economics and Statistics, 5*(1), 10–16.

Kleinknecht, A. (1987). *Innovation patterns in crisis and prosperity:Schumpeter's long cycle reconsidered*. London: Macmillan.

Knyazeva, Y. N., & Kurdyumov, S. P. (2002) Osnovaniya sinergetiki. Reghimy s obostreniem, samoorganizatsiya, tempomiry [Foundations of synergetic: Blow-up regimes, self-organization, tempoworlds] (414p.). Saint Petersburg: Aletheia.

Kondratieff, N. D. (1922). Mirovoe khozyaystvo i ego kon'yunktura vo vremya i posle voyny [The world economy and its conjunctures during and after the War]. Vologda: Regional Branch of the State Publishing House (in Russian).

Kondratieff, N. D. (1925). Bol'shie tsikly kon'iunktury [Major cycles in conjuncture]. *Voprosy kon'yunktury* [Issues of Conjuncture], *I*(I), 28–79 (in Russian).

Kondratieff, N. D. (1928). Bol'shie tsikly kon'yunktury [The major cycles of the conjuncture]. Moscow: Economika (in Russian).

Kondratieff, N. (1935). The long waves in economic life. *Review of Economic Statistics, 17*(6), 105–115.

Korotayev, A. V., & Tsirel, S. V. (2010). Kondrat'evskie volny v mirovoi ekonomicheskoi dinamike [Kondratieff waves in world economic dynamics]. In *Sistemnyi monitoring. Global'noe i regional'noe razvitie* [System monitoring. Global and regional development] (D. A. Khalturina, A. V. Korotaev, Eds.) (pp. 189–229). Moscow: Librokom/URSS (in Russian).

Korpinen, P. (1985). *A monetary model of long cycles*. In T. Vasko (Ed.), *The long wave debate*. Berlin: Springer.

Kuznets, S.. (1926). *Cyclical fluctuations, retail and wholesale trade. United States. 1919–1925*. New York: Adelphi.

Lucas, R. E. (1977). Understanding business cycles. In A. H. Meltser (Ed.), *Stabilisation of the domestic and international economy*. Amsterdam: North Holland.

Lucas, R. E. (1987). *Models of business cycles*. Oxford: Basil Blackwell.

Maddison, A. (2010). Historical statistics of the world economy. Retrieved January 8, 2015, from http://www.ggdc.net/MADDISON/oriindex.htm

Mensch, G. (1973, January). *Theory of innovation*. Berlin: International Institute of Management.

Mensch, G. (1979). *Stalemate in technology—Innovations overcame the depression* (G. Mensch, Ed.). New York: Ballinger.

Metcalfe, S. (2007) Replicator dynamics, Elgar companion to Neo-Schumpeterian economics (H. Hanusch, A. Pyka, Eds.) (pp. 440–452). Cheltenham: Edward Elgar.

Mitchell, W. C. (1959). Business cycles and their causes. University of California Press, Los Angeles.

Modelski, G. (1987). *Long cycles in world politics*. London: Macmillan.

Nakicenovic, N. (1987). *Transportation and energy system in the US//WP-87-1*. Laxenburg: IIASA.

Osipov, Y. M., & Shurgalina, I. N. (1994). *Perekhody i katastrofy: opyt sotsialno-ekonomicheskogo razvitiya* [Transitions and catastrophes: Social and economic development] (192 p). Moscow: Izdatelstvo Moskovskogo Universiteta.

Ostapenko, V. M. (2013). Sovremenniy meinstrim i politicheskaya ekonomiya: spetsifika issledovaniya ekonomicheskikh krizisov, Natsionalniye ekonomiki v globalnom mire: politicheskaya ekonomiya i ekonomicheskaya politika. In A. V. Buzgalin, M. I. Voeykov, M. O.Yu., V. T. Ryazanov (Eds.), *Materialy Pervogo Mezhdunarodnogo politekonomicheskogo kongressa* [Modern mainstream and political economy: Specifics of studies of economic crises], National economies in the global world: political economy and economic policy. Materials of the First international political and economic congress (Vol. 2, 552 p., pp. 490–493). Moscow: Lenard (in Russian).

Preskott, E., & Kudland, F. (1982). Time to build and aggregate fluctuations. *Econometrika, 50*, 1345–1371.

Prescott, E., & Kudland, F. (1990). Business cycles: Real facts and monetary myth. FRB Minneapolis Quarterly Review, Available at: http://www.nber.org/

Prigogine, I. & Stengers, I. (1986) Order out of chaos: Man's new dialogue with nature (V. I. Arshinov, Y. L. Klimontovich, Y. V. Sachkov, Trans., Ed.) (432 p). Moscow: Progress.

Rumyantseva, S. Y. (2003). Long wave in economy: Multifactor analyses. Saint-Petersburg: St. Petersburg University Publishing House (in Russian).

Sargent, T. J. (1986). *Rational expectations and inflation*. New York: Harper and Row.

Shumpeter, J. (1912). Theorie der wirtschaftlichen Entwicklung.: Nachdruck der 1. Auflage von.

Shumpeter, J. (1939). *Business cycles: A theoretical, historical and statistical analysis of the capitalist process* (Vol. 1). New York: McGraw-Hill.

Solomou, S. (1990). *Phases of economic growth 1850–1973*. Cambridge: Cambridge University Press.

Solomou, S. (2001). Economic cycles since 1870, cycles, growth and structural change (L. F. Punzo, Ed.) (pp. 3–26). London: Routledge.

Soto, J. H. (2008). Money, bank credit, and economic cycles (Kuryaev, A. V., Trans.) (663 p). Chelyabinsk: Sotsium.

Stigler, G. J. (1961). The economics of information. *Journal of Political Economy, 69*(3), 213–225.

The Conference Board Total Economy Database. (2014). Retrieved January 8, 2015, from http://www.conference-board.org/data/economydatabase

Tsvetkov, V. A. (2013). Tsykly i krizisy. Teoretiko-metodologicheskiy aspekt [Cycles and crises. Theoretical and methodological aspects] (in Russian). Saint Petersburg: Valeri.

Tugan-Baranovsky, M. I. (1997). *Periodicheskiye pormyshlenniye krizisy. Istoriia angliiskikh krizisov. Obshchaia teoriia krizisov* [Periodic industrial crises. History of english crises. General theory of crises]. Moscow: Nauka.

Van Duijn, J. J. (1983). *The long wave in economic life*. London: George Allen & Unwin.

Vasko, T. (1985). The long wave debate. Berlin: Springer.

Verick, S., & Islam, I. (2010) The great recession of 2008–2009: Causes, consequences and policy response. Discussion paper No. 49–34. Bonn: IZA.

Woodford, M. (2010). Convergence in macroeconomics: Elements of the new synthesis. *Voprosy economiki,* No. 10, pp. 17–30.

Yakovets, Y. V. (1999). *Tsycly. Crizisy. Prognozy* [Cycles. Crises. Forecasts]. Moscow: Nauka.

Zamulin, O. (2005) Kontseptsiya realnikh ekonomicheskikh tsyklov i eyo rol' v evolutsii makro-ekonomicheskoy teorii [The concept of real economic cycles and its role in evolution of macroeconomics]. *Voprosy economiki,* No. 1.

Modern Trends in Evaluation of Macroeconomic Structural Changes

Igor A. Yevsikov, Konstantin O. Korovin, and Askar I. Sarygulov

Abstract The evolution of social-economic systems is nonlinear, containing both periods of gradual changes and swift transformational upswings. Seven years after the end of the last crisis of 2007–2009, recovery rates of the US and European economies remain extremely low. One of the reasons is structural-economic crisis, requiring considerable time and substantial financial resources to overcome. By the start of the crisis, all developed economies of the world, particularly US economy, had well pronounced structural allometry with disproportionally high share of the financial and banking sectors.

In this paper we propose new approaches for the evaluation of structural shifts, based on the attractor theory, the use of proportionality factor and Wold-Mensch' nonlinear model. Our results allow the construction of structural changes monitoring system and also demonstrated, that there always exist certain states in economic system, in which the volume of production (GDP) exceeds their empirical value. The proposed system for the monitoring of structural changes can be used not only for research purposes, but also for structural shift management.

Keywords Industrial and technological structures • Nonlinear attractor • Structural monitoring systems

I.A. Yevsikov
Department of Applied Mathematics and Informatics of the Saint-Petersburg State University of Architecture and Civil Engineering, 4 Vtoraya Krasnoarmeyskaya, Saint-Petersburg, Russia
e-mail: evsikov.i.a@lan.spbgasu.ru

K.O. Korovin
Radiophysics Department, Saint Petersburg Peter the Great Polytechnic University, Saint-Petersburg, Russia
e-mail: konstkor@mail.ru

A.I. Sarygulov (✉)
Saint Petersburg State University of Architecture and Civil Engineering, Doctor of Economic Sciences, 4 Vtoraya Krasnoarmeyskaya, Saint-Petersburg, Russia
e-mail: asarygulov@lan.spbgasu.ru

T. Devezas et al. (eds.), *Industry 4.0*, Studies on Entrepreneurship, Structural Change and Industrial Dynamics, DOI 10.1007/978-3-319-49604-7_5

1 Introduction

Economic progress of industrially developed countries in the last decade of the twentieth century and in the beginning of this century is characterized by extremely low rate of economic growth. The so-called "golden years" of rapid global economic growth (1947–1982), when the highest growth rates of labor productivity were observed, are in the past. China and India—the countries with largest domestic markets and fast growing new high-tech industries of economy increasingly became the engines of world development. Globalization process not only facilitated opening of economic borders, but have also made the latest achievements in information and communication technologies available to many countries. The latter significantly changed the speed of information exchange and the pattern of cooperation and trade relations. Qualitatively new role of technological and industry structure of economy becomes another feature of modern economic development. Basic innovations of new industrial and technological platforms have not yet become an object of close attention of investors, while national economies are in various phases of structural commitment. All this occurs against the background of recurrent crises and structural imbalance.

Structural changes accumulate at micro level, since the innovations and new technologies continuously change industrial structure and, as consequence, the whole industrial tissue of the economy. The so-called "invisible hand of the market" doesn't always provide "correct" signals to economic entities on the place, time and volume of new investments, thereby creating prerequisites for accumulation of structural disproportions. In this regard, the definition of correct time for change in investment strategy of the companies and "structural readiness" of industries for development under new conditions (new products, new customer requirements, new technologies, new methods of management, etc.) is crucial.

The economy of the USA only during post-war years (1945–2009) has come through 11 crises of different durations. The most destructive were the crisis in the mid-1970s and the last crisis (2007–2009), when the recession phase was aggravated by the structural imbalance of economic system in general. Thus, the longest period of growth of the American economy in the twentieth century has hit the period from 1991 to 2001 and comprised exactly 10 years. Such stability of development was largely due to technology factor—the boom in the field of information and communication technologies facilitated the development of financial markets, provided growth of labor productivity and creation of jobs. The subsequent crisis in the field of IT technologies has put a question of the reasons of sharp recession and character of a new wave of development. The last crisis (2007–2009), has been predefined not only by the existence of "bubbles" in the financial sphere, but also by allometric nature of industry development when disproportionately high shares in national economies accounts for certain industries (finance, services), while the share of other industries (first of all manufacturing ones) has catastrophically reduced.

2 Structural Changes Research Technique

The main feature of modern economic investigation is the strict separation of macro— and microeconomic studies. In fact, both systems work together; mutually influencing each other and it is only their joint interaction that defines trends and system oscillations in general. In the long chain of economic processes, one has to find a key link, which connect the two levels of economy in a natural way. The structure of economic systems could serve as such connection link. Since the dynamics of structural changes is predetermined by economic activity of economic agents themselves, entrepreneur's demand for new ideas and their implementation rate or, on the contrary, the complete absence of them, will act as the key factors of the changes. In the first case the creation and development of new industries is observed, while in the second—stagnation and dying out—that is not the case, because the recession phase is "triggered" by the sharp decline in business activity of economic agents, while growth phase is provided by two factors—successful commercialization of new ideas by the entrepreneurs and interference of state (institutional reforms).

Structural changes have always been of interest to economists, with pioneering studies by Clark (1957), who was the first to single out and predict explosive growth of service sector. In the 1940s and 1950s, V. Leontief described the dynamics of structural changes in the US economy based on the developed "cost-output" models (Leontief 1941, 1953). One should point out that the technique, proposed by him, was widely applied later (Carter 1970; Sonis et al. 1996; Guo and Planting 2000). Industrial shifts in the economy of India in 1950–2000 were studied with the help of "cost-output" models (Dasgupta & Chakrobarty 2005). On the basis of aggregation of 72 industries in three-sector model (high intensive primary, high intensity technology, high field of structural changes in economy belonging to the English-Australian economist C. Clark-intensity "capital-labor"), interrelations and mutual influence of industrial structural shifts were estimated. Matrix models also comprise the basis of structural changes' studies in energy sector of China (Kahrl and Roland-Horst 2009) and estimation of the diffusion process of new technologies (nuclear energy) based on logistic functions and their connection to technology coefficients in "cost-output" matrix in Chinese economy. This class of models is also widely used for the analyses of structural transformation processes in economies of different type (Kei-Mu and Zhang 2010), analysis of relations between technology advance and saving ratio for two-sector economy (Laitner 2000). The most recent studies of structural and sectorial changes in the economies of five countries of South-East Asia (Indonesia, South Korea, Malaysia, Philippines, Thailand) for the time interval of 1975–2000, carried out by UNIDO, were also based on the models of inter-industrial balance (Haraguchi and Rezonja 2010).

Evolutionary changes in economic structure in Western Europe were analyzed by a group of Soviet economist in the late 1980s on the basis of "indicator of difference" (Kuznetsov 1988). Similar approach was proposed for the analysis of structural changes in the Japanese economy on the basis of "measure of structural changes" indicator (Natuhara 2008). Spatial geometric model for estimation of industrial structural changes was proposed by V.V. Kossow (Kossow 1975) and later developed in several works (Butina 1980; Minasyan 1983; Spasskaya 2003). Similar approaches were widely used in the UN practice (Industry and Development 1985). In the early 1980s, a model, suggesting new approaches to measuring structural changes over ling time intervals was suggested (Barkhin and Chesnokov 1983).

Modern stage of investigation of structural changes is determined by a high degree of detail in objects of interest, when significant efforts of scientists are directed at investigation of individual factors, defining the pattern and direction of changes. It is not accidental that many studies are aimed either at the search of empirical data, proving the validity of certain approaches, or, at the construction of models, capable of extending selection of research tools. A separate direction of studies is the retrospective structural analysis of national economies and compilation of industrial and technological structure forecast for groups of countries (OECD countries, BRICS, G-20). The studies of the impact of foreign trade on structural change, variation of cycles' length due to structural change, dynamics of labor market and its impact on structural change, degree of openness of national economy and structural changes in it singled out as independent fields of studies (Gaston and Nelson 2004; Oulton and Srinivasan 2005; Iiduka 2007; Industrial Development (2007); Nickell et al. 2008; Memedovic 2009; Robson 2009; Varum et al. 2009; Carroll 2012; Gawrycka et al. 2012; Thompson et al. 2012; Wang and Zhou 2014; Green and Stern 2015; Naude and Szirmai 2015). When studying structural changes, it is commonplace to assume inhomogeneity of technological development in different industries and unequal preferences of consumers to be the main reasons. This is confirmed by a huge volume of studies of microeconomic type, where main object of studies are innovations, management systems, and development strategies at the level of firms and companies.

From our point of view, high number of studies doesn't guarantee possibility of wide applications of the obtained results. In this connection, we argue that from the practical point of view it is important to create a system of macroeconomic monitoring of structural changes, which would implement a three-stage procedure:

- estimation of direction of structural changes;
- estimation of industrial and technological structure of economy;
- estimation of degree of imbalance of existing structure of economy and development of measures for its correction;

Let us briefly describe such system.

3 Estimation of Direction of Structural Changes Based on Attractors

Non-equilibrium is just a fundamental property of economic system, like equilibrium: it allows determination of a free set of optimization synthesis from the whole spectrum of possible directions. While equilibrium state is a necessary condition of stationary existence of economic systems, non-equilibrium state is an essential moment of transition to a new state, in which economic system acquires higher level of organization and productivity. It is only then, when the system loses functional stability, that self-organizing processes of the formation of new efficient structures appear.

When studying economic systems, which are not only complex, but also dynamic, determination of stable points, which are most typical of given dynamic system and the set of which determines the direction of development, is of extreme importance. Attractors are well developed mathematical tool for such studies.[1]

Dutch mathematicians (Ruelle and Takens 1971) have shown that there exist strange attractors, which are mathematical image of complex systems. In this case, system trajectories perform arbitrary walks, not subject to regular description, within certain domain.

Any self-organizing comprehensive system has its own attractor, which is the state, created together with environment, which could have been reached, had all initial conditions of internal and external media been stable during movement of the system towards its goal. Under natural conditions, however, on their way to attractors, systems undergo some random or certain events, which immediately change the attractor—attractive goal.

The main peculiarity of attractor's action is that it makes the future or previous states of the systems determinate. The language of attractors allows insight into phenomena of predictability and fundamental unpredictability, provides understanding of probabilistic, chaotic behavior of systems, determined not by the scarcity of research possibilities, but by the nature of nonlinear system itself.

In order to study structural changes of macroeconomic systems we used attractors, built for the simplest case of two equations (volume of added value and sector's share in added value over long time interval).

This logistic functions can be written as follows:

– for the dynamics of volume of added value:

$$\widetilde{y} = \frac{A}{1 + b \cdot e^{-c \cdot t}} \tag{1}$$

– for the dynamics of shares:

$$\widetilde{\delta} = \frac{A}{1 + b \cdot e^{-c \cdot t}} + \lambda \tag{2}$$

The following step is the construction of phase plane and, attractors of dynamic systems, respectively.

[1]Finite domain of unavoidable convergence of phase trajectories of complex system movement is called attractor in synergetic. Point (stable focus) or another system that is more complex can stay in the capacity of attractor.

Figure 1 present the following charts: (a) is the variation of production volume, (b) is the dynamics of sector's share in added value, (c) is the phase plane itself with attractors. Charts for manufacturing industries of Japan and Korea are presented in Fig. 1.

By making use of such approach, we performed the analysis of industrial structure of GDP of ten OECD countries—members (US, Canada, Japan, Korea, Finland, Italy, Sweden, Austria, Spain, Germany)[2] between 1970 and 2010. Some industries have clearly visible trend of decreasing their shares in GDP (first of all, agriculture), while other, on the contrary—sustainable growing trend (for example, finances). At the same time, there are sectors, where their GDP shares are subject to insignificant fluctuations during the time period at issue (trade, construction, transportation, electric-, natural gas-, water supply, mining industry).

Main result, obtained at given stage of research comes down to the fact, that phase planes demonstrate simultaneous changes of three indicators: limits of sectors' share in added value and limits of production volume; direction of change (sector's share can increase or decrease) and rate of change (attractor's tilting angle in phase plane). Of course, the nature of process is determined by retrospective data, but it is highly important for the compilation of short-term and middle-term forecasts of industrial dynamics. Since structural reforms require long time, correct estimation of initial state of the system and time intervals, when structural reforms required for sustainable development can be carried out, is exceptionally important.

In this study, we have chosen 1970 as starting point as the one preceding deep structural crisis, which started with sharp increase in oil prices in 1973. It is important that the world economic crisis of the 1970s was the first call about need look insight into the existing policy of energy consumption. Sharp growth of oil prices has led to considerable economic recession, which has been overcome, by many countries only at the expense of measures of structural type when individual power-intensive industries of economy were closed and new, less power-intensive were created. Japan is spectacular example. The energy crisis of 1973 was extremely painful for its economy, first of all due to the fact that its manufacturing industry had the highest energy consumption among industrially developed countries. The government of the country has quickly developed the plan for scraping (elimination) of excess capacities in manufacturing industry. For example, open-hearth steelmaking and production of electrical furnace steel in ingots were subject to such scraping (16 % or 2.3 million tons), as well as aluminum production (24 % or 390 thousand tons), production of nitrogen fertilizers (30 % or 2.5 million tons) and some others, and for implementation of this plan special fund has been founded with the guarantee limit of 100 billion yens (Lebedeva 1986). All this has allowed to make important changes in industrial structure of production and to provide transition to new model of growth when the emphasis has been placed on development of the knowledge-intensive productions and so-called system industries—systems of cars of trade information, systems of information service of

[2]In the capacity of statistical base we used database from http:/www.oecd.org/statisticdata/

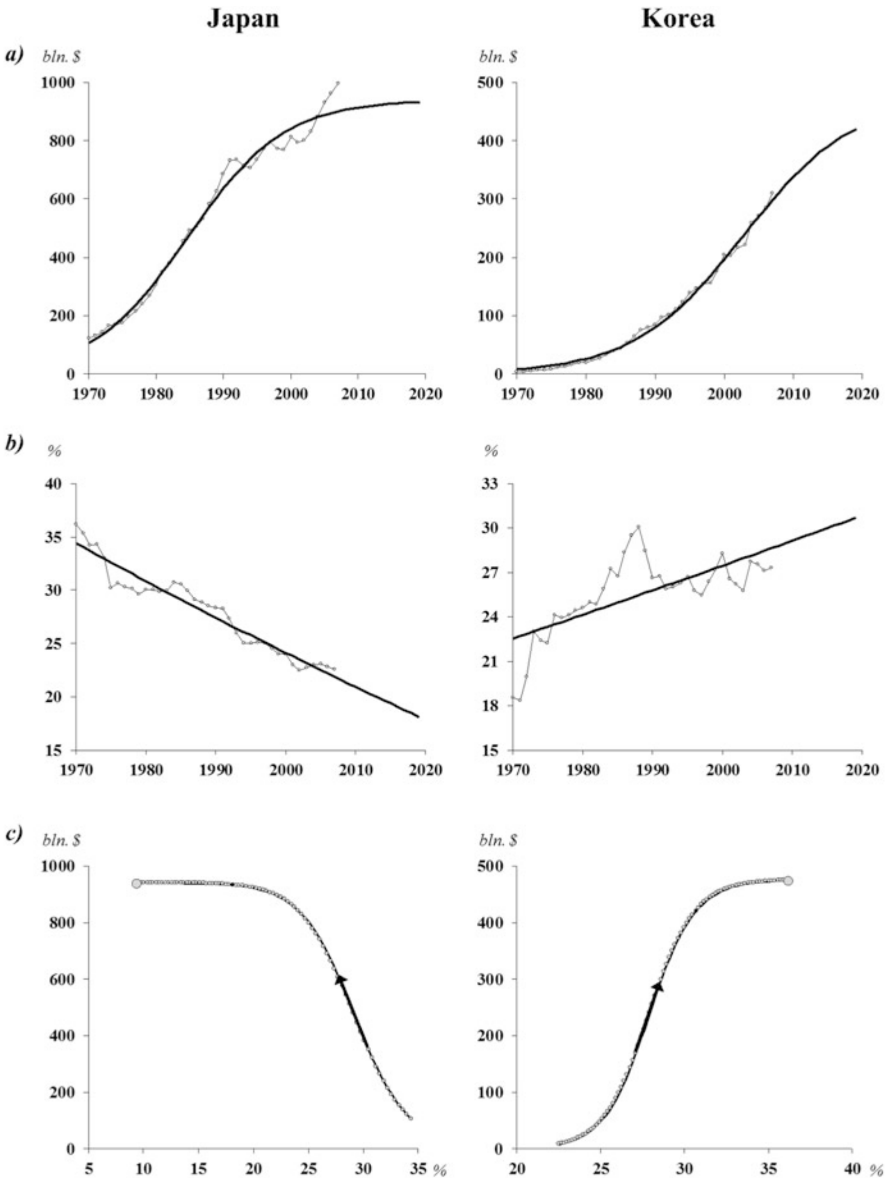


Fig. 1 Manufacturing sector VA and share dynamic, non-linear attractors (Japan and Korea)

hotels, systems of the control equipment for the subway, etc. Altogether, it has allowed Japan to begin to become in the early 80s the world's largest producer of electronic components and semiconductors, with allowance for the production at foreign branches of the Japanese firms its share in world production has comprised 60–70 %.

Nevertheless, transition to a new step of technological development hasn't solved all problems of manufacturing industry of Japan. The structural disproportions which have accumulated for two decades of rapid growth (1955–1973) have turned out so to be deep that to the middle of the 1980th "they haven't been completely overcome yet and continued to complicate development of the Japanese industry" (Lebedeva 1986).

The development path of the Korean industry was to an extent similar to the Japanese, but with a lag of 10–15 years. The starting conditions preceding economic reforms in the early 1960s were extremely hard, for example, per capita GDP comprised 87 dollars in 1962. High growth rates of GNP in 1962–1989 (8.5 % per year) to a great extent have been reached at the expense of manufacturing industry, where growth rate of production exceeded that of GDP. Thus, in 1962–1966 they comprised 13.5 %, in 1967–1971—22.5 %, and in 1972–1974—26.8 % (Harvie and Lee 2003; Lebedeva 1986). The government of South Korea carried out economic reforms rationally combining market mechanisms and methods of state regulation, paying special attention to achievement of the balanced industrial structure. Consistently developing industries from the mining and consumer goods manufacturing to heavy and microelectronics, increasing the volume of export and increasing competitiveness of the goods, introducing advanced technologies and diversifying production, Korea has achieved impressive progress: in 1981 the volume of export has made 20 billion dollars or one third of GNP of the country, the share of industrial production has reached 40.9 % of GNP, while average annual growth rates of manufacturing industry comprised 17.6 % in 1960–1970, and 17.8 % in 1970–1979. Peculiarity of the Korean way of industrialization were the high volumes of investment which averaged 27–29 % of GNP in 1970–1985, while the share of domestic investments steadily grew over the years, reaching almost 90 % by 1985 (Lebedeva 1986).

The assessment of a direction and nature of such changes is extremely important, but only the initial step in studying structural shifts. The economic analysis of individual organizational, institutional, technological decisions allows receiving fuller picture of structural transformations. However, to overcome "pattern effect" of structural changes we need one more level of research—an assessment of qualitative characteristics of structural and their influence on resultant indicators of activity of macroeconomic systems, which is GDP growth rate.

4 Estimation of Quality of Industrial and Technological Structure of Economy

The remark of the American economist F. Fischer that ". . . overwhelming part of empirical studies in the field of econometrics consists of estimation of structures; in other words, estimations of parameters of the dependences expressing technology or economic behavior, general content of which follows from the theory of mathematical economy" (Fisher 1966). On the basis of such approach Herfindahl-Hirschman index has been constructed (Herfindahl 1950; Hirschman 1945).

During construction of this index, one starts from a number of the following assumptions:

- probability that any specific firm will start successful project is proportional to the share of its R&D expenses in total expenses of industry;
- total spending of industry is proportional (corresponds) to the total capital of the industry;
- share of R&D of the companies corresponds to the share of fixed capital of the companies in the total capital of industry;
- percent of industrial capital of branch which can be modernized, is also proportional to percent of company's capital in total capital of industry (Nelson and Winter 1982).

Starting from all these assumptions the index of concentration of the capital (Herfindahl-Hirschman) is calculated, which is written as follows:

$$H = \sum_i \left(\frac{K_i}{K} \times \frac{K_i}{K} \right) = \sum_i \left(\frac{K_i}{K} \right)^2 \tag{3}$$

The construction of this index has turned out to be the most fruitful for studying structural dynamics within the evolutionary theory of economic changes, because it allowed by means of computer modeling provide forecasting characteristics of innovative and technological changes in virtually of any branch of economy (manufacturing, agriculture, transport, services, finance etc.).

As the economic statistics reveals, in real economy, different industries have essentially different specific weights and different impact on the formation of resulting indicators. As a result, with universality this index is almost not suitable for an assessment of macroeconomic structural dynamics, when several branches are taken into account (the international standard system of statistics includes no less than nine industries on the basis of which GDP of a country is calculated).

More universal, in our opinion, the method of an assessment of various structures (sets) on the basis of the principle of measurement of proportionality was proposed in paper (Vatnik 2011). This approach allows to create a certain reference assessment which can be later used in the research of dynamics of change of the structure itself. The coefficient of proportionality introduced in such way has the following form:

$$\mathrm{Prop}[X, Y] = \frac{\overline{XY}^2}{\overline{X^2}\ \overline{Y^2}}, \text{or}\ \mathrm{Prop}[X, Y] = \frac{\left(\sum_i X_i Y_i \right)^2}{\left(\sum_i X_i^2 \right) \left(\sum_i Y_i^2 \right)} \tag{4}$$

Table 1 Sample value of GDP industrial structure

Industry name	Share, %
Agriculture, hunting, forestry and fishing	2.2
Mining	0.9
Electricity-, gas- and water supply	2.5
Construction	6.2
Wholesale and retail, restaurants and hotels	14.1
Transportation, warehouses and communications	7.5
Finances, insurance, real-estate and business services	25.4
Services: individual, social and public	21.4
Processing industry	19.8
Total	100.0

is structural characteristic, defining degree of mutual proportionality of variables, which has the following basic properties:

- irrespectively of the dimension of variables, it is always dimensionless;
- at any a, b > 0 equation Prop [aX, bY] $=$ Prop [X, Y] holds true;
- possible values of proportionality coefficient falls within the range$0 \leq$ Prop $[X, Y] \leq 1$.

The meaning of proportionality coefficient of proportionality represents a numerical measure of proximity between the estimated structure and reference. At their full coincidence Prop [X, Y] $=$ 1. Unlike Herfindahl-Hirschman index, it is not focused on equality of shares of all components, and allows setting of reference structure proceeding from qualitative reasons.

Owing to the above reasons we adopted proportionality coefficient as the main indicator which is most precisely reflecting the dynamics and nature of structural changes in macroeconomic level.

Application of proportionality coefficient for an assessment of qualitative characteristics of the structures at issue, requires the construction of a certain reference, where to the estimated structure should tend.

For industrial structure of GDP, we have chosen average value of share of each industry in GDP over the studied sample at the end of the period at issue—2010 as a reference (Table 1).

The choice of "average" is due to inability to range industries with regard to their importance, since it contradicts the logic of economic development: all industries are equally important for daily activities. At the same time, we cannot choose a country's GDP industrial structure as an example (e.g. that of USA, Japan or Germany), because then we disregard peculiarities of national economies and diversity of economic environment surrounding us. Calculation results of GDP industrial structure are presented in Fig. 2.

One can see in Fig. 2, that most impressive quality of growth and, respectively, structural dynamics, is demonstrated by Korea. With the lowest starting indicators, the country managed in three decades to approach reference indicators of GDP

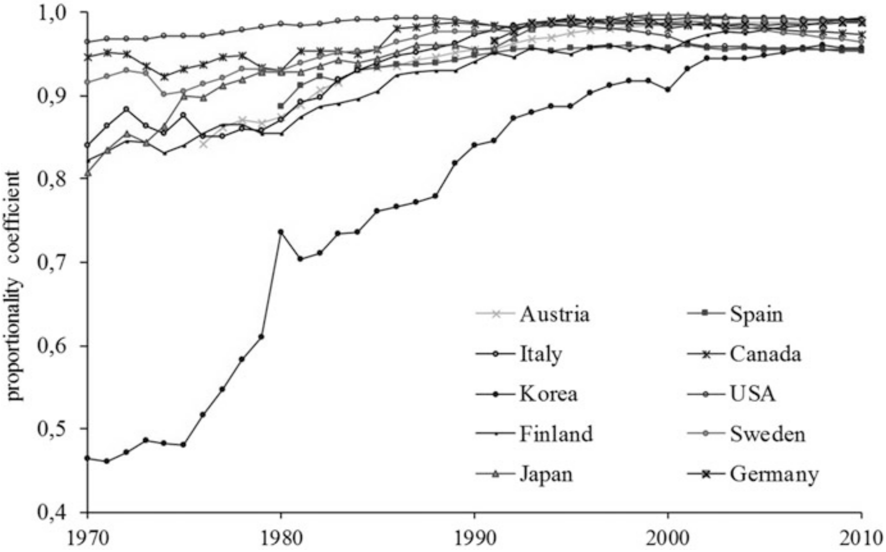


Fig. 2 The dynamic of national economy structure quality

industrial structure. Other countries improved industrial structure of GDP, but at lower rates. The best starting position was in the US, the economy of which could have been considered an example for over 20 years.

While in proving the standard for branch structure we were based on industrial average shares in GDP, but when choosing the example of technological structure, we, most likely, need different approach. In 2007, European Commission has adopted the system of classification of processing industries by the level of their technological development. According to the system, all manufacturing industries can be divided into four groups:

A. High-tech: (office machinery and computers; radio, television and communication equipment and apparatus; medical, precision and optical instruments, watches and clocks; aircraft and spacecraft; pharmaceuticals, medicinal chemicals and botanical products).
B. Medium-high-tech industries: (machinery and equipment; electrical machinery and apparatus; motor vehicles, trailers and semi-trailers; other transport equipment; chemicals and chemical products excluding pharmaceuticals, medicinal chemicals and botanical products).
C. Medium-low-tech: (coke; refined petroleum products and nuclear fuel, rubber and plastic products, non-metallic mineral products; basic metals; fabricated metal products except machinery and equipment; building and repairing of ships and boats).
D. Low-tech: (food products and beverages; tobacco products; textiles; wearing apparel, dressing and dyeing of fur; tanning and dressing of leather; manufacture of luggage, handbags, saddlery and harness; wood and products of wood

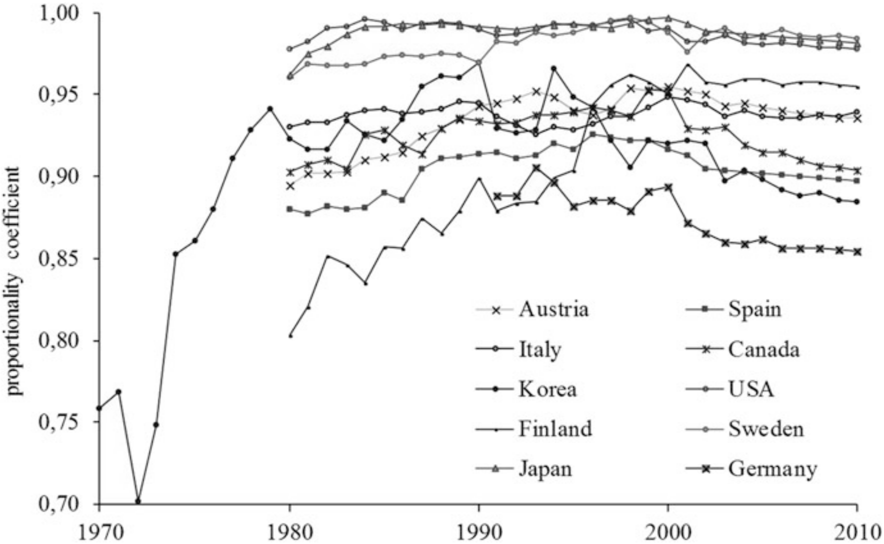


Fig. 3 The dynamic of manufacturing industry's technological structure quality

and cork, except furniture, pulp, paper and paper products; publishing, printing and reproduction of recorded media; furniture and other manufacturing; recycling).

First of all, common sense suggests that structure with high share of high-tech and medium-high-tech industries should be an example. At the same time, we can't completely eliminate medium-low-tech and low-tech industries from economy, since they are present in every economy, including developed ones, and make significant contribution to the creation of national wealth. Under these assumptions, we adopted the following ratio as sample technological structure: share of high-tech and medium-high-tech industries should aim to 50 %, with high-tech industries comprising about 20 %. As far as medium-low-tech and low-tech industries are concerned, their share should amount to about 50 %, with dominance of low-tech industries (Fig. 3; Table 2).

First of all, we should point out dynamic development of technological structure of Korean industry (as in the case of industrial structure). Technological structure of Japan is virtually perfect. As far as other countries are concerned, they obviously have unstable rates of modernization of technological structure. This is what we observe in Italy, Finland and Germany (after reunification). The industry of Russia corresponds to the sample only to a slight extent.

Table 2 Reference technological structure

Group of industries	Share, %
High-tech production	19
Medium-high-tech production	28
Medium-low-tech production	21
Low-tech production	32
Total	100

5 Evaluation of the Imbalance of Economic Structure

Upon analysis of processes of economic dynamics, it is crucially important to identify the period of changes in the trend of economy development. As macroeconomic processes are particularly inertial, there is always a significant lag between the beginning of changes in the trend and its representation in the form of changes in the key economic indicators. It manifests most intensely during recessions and crises when structure imbalance deteriorates general economic conjuncture. One of approaches to timing of changes in negative economic trends is proposals developed at the intersection of economics and management practice. In 1975, German professor Gerhard Mensch proposed a comprehensive theory of basic innovations (Mensch 1975) as a trigger to overcome depressions. As H. Wold and K. Kaasch reasonably mentioned: "Mensch's historical studies of economic evolution give evidence of Schumpeterian fits and spurts. They show that the social inventions and scientific discoveries that underlie basic innovations flow in a more or less continuous stream, whereas the basic innovations themselves cluster in relatively short "windows in time". There is gradual (continuous) emergence of possibilities for innovation, going hand in hand with speed-up and slow-down (discontinuity) in the rate and direction of basic innovation." (Wold and Kaasch 1985). As a follower of ideas of J. Schumpeter on the key role of innovations and wave nature of economic processes, G. Mensch proposed a series of models characterizing cascade character and metamorphosis nature of economic processes. He developed one of such models together with Hermann Wold and described it in two papers published in 1983 and 1985 (Wold and Mensch 1983; 1985). Mensch's model of long-term fluctuations represents an alternative to Kondratieff's wave model and includes the system of bi-equilibrium (Mensch et al. 1980; Mensch 1982). Two equilibriums of Mensch's BIEQ model represent the phase of production and employment growth expansion, as well as the capital-intensive phase of production growth per one worker, and the phase of decline in employment per production unit. The main contribution is that during technology stalemate development turns into bifurcation. It followed by a series of innovations, which tends to provide the worst results and less and less benefit, while companies and plants in established industries mature and become less flexible. If lack of flexibility is represented by excess amount of industries (structure setting), the economy is structurally ready for breakthrough of a new cluster of innovations. It is necessary to mention that at the time these models were only formulated and had not been checked against particular economic data (Bergstrom and Wold 1983).

In 1985, a paper was published (Wold and Kaasch 1985), in which this model was checked against real data of the US economy for the period of 1947–1977 for the following three industries: manufacturing, mining and transport.

Let us examine the BIEQ economic model in detail.

$$X_t^3 = R_t X_t + E_t \tag{5}$$

$$\begin{bmatrix} A_t \\ I_t \end{bmatrix} = \begin{bmatrix} \cos\gamma & \sin\delta \\ \sin\gamma & \cos\delta \end{bmatrix} \begin{bmatrix} E_t \\ R_t \end{bmatrix}; \tag{6a}$$

$$\begin{bmatrix} E_t \\ R_t \end{bmatrix} = \frac{1}{\cos(\gamma-\delta)} \begin{bmatrix} \cos\gamma & -\sin\delta \\ -\sin\gamma & \cos\delta \end{bmatrix} \begin{bmatrix} A_t \\ I_t \end{bmatrix} \tag{6b}$$

X_t, the solution of Eq. (5), evaluates P_t production output.

The E_t—expansive investment and R_t—rationalization investment are obtained by oblique rotation from the known values of A_t labor and I_t investment cost (6a, b).

P_t, A_t, I_t values are preliminarily reduced to a zero mean and unit square deviation according to the following formula:

$$\xi_t^* = \frac{\xi_t - M(\xi_t)}{\sqrt{D(\xi_t)}} \tag{7}$$

$M(\xi_t)$—expected value, $D(\xi_t)$—variance.

Let use numerical methods with the aid of Matlab package and find γ and δ, upon which solution[3] of Eq. (5), X_t, has the minimum square deviation from P_t value:

$$\Phi(\gamma,\delta) = \frac{1}{T}\sum_t (P_t - X_t)^2 \to min_{\gamma,\delta} \tag{8}$$

The data are based on the US statistics for 1947–1997. Let us change γ and δ values from 0 to 2π with increments of $\pi/30$.

In the first experiment let us take values for P_t, A_t, I_t from the data on Value Added, Compensation of employees, Investment for three industries in the whole (manufacturing, mining, transport). The obtained values $\Phi(\gamma,\delta) < 0.3$ are shown in Fig. 4.

Let us find such γ, δ upon which the solution of Eq. (5) provides higher values of X_t theoretical production output with comparison to P_t real values. For each pair γ, δ let us find coefficient $k \in [0, 1]$, where k is an indicator characterizing frequency of time periods when in a found solution a theoretical value is higher than a real value:

[3] Let us find solution of Eq. (5) with the aid of **fsolve** optimization function in proximity to P_t P_t known values.

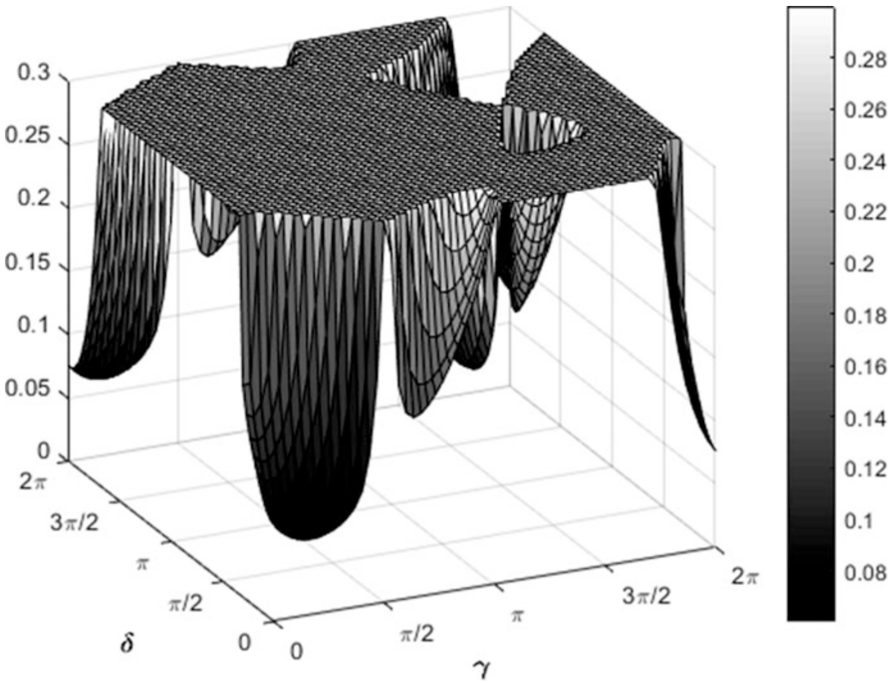


Fig. 4 $\Phi(\gamma, \delta)$ for three industries in the whole

$$k = \frac{N}{T}, \text{ where N is amount of times when } X_t \geq P \tag{9}$$

If $k = 1$, then an obtained solution is better than real values on the whole time interval. Upon $k = 0.9$ each solution has only 1 from 10 time intervals when a theoretical value is lower than a real value.

Let us examine development dynamics of the manufacturing industry of the USA for the period of 2003–2012 in detail. Firstly, let us find k coefficients for all pairs γ, δ varying from 0 to $2\pi^4$ and represent those graphically in a table with different colors. Values close to 0 are highlighted in blue, values close to 1 are highlighted with red. In Fig. 5, at the right, we observe a set of obtained solutions upon $k = 0.9$ and real values of P_t production output (red line). Unfortunately, we could not manage to find solutions with the coefficient of 1 in the examined decade.

5.1 Disaggregation and Structural Technological Shifts

In the paper of Wold and Kaasch (Wold and Kaasch 1985) it was noted that their data cover only three industries, while the formal model itself classifies the

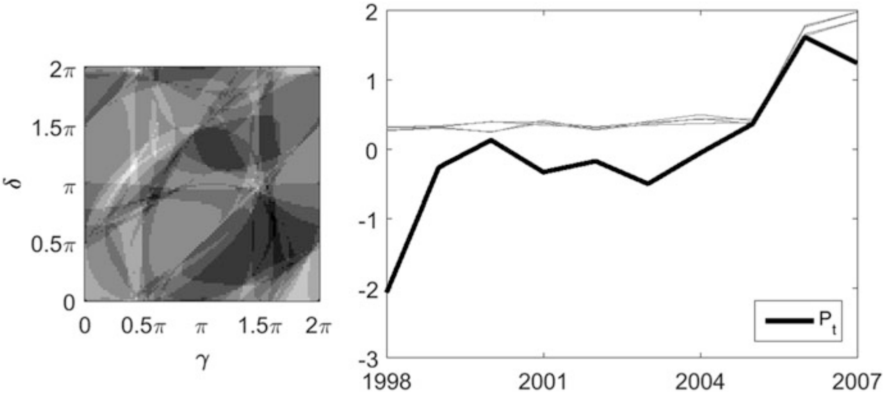


Fig. 5 Chart of k coefficients for the manufacturing industry for 2003–2012, and solutions of Eq. (5) upon $k_{max} = 0.9$

production sector into N industries, allowing for analysis of structural changes at more disaggregated levels.

As a rule, structural changes are initiated at a microlevel when companies make decisions related to new investment and modification of technological processes or management system. Therefore, the next step in our studies, i.e. evaluation of the technological development of the manufacturing sector, seems to be logical. This level represents a part of a vertically-integrated technological structure. Despite the fact that statistical authorities from time to time review the classification of industries by technological development, we observe certain consistency. Despite the fact that these classification systems vary among the EU and North America, division into four levels of the technological development (*high-tech, medium-high-tech, medium-low-tech ans low-tech*) is common for them. We use the classification approved by member states of the EU (EC 2007). With regard to our study of the US manufacturing industry, this classification is shown in Table 3.

Tables 4 and 5 shows that all technological groups (excl. *high-tech*) have the coefficient of 1 for the period of 1998–2007. Thus, in this groups could be theoretically obtained results exceeding real indicators of Value Added for the whole decade. It can also be observed that in the *low-tech* group of industries the coefficient equals to 1 for all decades. This means that higher production output could be obtained in this group within the whole considered period. The same applies to the *high-tech* group, excl. the last decade.

Let us examine the last decade (1997–2008) as the decade preceding to the current economic crisis in detail. Let us construct charts of k coefficients distribution and solutions of Eq. (5) upon maximum k values (Fig. 6) for high-tech industries group.

Table 3 The classification of industries by technological development approved by member states of the EU

High-tech	Medium-high-tech	Medium-low-tech	Low-tech
Computer and electronic products; Electrical equipment	Machinery; appliances, and components; Motor vehicles, bodies and trailers, and parts; Other transportation equipment; Chemical products	Nonmetallic mineral products; Primary metals; Fabricated metal products; Petroleum and coal products; Plastics and rubber products	Wood products; Furniture and related products; Miscellaneous manufacturing; Food and beverage and tobacco products; Textile mills and textile product mills; Apparel and leather and allied products; Paper products; Printing and related support activities

Table 4 Maximum values of k coefficient for each group in various decades

	High-tech	Medium-high-tech	Medium-low-tech	Low-tech
1948–1957	1	0.9	0.9	1
1958–1967	1	1	1	1
1968–1977	1	1	0.9	1
1978–1987	1	0.9	0.8	1
1988–1997	1	1	0.9	1
1998–2007	0.8	1	1	1

5.2 Analytical Solution

All the results were used in previous experiments can be achieved faster and more accurately, if we find the analytical solution of the Eq. (5). Let's do it using Vieta's trigonometric formulas. For this we convert Eq. (5) to (5b) and do some substitution:

$$X_t^3 - R_t X_t - E_t = 0 \tag{5b}$$

$$Q = \frac{R_t}{3} \quad K = \frac{-E_t}{2}$$

$$S = \frac{R_t^3}{27} - \frac{E_t^2}{4}$$

Equation (5b) will have different real roots depending on S. Consider them:

If $S > 0$, then 3 real roots x:

Table 5 γ, δ, upon which $k = 0.9$ for all manufacturing industry from 2003 to 2012 and corresponding values of Δ and Δ^*

No.	γ	δ	Δ	Δ^* (%)	No.	γ	δ	Δ	Δ^* (%)	No.	γ	δ	Δ	Δ^*(%)
1	1.068	6.158	0.583	4.92	23	1.131	0.094	0.513	4.50	45	1.194	0.157	0.476	4.25
2	1.068	6.189	0.580	4.93	24	1.194	5.969	0.510	4.11	46	1.257	0.094	0.473	4.19
3	1.257	3.204	0.576	5.03	25	1.162	0.031	0.510	4.43	47	1.100	5.906	0.473	3.84
4	1.257	3.236	0.567	4.98	26	1.131	0.126	0.509	4.50	48	1.225	0.126	0.472	4.20
5	1.100	6.189	0.560	4.76	27	1.131	0.157	0.507	4.50	49	1.194	0.188	0.471	4.24
6	1.100	6.220	0.553	4.72	28	1.131	0.220	0.507	4.54	50	1.194	0.220	0.467	4.23
7	1.100	0.094	0.551	4.81	29	1.131	0.188	0.506	4.51	51	1.225	0.157	0.466	4.18
8	1.100	6.252	0.548	4.70	30	1.131	0.251	0.505	4.55	52	0.974	5.906	0.464	3.90
9	1.100	0.126	0.545	4.79	31	1.162	0.063	0.503	4.40	53	1.005	6.000	0.459	3.93
10	1.100	0.000	0.544	4.69	32	1.225	6.000	0.503	4.07	54	0.565	2.702	0.424	3.82
11	1.100	6.283	0.544	4.69	33	1.194	0.031	0.502	4.37	55	0.597	2.733	0.418	3.77
12	1.131	6.220	0.542	4.63	34	1.162	0.094	0.498	4.38	56	0.565	2.670	0.416	3.74
13	1.100	0.031	0.542	4.70	35	1.194	0.063	0.494	4.33	57	0.597	2.702	0.412	3.71
14	1.100	0.063	0.542	4.72	36	1.162	0.126	0.493	4.36	58	0.597	2.670	0.407	3.66
15	1.131	6.252	0.534	4.59	37	1.162	0.157	0.488	4.35	59	1.131	6.032	0.401	3.42
16	1.131	0.000	0.527	4.55	38	1.225	0.063	0.487	4.27	60	1.162	6.063	0.392	3.36
17	1.131	6.283	0.527	4.55	39	1.194	0.094	0.487	4.29	61	1.194	6.095	0.383	3.30
18	1.162	6.252	0.526	4.52	40	1.162	0.188	0.484	4.34	62	1.194	6.126	0.374	3.27
19	1.131	0.031	0.521	4.53	41	1.162	0.220	0.481	4.34	63	1.225	6.126	0.374	3.25
20	1.162	6.283	0.517	4.47	42	1.194	0.126	0.481	4.27	64	4.335	2.576	0.211	1.78
21	1.162	0.000	0.517	4.47	43	1.225	0.094	0.479	4.23	65	4.335	2.545	0.180	1.55
22	1.131	0.063	0.517	4.51	44	1.162	0.251	0.478	4.34					

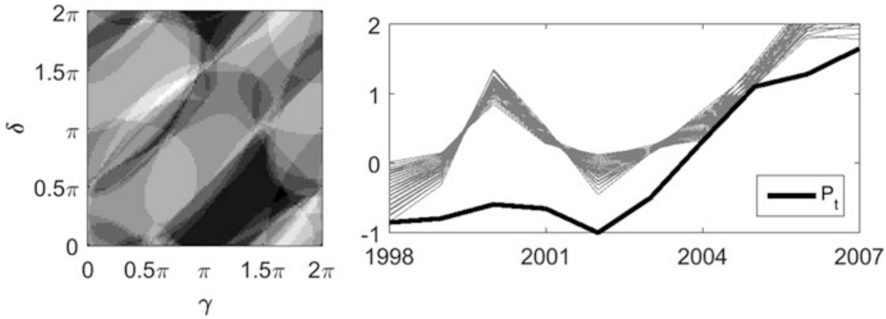


Fig. 6 Chart of k coefficients for high-tech industries for 1998–2007, and solution of Eq. (5) upon $k_{max} = 0.8$

$$\varphi = \frac{1}{3}\arccos\left(\frac{K}{\sqrt{Q^3}}\right)$$

$$x_1 = -2\sqrt{Q}\cos(\varphi)$$

$$x_2 = -2\sqrt{Q}\cos\left(\varphi + \frac{2}{3}\pi\right)$$

$$x_3 = -2\sqrt{Q}\cos\left(\varphi - \frac{2}{3}\pi\right)$$

If $S = 0$, then 2 real roots:

$$x_1 = \sqrt[3]{K}$$

$$x_2 = -2x_1$$

If $S < 0$, then 1 real root:
for $Q > 0$:

$$\varphi = \frac{1}{3}Arch\left(\frac{|R|}{\sqrt{Q^3}}\right)$$

$$x_1 = -2sgn(R)\sqrt{Q}ch(\varphi)$$

for $Q < 0$:

$$\varphi = \frac{1}{3} Arsh\left(\frac{|R|}{\sqrt{|Q|^3}}\right)$$

$$x_1 = -2sgn(R)\sqrt{|Q|}sh(\varphi)$$

for $Q = 0$:

$$x_1 = \sqrt[3]{E_t}.$$

The solutions are ambiguous. Sometimes it gives 3 real roots, sometimes 2, -sometimes 1. In the event that we obtain more than one solution will choose something that is closer to the actual value.

6 Results

1. Figure 6 show that there always exist areas (intervals) where all solutions converge to one point. We may conclude that upon any γ, δ parameters we will obtain the same production output for these intervals.

2. Table 4 and Fig. 6 allow to conclude that during the last decade before the crisis of 2008 it was possible to get higher production output in all technological groups of manufacturing, except the high-tech group, as they have $k_{max} = 1$. This paradox shall be examined further, Nature of high technologies can be considered as the most probable argument in favor of domination of empirical (real) trajectories over theoretical (model) trajectories: low share in overall production output, large investment, relatively narrow segment of use and more qualitative management in general.

3. Unexpectedly low percentage of theoretical trajectories of output, which provide better results with comparison to empirical data. They are present but there are few of them. Only 65 from 40,401 analyzed pairs of γ, δ gave solutions with coefficient $k = 0.9$.

Parameters Δ and Δ^* show average difference between production output obtained with the aid of the BIEQ model and a real value:

Absolute mean variation according to standardized data:

$$\Delta = \frac{\sum_{t} (\mathrm{X_t} - \mathrm{P_t})}{\mathrm{T}} \tag{10}$$

Δparameter can be interpreted in relation to relative mean change of P_t production output. As all data are standardized and have unit variance, then Δ parameter represents a fraction of P_t variance. Δ represents objective evaluation of a solution within the same decade for various γ, δ. It is defined in relation to the value, constant within the same decade, i.e. variance.

Relative mean variation according to initial data:

$$\Delta^* = \frac{\sum_t \frac{(X_t - P_t)\sqrt{D(P_t^*)}}{P_t^*}}{T} * 100\% \tag{11}$$

P_t^*—initial, non-standardized value of production output.

Δ^*parameter represents average difference between theoretical value and initial production output. Let us examine this indicator in detail. Its values vary from 1.55 to 5.03 %, i.e. annual theoretical (calculated on the basis of the model) production output could be more by 1.55 as minimum and 5.03 as maximum. And this is only for the account of more balanced structure of the industry.

7 Conclusion

The inconsistency of economic development, mainly characterized by recurrent (cyclic) crises, is not disputed by economists. However, the attitude of economists to structural disproportions, which are manifested with regular frequency, is more complex. If one considers two big economic crises which occurred in the last 100 years—the Great Depression of 1929–1933 and Great Recession of 2008–2009 (with oil price shock of 1973–1974 between them), it is obvious that both include two components: both cyclic and structural. Like Siamese twins, economic structures and cycles are inseparable and always go hand-by-hand. The economic theory and practice is in continuous search of various tools the use of which could be efficient in prediction of cyclic crises or in overcoming of their consequences. In our opinion, the time has come for the creation of a structural crises monitoring system. Unfortunately, the frequency at which analytical materials about the state of macroeconomic structures are published comprises about 10–12 years (for example, reports of UNIDO) doesn't allow to notice negative trends in formation of economic structures in due time. In our opinion, losses of economic systems from structural imbalance, are not only comparable with losses from cyclic crises, but considerably exceeds them due to the inertia of processes of structural reforms. The approach proposed in this work, when the direction of structural changes is determined, their quality is estimated and disproportions in structural development are calculated, can be the basis for such system of monitoring. Such system has also to take into account the influence of technical progress and its uneven character when structural readiness (or unreadiness) can accelerate or slow down diffusion of new technologies into the industry and thereby to predetermine the rates of development of economy in general. We hope that creation of such monitoring system will stimulate economic studies relating to the development of new instrument of structural changes evaluation, especially at the micro-level.

References

Barhin, G. I., & Chesnokov, A. S. (1983). O metodologii izmerenija strukturnyh sdvigov [About the methodology for measuring structural changes]. *Ekonomika i matematicheskie metody, 19* (2), 251–258.

Bergstrom, R., & Wold, H. (1983). *Fix-point estimation in theory and practice*. Gottingen: Vandenhoeck & Ruprecht.

Butina, M. A. (1980). Metody izmerenija strukturnyh sdvigov i ih prilozhenija [Methods of measuring structural changes and their applications]. *Ekonomika i matematicheskie metody, 16*(4), 687–695.

Carroll, N. (2012). Structural change in the New Zealand economy (1974–2012). Draft Paper for the Long-term External Panel, Wellington.

Carter, A. P. (1970). *Structural change in the american economy*. Cambridge: Harvard University Press.

Clark, C. (1957). *The conditions of economic progress*. New York and London: Macmillan and St. Martin's Press.

Dasgupta, P., & Chakrobarty, D. (2005). The structure of the indian economy. Fifteenth International Input–Output Conference. Renmin University in Beijing, China

Fisher, F. M. (1966). *The identification problem in econometrics*. New York: McGraw Hill.

Gaston, N., & Nelson, D. (2004). Structural change and the labor-market effects of globalization. *Review of International Economics, 12*(5), 769–792.

Gawrycka, M., Sobiechowska-Ziegert, A., & Szymczak, A. (2012). The impact of technological and structural changes in the national economy on the labour-capital relations. *Contemporary Economics, 6*(1), 4–12.

Green, F., & Stern, N. (2015, June). *China's "new normal" : Structural change, better growth, and peak emissions*. Policy brief. London: Centre for Climate Change Economics and Policy.

Guo, J., & Planting, M. A. (2000). Using input-output analysis to measure U.S. economic structural change over a 24 Year period. BEA WP2000-01.

Haraguchi, N., & Rezonja, G. (2010). *Structural change and sectoral growth in selected East Asian countries*. Working paper 18/2009. Vienna: UNIDO.

Harvie, C., & Lee, H. H. (2003, January). Export led industrialization and growth Korea's economic miracle 1962–1989. Department of Economics, University of Wollongong. http://ro.uow.edu.au/commwkpapers/67/

Herfindahl, O.C. (1950). *Concentration in the steel industry*. Dissertation, Columbia University. Retrieved 6 June 2015.

Hirschman, A. O. (1945). *National power and the structure of foreign trade 1980 expanded ed.* Berkeley: University of California Press.

Iiduka, N. (2007). *Examining Japan's expansion from the viewpoint of industrial structure*. Japan Center for Economic Research (JCER) Staff Report.

Industrial Development for the 21st Century: Sustainable Development Perspectives. (2007). New York: United Nations.

Industry and Development. (1985). Global Report. New York.

Kahrl, F., & Roland-Horst, D. (2009). Growth and structural change in China's energy economy. *Energy*, 1–10. doi:10.1016/j.energy.2009.03.009.

Kei-Mu, Y., & Zhang, J. (2010). Structural change in an open economy. Working paper 595. Research Seminar in International Economic, University of Michigan.

Kossov, V. V. (1975). Pokazateli rosta i razvitija ekonomiki [Indicators of economic growth and development]. *Voprosy jekonomiki, 12*, 37–39.

Kuznetsov, V. I. (ed) (1988). *Zapadnaja Evropa—evoljucija ekonomicheskoj struktury* [Western Europe—The evolution of the economic structure]. Moscow: Nauka.

Laitner, J. (2000). Structural change and economic growth. *Review of Economic Studies, 67*, 545–561.

Lebedeva, I. P. (1986). *Strukturnye izmenenija v yaponskoy promyshlennosti* [Structural changes in the Japanese industry]. Moscow: Nauka.

Leontief, W. (1941). *Structure of the American economy, 1919–1929*. Cambridge: Harvard University Press.

Leontief, W. (1953). *Studies in the structure of the american economy*. New York: Oxford University Press.

Memedovic, O. (2009). *The Impact of institutions on structural change in manufacturing: The case of international trade regime in textiles and clothing*. Research and statistics branch working Paper 25/2009. Veinna: UNIDO.

Mensch, G. O. (1975). *Das Technologische Patt: Innovationen uberwingen die Depression*. Frankfurt: Umschau. (G. O. Mensch, Trans., 1979, *Stalemate in technology: Innovations overcome the depression,* Cambridge, MA: Ballinger)

Mensch, G. O. (1982). Basic innovations and industrial growth. In *Proceedings of the 10th International Conference on the Unity of the Sciences (lCUS) 1981* (Vol. 2, pp. 499–515). New York: ICF Press.

Mensch, G. O., Kaasch, K., Kleinknecht, A., & Schnopp, R. (1980). *Innovation trends and switching between full and underemployment equilibria, 1950–1978. Discussion paper IIM/dp 80–5*. Berlin: International Institute of Management.

Minasyan, G. (1983). *K izmereniju i analizu strukturnoj dinamiki* [About measuring and analyzing of structural dynamics]. Ekonomika i matematicheskie metody 19(2):264–265.

Natuhara, K. (2008). Structural change and economic growth: evidence from Japan. *Economic Bulletin, 15*(9), 1–11.

Naude, W., & Szirmai, A. (2015). *Structural change and industrial development in the brics*. Oxford: OUP.

Nelson, R., & Winter, S. J. (1982). *An evolutionary theory of economic change*. Cambridge, MA: The Belknap Press of Harvard University.

Nickell, S., Redding, S., & Swaffield, J. (2008). The uneven pace of deindustrialization in the OECD. *World Economy, 31*(9), 1154–1184.

Oulton, N., Srinivasan, S. (2005) Productivity growth in UK industries, 1970–2000: Structural change and the role of ICT. Working Paper No. 259. Bank of England.

Robson, M. T. (2009). Structural change, employment specialization and regional labour market performance: Evidence for the UK. *Applied Economics, 41*(3), 275–293. doi:10.1080/00036840601007278.

Ruelle, D., & Takens, F. (1971). On the nature of turbulence. *Communications of Mathematical Physics, 20*, 167–192.

Sonis, M., Hewings, G. J. D., & Guo, J. (1996). Sources of structural change in input–output systems: A field of influence approach. *Economic Systems Research, 8*(1).

Spasskaya, O. V. (2003). Makrojekonomicheskie metody issledovanija i izmerenija strukturnyh izmenenij [Macroeconomic research methods and measurement of structural changes] (Vol. 1, pp. 20–39). Working papers, Institute of Economic Forecasting of the Russian Academy of Sciences.

Thompson, G., Murray, T., & Jomini, P. (2012). Trade, employment and structural change: The Australian experience. In D. Lippoldt (Ed.), Policy priorities for international trade and jobs. OECD, e-publication. http://www.oecd.org/trade/icite

Varum, C. A., Cibrão, B., Morgado, A., & Costa, J. (2009). R&D, structural change and productivity: The role of high and medium-high technology industries. *Economia Aplicada, 13*(4), 399–424.

Vatnik, P. A. (2011). Izmerenie proporcional'nosti [The measurement of proportionality]. *Vestnik Inzhekona. Seriya: Ekonomika, Vypusk, 5*(48), 172–174.

Wang, X., & Zhou, Y. (2014). Structural imbalance, inequality and economic growth. In L. Song, R. Garnaut, & C. Fang (Eds.), *Deepening reform for China's long-term growth and development*. Canberra: ANU Press.

Wold, H., & Kaasch, K. (1985). Transfers between industrial branches in the course of Schumpeter-Mensch Long Swings. In T. Vasko (Ed.), *The long wave debate. Selected papers. Weimar, GDR*. Berlin: Springer.

Wold, H., & Mensch, G. O. (1983). Nonlinear extensions in pls soft modeling. Working Paper WSOM WP-B3-D11. Cleveland, OH: Weatherhead School of Management, Case Western Reserve University.

Wold, H., & Mensch, G. O. (1985) The BIEQ-PLS model of a Schumpeterian wave. Working Paper WSOM WP 85–017. Cleveland, OH: Weatherhead School of Management, Case Western Reserve University.

Distribution and Clusters of Basic Innovations

Michael Y. Bolotin and Tessaleno C. Devezas

Abstract This paper presents our recent results on the issue of the temporal distribution of basic innovations, which, as we know, it is the main driver of economic growth. Numerous scholars have addressed this issue, and we can say that the Poison distribution is the commonly accepted one. On the other hand, there is not a convincing substantiation of the fact. This article provides a retrospective analysis of statistical evidence, and present a proof of the Poisson distribution of basic innovations based on the theory of random processes. In the empirical part of the research we have used one unified supersample time series rather than smaller different datasets as previously done by other authors.

Keywords Basic innovation • Economic cycles • Random processes • Poisson distribution • Innovation clusters

1 Introduction

Basic or radical innovation is meant to be a disruptive technology resulting from scientific advances, which is adopted by and spreads in society, transforming substantially the socioeconomic realm. The majority of scientists regard the introduction of basic innovations as a condition of economic growth, through encouraging the creation of new enterprises. In the 1930s Schumpeter (1939) was the first economist to use the mechanism of basic innovations as the basis for the discussion on the long-term development of global capitalism.

After a lengthy halt Schumpeter's hypothesis became a matter for hot debate once again in the 1970–1980s. A significant part of the empirical argumentation of this debate was concentrated on historical time series of basic innovations.

M.Y. Bolotin (✉)
St. Petersburg State University of Economics, 21 Sadovaya street, St. Petersburg, Russian Federation
e-mail: bolotin91@mail.ru

T.C. Devezas
Department of Electromechanics, University of Beira Interior, Covilhã, Portugal
e-mail: tessalen@ubi.pt

T. Devezas et al. (eds.), *Industry 4.0*, Studies on Entrepreneurship, Structural Change and Industrial Dynamics, DOI 10.1007/978-3-319-49604-7_6

Different authors collected them (and obtained different results) on the basis of historical evaluation of the importance of innovations. Silverberg and Lehnert (1993) claimed that the statistical methods used by the researchers involved in this discussion were faulty because innovations are discrete random events and a description of the temporal distribution of innovations should consider this fact. T- and Z-tests as applied by Kleinknecht (1990) and Solomou (1986) should not be used, for the data have no normal distribution. Instead of using nonparametric tests Silverberg and Lehnert applied a corresponding null hypothesis of the Poisson distribution. Sahal (1974) proposed the usage of statistical methods based on the Poisson distribution and negative binomial distribution to describe time series of increasing number of innovations. The Poisson regression became then a standard part of econometric tools for evaluation of the temporal distribution of innovations (Greene 1995).

In this study we will assume that the distribution of basic innovations is invariable in time. The article provides evidence and empirical verification of Poisson-distributed basic innovations using the same data as Kleinknecht (1981) and Mensch (1979) used in their studies.

Should be noted that there are also debates about innovation clusters in time. We will consider the findings Mensch and make our conclusions.

2 Poisson Regression

We shall regard innovations as quantitative discrete characters and consider their appearing a point process like all researchers before us. While statistical properties of point processes were explored at the very beginning of the probability theory and mathematical statistics, many economists still use these methods without substantiation. It should be noted, however, that many modeling tools have been developed recently.

Our point of departure for the entire subsequent analysis is that a Poisson process is time-homogeneous. This allows us to make a simplifying assumption that the probability of introduction of innovations in a fixed period of time has no relation to previous innovations and does not depend on time. Let us define a Poisson process first.

A discrete random variable $Y = y$ is said to have a Poisson distribution with parameter $\lambda > 0$ (occurrence rate of the Poisson process), if for $y \in N$ its probability of occurrence (also called probability mass function of Y) is given by:

$$P_{rob}(Y = y) = \frac{e^{-\lambda}\lambda^{y}}{y!} \tag{1}$$

Note that λ is not necessarily an integer. It is not hard to show that the expected number of events that occur per unit time, λ, is also the dispersion of the distribution. It should be noted that time series generated from a time-homogeneous Poisson process will not reflect a completely uniform picture of occurrence of random events. The model introduced by Silverberg and Lehnert (1993) shows that such exogenous and unstructured innovative process is sufficient to generate long waves of economic growth. Parameter λ may be introduced exogenously as well, for instance, as $\ln\lambda = \beta x$, where x is a vector of independent variables, and β is a vector of parameters.

Sahal (1974) proposed to use Poisson models to explore characteristics of various time series of innovations in different industries. His conclusion was: "one can characterize an invention as a random Poisson process, but its rate depends on economic factors". Though Sahal actually evaluated some β-type parameters, they were based on ordinary least square methods. Nevertheless, when data contain many nulls and small integral values, a maximum likelihood approach based explicitly on the Poisson distribution is more appropriate. Hausman, Hall and Griliches introduced this procedure in 1984. They evaluated the model by relation between a number of patents of companies and their research and development costs. Crepon and Duguet (1997a, b), and Cincera (1997) have also used this approach. One of the problems with the Poisson model is that the positive real number λ is equal to the expected value of x and also to its variance $E(x) = V(x)$. Empirical variables often show greater variability than the dependent variable, a phenomenon called as *overdispersion*. Hausman et al. (1984) observed overdispersion in their database of patents level for various companies. It is possible to develop a model that fits overdispersion by adding a non-observable random component to the Poisson distribution (Hausman et al. 1984). This leads to a modified probability distribution of the type presented in Eq. (2)

$$P_{rob}(Y = y|u) = \frac{e^{-\lambda u}(\lambda u)^y}{y!} \tag{2}$$

where u is a random variable which requires some kind of distribution (for example, variable u may correspond to random noise). If we suppose that u has gamma distribution, we shall have the following marginal distribution (Greene 1995):

$$P_{rob}(Y = y|x) = \frac{\Gamma(\theta + y)}{\Gamma(y+1)\Gamma(\theta)} r^y (1-r)^\theta \tag{3}$$

where $r = \frac{\lambda}{\lambda+\theta}$.

This distribution is known as negative binomial distribution, which means that dispersion equals $\lambda\left(1 + \frac{\lambda}{\theta}\right)$, where $\theta > 0$. When $\theta = 0$, the model comes to the standard Poisson distribution with dispersion λ. One can evaluate negative binomial distribution using a maximum likelihood approach as well. We can test the Poisson distribution and negative binomial distribution by a likelihood ratio test or Wald test (Greene 1995) of a null hypothesis $\theta = 0$.

3 Mathematical Proof of the Poisson-distributed Basic Innovations

Before moving on to the proof we should note that despite we advocate the Markov theory of random processes, we will also admit non-Markov approaches to description of stochastic processes in this work. The fact that we need to understand the true nature of the processes requires that we take into account the opinion of part of the expert community that considers that the distribution of innovations is not random in fact, for there are actually some inner links in economic systems. Therefore, we should alternatively consider economic systems and basic innovations as non-Markov processes.

Consider the following process. Basic innovations occur at random time points one after another in the course of time (the introduction of a basic innovation is the event of the given process). We are interested in the number of occurrences during a time interval $[0; t]$. Let us denote this umber as X_t and introduce some assumptions as to the process of occurrence of basic innovations.

Stationarity and Independence in Increments This property means that the probability of occurrence of a certain number of events X_t in any time interval $[\tau; \tau + t]$ depends on t only and does not depend on τ. Secondly, all of this occurs with no relation to the number of events or the manner in which they occurred prior to the time point τ.

Ordinariness Let us denote $P_k(t) = P(X_t = k)$, $k = 0, 1, \ldots,$ and suppose that while $h \to 0$:

$$P(X_n \geq 2) = \sum_{k=2}^{\infty} P_k(h) = o(h)\,(4) \Leftrightarrow \frac{\lim_{h \to 0} P(X_n \geq 2)}{h} = 0 \tag{5}$$

This condition means practical impossibility that two or more events occur in a short time interval$h$.

Statement 1. [See Proof at Appendix 1]
If the above is satisfied,

Then:

$$P_k(t) = \frac{(\lambda t)^k}{k!} e^{-\lambda t}, k = 0, 1, \ldots, \tag{6}$$

while $\lambda > 0$.

This was made possible due to the proof of the statement 2

Statement 2. [See Elegant Proof at Appendix 1]
If $f(x + y) = f(x) * f(y)$ for continuous f, then $f(x) = a^x$.

Thus, we obtain probabilities corresponding to the Poisson distribution with parameter λt. Due to this reason the process under consideration is called a Poisson process. Researchers often call it Poisson flow of events.

4 Empirical Results

Like Kleinknecht (1990) we used time series for basic innovations composed of other time series introduced earlier by other researchers. We developed this new time series by merging two long time series found in literature: the series presented by Haustein and Neuwirth (1982), and the data obtained by Van Duijn (1983). Kleinknecht's time series is too short when compared with these two series of basic innovations. Another well-known data series of innovations developed some years earlier is that presented by Baker (1976), but we did not include Baker's data in our series because he considered patent data and not properly innovations.

We used a merging technique for innovations time series which differed significantly from Kleinknecht's approach. The main difference lies in duplication, i.e. we considered just one time some innovations included in both sources. Kleinknecht summed up the number of innovations from different time series, so he counted many innovations more than once. Thus, he took into account the weight of each innovation.

It is clear that such an implicit weighing procedure (and repetition as well) is not adequate in the context of the Poisson regression approach, and this is the main reason for counting even very important innovations only one time. If the dates of introduction of innovations differed, we chose the earliest one. The combined series consists of 88 innovations present in both databases, 90 innovations listed only by Haustein and Neuwirth, and 70 innovations mentioned only by Van Duijn. Thus, the combined series contains 248 innovations from 1764 to 1976. See the complete list of all innovations of the combined series in Appendix 2. Following Kleinknecht's terminology, we shall call this combined time series a "supersample" time series.

As for the data itself, there is, of course, the issue of validity of the expert estimation. In addition, there are different opinions in regard to the dates of basic innovations. The case of the Maser is a good example. Some authors consider that the Maser was invented by British scientists in 1960, other consider the Maser invented by the Rumanian physicist Theodore V. Ionescu in the USSR in 1946. In our research we considered the former.

Figure 1 illustrates diagrams of Baker's patent time series, two time series used to form the "supersample time series" (data by Haustein-Neuwirth and Van Duijn), as well as the "supersample time series" itself). Baker's time series (Fig. 1a), illustrates patents between 1769 and 1970 and, thus, reflects 202 years. Haustein and Neuwirth (1982) (Fig. 1b), present the data between 1764 and 1975 (212 years), while Van Duijn's data (1983) (Fig. 1c), covers the period from 1811 to 1971 (161 years). The "supersample time series" (Fig. 1d), covers the period between 1764 and 1976 (213 years). In all these graphs the horizontal axis expresses the number of innovation per annum, and the vertical axis expresses the percentage of the total sample. These diagram are based on the original data and they differ from the diagram of other researchers.

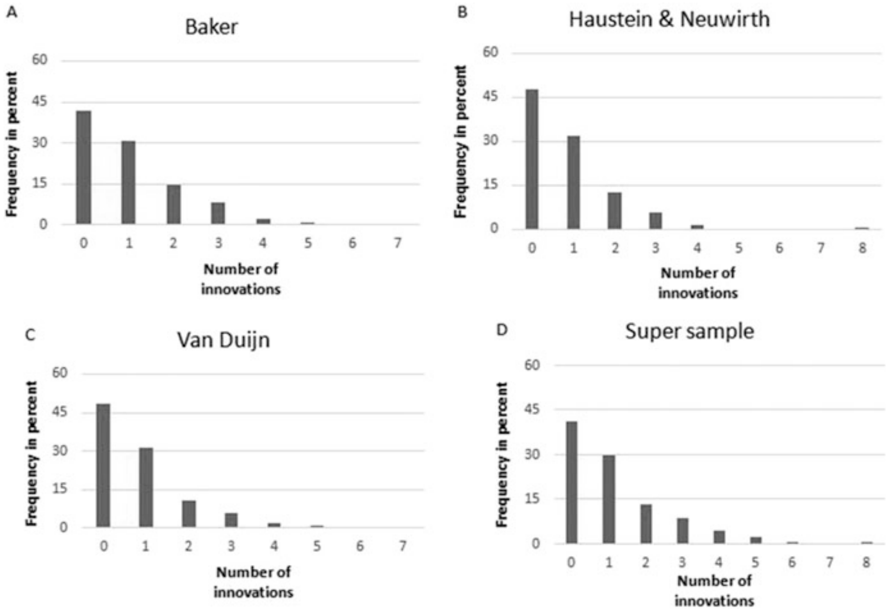


Fig. 1 Frequency distribution diagram of the number of innovations

All three diagrams show that the highest frequency is at zero values. Moreover, all diagrams show the reduction of frequencies while the number of innovation per annum grows. Constructing similar diagrams Silverberg and Lehnert (1993) came to conclusion that the assumption of normal data distribution should be rejected as false, thus devaluing the statistical methods used by Kleinknecht (1990) and suggested that the Poisson distribution or negative binomial distribution of time series may be more appropriate.

We used Pearson's test χ^2 to verify the empirical data. Using statistics package Statistica we found that this series may satisfy only one of the known distributions: the Poisson distribution or geometric distribution. Lets start with the analysis of the original data series: the geometric and Poisson distribution diagram as presented in Fig. 2. On Figs. 2 and 3 the number of basic innovations per year is represented on the horizontal axis, and the frequency of the variations is presented on the ordinate axis.

As we can see, Pearson's test fitted geometric distribution well, thus implying that our distribution explicitly follows the geometric distribution. The substantiation is good too, as we instinctively understand that the probability of occurrence of new basic innovations is equivalent to the classic play of the probability theory before the first event occurs. We remind that the probability of an event in geometric distribution is given by $p(x = k) = pq^{k-1}$. But there is a drawback: the probability of occurrence of a new basic innovation during a year does not depend on introduction of a previous innovation during the same year. This contradicts the choice of geometric distribution. This distribution requires taking the probability of

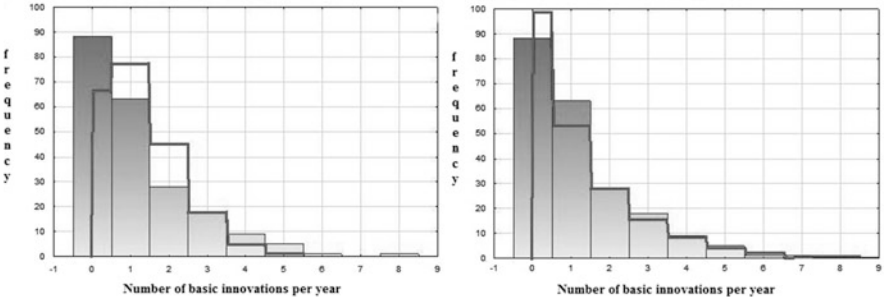


Fig. 2 Diagram comparing the geometric and Poisson distribution performed with Statistica package (*bold line*) and our results

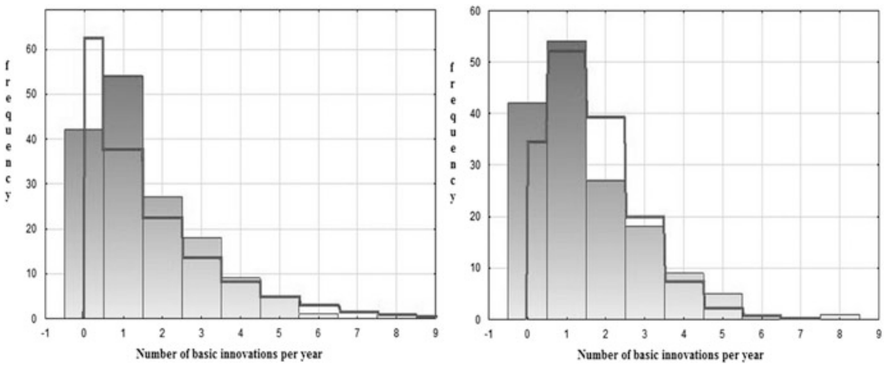


Fig. 3 Diagram comparing the geometric and Poisson distribution performed with Statistica package (*bold line*) and our results for cut data

occurrence or absence of the previous event into account. Also we have a theoretical proof of the fact that the geometric distribution is not appropriate [Statement 1].

Let us now examine the Poisson distribution presented in Fig. 2. The chi-squared test statistic is high, but this may be the result of the fact that some basic innovations are interrelated: cinema and color television, nuclear energy and atomic ice breaker, etc.

Therefore, we conducted one more research with cut off the data of 1764–1819 (prior to the Industrial Revolution) from our original series. We obtained the following results: $\chi^2_{Poisson} \approx \chi^2_{geometric} \cong 30$.

No other known distribution fits at all. For un-cutoff data series $\chi^2_{geometric}$ distribution is substantially less than the Poisson distribution but the geometric distribution is not appropriate for the reasons given above. We consider independent innovation in the sense of time. For cut data $\chi^2_{Poisson}$ not too large.

Thus, in the future we shall use the Poisson distribution as the most probable temporal distribution of basic innovations.

5 Clusters of Basic Innovations

Let's focus on the problem of cluster of basic innovations.

Basic innovations have a dominant role in almost all long-wave economic theories. Supporters of this opinion, for example, Schumpeter, Mensch, Van Duijn, believe that basic innovations explicitly manage economic development. Schumpeter was the first to assume the economy advanced by means of entrepreneurial innovations. It is assumed that economic development depends on the entrepreneurs' ability to implement new basic innovations in production. Such implementation leads to clustering of basic innovations which, in their turn, promote new leading industries. Mensch dwelt upon this notion and proposed an idea constituting that clusters of basic innovations develop during economy depression phases as only in these periods companies are desperate enough to try and recover profitability by means of investment in extremely new technologies.

According to Mensch, transfer from one technological stalemate to another occurs through transfer from basic innovations to improving innovations and then to pseudo innovations. He notes that a technological stalemate corresponds to the long wave recession phase, and that in the phase of depression economy is structurally ready for transfer to new basic innovations. At this phase clusters of basic innovations form. This phase includes development of innovation priorities and their formalization in regulatory documents and innovation programs; government support of top-priority basic innovations with the aid of budgetary appropriations on non-repayable and repayable terms, tax advantages and customs privileges; small business support, promotion of innovation infrastructure (technological centers, technology parks), training of innovation managers; development of legal framework for innovation activity.

Thus, the level of innovation and investment activity is the lowest at the phase of an economic crisis. It begins to grow at the end of the depression phase, and reaches the peak in the reanimation and recovery phases (upon mass capital renewal using basic innovations), then stabilizes at the phase of maturity (when improving innovations prevail), and slumps as there is no sense in improvement of obsolete equipment.

Let us then compare Kodratieff cycles of basic innovations with US GDP cycles.

As can be seen from Fig. 4, the cycle of basic innovations predetermines economy development. Initially, the lag between innovations and GDP amounted to approximately 25 years, however, it tends to decrease with time. This result corresponds to conclusions of Japanese researcher Masaaki Hirooka (Hirooka 2006) who established that life cycle of innovations had been gradually contracting since the first industrial revolution.

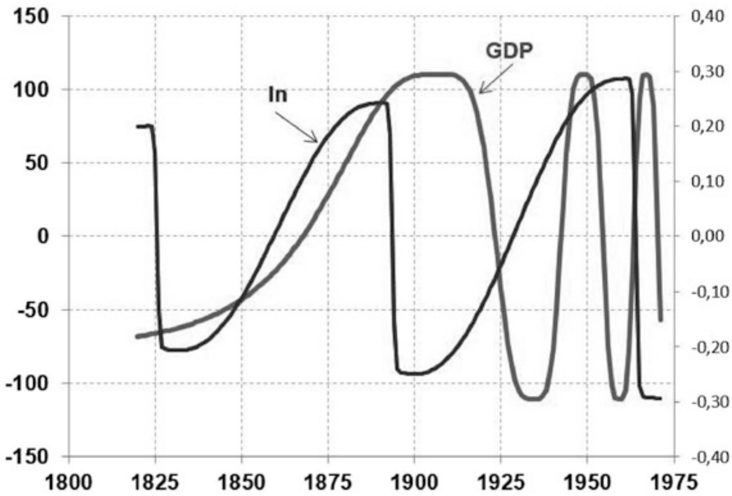


Fig. 4 Kondratieff Cycles in US GDP and basic innovations

6 Conclusions

As Simon Kuznets noted the major question of the discussion on the distribution of basic innovations is "What should we consider a basic innovation?" The final conclusions as to the distribution of basic innovations depend on the answer to this question. But, as demonstrated in this paper, there is strong evidence favoring the Poisson distribution of innovations. This provides us with an excellent foundation for further research of cyclic processes in economy.

The fundamental changes with regard to occurrence rate of basic innovations took place in the history of capitalism. However, these changes cannot be considered as frequent and regular enough, as well as proving the pattern of long waves in the Schumpeterian economy. The results point at insignificant correlations proposed by Schumpeter, Mensch and Kleinknecht with regard to the fact that basic innovations occur more often during economy recession. However, it was true 150 years ago. Long-term changes of appearance of basic innovations occur less frequently than expected by the theory of long waves. Therefore, long-term changes in occurrence of basic innovations are not strictly periodic. In dynamics of waves of global economies we shall look for its correlation with innovation diffusion rather than with their clustering.

Appendix 1

A. Proof of Statement 1

We divide the proof into several stages:

1. We show that, except for some simple situations, at some $\lambda > 0$, $P_0(t) = e^{-\lambda t}$ is satisfied.

Indeed, let p = $P_0(1)$ (14). We divide a time interval [0; 1] into n equal parts. Absence of events per time unit means that each short time segment of length $1/n$ has 0 events. Due to independence we have $p = (P_0(1/n))^n$ (15), hence $P_0(1/n) = p^{1/n}$ (16). It immediately follows that $P_0(k/n) = p^{k/n}$ (17) at any k $\geq$ 1.

We show now that generally $P_0(t) = p^t$ (18) at all $t \geq 0$. For each such number and random natural n there is such a number $k \geq 1$ that:

$$\frac{k-1}{n} \leq t < \frac{k}{n} \tag{19}$$

Function $P_0(t)$ does not increase on t, therefore:

$$P_0\left(\frac{k-1}{n}\right) \geq P_0(t) \geq P_0\left(\frac{k}{n}\right) \tag{20}$$

or

$$p^{\frac{k-1}{n}} \geq P_0(t) \geq p^{\frac{n}{k}} \tag{21}$$

Let $n \to \infty$, so that $k/n \to t$, we obtain $P_0(t) = p^t$.(22)

There can be three cases: (a) $p = 0$, (b) $p = 1$, and (c) $0 < p < 1$.

In the first case $P_0(t) = 0$ at any t > 0, i.e. at least one event occurs with a probability of 1 in any time interval, however short, which is equivalent of the case when an infinite number of events occur in time intervals of any length. This resembles chain reaction of a nuclear explosion, so we shall not consider such an extremity. In case (b) $P_0(t) = 1$, i.e. events never occur. Thus, we are interested only in case (c). Assume $p = e^{-\lambda}$ (23), here $0 < \lambda < \infty$. Hence, we obtain $P_0(t) = e^{-\lambda t}$ (24). (Let us prove later, that other cases are impossible using statement 2).

2. We prove that at $h \to 0$:

$P_1(h) = \lambda h + o(h)$(25), where

o(h) is a probability that more than one change occurs during (t;t+h).

To do so we use the fact that $P_0(h) = e^{-\lambda h} = 1 - \lambda h + o(h)$ (26) and $P_0(h) + P_1(h) + \sum_{k=2}^{\infty} P_k(h) = 1$.(27)

Therefore: $P_1(h) = 1 - P_0(h) - \sum_{k=2}^{\infty} P_k(h) = \lambda h + o(h)$. (28)

3. In this section we show that probabilities $P_k(t)$ satisfy a certain system of differential equations.

For t $\geq$ 0 and h > 0 we have

$$P_k(t+h) = \sum_{j=0}^{k} P_j(t)P_{k-j}(h) \tag{29}$$

(Here we go through some variants of a number of events that occur during time t and during a subsequent time interval of length h).

If h $\to$ 0, then

$$\sum_{j=0}^{k-2} P_j(t)P_{k-j}(h) \leq \sum_{j=0}^{k-2} P_{k-j}(h) = \sum_{i=2}^{k} P_i(h) = o(h) \tag{30}$$

Hence, there may occur even k changes at $k \geq 1$ in time interval $(0; t+h)$ by three mutually exclusive means:

1) No changes in time $(t; t+h)$ and k changes in $(0; t)$.
2) One change in time $(t; t+h)$ and $k-1$ changes in $(0; t)$.
3) x changes in time $(t; t+h)$ and $k-x$ changes in $(0; t)$.

Let us find probabilities for each of these means:

P(1) = $P_k(t)(1-\lambda h - o(h))$ (31)

P(2) = $P_{k-1}(t)\lambda h + o(h))$(32)

P(3) decreases faster than h

or

$$P_k(t+h) = P_k(t)P_0(h) + P_{k-1}(t)P_1(h) + o(h) =$$
$$= P_k(t)(1-\lambda h + o(h)) + P_{k-1}(t)(\lambda h + o(h)) + o(h) =$$

$=P_k(t)(1-\lambda h) + P_{k-1}(t)\lambda h + o(h)$.(33)

Hence, we obtain:

$\frac{P_k(t+h)-P_k(t)}{h} = -\lambda P_k(t) + \lambda P_{k-1}(t) + \frac{o(h)}{h}$, $k=1,2,\ldots,$ (34)

Let $h \to 0$. As there is a limit for the right part, there is a limit for the left part as well. As a result, we obtain:

$P_k^{'}(t) = -\lambda P_k(t) + \lambda P_{k-1}(t)$, $k=1,2,\ldots,$ (35)

At k = 0 probabilities (2) and (3) do not occur, hence:

$P_0(t+h) = P_0(t)(1-\lambda h) + o(h)$, (36) remove parentheses, divide by h:

$$P_0^{'}(t) = -\lambda P_0(t) \tag{37}$$

Initial conditions:

$P_0(0) = 1$, $P_k(0) = 0$ we have the answer:

$$P_0(t) = e^{-\lambda t} \tag{38}$$

$P_0(0) = 1$, $P_k(0) = 0$ at $k \geq 1$.

4. Solution of a system of equations.

Let us present this differential equation as follows:
$\dot{f} = -\lambda f + \lambda f_{-1}$, (39) where
$f - P_k(t), f_{-1} - P_{k-1}(t)$, while $P_0(t) = e^{-\lambda t}$(40)
Solve it using method $f = uv$(41)
Let $f = uv$, then $\dot{f} = u' v + uv'$.(42)
Rewrite our differential equation:

$$f' + \lambda f = \lambda f_{-1} \tag{43}$$

$$u'v + uv' + \lambda uv = \lambda f_{-1} \tag{44}$$

$$uv' + \lambda uv = 0 \tag{45}$$

$\frac{v'}{v} = -\lambda \Rightarrow v = e^{-\lambda t}$(46)

$$u' e^{-\lambda t} = \lambda f_{-1} \tag{47}$$

Let k = 1 $\Rightarrow u' e^{-\lambda t} = \lambda e^{-\lambda t} \Rightarrow u' = \lambda \Rightarrow u = \lambda t \Rightarrow f = \lambda t e^{-\lambda t}$(48)
Let k = 2 $\Rightarrow u' e^{-\lambda t} = \lambda^2 t e^{-\lambda t} \Rightarrow u' = \lambda^2 t \Rightarrow \ u = \frac{\lambda^2 t^2}{2} \Rightarrow$ (49)$f = \frac{(\lambda t)^2}{2} e^{-\lambda t}$(50)
Let k = 3 $\Rightarrow u' e^{-\lambda t} = \lambda^3 t^2 e^{-\lambda t} \Rightarrow u' = \lambda^3 t^2 \Rightarrow \ u = \frac{\lambda^3 t^3}{2 \cdot 3} \Rightarrow \ f = \frac{(\lambda t)^3}{3!} e^{-\lambda t}$(51)
Then let us use the method of mathematical induction to prove that:
For random k: $u' e^{-\lambda t} = \lambda^k t^{k-1} e^{-\lambda t} \Rightarrow u' = \lambda^k t^{k-1} \Rightarrow \ u = \frac{\lambda^k t^k}{2 \cdot 3} \Rightarrow \ f = \frac{(\lambda t)^k}{k!} e^{-\lambda t}$.(52)
B. Proof of Statement 2
Let there be a random function f which satisfies the conditions.
Assuming that $y = x_0 - x$ in (8), we obtain $f(x)f(x_0 - x) = f(x_0) \neq 0$(53)

Therefore, we see that $f(x)$ differs from 0 at all x. Moreover, replacing x in (1) and y with $\frac{x}{2}$ we find: $f(x) = \left[f\left(\frac{x}{2}\right)\right]^2$ (54), so $f(x)$ is always strictly positive. Using this, take a logarithm

$$\ln f(x + y) = \ln f(x) + \ln f(y) \tag{55}$$

If we assume $\phi(x) = \ln f(x)$ (56), then we have a function which is continuous and satisfies the condition:

$\phi(x + y) = \phi(x) + \phi(y)$(57), but this is a well-known fact (this function can be linear only). $\Rightarrow \phi(x) = \ln(f(x) + cx, (c = const)$(58)

Therefore, finally, $f(x) = e^{cx} = a^x$.(59).

Appendix 2

List of basic innovations

1764	Spinning machine
1775	Steam engine
1780	Automatic band loom
1794	Sliding carriage
1796	Blast furnace
1809	Steam ship
1810	Whitney’s method
1811	Crucible steel
1814	Street lighting (gas)
1814	Mechanical printing press
1819	Lead chamber process
1820	Quinine
1820	Isolated conduction
1820	Rolled wire
1820	Cartwright’s loom
1824	Steam locomotive
1824	Cement
1824	Puddling furnace
1827	Pharma fabrication
1831	Calciumchlorate
1833	Telegraphy
1833	Urban gas
1835	Rolled rails
1837	Electric motor
1838	Photography
1839	Bicycle
1840	Vulcanized rubber
1841	Arc lamp
1844	Jacquard loom
1845	Lathe
1846	Inductor
1846	Electrodynamic measuring
1846	Rotary press
1846	Anesthetics
1849	Steel (puddling process)
1851	Sewing machine
1852	Plaster of paris
1854	Aluminum
1855	Safety match
1855	Bunsen burner

(continued)

1856	Refined steel/Bessemer steel
1856	Steel pen / Fountain pen
1856	Tare colors industry
1856	Baking powder
1857	Elevator
1859	Lead battery
1859	Drilling for oil
1860	Internal combustion engine
1861	Soda works
1863	Aniline dyes
1864	Siemens-Martin steel
1865	Paper from wood
1866	Deep sea cable
1867	Dynamite
1867	Dynamo
1869	Commutator
1870	Typewriter
1870	Celluloid
1870	Combine harvester
1871	Margarine
1872	Thomas steel
1872	Reinforced concrete
1872	Drum rotor
1873	Preservatives
1875	Sulphuric acid
1876	Four-stroke engine
1877	Telephone
1878	Nickel
1879	Electric Railway
1880	Incandescent lamp
1880	Water turbine
1880	Jodoforme
1880	Half-tone process
1881	Electric power station
1882	Veronal
1882	Cable
1883	Antipyrin
1883	Coals whisks
1884	Steam turbine
1884	Chloroform
1884	Punched card
1884	Cash register
1885	Synthetic fertilizers
1885	Transformer

(continued)

1885	Synthetic alkaloids
1886	Magnesium
1886	Electric welding
1886	Linotype
1887	Phonograph
1887	Electrolyze
1888	Motor car
1888	Pneumatic tire
1888	Electric counter
1888	Portable camera
1888	Alternating-current generator
1890	Man-made fibers
1890	Chemical fibers
1891	Melting by induction
1892	Acetylene welding
1892	Accounting machine
1894	Cinematography
1894	Antitoxins
1894	Motor cycle
1894	Monotype
1895	Diesel engine
1895	Drilling machine for mining
1895	Electric automobile
1896	X-rays
1898	Aspirin
1898	Arc welding
1900	Air ship
1900	Synthesis of indigo
1900	Submarine
1901	Holing machine
1902	Electric steel making
1903	Safety razor
1905	Viscose rayon
1905	Vacuum cleaner
1906	Acetylene
1906	Chemical accelerator for rubber vulcanization
1907	Electric washing machine
1909	Gyro compass
1910	Airplane
1910	Bakelite (Phenol plastics)
1910	High voltage isolation
1913	Vacuum tube
1913	Assembly line
1913	Thermal cracking

(continued)

1913	Domestic refrigerator
1914	Ammonia synthesis
1914	Tractor
1914	Stainless steel
1915	Tank
1916	Synthetic rubber
1917	Cellophane
1918	Zip fastener
1920	AM Radio
1920	Acetate rayon
1920	Continuous thermal cracking
1922	Synthetic detergents
1922	Insulin
1922	Synthesis of methanol
1923	Continuous rolling
1924	Dynamic loudspeaker
1924	Leica camera
1925	Deep frozen food
1925	Electric record player
1927	Coal hydrogenation
1930	Power steering
1930	Polystyrene
1930	Rapid freezing
1931	Freon refrigerants
1932	Crease-resistant fabrics
1932	Gas turbine
1932	Polyvinylchloride
1932	Antimalaria drugs
1932	Sulfa drugs
1934	Fluor lamp
1934	Diesel locomotive
1934	Fischer-Tropsch procedure
1935	Radar
1935	Ballpoint pen
1935	Rockets/guided missiles
1935	Plexiglas
1935	Magnetophone
1935	Catalytic cracking
1935	Color photo
1935	Gasoline
1936	Television
1936	Photoelectric cell
1936	FM radio
1937	Vitamins

(continued)

1937	Electron microscope
1938	Helicopter
1938	Nylon
1939	Polyethylene
1939	Automatic gears
1939	Hydraulic gear
1940	Antibiotics (penicillin)
1941	Cotton picker
1942	Jet engine/plane
1942	DDT
1942	Heavy water
1942	Continuous catalytic cracking
1943	Silicones
1943	Aerosol spray
1943	High-energy accelerators
1944	Streptomycin
1944	Titan reduction
1945	Sulzer loom
1946	Oxygen steelmaking
1946	Phototype
1948	Numerically controlled machine tools
1948	Continuous steel making
1948	Orlon
1948	Cortisone
1948	Long-playing record
1948	Polaroid land camera
1949	Thonet furniture
1949	Polyester
1950	Computer
1950	Transistor
1950	Xerography
1950	Terylene
1950	Radial tire
1951	Double-floor railway
1953	Cinerama
1953	Color television
1954	Nuclear energy
1954	Gas chromatograph
1954	Remote control
1954	Silicon transistor
1956	Air compressed building
1957	Atomic ice breaker
1957	Space travel
1958	Stitching bond

(continued)

1958	Holography
1958	Transistor radio
1958	Diffusion process
1958	Fuel cell
1959	Quartz clocks
1959	Polyacetates
1959	Float glass
1960	Maser
1960	Micro modules
1960	Polycarbonates
1960	Contraceptive pill
1960	Hovercraft
1961	Integrated circuit
1961	Planar process
1962	Laser
1962	Communication satellite
1963	Implementation of ions
1963	Epitaxy
1964	Synthetic leather
1964	Transistor laser
1966	Optoelectronic diodes
1967	Wankel-motor
1968	Video
1968	Light emitting fluor display
1968	Minicomputers
1970	Quartz watches
1971	Microprocessor
1971	Electronic calculator
1972	Light-tunnel technology
1975	16-bit microprocessor
1976	16,384 bit RAM
1976	Microcomputer

References

Baker, R. (1976). *New and improved: Inventors and inventions that have changed the modern world*. London: British Museum.

Cincera, M. (1997). Patents, R&D and technological spillovers at the firm level: Some evidence from econometric count models for panel data. *Journal of Applied Econometrics, 12*, 265–280.

Crepon, B., & Duguet, E. (1997a). Estimating the innovation function from patent numbers: GMM on count panel data. *Journal of Applied Econometrics, 12*, 243–263.

Crepon, B., & Duguet, E. (1997b). Research and development, competition and innovation pseudo-maximum likelihood and simulated maximum likelihood methods applied to count data models with heterogeneity. *Journal of Econometrics, 79*, 355–378.

Greene, W. (1995). *Econometric analysis*. Upper Saddle River, NJ: Prentice Hall.

Hausman, J., Hall, B. H., & Griliches, Z. (1984). Econometric models for count data with an application to the patents-R&D relationship. *Econometrica, 22*, 909–938.

Haustein, H. D., & Neuwirth, E. (1982). Long waves in world industrial production, energy consumption, innovations, inventions, and patents and their identification by spectral analysis. *Technological Forecasting and Social Change, 22*, 53–89.

Hirooka, M. (2006). *Innovation dynamism and economic growth: A nonlinear perspective*. Kyoto: Institute of Technoeconomics.

Kleinknecht, A. (1981). Observations on the Schumpeterian swarming of innovations. *Futures, 13*, 293–307.

Kleinknecht, A. (1990). Are there Schumpeterian waves of innovations. *Cambridge Journal of Economics, 14*, 81–92.

Mensch, G. (1979). *Stalemate in technology. Innovations overcome depression*. Cambridge: Ballinger.

Sahal, D. (1974). Generalized poisson and related models of technological innovation. *Technological Forecasting and Social Change, 6*, 403–436.

Schumpeter, J. A. (1939). *Business cycles: A theoretical, historical and statistical analysis of the capitalist process*. New York: McGraw-Hill.

Silverberg, G., & Lehnert, D. (1993). Long waves and 'evolutionary chaos' in a simple Schumpeterian model of embodied technical change. *Structural Change and Economic Dynamics, 4*, 9–37.

Solomou, S. (1986). Innovation clusters and Kondratieff long waves in economic growth. *Cambridge Journal of Economics, 10*, 101–112.

Van Duijn, J. J. (1983). *The long wave in economic life*. London: Allen & Unwin.

Financial Instability Under Innovation Development: Reasons and Regulation Within the Model of Evolutionary Processes

Yuri Ichkitidze

Abstract In this paper, we suggest an approach to the study of the financial instability based on the model of evolutionary processes. In the first place, we present some empirical facts that confirm that the stock's price dynamics is better described by the Markov switching model rather than by the pure random walk. Further, using the equilibrium model of price formation, we show that the temporary price trends on stock market are evolutionary processes that occur in the conditions of a duality of the equilibrium between the market price and the fair value. Then, within the framework of the constructed model, we analyze the causes of the financial market instability and its impact on the real sector, and show how the financial markets create a destructive impulse under the economic growth slowdown, and therefore adversely affect the process of innovations diffusion into the market. The conducted study shows that the causes of the financial instability are the capital concentration in the narrow circles of society and the lack of investment opportunities, as compared with the available financial resources, whereas the symptoms are frequently recurring financial bubbles and crises.

Keywords Markov switching • Multiple equilibria • Feedback effects • Evolutionary processes • Cycles

1 Introduction

The financial instability hypothesis proposed by Minsky (1972, 1992) became widely discussed after the crisis of 2008. Main discussions focus at the sources of negative influence of the asset market over economic development, as well as the shifting mechanisms in financial systems. The crux is whether financial crises are the result of economic activities of a society, or they perform independent external shocks potentially existing under any condition of the economic system. Minsky

Y. Ichkitidze (✉)
Department of Finance, National Research University Higher School of Economics,
Kantemirovskaya Street 3, St-Petersburg, Russia
e-mail: ichkitidze@reflexivity.ru

T. Devezas et al. (eds.), *Industry 4.0*, Studies on Entrepreneurship, Structural
Change and Industrial Dynamics, DOI 10.1007/978-3-319-49604-7_7

himself followed Schumpeter's traditions, i.e. considered preceding period of expansion and business growth to be the source of structural changes in the financial system, which cause instability in the future.

To speak in simple terms, his idea lies in the fact that risks assumed by the economic system increase during continuous economic development, so assets value is increasingly maintained by expectations for further prices growth. He includes financial intermediaries seeking profit (he calls them "merchants of debt") in his model and emphasizes that they "strive to innovate in the assets they acquire and the liabilities they market" (Minsky 1992, p. 6). According to him, this particular bankers' striving for innovations is the reason of further negative structural changes.

Basic financial system models studying mechanisms of financial markets instability at length are the models of investment bubbles and bank panics. Bubbles tend to appear under financial deregulation and increase in liquidity, and can be explained mainly by nonrationality of the participants (see Kindleberger 1978; Shiller 1984, 2000; De Long et al. 1990a, b; Shleifer and Summers 1990; Shleifer 2000). Bank panics feature conditions of multiple equilibriums (see Diamond and Dybvig 1983; Morris and Shin 1998; Goldstein and Pauzner 2005), due to which negligible external variables can easily cause swing in the financial system. Separate models of bank panics describe endogenous changes in the financial system (Wallace 1988, 1990; Chari 1989; Hellwig 1994; Alonso 1996; Allen and Gale 1998), in particular, financial crisis caused by increased risks of recession.

Leaving the results of the above mentioned models unquestioned, we must note that they rather describe symptoms of economic instability, but not its origin; first of all, they ignore peculiarities of the innovation process supposing its continuity, and also do not take into consideration interconnection between the asset price and its fair value, which causes conditions of dual equilibrium. In the present article we include these factors into our study and come to conclusion that financial instability regulation applying monetary policy tools (including limitations imposed for profiting and measures aimed at centralized control over financial risks amount) turns to show poor efficiency; it only postpones issues for an uncertain period of time, but not provides solution. We provide the following argumentation in order to defense our point of view. Firstly, we determine temporary trends for USA stock price index empirically and demonstrate their connection to diffusion of innovations and periods of high inflation. Secondly, considering the conditions of dual equilibrium between the asset price and the fair value, we design a model of balanced price setting, which reflects that distortion of asset prices (and swings of systems related to it) happens through striving of rational participants of the market to the profit. Thirdly, we demonstrate ways of regulation of financial instability in the framework of the proposed pricing model and show that effective regulation of financial instability must involve elimination of excessive level of income inequality and development of socially significant investment.

2 Evolutionary Processes in Financial Markets

Understanding of financial instability mechanisms greatly depends on the results of discussion on share price performance. In case changes in stock indexes are a proper random walk (Bachelier 1900; Osborne 1959), and markets themselves meet efficiency concept (i.e. "the price is a total reflection of all available information"), then the need in regulation of markets (including financial instability), as well as possibility to perform effective regulation, seem to be doubtful. Everything changes, if because of irrational behavior of separate participants (see models proposed by Shiller 1984; De Long et al. 1990a, b; Shleifer and Summers 1990; Shleifer and Vishny 1997; Daniel et al. 1998; Barberis et al. 1999), the share price performance sees systematical deviations from fair values from time to time. In such case, the share price performance shall meet the model with slightly decomposing permanent component part, proposed by Summers (1986). Characteristic feature of this model is that, as the random walk model, it provides independence of gradual price increments, nevertheless, it explains cyclical swings of stock indexes through not simple accumulating of random events, but evolving the general run. Within this model, financial instability is considered as the result of investor's failure to act in a rational manner, and the only way to improve the situation is to limit their ability to influence the prices destructively; for instance, through excessive liquidity in the financial system (see Yuan 2005) or by imposing a tax on transactions (like Tobin tax). Nevertheless, a third, alternative model exists. This model is based on Hearst's imitation (see Peters 1989, 1994), and within wider class of models it is a variant of the swing model of Markov processes (Hamilton 1989). According to this model, there exist temporary-constant trends in share prices, which rarely change course, moreover, independence of gradual price increments is constructed through stationary fluctuations around the trend. These temporary-constant trends at first approximation can be used in order to distinguish evolutionary processes in financial markets. If we demonstrate that such trends exist in fact, it will make no difficulty at the next stage of the research to prove that these trends are evolutionary processes. In such case, we can regard the trend-proof model of share price performance as the stochastic model of evolutionary dynamics.

These three alternative models of share price performance are represented in mathematical formula 1.

$$p(t) = p(0) + \sum_{i=1}^{t} \lambda \cdot \mu_i + \sum_{i=1}^{t} \omega^{t-i} \cdot \upsilon_i(0, \sigma) \tag{1}$$

If $\lambda = 0$ and $\omega = 1$ we have a proper random walk, if $\lambda = 0$ and $0 < \omega < 1$ we receive the model proposed by Summers (1986), and finally, if $|\lambda| > 0$ and $0 < \omega \leq 1$,[1] the trend proof model comes out.

[1] If $|\lambda| > 0$ and $\omega =$ we have Markov switching model, proposed by Hamilton (1989) and widely discussed nowadays.

Econometric problem is that all three variants of this model, even at realistic evaluation of the parameters, show similar observable outcome, namely: (i) independence of gradual price increments, i.e. low (statistically non-significant) values of autoregression index; (ii) a posteriori normality; and (iii) similar indexes of efficiency if using filter rule. This implies low power of common statistical tests—even if the test does not allow to refuse the basic hypothesis, there remains high probability that the alternative hypothesis is true. Hence, it is not surprising that each model receives telling arguments, and the issue of share price performance is unclear. At the same time, the value of the matter is very high. Our understanding of true reasons (mechanisms) of financial instability depends on the results of this discussion, i.e., in fact, on the fact, which variant of the model is more consistent with reality. This fact, in its turn, determines regulation methods.

Further in this paper we bring arguments, which prove that the trend-proof model describes share price performance in a more accurate way as compared with alternative models. We distinguish long-term trends related to innovative development of the society and interrupted by periods of anomalously high inflation using the example of monthly data for USA share price index. The fact that they constitute trends and not just a special set of circumstances accepted within the random walk model is proved by statistical tests.

Let us consider the price dynamics in the US market in the period since 1871 (see Fig. 1). In the first approximation, we can distinguish four periods of growth (from 1871 until September 1906, from December 1920 to October 1929, from June 1949 to December 1968 and from March 1982 until now), as well as three downward trends (the rest of the time). From the point of view of the efficient market hypothesis, these price fluctuations were explained by a random change in the risk premium that is inconstant in time (see Fama 1991), whereas the model of behavioral finance (Shiller 1984; Shleifer 2000; Shiller 2000) interpreted them as fluctuations caused by fashion or mood—by irrational actions of non-professionals. Adhering to the hypothesis of temporal trends existing in the stock prices, noteworthy is the fact that each downward trend in Fig. 1 was accompanied by an unusually strong inflation. Thus, in the period from March 1916 to July 1920, the level of consumer price growth exceeded 15 % per annum; during the same period, the real return on stocks was −7.0 %. Then, from March 1941 to July 1948, inflation was slightly over 7.6 %, at the same time, the real return on stocks was −7.0 %. Finally, from March 1973 to August 1982, with an inflation of 9.0 %, the real return on stocks was −6.0 %. During all other periods of time, such a high and long-term inflation was not observed, whereas the real return on stock was 6 % per annum on average.

Let us study regressive model of first-order difference performance based on observed correlation indexes as the null hypothesis. According to it, we can represent the prices level as follows

$$\Delta P_t = \sum_i a_i \cdot \Delta P_{t-i} + b + \varepsilon,$$

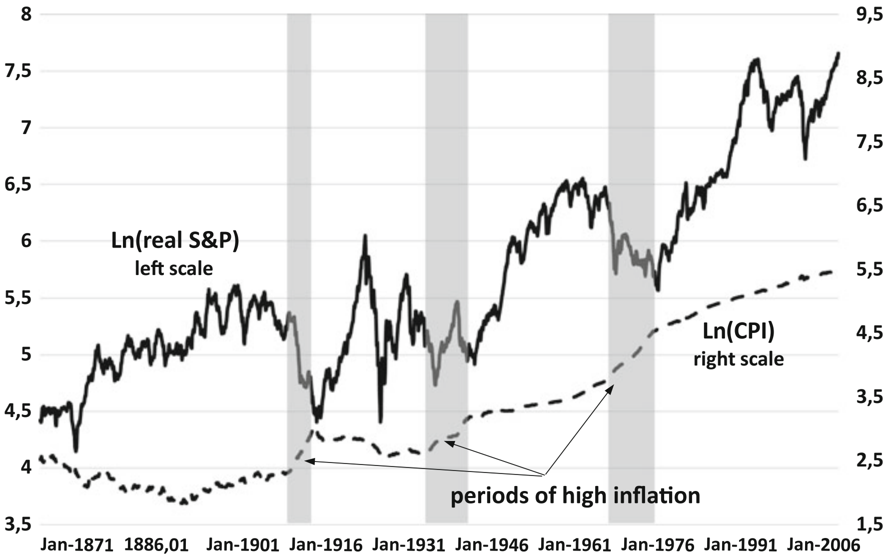


Fig. 1 S&P500 in constant prices and the consumer price growth

and S&P500 in the current prices from the Schiller's base as[2, 3]

$$\Delta Y_t = \sum_i c_i \cdot \Delta Y_{t-i} + \sum_i d_i \cdot \Delta P_{t+i} + f + \varepsilon$$

Trend-proof model is an alternative hypothesis. It means that there exists a connection between trends between USA inflation level and share price index, i.e. share prices fall mainly during high inflation and, on the contrary, gain mainly during low inflation.

In order to check null hypothesis in the given alternative, we will perform the following statistical test: select period, when average inflation in the USA for the last 6 months exceeded $k\%$ per annum, for k we will choose values ranging from -5 to 20 %, and, determine, how share index value was changing during this period. The results are shown in Fig. 2, we can observe, that if k ranges from 10 to 17 %, average change in share indexes during high inflation was less than 1 %. As fare as the statistics for share index change during high inflation (when inflation for the latest 6 months exceeds $k\%$ per annum) in accordance with the null hypothesis is significantly higher (lower quantile of 1 % for k ranging from 10 to 17 % lies at range from -0.7 % to -1), than we can refuse null hypothesis with reliability of 0.99. Thus, the conducted experiment proves that inflation acts as the predictor of

[2]This regression, but not a random walk, take place because the initial data in the Schiller's base include autocorrelation coefficient of 0.28, as they perform data smoothed as for daily sliding averages.

[3]All model parameters are evaluated applying least square method.

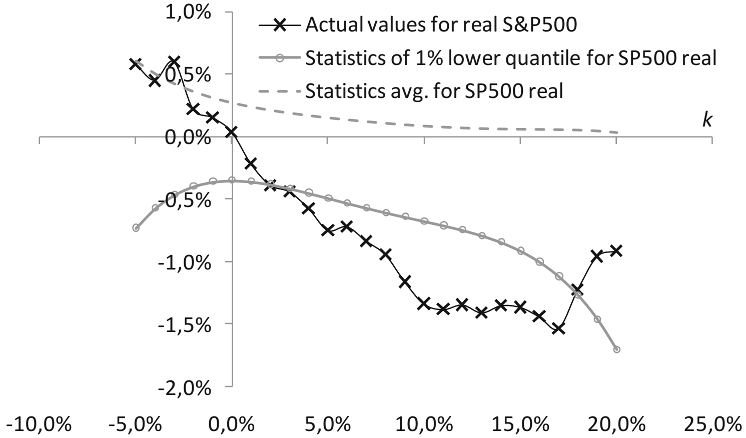


Fig. 2 The average profitability of stock price trends when the inflation during the last 6 months exceeded *k*% per annum, and the statistics of the average trend profitability in difference-stationary models

the future share indexes profitability, regardless of not high values of the correlation index, which enables us to gain additional profit on the average, provided that we have cut down the share fraction in the portfolio at inflation exceeding a certain level (*k*% per annum), and increased the share fraction at the inflation dropping lower this level.[4]

Apparently, long-term trends of the market shares revealed in the course of the experiment are associated with technological (innovative) development of the society, primarily with consequences of diffusion of innovation products (and technologies) on the market. Hirooka (2003, 2006) has defined the notion of mainstream innovations. It is easy to notice that innovation development waves, he distinguished (1897–1929, then 1950–1974, and after 1980), exist in tight correlation with the indicated share market trends. Inflation only acts as a constant diffusing the mainstream innovations' impact on the share market growth. This established fact can be considered as empirical evidence of the financial instability hypothesis (Minsky 1972, 1992).

3 Mathematical Model of Evolutionary Processes

There are not so many share price performance models explaining the appearance of temporary trends. They mostly have behavior-based approach in its basis, i.e. suggest non-rational reaction of market participants to the information they

[4]Unfortunately, achieving positive results from implementation of this strategy may require decades.

receive (see Barberis et al. 1998; Hong and Stein 1999; Frazzini 2006). Nevertheless, evolutionary approach can also be applied in interpreting temporary trends mechanisms. It features endogenous character of dynamic changes; which becomes possible due to situations of multiple equilibrium. Endogenous changes in the market, applied to share prices, take place, if multiple equilibrium between the market price and the fair price exists (Soros 1987; Ozdenoren and Yuan 2008; Ichkitidze 2013, 2016). For banks and corporate bonds, multiple equilibrium appears between the expected and real probability of default (Diamond and Dybvig 1983; Ichkitidze and Zvontsov 2016). Macroeconomic conditions can become the source of multiple equilibrium, as well, firs of all, equilibrium between investment and savings that depends on economic activity level (Kaldor 1940; Chang and Smyth 1972). Another advantage of the evolutionary approach is that it helps explain observed behavior of the agents within the rational behavior model.

In order to understand the mechanism of price setting under conditions of dual fundamental equilibrium, we will study the following mathematical model. Let us suppose the market has two groups of homogeneous competent investors, each one having expectations of the fair price connected only with one equilibrium (P_{f_a} and P_{f_b}, correspondingly). Each participant presents demand on shares that can be described by the demand function of the following type

$$Q^d(P) = \frac{a(P - p_0)^2 + b(P - p_0)}{(P - p_u)^k} + c \tag{2}$$

where,

$$p_0 = \frac{P_f + P_m}{2}$$

$$p_u = (1 + \gamma) \cdot P_f$$

a, b, c—parameters estimated for each individual case applying the least square method.

The graphical form of the given demand function is shown in Fig. 3 and includes two sectors: classical and reflected. Empirically such demand function was defined in the work of Yuan (2005), detailed theoretical background had been given earlier in Ichkitidze (2006).

Based on the demand individual function, we can construct the joint function of all investors' demand, which shall be

$$Q^D(P_t) = \alpha \cdot Q^d_\alpha(P_t) + \beta \cdot Q^d_\beta(P_t)$$

where,

α, β—number of investors in A and B groups, correspondingly (equilibrium expectations are P_{f_a} and P_{f_b}, correspondingly).

In such case, the condition for the market equilibrium is represented as

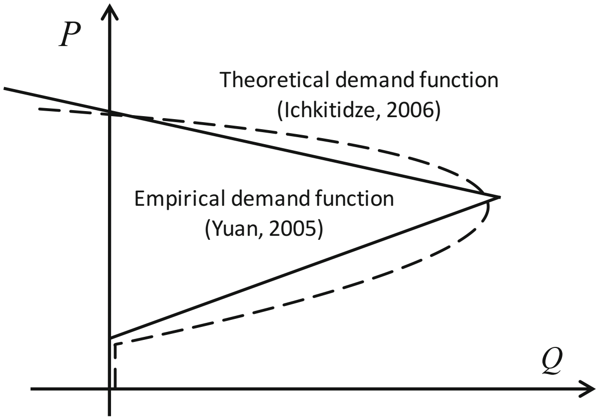


Fig. 3 Individual function of the optimal investor's demand for shares

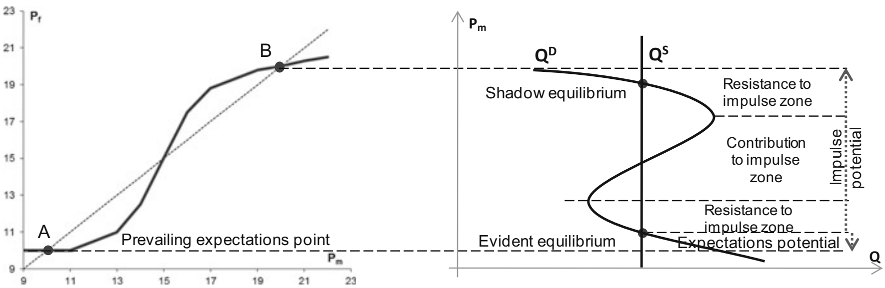


Fig. 4 Market equilibrium under dual nature of the fundamental equilibrium

$$Q^D(P_t) = N$$

where,

N—number of placed securities.

Peculiarities of demand individual function (2) and conditions of the equilibrium duality shape the joint demand function as a Z-type function. Conditions of the market equilibrium in the case of dual equilibrium between the price and the fair price are schematically shown in Fig. 4.

Shall no exogenous changes exist, shift in the balance price within the model is possible due to change in the participants' choice, i.t. due to impulse. Mathematically it is characterized by decrease in one group size and increase in the second one. Where endogenous changes take place, condition of the market equilibrium in time is

$$(\alpha + \lambda) \cdot Q_\alpha^d(P_t + \Delta P) + (\beta - \lambda) \cdot Q_\beta^d(P_t + \Delta P) = \alpha \cdot Q_\alpha^d(P_t) + \beta \cdot Q_\beta^d(P_t)$$

where

λ is impulse, i.e. number of participants moved from one group to the other,

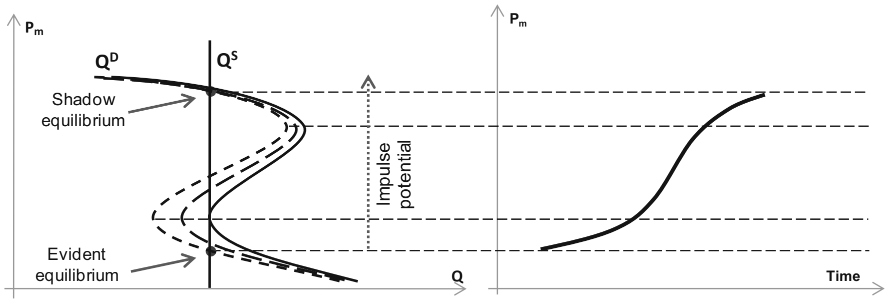


Fig. 5 Shift in the market equilibrium due to endogenous changes and appearance of a temporary trend

From which it is clear that shift in the price (ΔP) is the function of impulse. Graphically, endogenous changes in the market equilibrium within the model are represented in Fig. 5.

It is not difficult to notice that in conditions of the model studied above shift in the equilibrium caused by endogenous changes represent evolutionary processes. To talk about evolution, we must rely on three arguments: (1) change of an index of the system (market share price) occurs under fixed external conditions of the model; (2) in motion the system undergoes qualitative changes, i.e. its shift from one rational equilibrium condition into another; (3) such change is irreversible, i.e. it cannot return into its initial state. All these arguments are true for the model studied above. Now we can claim that temporary trend in the market share prices is the result of the evolutionary process (and inversely, we can describe the evolutionary process using the trend of market price per share).

4 Instability of Financial Markets and It Impact on the Macroeconomic Development

Now, in the framework of the proposed model, we have any ground to study the origin of financial instability. Under dual equilibrium the impulse from financial markets is directed to that equilibrium type, which includes the higher potential for profit, that is why in the beginning, when investment opportunities appear, the trend for increase in the assets value arises. The speed of this trend may vary. Two situations may possibly take place: (a) the speed is high, and the potential for growth shall exhaust itself quickly, and (b) the speed is slow, and the potential for growth shall have continuous character. Financial instability, i.e. destructive influence of financial markets on the real economy, shall arise only in the first case. It will be triggered by participants' striving for profit under conditions, when all opportunities to make profit through the ascending trend have been exhausted already. Market participants have an opportunity to make extra profit through the shift of the financial system.

Such situation implies that, at the takeoff stage, regulation of assets value growth is the key to regulation of financial instability. Let us have a closer look at the market participants' structure within the model. To exhaust trend potential for growth in short terms we shall require: (1) accurate knowledge of the market participants about the "good" balance; (2) significant volumes of assets in possession of market participants. On the contrary, if prospects for the "good" equilibrium are dim, i.e. its implementation is associated with considerable difficulties, and the market participants' wealth is insufficient as opposed to the wealth potential, the future investment contain, then the price will be shifting to the "good" equilibrium slowly. The latter means the financial instability is not to arise: financial markets and financial system will self-regulate without impact of the government authorities, in accordance with the classical theory predictions, and doubts in its capability to arrange its activities in a rational manner shall not arise.

Regulation is the conclusion from this. Increase in income inequality caused by innovative development leads to increase in savings (financial market participants' wealth), which contributes to faster decrease of the potential. At the same time, consumption possibilities of other members of the society become limited. At the moment, when the potential of the descending trend starts to dominate, systems shift and financial markets begin to influence the real economy destructively, both slowing down innovations, and easing consumer (and investment) demand. Increasing of income redistribution during takeoff (therefore decreasing savings and wealth available for financial markets participants), will result in avoiding such situation, as considerable part of the potential for growth will remain favorable system of financing. Excess funds also can be removed from financial markets by means of social (infrastructure) investment, as they make low return, but social benefit in a long-term perspective is significant. The graphic of the system evolution caused by technological innovations at a stable financial market (a) and at financial instability (b), is shown in Fig. 6.

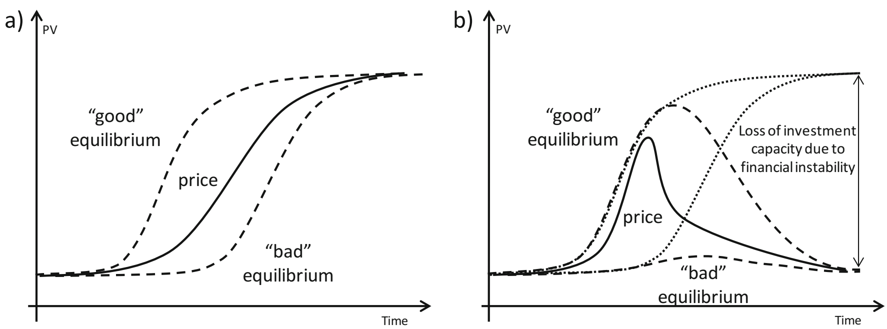


Fig. 6 System evolution due to technological innovations in a stable financial market (a) and at financial instability (b)

The above said provides general features of the financial instability mechanisms within the evolution processes model, however, it does not constitute a complete concept. For its completion and understanding of the whole impact from these mechanisms on the real economy, it is necessary to overview two suppositions in addition to the above described model, which lie not far from the empirical facts:

1. Financial markets system influences dynamics of equilibrium points;
2. While the economy is flourishing (gaining wealth), issues contributing to lowering of the "bad" equilibrium point are being accumulated.

Taking these suppositions into consideration, from the perspective of the preceding mathematical model, financial instability concept may be stated in words as a sequence of the following six stages.

1. Harmony in the society and stability (including small inequality of incomes) create potential for economic growth.
2. Through potential innovations are created and significant investment opportunities appear (on the visible horizon they are countless).
3. For the fastest economic growth driven by these investment opportunities minimal income redistribution is required, as well as deregulation of the financial markets and liquidity.
4. Period of economic prosperity exhausts the potential; inequality of incomes increases, motivation gets interrupted, social classes with strong division appear, education system falls behind noticeably. More and more pseudo-innovations appear among new investment opportunities. These effects are smoothed over at the first stage, until the financial system remains unchanged. Flourishing continues, regardless of the increasing problems.
5. The shift in the financial system becomes the key time point of the oncoming crisis. Change in mood at the financial market no longer can compensate increasing problems. Now financial markets are not only neutral in respect of economy, but influence it destructively, aggravating the problems, and new innovations as well.
6. The system gets out of control: unemployment level increases, the economy is far off its potential opportunities, provided by the level of scientific and technical progress, inequity in the society goes over-the-top, destructive social phenomena appear. The solution in such situation may be a long lasting inflation. In the absence of such scenario, introduction of authoritarian governing system can be an option.

These six stages of financial instability are schematically represented in Fig. 7. They represent theoretical illustration of empirical cycles of the USA stock market, shown in Sect. 1 of the present article. They can be regulated through more intensive redistribution of income (and large social investment). Suh measures will limit growth at the first stage and will lead to refusal from several investment opportunities, however, will make the system more stable, dynamically as well; the harmony in the system will remain, which will allow the potential to renew itself timely. Negative effects from structural changes will show themselves in a smaller

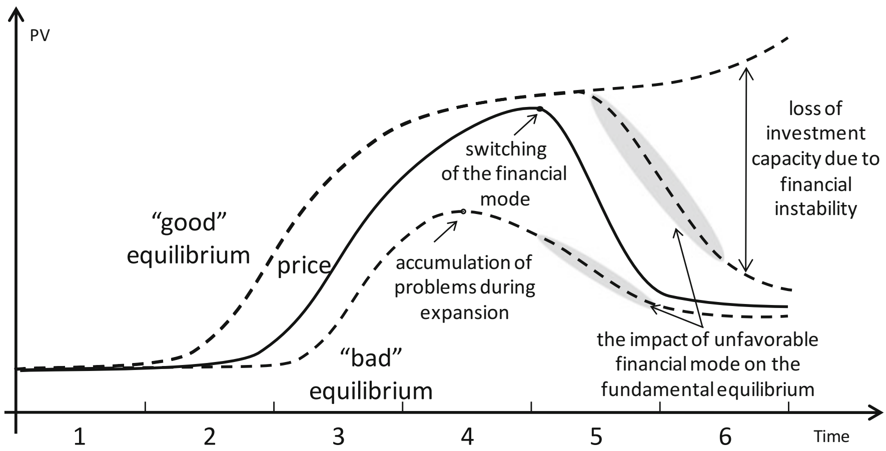


Fig. 7 Stages of financial instability

extent during growth, and the following shift in the financial system will be whether temporary and short, or will not occur at all. While change in waves and mainstream of innovations (Glazyev 1993; Hirooka 2006; Akaev and Rudskoy 2014), financial system will not have destructive impact on innovations.

5 Empirical Analysis of the Existing Trends

Dual nature of equilibrium and, correspondingly, capital asset price movement, according to the scheme proposed above, mean that such phenomenon as "spirit of time", i.e. prejudicial subjective distortion of the environment by the participants, exists, so we can realize, what kind of work such "spirit of time" performs, and how we perform: support it or resist it? In the latter case, all our attempts to reach the goal are doomed to failure, that is why here exists only deferring of the end.

Foregoing arguments demonstrated that the "spirit of time" acts towards the greater potential, and income inequality and concentration of capital within the limited groups of people are the major reasons of financial instability. In order to balance the situation and let the "spirit of time" lead us in a prosperous direction, it is necessary to pay more attention to social programs and redistribute resources in favor of less wealthy social groups.

The current stage of the financial system evolution started from 1980–1982, when the core countries (countries-G7 members) overcame last evidences of high inflation, and development of computer and informational technologies speeded up the rates of economic growth. Starting from the same period of time structural changes, later turned to be sources of financial instability, appeared; its first vivid symptoms revealed itself during stock-market crisis of 1987 and Japanese crisis of 1989, as well as in the form of series of crises in developing markets in the 90s.

Full-sized malady became vivid in the 2000th, Nasdaq sector crisis in 2000 turned to be its symptoms, as well as the mortgage crisis of 2008.

The present stage of the system evolution is different from the previous periods in the international character of the financial instability, as unprecedented integration of the global economy occurred between periods after 1990. This event produced an interesting effect, which was not mentioned in the previous model. When countries compete for investment projects, saddle-focus is that they do not need to regulate inequity of incomes, as far as at the first stage such actions will make the biggest profit, and the future consequences remain undecided. Those, who offer the best conditions to the proprietaries (project initiators), will indeed win the competition, on the contrary, those who set high level of income redistribution, will lose. Such feature enables several countries to profit from financial instability at the first stage by letting it exist and affecting the rest of the world as the hegemon afterwards.

This theoretical thesis can be illustrated empirically through the upward trend in the US deficiency in the current account balance, which started from the 1980s (see Fig. 8).

The USA became the attractor of the capital from the whole word, because it offered the best conditions (institutional as well), at the same time, they won competition for investment projects: academic sphere and testing laboratories, along with acceptability of the financial sources and development of consumer markets, all this made the USA attractive for entrepreneurs of the whole world. In this context higher level of income redistribution would look illogical, but lowered significantly those profits, which the USA provided for profitable entrepreneurs. Over time, as the innovations were developing, the inequity continued to increase (Gini index rose from 0.38 in 1980 to 0.45 in 2010), but this problem was covered up at the first stage by increase in assets price (wealth effect). Saving rate fell to the

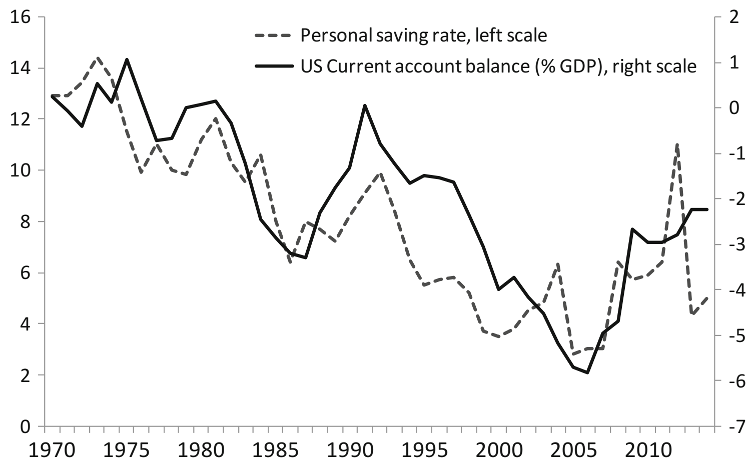


Fig. 8 Deficiency in the current account balance and personal saving rate in the USA

historically low level (see Fig. 8); those, who did not keep up with the common national growth of the living standard, ran into debt over time (upon assets as well). All these structural changes of the financial system during expansion constitute the sources of future financial instability, the one Minsky talked about.

The crisis of 1987 did not have significant impact on economic development and was caused by specific character of market risks management (for more information on the reasons see Leland and Rubinstein 1988; Black 1988; Gennotte and Leland 1990; Romer 1993); it was, in fact, natural result of the swing of the economic system from one fundamental equilibrium into another shifted by innovation process impact. And the Japanese financial crisis, which came next in 1989, impacted the real economy greatly, however, it had only local, sub-national character. Then the 90s started with their range of regional crises of the developing markets. Political transformation of 1991 triggered development of international trade: the core countries got access to new product markets, while the fringe countries—to the capital and new technologies; as the result synergistic effect appeared, predefining direction of the capital flow. However, the financial capital, which flew to the developing countries through stock markets, did not ensure sustainable improvement of the living standard; that is where the above studied scheme of the financial instability worked, and after potential for growth turned to be exhausted, the system swung themselves. The Mexican crisis of 1995, Asian crisis of 1997, followed by the Russian crisis of 1998 occurred successively.

Incapability of the fringe countries to ensure sustainable improvement of the living standard under financial markets instability resulted in the era of economic development of the fringe countries started from 1999–2000 maintained by accumulation of gold and forex reserves. It introduced a certain degree of stability to the shaping structure of the global economy and made it possible to implement significant part of the synergistic potential. However, since accumulated foreign exchange reserves flew back to the core countries (first of all, to the USA) in the form of investment into the government bonds under minimum interest, the core countries still made considerable profit from the situation created. As a matter of fact, the reason was that great part of resources was exported for the fringe countries almost for free.

The period of the core countries' expansion in 2001–2008 was characterized by consumer demand increased due to the boom at the real property market. Same with the situation in the 1990s, increasing inequality of incomes could be ignored due to wealth effect, but the crisis of 2008 showed that such way of development is futureless. Nowadays, within the former technological waves, it turned out that the core countries have nothing to offer to the rest of the world. Moreover, financial instability now constitutes a threat for them.

The crisis of 2008 was the turning point in the modern history of economic evolution, the degree of its influence on social and economic processes remains unclear. Financial instability performs a problem, as it is accumulating all the time; only its external evidences are solved during crises—symptoms, while the reason of the malady—excessive inequality of incomes, remains unchanged. During 2009–2012 the symptoms were relieved through unprecedented procedure of quantitative easing, resulted in considerable growth of the money base, while impaired illiquid assets had stored on the FRS balance, and the US banking system experienced excess liquidity. In such way the system was balanced and the time was gained. However, now the system must provide innovations in order to ensure the balance of the bloated financial sector corresponding with the real economy content. The fact that it is an obligation, but not a right signifies financial instability in the present crisis-free period. The issue of innovations diffusion is to become the key one. While the external demand is quite weak because of the backwardness of the rest world in technologies and social and economic institutions, internal demand is limited by excessive household debt, so if the malady is going to develop depends on opportunities for the joined demand broadening. The experience of the preceding periods shows that the problem of inequity and excessive household debt in the core countries may be solved by high inflation rate. However, due to the peculiarities of the modern structure of the global economy this time interests of other large countries will be affected, and it is still the question, if such solution is advantageous. That integrated global economy of the 2000s may no longer exist.

6 Concluding Remarks

In fact, the financial instability issue under dual fundamental equilibrium resides in the fact that short-term structural changes look attractive, while their long-term consequences are not considered to be worth of attention. Excessive inequality of incomes is "veiled" successively as long as favorable financial system remains active, however, as soon as the financial system changes, even more sever crisis, than it had been expected, occurs.

In this work we suggest an approach to financial instability analysis based on the evolution processes model. Many of its theses represent only a set of ideas that still must be discussed and better proved. The one thing is clear for now, assets price performance shows temporary trends connected with diffusion of the mainstream innovations and interrupted by periods of high inflation. The suggested model of price setting under dual fundamental equilibrium proves that the true reason of the financial instability is the increase in inequity of incomes caused by innovation process, but not the financial system structure. This conclusion gives way for a new approach to development of stable financial systems.

Our forecast for the next two decades is related to the fact that the financial instability continues exerting a strong influence on the development of the global economy towards the transformation of international trade relations. In particular, except for the series of regional financial crises similar to the Russian foreign exchange crisis in 2014, a possible scenario is an abnormal increase in the inflation rates in the developed countries (an average of 10 % per annum during 5–10 years), the preconditions for such scenario being created by the consequences of the crisis in 2008. This scenario determines the fact that the developing countries, to ensure their economic growth, will seek to enter into regional alliances and create a monetary and financial system independent from the core system, that will ensure their development at the achieved scientific and technological level.

Acknowledgments The research and its findings has been supported by the Grant of the Russian Science Foundation (Research project No. 14-28-00065 "Structural and cyclical paradigm of economic and technological renovation of macro-systems (World and Russia in the first half of the XXI century)").

References

Akaev, A. A., & Rudskoy, A. I. (2014). Synergetic effect of NBIC—technologies and economic growth in the first half of the XXI century [Sinergeticheskij effekt NBIC—tekhnologij i mirovoj economicheskij rost v pervoj polovine XXI veka (in Russian)]. *Ekonomicheskaja politika, 2*, 25–46.

Allen, F., & Gale, D. (1998). Optimal financial crises. *Journal of Finance, 53*, 1245–1284.

Alonso, I. (1996). On avoiding bank runs. *Journal of Monetary Economics, 37*, 73–87.

Bachelier, L. J. B. A. (1900). *Théorie de la speculation*. Paris: Gauthier-Villars.

Barberis, N., Huang, M., & Santos, T. (1999). *Prospect theory and asset prices* (NBER Working Paper No. 7220).

Barberis, N., Shleifer, A., & Vishny, R. (1998). A model of investor sentiment. *Journal of Financial Economics, 49*(3), 307–343.

Black, F. (1988). An equilibrium model of the crash. *NBER Macroeconomics Annual, 3*, 269–275.

Chang, W. W., & Smyth, D. J. (1972). Stability and instability of IS-LM equilibrium. *Oxford Economic Papers, 24*(3), 372–384.

Chari, V. (1989). Banking without deposit insurance or bank panics: Lessons from a model of the U.S. national banking system. *Federal Reserve Bank of Minneapolis Quarterly Review, 13* (Summer), 3–19.

Daniel, K., Hirshleifer, D., & Subrahmanyam, A. (1998). Investor psychology and security market under-and overreactions. *The Journal of Finance, 53*(6), 1839–1885.

De Long, J. B., et al. (1990a). Positive feedback investment strategies and destabilizing rational speculation. *The Journal of Finance, 45*(2), 379–395.

De Long, J. B., et al. (1990b). Noise trader risk in financial markets. *Journal of Political Economy, 98*(4), 703–738.

Diamond, D., & Dybvig, P. (1983). Bank runs, deposit insurance, and liquidity. *Journal of Political Economy, 91*, 401–419.

Fama, E. (1991). Efficient capital markets: II. *The Journal of Finance, 46*(5), 1575–1617.

Frazzini, A. (2006). The disposition effect and underreaction to news. *The Journal of Finance, 61* (4), 2017–2046.

Gennotte, G., & Leland, H. (1990). Market liquidity, hedging, and crashes. *The American Economic Review, 80*(5), 999–1021.

Glazyev, S. Y. (1993). The theory of long-term technical and economic development [Teoriya dolgosrochnogo thechniko-economicheskogo razvitiya (in Russian)]. Moscow: VlaDar.

Goldstein, I., & Pauzner, A. (2005). Demand deposit contracts and the probability of bank runs. *Journal of Finance, 60*(3), 1293–1327.

Hamilton, J. B. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica, 57*(2), 357–384.

Hellwig, M. (1994). Liquidity provision, banking, and the allocation of interest rate risk. *European Economic Review, 38*, 1363–1389.

Hirooka, M. (2003). Nonlinear dynamism of innovation and business cycles. *Journal of Evolutionary Economics, 13*(5), 549–576.

Hirooka, M. (2006). *Innovation dynamism and economic growth*. Cheltenham: Edward Elgar.

Hong, H., & Stein, J. C. (1999). A unified theory of underreaction, momentum trading, and overreaction in asset markets. *The Journal of Finance, 54*(6), 2143–2184.

Ichkitidze, Y. R. (2006). On reflexivity of the financial market [O refleksivnosti finansovogo rynka (in Russian)]. *Finance and Business, 1*, 25–38.

Ichkitidze, Y. R. (2013). Reflexive model of stock market [Refleksivnaya model fondovogo rynka (in Russian)]. *Vestnik Engecona*, No. 6, 73–84.

Ichkitidze, Y. R. (2016). The duality of equilibrium between the stock price and the fair value for a fast-growing company. *Proceedings of International Conference on Soft Computing and Measurements, SCM 2016*.

Ichkitidze, Y. R., & Zvontsov, A. V. (2016). The optimal debt size under instability of financial markets. *Proceedings of International Conference on Soft Computing and Measurements, SCM 2016*.

Kaldor, N. (1940). A model of the trade cycle. *The Economic Journal, 50*(197), 78–92.

Kindleberger, C. (1978). *Manias, panics, and crashes: A history of financial crises*. New York: Basic Books.

Leland, H., & Rubinstein, M. (1988). Comments on the market crash: Six months after. *The Journal of Economic Perspectives, 2*(3), 45–50.

Minsky, H. P. (1972). Financial stability revisited: The economics of disaster. In *Reappraisal of the federal reserve discount mechanism* (Vol. 3, pp. 95–136). Washington, DC: Board of Governors of the Federal Reserve System.

Minsky, H. P. (1992). *The financial instability hypothesis* (Working Paper No. 74). Annandale-on-Hudson, NY: The Levy Economics Institute.

Morris, S., & Shin, H. (1998). Unique equilibrium in a model of self-fulfilling currency attacks. *American Economic Review, 88*, 587–597.

Osborne, M. F. M. (1959). Brownian motion in the stock market. *Operations Research, 7*(2), 145–173.

Ozdenoren, E., & Yuan, K. (2008). Feedback effects and asset prices. *The Journal of Finance, 63* (4), 1939–1975.

Peters, E. (1989). Fractal structure in the capital markets. *Financial Analysts Journal, 45*(4), 32–37.

Peters, E. (1994). *Fractal market analysis: Applying chaos theory to investment and economics*. New York: Wiley.

Romer, D. (1993). Rational asset-price movements without news. *American Economic Review, 83*, 1112–1130.

Shiller, R. (1984). Stock prices and social dynamics. *Brookings Papers on Economic Activity, 2*, 457–510.

Shiller, R. (2000). *Irrational exuberance*. Princeton, NJ: Princeton University Press.

Shleifer, A. (2000). *Inefficient markets: An introduction to behavioral finance*. New York: Oxford University Press.

Shleifer, A., & Summers, L. H. (1990). The noise trader approach to finance. *The Journal of Economic Perspectives, 4*(2), 19–33.

Shleifer, A., & Vishny, R. (1997). Limits of arbitrage. *The Journal of Finance, 52*(1), 35–55.

Soros, G. (1987). *The alchemy of finance. Reading the mind of the market*. New York: Wiley.

Summers, L. (1986). Does the stock market rationally reflect fundamental values? *The Journal of Finance, 41*(3), 591–601.

Wallace, N. (1988). Another attempt to explain an illiquid banking system: The Diamond and Dybvig model with sequential service taken seriously. *Federal Reserve Bank of Minneapolis Quarterly Review, 12*(Fall), 3–16.

Wallace, N. (1990). A banking model in which partial suspension is best. *Federal Reserve Bank of Minneapolis Quarterly Review, 14*(Fall), 11–23.

Yuan, K. (2005). Asymmetric price movements and borrowing constraints: A rational expectations equilibrium model of crises, contagion, and confusion. *The Journal of Finance, 60*(1), 379–411.

Part II
Technological Change

The Challenges in the Transition from Fossil Fuel to Renewable Energy

Unurjargal Nyambuu and Willi Semmler

Abstract As is currently widely discussed, the implementation of climate stability requires a change in energy technologies. There is active ongoing research undertaken on the future of fossil fuel technology, to produce energy, as well as on alternative technologies, such as renewable energy, to provide energy for the future. Fossil fuel has strong externality effects, extensive CO_2 emissions, which are largely responsible for triggering global warming in the long run. Yet, those externalities are not incorporated into the cost and price of fossil fuel when currently supplied. Renewable energy exhibits no externality when produced and is of infinite supply; yet, it can only be harvested with some costs. Many countries have already undertaken projects to phase out fossil fuel and to phase in renewable energy. Our paper studies the cost and price trends of fossil versus renewable energy and the effects of those industries' performance on stock markets. We sketch a growth model that explains the dynamics of how fossil fuel is phased out and renewables are phased in. Finally, we elaborate on how this can be supported by industrial policies, and discuss what employment effects this transition from fossil fuel to renewable energy may have.

Keywords Renewable energy • Alternative technology • CO_2 emissions • Externality

U. Nyambuu (✉)
Department of Social Science, The New York City College of Technology, The City University of New York, 300 Jay Street, Brooklyn, NY 11201, USA
e-mail: UNyambuu@citytech.cuny.edu

W. Semmler
Department of Economics, The New School for Social Research, 65 Fifth Avenue, New York, NY 10003, USA
e-mail: SemmlerW@newschool.edu

T. Devezas et al. (eds.), *Industry 4.0*, Studies on Entrepreneurship, Structural Change and Industrial Dynamics, DOI 10.1007/978-3-319-49604-7_8

1 Introduction

Non-renewable energy still remains the main source of energy in many countries and is widely used in production, transport and households. However, there is a large amount of literature and evidence highlighting the negative externality effect of fossil fuel resulting in air pollution and environmental degradation. These hidden and indirect costs underscore the importance of using alternative energy and the transition from non-renewable energy technology to renewable energy technology. Recent data demonstrate a rise in renewable energy in power generation indicating a significant transition from fossil fuel to renewables. According to the Bloomberg New Energy Finance, renewable energy accounts for almost 70 % of total new U.S. power capacity with an increase of renewables to around 16,000 MW (from around 13,000 MW in 2014) in 2015. Natural gas showed a decline to around 6000 MW in 2015 from around 8000 MW in 2014. In addition, new nuclear capacity was around 1500 MW in 2015.

In this chapter, we study the cost and price trends of fossil versus renewable energy and the effects of those industries on economic performance. According to the International Panel on Climate Change (2011), electricity generating technologies through renewable energy produces low greenhouse gas emissions compared to the non-renewable technologies. Delucchi and Jacobson (2010b) address the issues of generation and transmission costs of wind, water, sun (WWS) power, and the variability of WWS energy supply. They estimate the cost of supergrid and V2G storage would cost and what the saving of cost of WWS power particularly wind power, would be (also the lower social costs) than fuel. According to Delucchi and Jacobson (2010b, p. 1175), estimated future costs for fossil fuel in the U.S. would rise to \$0.08 with higher social or externality costs to \$0.14, whereas the cost of wind (onshore) power would fall to <\$0.04. The cost of solar PV would drop to \$0.1, and tidal energy would cost between \$0.05 and \$0.07.

We also study fossil versus renewable energy industries with respect to their performance in the stock markets by referring to different indexes including S&P Global Clean Energy Index, S&P Global Oil Index, S&P Global Alternative Energy Index, Global Nuclear Energy Index, WilderHill New Energy Global Innovation Index, and NYSE Bloomberg Global Solar Energy Index that track leading companies engaged in different energy sectors.

We sketch a growth model that explains the dynamics of how fossil fuel is phased out and renewables are phased in. The canonical control model shows how the welfare of the household is affected by a negative externality in terms of CO_2 emissions. Following Greiner et al. (2014), we explain the optimal extraction rate of fossil fuel, its negative externalities, and the capital stock used for renewable energy production. This stock is crucially dependent on various initial conditions and the dynamic paths. An essential result is that it is not optimal to extract all the fossil energy in the ground.

There is a large literature on growth models with natural resources. Early studies that introduced non-renewable resources into a growth model include Dasgupta and

Heal (1974), Solow (1973), and Stiglitz (1974). Heal (1993) modified the growth model by adding a backstop technology. Other variations of growth models incorporate constraints such as global temperature, carbon dioxide emissions, and technical change, moving from a dirty technology to a clean technology. These models are summarized in Nyambuu et al. (2014). Using "endogenous and directed technical change", Acemoglu et al. (2012, p. 131) present the distinction between "clean" and "dirty" (using non-renewable resource) production inputs, and state that "when inputs are sufficiently substitutable, sustainable growth can be achieved with temporary taxes/subsidies that redirect innovation toward clean inputs."

New investment in clean energy has been growing in recent years with large investments in Asia. According to Bloomberg's Clean Energy Investment, new investment surged, to $329 billion in 2015 from $60 billion in 2004, in project investment, equipment manufacturing, and technology development with a majority investment in solar and wind power. In 2015, China's wind power investment has surpassed EU and China became the top investor in wind turbine installations. Developing countries with high growth, e.g., China and India, also have become major polluters in the world. These countries' investment in clean energy technology has shown a significant increase. Popp (2015) discusses climate friendly technological change in developing countries in the form of adaptation and diffusion of technology and how clean technology can be introduced and spread worldwide. He notes that China plays an important role in off-grid technologies as it has the largest number of patent counts on all types of indoor cooking stoves (solar, biomass, LPG, kerosene or butane stoves). Popp (2015, p. 342) stresses the following five lessons:

> First, the fact that high-income countries dominate R&D activities, both for green technologies and more generally, suggests that technology transfer is important. . . . Second, even when technologies are available, adaptive R&D can improve the fit of new technologies to local market conditions. . . . Third, technology transfer will be most likely to promote cleaner growth when it promotes knowledge spillovers. Fourth, financial constraints and ease of use play important roles determining diffusion of green technologies within developing countries. Fifth, policy incentives are important for creating markets for clean technologies, as market forces typically do not reward pollution prevention completely.

As to the solution method of our proposed model, we want to note that the growth model, with shifts into a clean technology, is numerically solved by a finite horizon procedure, using Nonlinear Model Predictive Control (NMPC) which is presented in Grüne et al. (2015). The NMPC method is as an approximation of an infinite-horizon optimal control problem whereby the error exponentially decreases with longer time horizon (Grüne and Pannek 2011).

The chapter is organized as follows. Second section presents a more extensive literature review. In third section, we present data trends in both renewable and non-renewable sectors covering consumption, supply, net generation of electricity, receipts and costs, prices, stock market performance, and carbon dioxide emissions. Fourth section sketches a theoretical model. Fifth section presents the numerical results using NMPC procedure. Finally, sixth section concludes the paper.

2 Literature Review

There exists large body of literature on economic growth where natural resources are used in the production process. In the basic growth models, the state constraints are capital accumulation as well as an exploitable finite resource (see Solow 1973; Stiglitz 1974; Dasgupta and Heal 1974). Furthermore, Heal (1933) assuming unlimited quantities in resources with various grades and costs introduces a back-stop technology into the growth model. Heal (1976, p. 371) states that resource can be "inexhaustible but available at various grades and at various costs" and "extracted from marine sources or crustal rocks" that "cost is assumed to increase with cumulative extraction up to a point, but then to remain constant as a "back-stop" supply is reached."

For the measurement of the scarcity of resources, we often use extraction cost, user cost and prices. Prices of non-renewable resources often rise as they become scarcer. Early papers by Hotelling (1931) and Solow (1974) discussed the increase in prices. The basic Hotelling model was criticized for its oversimplification and for failing to consider extraction cost (see Livernois 2009). Many studies indicate that technological progress and market structure might have an important role for changes in the price. Other studies, such as by Pindyck (1978), Liu and Sutinen (1982), Slade (1982), Livernois and Uhler (1987), and Nyambuu and Semmler (2014) show a U-shaped price trend for non-renewable resources, especially in the long run, depending on the initially proved reserves. In addition, high prices of fossil energy promote its supply leading to more borrowing to finance its production (see Nyambuu and Bernard 2015).

Many studies emphasize the importance of the technical change for the sustainability and economic growth. Weitzman (1997, p. 1) analyzes the relationship between sustainability, Green Net National Product (depreciation of capital and depletion of natural resources at current market prices are deducted from GNP), and technological progress.

As the extension of the growth model, climate change effects can be introduced by building in constraints on temperature and greenhouse gases (GHGs). Social welfare is maximized including an effort to limit environmental degradation. Greiner et al. (2014) introduce a rise in GHGs as compared to pre-industrial level of GHGs. Bondarev et al. (2013) consider a dynamic growth model with environmental variables with technological change and highlight the role of technological efficiency in increasing welfare. They stress the importance of endogenous technical change compared to exogenous technical change, and show how environmental damages or GHGs emissions can be reduced by a lower temperature increase. They use GHG concentration in the atmosphere that is increased by economic activity. Controlled abatements and exogenous clean energy improvements are stressed. The variables used in the model are rate of recovery of the atmosphere because of natural absorption, abatement rate, reduction of intensity of emissions, and a rise in temperature from pre-industrial levels.

A dynamic Integrated Model of Climate and the Economy (DICE) is introduced by Nordhaus (2008). Variables in this growth model include natural capital, climate change impacts, CO_2 emissions, climate damages and policies. DICE-2007 model considers equations of damages, total carbon emission from industry and land, abatement-cost function, mass of carbon in earth's "reservoir," global mean surface temperature, total radiative forcing, and temperature of lower ocean, a variable describing effects and policy of climate change, and abatement cost.

Acemoglu et al. (2012) adopt "endogenous and directed technical change" and make distinction between "clean" and "dirty" technologies. They are used as inputs for production. Acemoglu et al. (2012, p. 131) note that "when inputs are sufficiently substitutable, sustainable growth can be achieved with temporary taxes/ subsidies that redirect innovation toward clean inputs". Their results indicate an importance of innovation in clean inputs that lead to sustainable growth and "use of an exhaustible resource in dirty input production helps the switch to clean innovation under laissez-faire." They use evolution of the environmental quality, with the rate of "environmental degradation" and rate of "environmental regeneration." Empirical results of Acemoglu et al. (2012, p. 159) suggest that "immediate switch of R&D resources to clean technology, followed by a gradual switch of all production to clean inputs" is necessary if two production inputs are substitutes with high elasticity.

Acemoglu et al. (2016) further studies the transition to clean technology using US energy sector's microdata with dirty and clean technologies. Implications for taxes and subsidies on optimal policy are discussed. Acemoglu et al. (2016, pp. 100–101) study "the structure and time path of optimal policies, how rapidly they will be able to secure a transition to clean technology and slow down the potential increases in temperatures, and what the costs of alternative, nonoptimal policies are" and conclude that "qualitative, and even quantitative, patterns of optimal policy, including the heavy reliance on research subsidies, are fairly robust".

Aghion et al. (2016) study the auto industry that is a major GHG emitter covering 80 countries using firm-level panel data on dirty and clean patents. They emphasize the role of innovation for clean technologies when fuel prices are high. Their empirical findings show "path dependence in the type of innovation (clean/ dirty) both from aggregate spillovers and from the firm's own innovation history" (p. 2).

Delucchi and Jacobson (2010a) discuss changes in the infrastructure of energy and assess the global energy sources from wind, water, and sunlight (WWS) energy system focusing on demand for energy, and other issues pertaining to WWS resources, e. g., devices, supply of material, quantities and grid. Their findings indicate that today's cost would be similar to what the fully developed WWS energy system would cost. They show that by powering the world using WWS by 2030 global demand for power would be reduced by 30 %, thus pre-existing fossil energy could be replaced by 2050 whereby all new energy could be supplied by 2030 through the use of WWS. According to the estimation by Delucchi and Jacobson (2010a, p. 1), for 2030 WWS we would need the following energy sources, where the supply of wind power would dominate (see Table 1).

Table 1 2030 WWS infrastructure

Technology for energy	Power	Number
Wind turbines	5 MW	3,800,000
Concentrated solar plants	300 MW	49,000
Solar PV power plants	300 MW	40,000
Rooftop PV systems	3 kW	1.7 billion
Geothermal power plants	100 MW	5350
New hydroelectric power plants	1300 MW	270
Wave devices	0.75 MW	720,000
Tidal turbines	1 MW	490,000

Data Source: Delucchi and Jacobson (2010a)

Furthermore, Delucchi and Jacobson (2010b) study the generation and transmission cost of WWS, the variability of this energy and needed policy measures. These cover interconnection of dispersed resources, hydroelectricity, storage of electric power, generation capacity, production of hydrogen, and projection of energy supplies. Delucchi and Jacobson (2010b, p. 1176) estimate that the usage of a supergrid and V2G storage would cost <$0.02/kWh-generated. They conclude that "future wind power is likely to have a lower private cost than future conventional fossil generation, and all WWS alternatives are likely to have a lower social cost than fossil-fuel generation". Delucchi and Jacobson (2010b, p. 1175) summarize estimated costs for different energy technology (generation and conventional transmission) and show that during the period of 2005–2010 fossil fuel in the U.S. costs $0.07 but its social or externality costs (environmental damage and air pollution) is $0.12, on the other hand, wind costs $0.04–$0.07 for onshore and $0.1–$0.17 for offshore, geothermal costs $0.04–$0.07, hydroelectric costs $0.04, solar PV costs >$0.2, tidal costs >$0.11. Estimated future costs in 2020 are expected to rise for fossil fuel but are predicted to decline for renewable energy.

Toll (2015) studies how climate change affect welfare by estimating the social cost of carbon and the growth rate of marginal damage cost, given the uncertainty of the effect of climate change. He summarizes the welfare effect of different studies predicting (pp. 298–299):

1. A small effect of a "doubling of the atmospheric concentration of GHG emissions on the current economy".
2. A small rise in temperature is likely to have a positive initial benefits. But further rise in temperature may lead to losses.
3. Right-skewed and vast uncertainty.
4. Vulnerability to climate change rises for poorer countries.

Toll (2015, p. 300) discusses several factors that affect the social cost of carbon including the growth rate of per capita consumption, the elasticity of the marginal utility of consumption, and the rate of pure time preference. He shows an uneven distribution and large uncertainty in social cost of climate change. Based on the estimation of the carbon dioxide emissions' marginal damage costs, Toll (2015, p. 308) suggests that "the optimal carbon tax would be about US$34/tC, and growth by 2.1 % per year. A carbon tax of US$34/tC would accelerate energy efficiency improvements and induce a switch from carbon-intensive to carbon-neutral fuel." He also argues that "the optimal course of action also proves to be a middle ground between policy rhetoric and policy action, and between those who favor drastic action and those who prefer no action" (Toll 2015, p. 297).

Based on US automobile industry, Mazzucato and Semmler (1999) explore how the value of stocks of such an industry are affected by industry specific factors when a new industry, such as renewable energy takes off an important issue is how the stock market of the industry performs. They show a strong connection between stock price volatility and market share instability. They conclude that "there are distinct patterns in the evolution of stock price volatility over the life-cycle, and hence that the degree of excess volatility might be influenced by industry specific phenomena. The finding that excess volatility is higher in the first phase of each firm's history suggests that market share instability produces a form of "uncertainty" which makes predictions of future growth rates more difficult and hence market values to be more turbulent than actual performance measures (dividends or earnings)" (p. 94). Furthermore, determinants of the changes in stock prices are studied in Mazzucato and Semmler (2002) pointing out short term volatility in stock prices as well as long swings in the U.S. auto industry. A similar study could be conducted for the recent development of the energy sector. Preliminary work is reported below.

3 Stylized Facts

Clean energy investments across the world have been rising. According to Bloomberg's Clean Energy Investment[1]: total of $425 billion clean energy investment was recorded in 2015 consisting of mergers and acquisitions (M&A) of $97 billion and new investment of $329 billion in project investment, equipment manufacturing, and technology development. In earlier times, new investments in clean energy were only around $60 billion mainly undertaken by Europe, but in 2015 more than half of the global investment came from Asia and Pacific region, only around 20 % came from Europe. China's new investment in 2015 accounted for over $100 billion (excluding R&D and EST asset finance). The majority of the new investments are in solar and wind power. Some of these funds were raised by issuing green bonds, for example, a total of almost $20 billion supranational and sovereign and another $20 billion corporate green bonds were issued in 2015. Other types of green bonds include project based, US municipal, and asset backed securities. The latest large equities raised on the stock market include transactions by Sanan Optoelectronics and by SunPower for wind power at the end of 2015.

[1]Q4 2015 Factpack (2016).

3.1 Consumption and Supply of Energy

Consumption of renewable energy has been rising significantly in recent years. In Fig. 1, U.S. consumption of energy from different sources obtained from the U.S. EIA (2015, 2016) are illustrated. Total consumption of renewable energy increased from 5.3 quadrillion Btu in 2001 to 10 quadrillion Btu in 2015 and is expected to rise further to the estimations by the U.S. EIA. Both wind and solar energy consumption have been rising faster since 2007. Consumption of wind energy is growing faster than the solar energy: wind energy consumption increased from 0.07 quadrillion Btu in 2001 to 1.8 quadrillion Btu in 2015 while consumption of solar energy increased from 0.06 quadrillion Btu in 2001 to 0.5 quadrillion Btu in 2015. In addition, energy from ethanol has also shown an increasing trend in the last decade; rising from 0.1 quadrillion Btu in 2001 to 1.2 quadrillion Btu in 2015.

Next, we look at the sectoral breakdown of the consumption of renewable energy in Fig. 2. The main contributing sector to the fast growth in total renewable energy consumption is electric power sector (doubled to 5.1 quadrillion Btu in 2015 compared to 2001). Other sectors have also shown growth in the usage of renewable energy, especially the industrial, transportation, and residential sectors.

Renewable energy supply in the United States has also grown substantially. In particular, wind power and solar are replacing other types of renewable energy. According the U.S. EIA's report on Short-Term Energy Outlook (2016), supply in total renewable energy has increased; wind's share increased to 20 % in 2015 from 8 % in 2008, solar energy's share increased to 6 % from 1 % during the same period (see Fig. 3).

Newly built wind tribune and PVs for solar power have been increasing in recent years. Especially China has invested in wind power a large sum. Table 2 shows the share of different regions in the wind turbine installations as well as in newly built

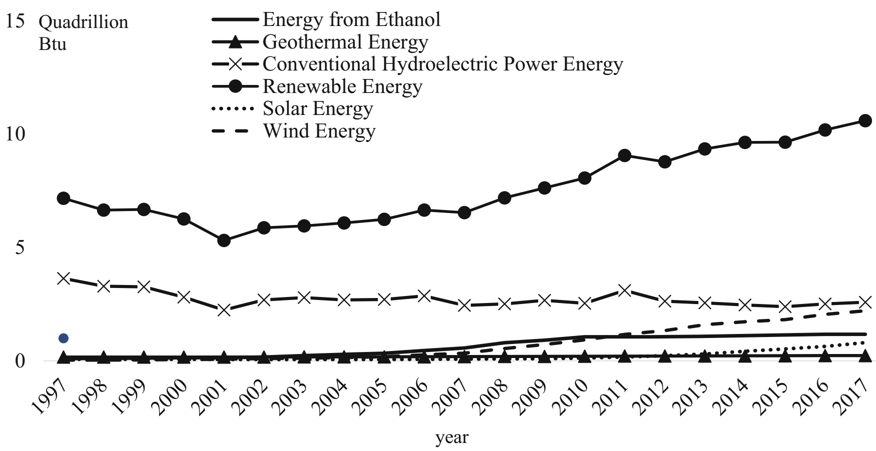


Fig. 1 U.S. Energy Consumption: Data Source: U.S. EIA

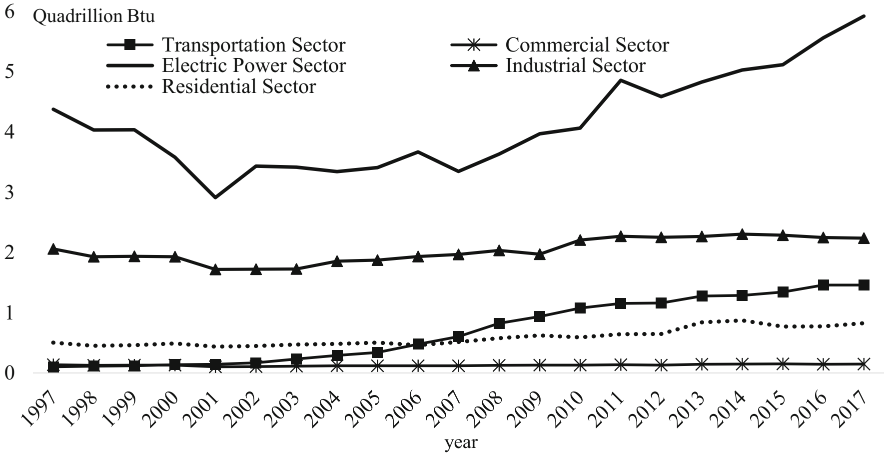


Fig. 2 U.S. total consumption of renewable energy by sectors. Data Source: U.S. EIA

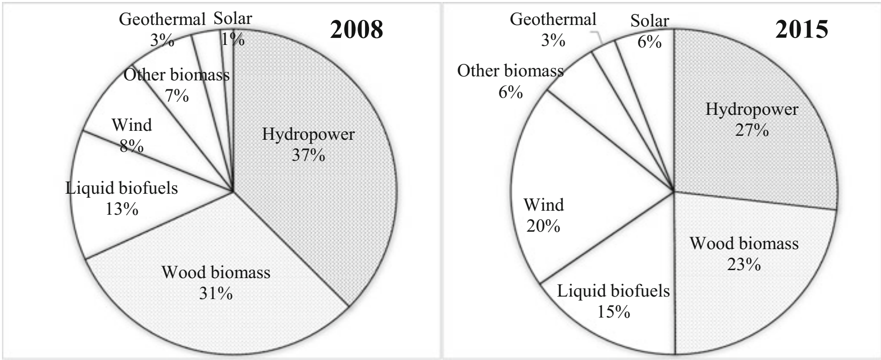


Fig. 3 U.S. renewable energy supply breakdown in 2008 and 2015. Data Source: U.S. EIA

Table 2 Wind turbine installations and newly built solar PVs

	Wind turbine install (GWEC)			Solar, global PV new built	
	2015	2006		2013	2006
Asia	53.3	24	APAC + CHINA	57	22
Europe	22	51	Europe	29	64
North America	17	21	Americas	14	7

Data Source: Bloomberg

solar PVs in 2015 for wind and 2013 for solar compared to 2006. While Europe's and North America's wind tribune installations in global installations have declined, in Asia it has increased significantly: Europe's share dropped to 22 % in 2015 from 51 % in 2006, North America fell to 17 % from 21 %, Asia's share increased to 53 % from 24 % during the same period. China in wind turbine installations surpassed the European Union (EU) in 2015. For solar PV, Europe had the largest share in global new built PV in 2006 but its share has declined to 29 %. In recent years Asia and the Pacific region dominates, accounting for 57 % in 2013.

3.2 Net Generation of Electricity

Electricity is generated by different energy sources, especially with a large generation by coal and natural gas followed by nuclear power. As shown in Fig. 4, net generation of electricity by coal has been declining (from 2 billion to 1.4 billion megawatthours) while it was increasing for natural gas (almost doubled to 1.3 billion megawatthours in 2015) since 2005. Meanwhile, solar generated 26 billion megawatthours of electricity in 2015 (from 0.6 billion megawatthours in 2005).

Next, we show a net generation of electricity from renewable sources in the following Fig. 5. Total renewable generation has shown an increasing trend in the last decade. Electricity is generated mostly from conventional hydroelectric, but its share in total renewable energy has declined to 46 % in 2015 from 76 % in 2005. This decline occurred due to an increase in wind and solar generated electricity; while wind's share increased to 35 % from 5 %, solar PV increased to 4.2 % from 0.005 % during the same period.

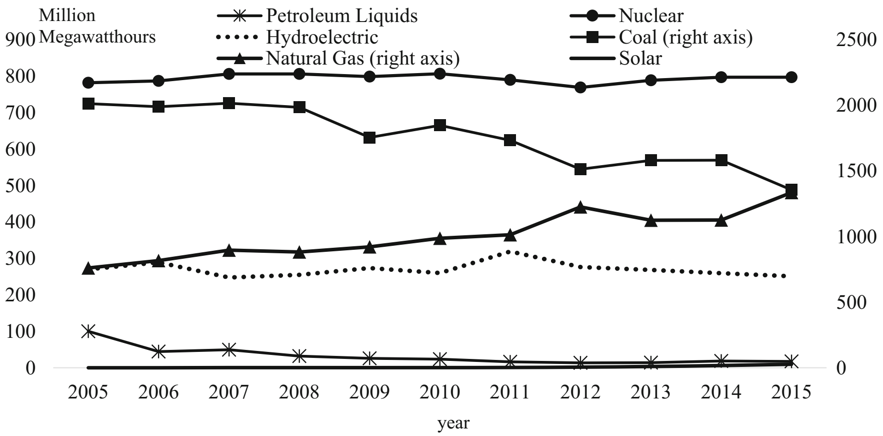


Fig. 4 US net generation of electricity by energy source. Data Source: U.S. EIA

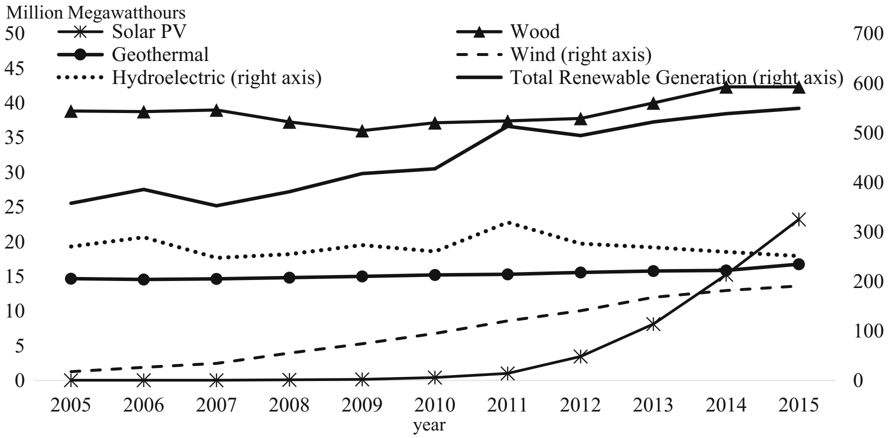


Fig. 5 US net generation of electricity from renewable sources. Data Source: U.S. EIA

3.3 *Receipts and Costs*

Receipts of fossil fuel is illustrated in Fig. 6. While receipts of coal as well as petroleum liquids have been declining for the last decade, receipts from natural gas has been increasing. On the other hand, costs of fossil fuel as shown in Fig. 7, indicate a sharp decline in natural gas and moderate changes for other types of fossil fuel. Costs of petroleum liquids have declined since 2012. Costs of all fossil fuel declined to $2.7 MMBtu in 2015 from $3.3 MMBtu in 2005.

On the other hand, costs of the clean energy have shown a declining trend in recent years. Table 3 shows different costs of wind and solar power. For example, BNEF Polysilicon Price for the solar cell has shown a considerable decline from $22/kg in 2012 to $16/kg in 2015. Similarly, different types of PV costs have shown a decline in the last years. PV insights Module cost declined from $1.2 to $1.1/watt, grid-tie system fell from $2.9 to $2.2/watt, grid-tie inverter from $0.5 to $0.4/watt in 2015. For wind cost, BNEF WTPI as Wind Turbine Index, has decline from Mln 2.3€/MW in 2012 to Mln 1.1€/MW in 2015.

3.4 *Prices of Fossil Fuel*

Prices of fossil fuel is illustrated in Fig. 8. Oil spot prices (Brent) reached a record high of $132.7 per barrel in July 2008, but plummeted to $32.1 per barrel in February 2016. Natural gas spot prices (Henry Hub) reached a high of $13.4 per million Btu in October 2005, but fell to $1.99 per million Btu in February 2016.

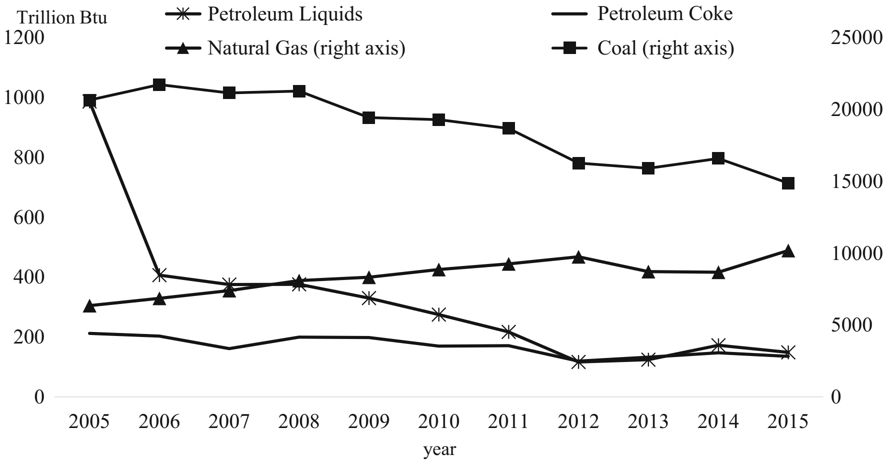


Fig. 6 Receipts of fossil fuels. Data Source: U.S. EIA

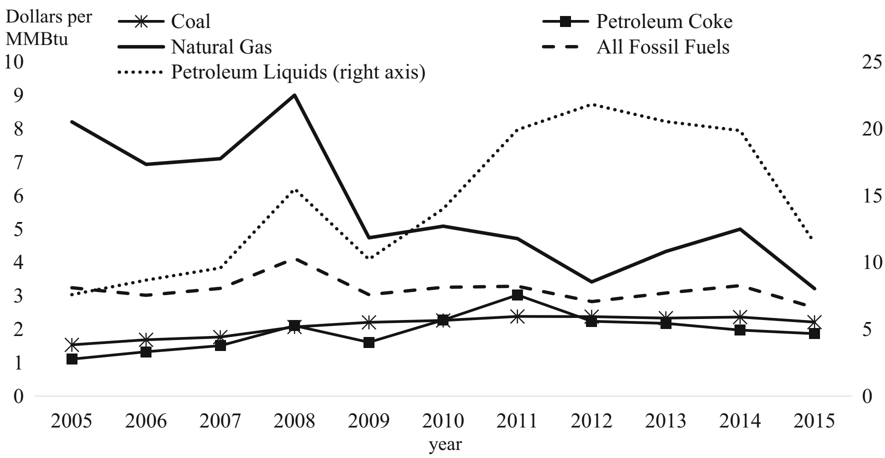


Fig. 7 Costs of fossil fuels. Data Source: U.S. EIA

Table 3 Costs of solar and wind power

	PV insights ($/watt)			BNEF polysilicon ($/kg)	BNEF WTPI (Mln €/MW)
	Module	Grid-tie system	Grid-tie inverter		
2012.12.31	1.2	2.9	0.5	22	2.3
2015.12.31	1.1	2.2	0.4	16	1.1

Data Source: Bloomberg (2016).

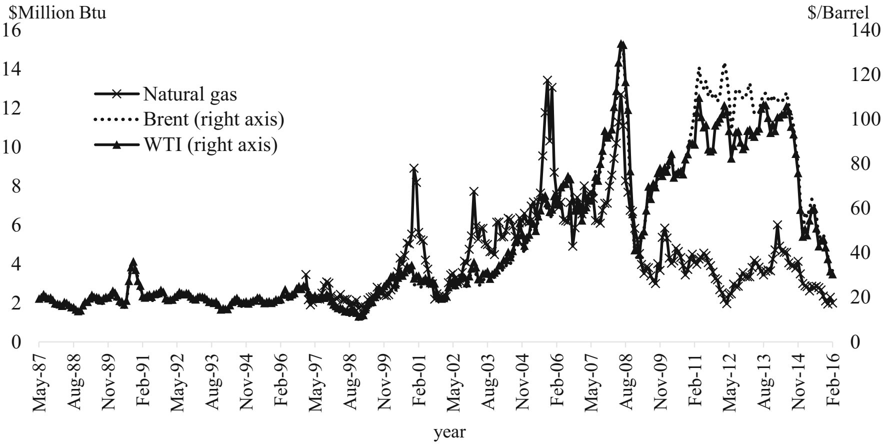


Fig. 8 Fossil fuel prices. Data Source: U.S. EIA

3.5 *Stock Market Performance*

Stock market performance of publicly-traded companies in clean energy and non-renewable resources such as crude oil index is shown in Fig. 9. Prices of these indexes are obtained from the S&P Dow Jones Indices. S&P Global Clean Energy Index tracks 30 companies engaged in clean energy not only in production of energy, but also in its equipment and technology. This index reached the peak in 2008 and declined until the end of 2009; after an increasing trend with moderate fluctuations, it started declining in the mid-2014 exactly when the oil price started dropping. S&P Global Oil Index tracks the performance of 120 of biggest companies operating in oil and gas. This index reached its peak price on 2007 and dropped sharply; it showed a moderately declining trend between 2009 and 2013. The rate of recent decline in the Oil Index has been much slower than the fall in the Clean Energy Index.

Next, we show the performance of an alternative energy index both Global and Asian (including 20 companies) as well as Global Nuclear Energy Index in Fig. 10. The S&P Global Alternative Energy Index is a combination of S&P Global Clean Energy and S&P Global Nuclear Energy. The S&P Global Nuclear Energy Index measures the performance of the leading nuclear energy companies in nuclear production, services, materials, and equipment. Similar to previous indexes, they also peak in the period between 2007 and 2008 and then drop sharply. The recent fall since the mid-2014 is attributable to the cheaper oil.

Other indexes for wind and solar power as well as WilderHill New Energy Global Innovation Index (or NEX) are illustrated in Fig. 11. NEX is comprised of 104 stocks in solar, wind, energy efficiency, biofuels and biomass, power storage and energy conversion. During the period of gains in NEX, especially at the end of 2015, best performing sector was solar (led by Canadian Solar Inc, Universal

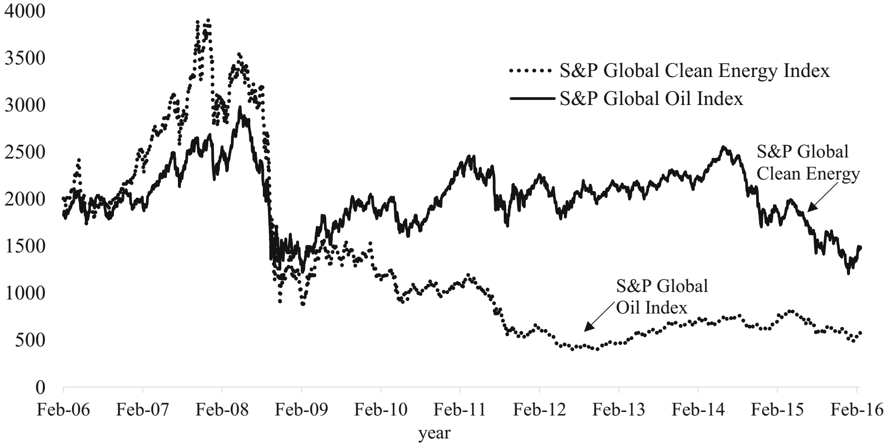


Fig. 9 Stock index performance. Data Source: S&P

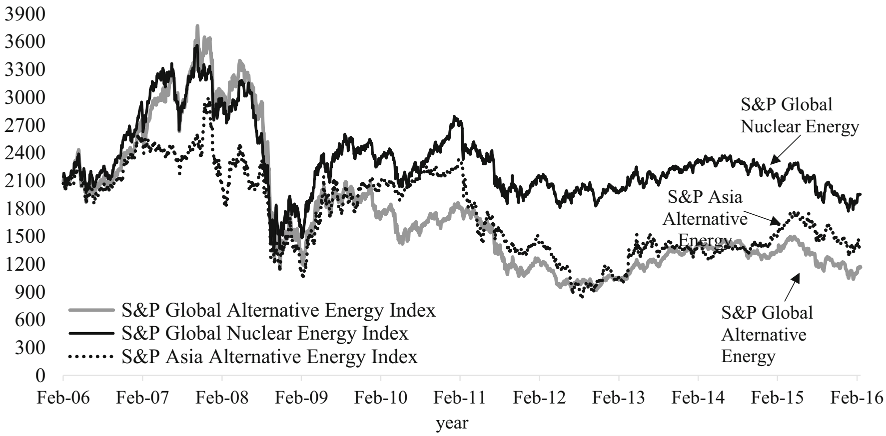


Fig. 10 Stock index performance (alternative energy, nuclear energy). Data Source: S&P

Display Corp, First Solar Inc., SunPower Corp.) followed by energy efficiency. According to The NYSE Bloomberg, Global Solar Energy Index is for solar energy producing and developing companies as well as equipment manufacturers. These indexes demonstrated a similar trends as above analyzed indexes (see Fig. 11). Their performance was affected by the recent oil price collapse.

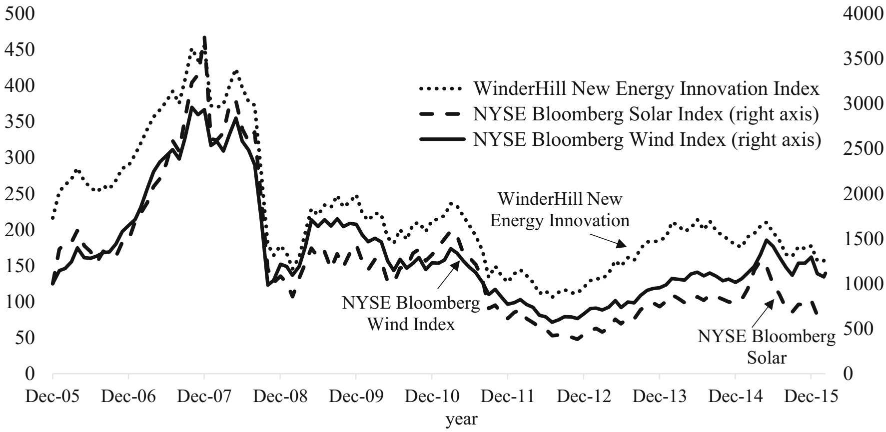


Fig. 11 Stock index performance (wind and solar energy). Data Source: Bloomberg

3.6 Carbon Dioxide Emissions

Carbon dioxide emissions from different sources are shown in Fig. 12. Fossil fuel produces massive emissions, especially petroleum and coal. Overall CO2 emissions from fossil fuel shows a slight reduction in the past 5 years compared to the level of emissions in 2007. CO_2 emissions per person by countries are shown in Fig. 13. While the emissions have shown a decreasing trend for European countries and the United States, it has risen sharply in China especially in the last decade. Because of China's expenditure of energy derived from fossil energy, its high growth rates of GDP are reached with an environmental cost (Bernard and Nyambuu 2015).

3.7 Jobs in Renewable Energy

In recent years, jobs in the renewable energy and energy efficiency sectors including clean transportation and green buildings as well as in other related sectors promoting sustainable development aiming at the restoration of the environment have been growing across the countries. The International Labor Organization (ILO) defines green jobs: "Jobs are green when they help reduce negative environmental impact ultimately leading to environmentally, economically and socially sustainable enterprises and economies."[2]

Estimations in International Renewable Energy Agency (IRENA) annual review on Renewable Energy and Jobs (IRENA 2015) indicate an increase in jobs in renewable energy which reached 7.7 million in 2014, up from 6.5 million in 2013

[2]http://www.ilo.org/global/topics/green-jobs/news/WCMS_220248/lang--en/index.htm

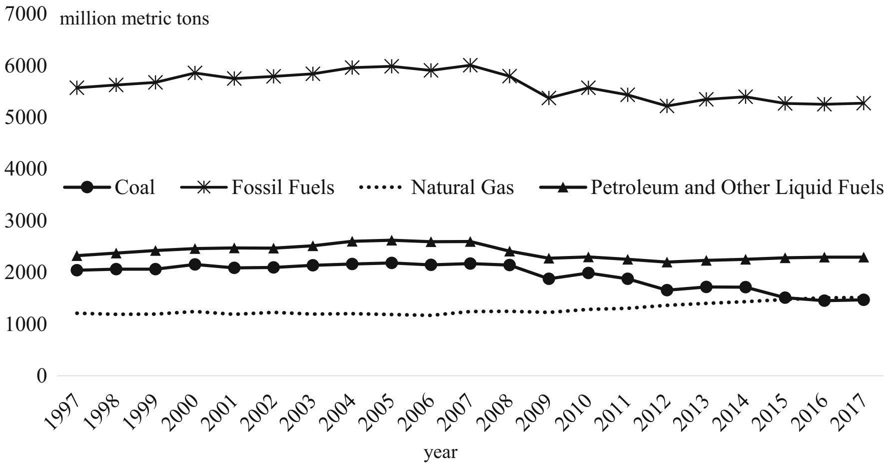


Fig. 12 Carbon dioxide emissions by sources. Data Source: U.S. EIA

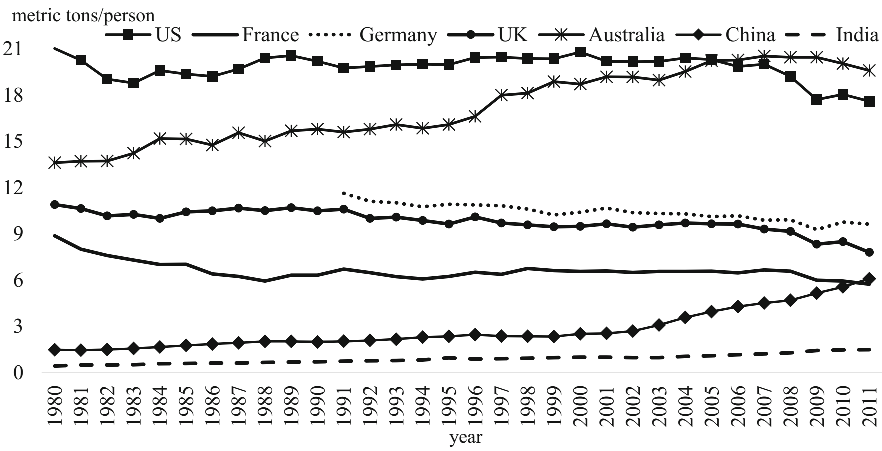


Fig. 13 Carbon dioxide emissions by countries. Data Source: U.S. EIA

(other than large hydropower). These jobs created in 2014 are dominated by solar PV, liquid biofuels, wind power, biomass, solar heating/cooling, and others as shown in Table 4.

China dominates in renewable energy employment in both manufacturing and installations in solar and wind power as well as in hydropower, followed by Brazil, the United States, India and Germany (see Fig. 14). Over two thirds of China's renewable energy job was in solar PV and solar heating/cooling industry. In 2014, almost 45 % of world renewable energy jobs were in China, 12 % in Brazil, 9 % in the United States, 6 % in India, 5 % in Germany, 3 % in Japan, 3 % in Indonesia, and 2 % in France.

Table 4 Jobs created in renewable energy in 2014

Industry	Jobs (million)	Industry	Jobs (million)
Solar PV sector	2.5	Biogas	0.4
Biofuels	1.8	Solar water heating and cooling	0.8
Biomass	0.822	Small hydropower	0.2
Wind	1	Large hydropower	1.5

Data Source: International Renewable Energy Agency (IRENA 2015)

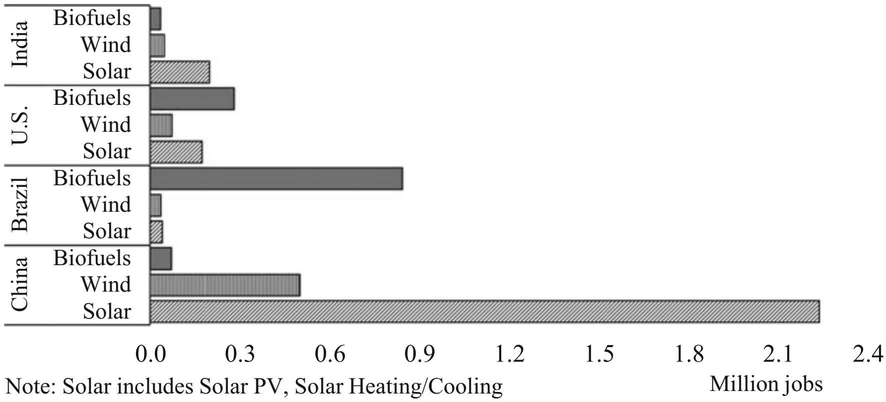


Fig. 14 Renewable energy jobs in 2014 by major employers. Data Source: International Renewable Energy Agency (IRENA 2015)

4 Dynamic Model

Our model follows Greiner et al. (2014) and others such as Byrne (1997), Smulders and Gradus (1996), and Greiner (2011). The production function is defined by two energy sources: non-renewable energy (q_t) and renewable energy. In addition, there is a state equation for the capital stock (K_t). Then, denoting efficiency indexes as E_1 and E_2, output (Y) can be produced using the following equation:

$$Y = E(E_1K + E_2q)^{\sigma},\ 0 < \sigma \leq 1,\ E > 0 \tag{1}$$

The following welfare function with consumption (C) and damages ($-v(D - D^*)^2$) is used for the decision problem of the social decision maker. With decision variables of consumption and non-renewable energy's extraction rate, the households' welfare is constrained by the dynamics of capital stock (K), damages resulting from CO_2 emissions (cumulative) (D), and the evolution of the stock of non-renewable energy (R).

$$\max_{C,q} \int_0^\infty e^{-\theta t}\left(\ln(C) - v(D - D^*)^2\right)dt \tag{2}$$

Subject to

$$\dot{K} = Y - C - \delta K - \phi q \tag{3}$$

$$\dot{D} = \psi q - \rho(D - \xi D^*) \tag{4}$$

$$\dot{R} = -q \tag{5}$$

where ϕ indicates a positive cost of extraction of non-renewable energy, and D^* is the pre-industrial level of damages (GHGs) caused by fossil fuel, ψ is the fraction of GHG not absorbed by the ocean with $0 < \psi < 1$, ρ denotes GHGs' atmospheric lifetime with $0 < \rho < 1$, ξ is a parameter associated with the pre-industrial level of GHG and shows the stabilization of GHGs with $\xi > 1$.[3]

In the next section, this dynamic decision problem will be solved using NPMC with different initial conditions and chosen parameters indicating different initial states of the variables and functions involved.

5 Solutions and Optimal Path

In this section, we present numerical solutions using the NMPC with different initial conditions of three state variables, R_0, K_0, and D_0, indicating different states of economies and environment. In the NMPC setting, we use the parameters: $\xi = 2$, $\theta = 0.03$, $\delta = 0.05$, $\phi = 0.1$, $\sigma = 0.5$, $\rho = 0.1$, $\psi = 0.5$, and $E = 1$, $E_1 = 1$, $E_2 = 1$.[4] Optimal paths of the non-renewable energy (*R*, regardless of its initial condition, demonstrate a declining trend and with eventual depletion in the absence of alternative sources of energy such as green or clean energy (see Nyambuu and Semmler 2014). But in this paper we analyze a specific case and its NMPC results where the optimal path of non-renewable energy are affected by the phasing in of the renewable energy. In contrast to other findings in the literature, our results show that stocks of non-renewable energy do not entirely deplete. There will still be left some fossil energy in the ground. Optimal paths of state variables are shown in Fig. 15 in the case of large initial deposit of non-renewable energy, $R_0 = 5$, $K_0 = 1$, and $D_0 = 1$.

Capital stock used for renewable energy demonstrate an increasing trend with a large rise at the beginning followed by modest growth. As observable a high capital stock is needed during the initial period when the high stock of non-renewable energy generates high negative externalities. The effect of damages is resulting from the changes in the CO_2 emissions (cumulative) that show a significant increase initially due to large extraction of non-renewable energy. But over time, the stock of

[3]For details see Greiner et al. (2014).

[4]Parameter values follow Greiner et al. (2014).

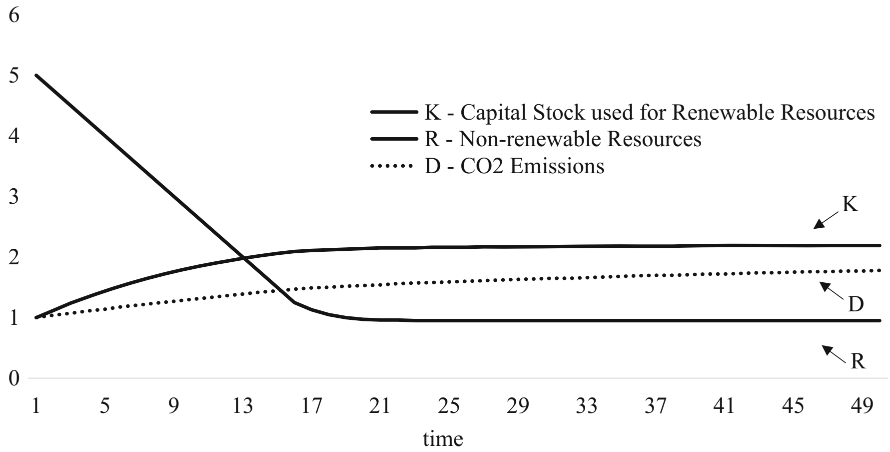


Fig. 15 Optimal path of state variables ($K_0 = 1$, $R_0 = 5$, $D_0 = 1$)

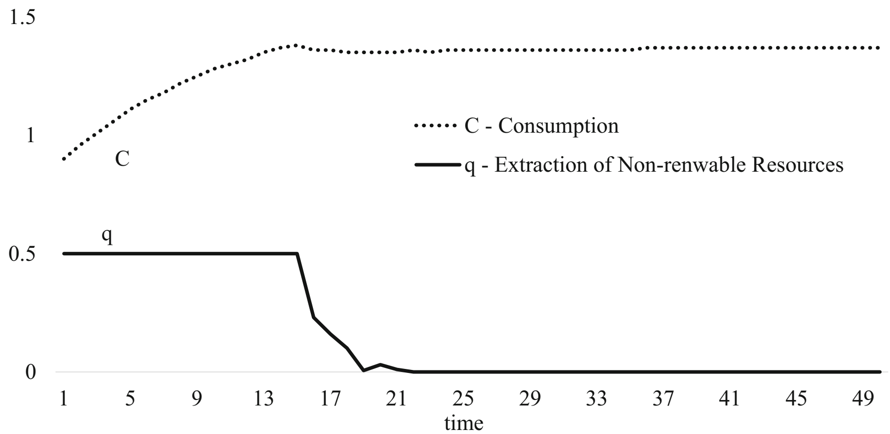


Fig. 16 Optimal path of decision variables ($K_0 = 1$, $R_0 = 5$, $D_0 = 1$)

non-renewable energy declines and reaches a very low level (with some un-extracted stocks left in the ground). As a result, growth rates of capital stock used for the production of renewable energy as well as changes in cumulative emissions slow down significantly.

Corresponding optimal paths of consumption as well as extraction rate of non-renewable energy are illustrated in Fig. 16. The extraction rate of non-renewable energy decline sharply and reached zero at period 21 indicating no more extraction and when the stock of non-renewable energy reached a steady state of 0.95. The optimal path of consumption shows a sharp increase initially when the capital stock, producing renewable energy, was growing relatively fast and when there was still a large deposit of fossil fuel available. But as capital stock growth this is accompanied by a significant decline in the stock of fossil fuel, and also consumption growth rate slows.

Next, we show the numerical results with different initial conditions for the state variables but with the same parameter values as stated above. In Fig. 17, a case of lower initial deposit of non-renewable energy is shown with initial conditions: $R_0 = 1$, $K_0 = 1$, and $D_0 = 1$. Non-renewable energy is extracted at faster rate than in the previous case, capital stock for renewable energy increases at much faster rate and CO_2 emissions also increase and takes a longer time to get stabilized. The corresponding paths for consumption and extraction rates are shown in Fig. 18.

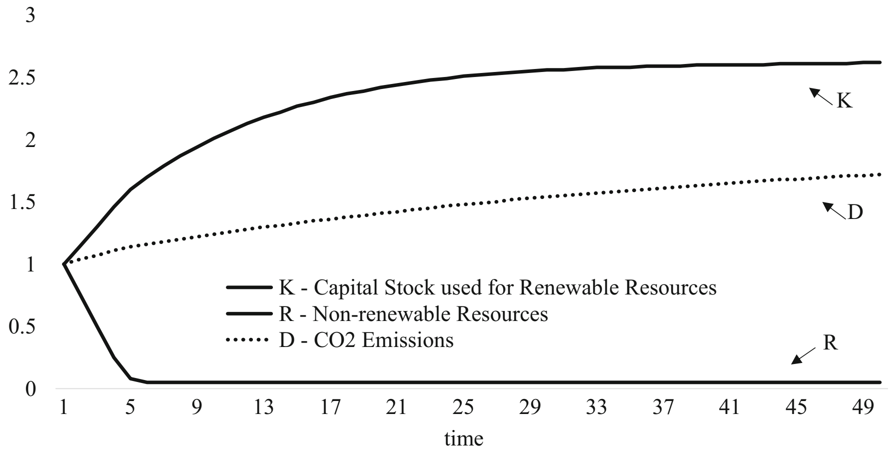


Fig. 17 Optimal path of state variables ($K_0 = 1$, $R_0 = 1$, $D_0 = 1$)

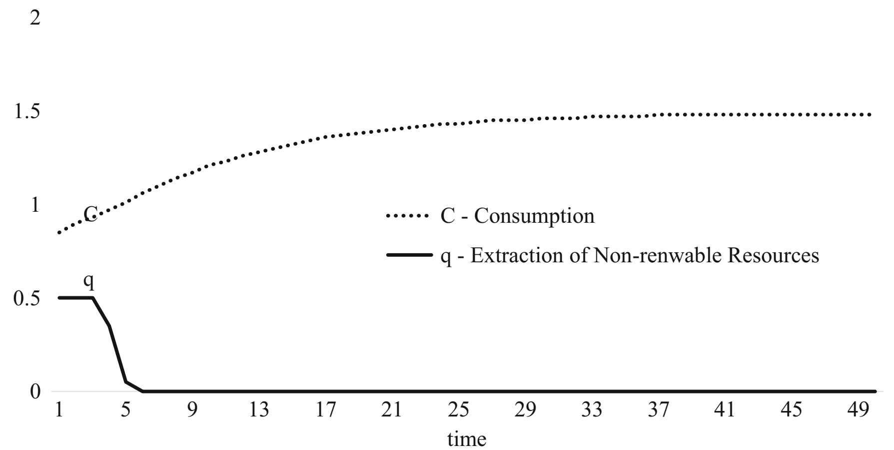


Fig. 18 Optimal path of decision variables ($K_0 = 1$, $R_0 = 1$, $D_0 = 1$)

In the next scenario, we take the same parameter values but different initial conditions of the state variables, especially with higher capital stock $R_0 = 1$, $K_0 = 2.5$, and $D_0 = 1$, is shown in Figs. 19 and 20. Similar to previous case, non-renewable energy is extracted fast, but capital stock for renewable energy increases slightly at first, but then stays roughly constant, CO_2 emissions rise at much faster rate. The corresponding path for consumption shows a declining trend as shown in Fig. 20.

Another scenario with the same parameter values but with different initial conditions of state variables, especially with a very high capital stock, $R_0 = 1$, $K_0 = 6$, and $D_0 = 1$, is shown in Figs. 21 and 22. Non-renewable energy is extracted at

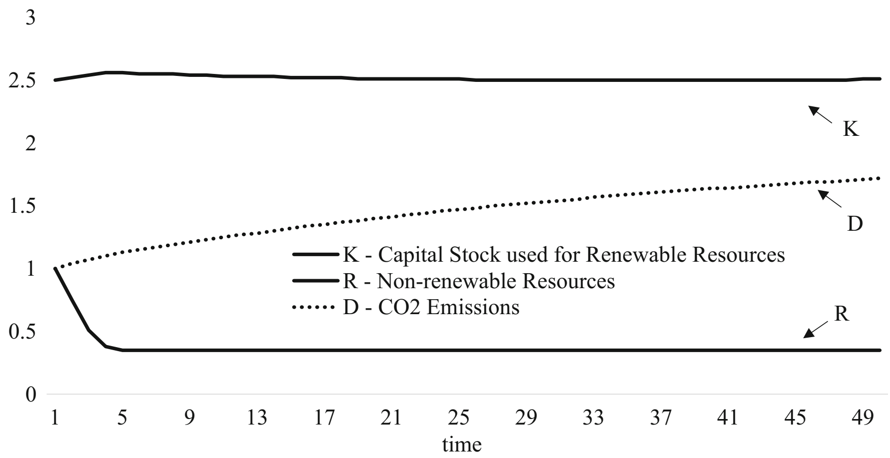


Fig. 19 Optimal path of state variables ($K_0 = 2.5$, $R_0 = 1$, $D_0 = 1$)

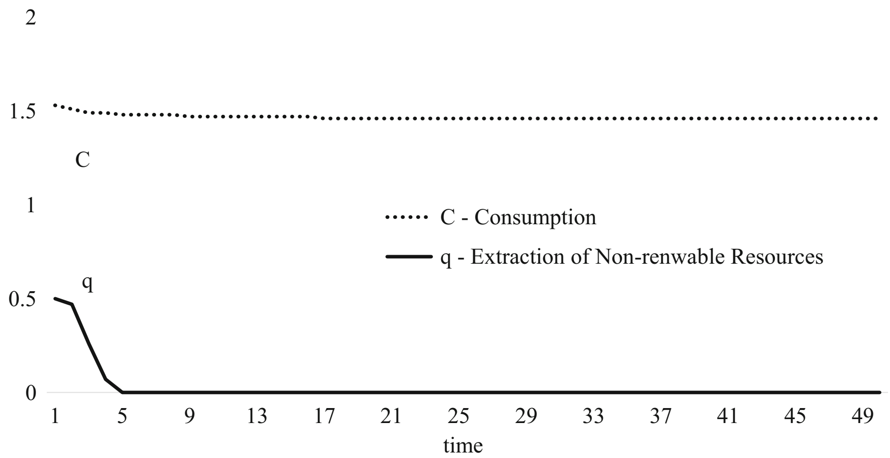


Fig. 20 Optimal path of decision variables ($K_0 = 2.5$, $R_0 = 1$, $D_0 = 1$)

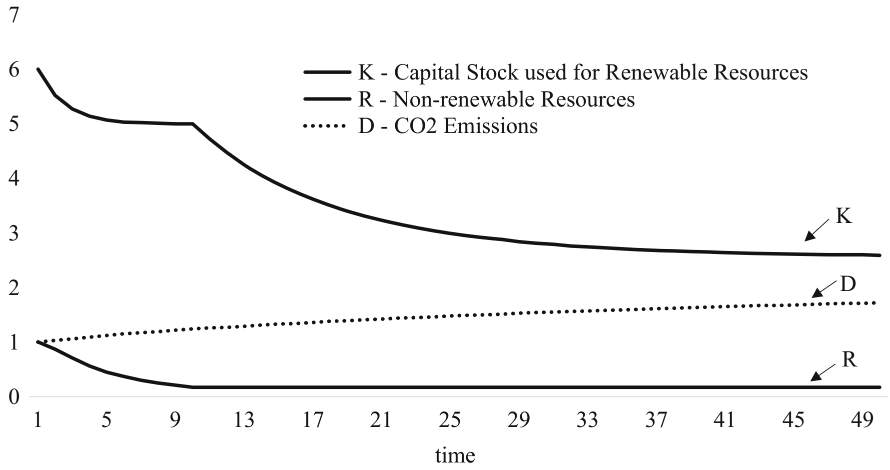


Fig. 21 Optimal path of state variables ($K_0 = 6$, $R_0 = 1$, $D_0 = 1$)

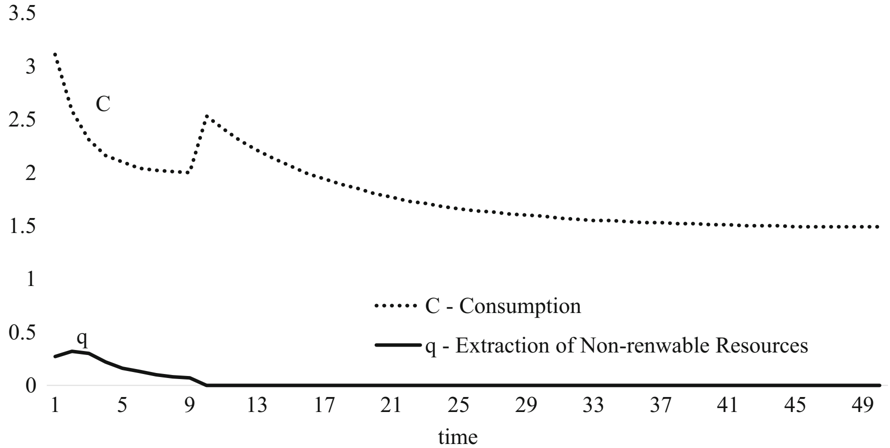


Fig. 22 Optimal path of decision variables ($K_0 = 6$, $R_0 = 1$, $D_0 = 1$)

slower rate than in the previous case, but the capital stock for renewable energy falls at much faster rate, and CO_2 emissions rise monotonically. The extraction rate falls at lower rate but consumption declines at much faster rate as compared to the previous scenario (see Fig. 22).

Overall, we can conclude that with a phasing in of an alternative energy, renewable energy, the fossil energy does not have to be completely extracted. In all cases, some fossil energy is left in the ground and the negative externality and the damages can be limited. For this effect, the initial size of green energy capital is also important.

6 Conclusion

Fossil fuel's negative externality has increased significantly due to large energy use. We showed that CO_2 emissions have shown a declining trend for many European countries and for the United States, but has risen largely in emerging countries especially in China with its fast growth in GDP growths. Estimated future costs in different studies indicate an expected rise in fossil fuel but decline in renewable energy cost. Estimated annualized costs of power generation and conventional transmission summarized in Delucchi and Jacobson (2010b) show that when the social or externality cost of the fossil energy is added, it becomes more costly than alternative forms of energy technology (wind, geothermal, and hydroelectric).

In this chapter, we also elaborated on how this transition can be supported by industrial policies. We study the ways renewable energy can be phased in, and its employment effects. The renewable energy sector is more labor-intensive than the non-renewable industry, requiring also higher skills. Due to growth in renewable technology, employment in the renewable energy sector has risen.

Our numerical solutions of the growth model with externality effects of the fossil energy shows an increasing trend in capital stock used for renewable energy for not very high initial condition. This is because the high stock of fossil energy and its extraction generating high negative externalities. In this case, cumulative CO_2 emissions rise initially due to large extraction of fossil energy. The stock of fossil energy declines eventually to a low level of un-extracted stocks. As a result, capital stock used for the production of renewable energy as well as changes in cumulative emissions grow very slowly.

Our theoretical study and empirical trends suggest, in contrast to what Hotelling had proposed based on his one resource model, renewable energy should be phased in before fossil fuel is completely exhausted. The benefits of this transition would allow for a decrease in emissions and reduction in the externalities from the fossil energy that eventually would reduce CO_2 emission and global warming. This effect improves the environment while benefiting public health, as well as preventing damages to the economic activities in the long run.

References

Acemoglu, D., Aghion, P., Bursztyn, L., & Hemous, D. (2012). The environment and directed technical change. *American Economic Review, 102*(1), 131–166.

Acemoglu, D., Akcigit, U., Hanley, D., & Kerr, W. (2016). Transition to clean technology. *Journal of Political Economy, 124*(1), 52–104.

Aghion, P., Dechezlepretre, A., Hemous, D., Martin, R., & Van Reenen, J. (2016). Carbon taxes, path dependency, and directed technical change: Evidence from the auto industry. *Journal of Political Economy, 124*(1), 1–51.

Bernard, L., & Nyambuu, U. (2015, July 8). Global warming and clean energy in Asia. *OUP Blog*. Oxford University Press's Academic Insights for the Thinking World. Accessed March 12, 2016, from http://blog.oup.com/2015/07/global-warming-clean-energy-asia/

Bloomberg. (2016). Bloomberg New Energy Finance. Accessed May 30, 2016.

Bondarev, A., Clemens, C., & Greiner, A. (2013). *Climate change and technical progress: Impact of on formational constraints* (SSRN Scholarly Paper ID 2207947). Rochester, NY: Social Science Research Network.

Byrne, M. M. (1997). Is growth a dirty word? Pollution, abatement and endogenous growth. *Journal of Development Economics, 54*, 261–284.

Dasgupta, P., & Heal, G. (1974). The optimal depletion of exhaustible resources. *The Review of Economic Studies, 41*(Symposium on the Economics of Exhaustible Resources), 3–28.

Delucchi, A. M., & Jacobson, M. Z. (2010a). Providing all global energy with wind, water, and solar power, Part I: Technologies, energy resources, quantities and areas of infrastructure, and materials. *Energy Policy, 39*(3), 1154–1169.

Delucchi, A. M., & Jacobson, M. Z. (2010b). Providing all global energy with wind water, and solar power, Part II: Reliability, system and transmission costs, and policies. *Energy Policy, 39*, 1170–1190.

Factpack. 2016. Bloomberg's clean energy investment 2015 quarter 4. Bloomberg Terminal.

Greiner, A. (2011). Environmental pollution the public sector and economic growth: A comparison of different scenarios. *Optimal Control Applications & Methods, 32*, 527–544.

Greiner, A., Grüne, L., & Semmler, W. (2014). Economic growth and the transition from non-renewable to renewable energy. *Environment and Development Economics, 19*, 417–439.

Grüne, L., & Pannek, J. (2011). *Nonlinear model predictive control: Theory and algorithms*. Berlin: Springer.

Grüne, L., Semmler, W., & Stieler, M. (2015). Using nonlinear model predictive control for dynamic decision problems in economics. doi:10.2139/ssrn.2242339

Heal, G. (1976). The relationship between price and extraction cost for a resource with a backstop technology. *The Bell Journal of Economics, 7*(2), 371–378.

Heal, G. (1933). The relationship between price and extraction cost for a resource with a backstop technology. In G. Heal (Ed.), *The economics of exhaustible resources*, *International library of critical writings in economics*. Brookfield, VT: Edward Elgar.

Hotelling, H. (1931). The economics of exhaustible resources. *Journal of Political Economy, 39*(2), 137–175.

International Panel on Climate Change. (2011). Renewable energy sources and climate change mitigation: Summary for policymakers and technical summary. https://www.ipcc.ch/pdf/special-reports/srren/SRREN_FD_SPM_final.pdf. Accessed 1 July 2016.

International Renewable Energy Agency. IRENA. (2015). *Renewable energy and jobs: Annual review 2015*. http://www.irena.org/DocumentDownloads/Publications/IRENA_RE_Jobs_Annual_Review_2015.pdf

Liu, P. T., & Sutinen, J. G. (1982). On the behavior of optimal exploration and extraction rates for non-renewable resource stocks. *Resources and Energy, 4*(2), 145–162.

Livernois, J. R. (2009). On the empirical significance of the Hotelling rule. *Review of Environmental Economics and Policy, 3*(1), 22–41.

Livernois, J. R., & Uhler, R. S. (1987). Extraction costs and the economics of nonrenewable resources. *Journal of Political Economy, 95*(1), 195–203.

Mazzucato, M., & Semmler, W. (1999). Market share instability and stock price volatility during the industry life-cycle: The US automobile industry. *Journal of Evolution Economics, 9*, 67–96.

Mazzucato, M., & Semmler, W. (2002). The determinants of stock price volatility: An industry study. *Nonlinear Dynamics, Psychology, and Life Sciences, 6*, 230–253.

Nordhaus, W. D. (2008). *A question of balance: Weighing the options on global warming policies*. New Haven, CT: Yale University Press.

Nyambuu, U., & Bernard, L. (2015). A quantitative approach to assessing sovereign default risk in resource-rich emerging economies. *International Finance and Economics Journal, 20*(3), 220–241.

Nyambuu, U., & Semmler, W. (2014). Trends in the extraction of nonrenewable resources: The case of fossil energy. *Economic Modelling, 37*(C), 271–279.

Nyambuu, U., Semmler, W., & Palokangas, T. (2014, July). *Sustainable growth: Modelling, issues, and policies* (International Institute for Applied Systems Analysis (IIASA) Interim Report, IR-14-019).

Pindyck, R. S. (1978). The optimal exploitation and production of nonrenewable resources. *Journal of Political Economy, 86*(5), 841–861.

Popp, D. (2015). Climate-friendly technological change for developing countries. In L. Bernard & W. Semmler (Eds.), *The Oxford handbook of the macroeconomics of global warming*. New York, NY: Oxford University Press.

Slade, M. E. (1982). Trends in natural-resource commodity prices: An analysis of the time domain. *Journal of Environmental Economics and Management, 9*(2), 122–137.

Smulders, S., & Gradus, R. (1996). Pollution abatement and long-term growth. *European Journal of Political Economy, 12*, 505–532.

Solow, R. M. (1973). Is the end of the world at hand? In A. Weintraub, E. Schwartz, & J. R. Aronson (Eds.), *The economic growth controversy*. New York, NY: International Arts & Sciences Press.

Solow, R. M. (1974). Intergenerational equity and exhaustible resources. *Review of Economic Studies, Symposium on the Economics of Exhaustible Resources*, 29–45.

Stiglitz, J. (1974). Growth with exhaustible natural resources: Efficient and optimal growth paths. *The Review of Economic Studies, 41* (Symposium on the Economics of Exhaustible Resources), 123–137.

Toll, R. S. J. (2015). *The social cost of carbon*. In L. Bernard & W. Semmler (Eds.), *The Oxford handbook of the macroeconomics of global warming*. New York, NY: Oxford University Press.

U.S. Energy Information Administration (U.S. EIA). (2015). Annual energy outlook. http://www.eia.gov/. Accessed 12 May 2016.

U.S. Energy Information Administration (U.S. EIA). (2016). Short-term energy outlook. http://www.eia.gov/forecasts/steo/. Accessed 22 May 2016.

Weitzman, M. L. (1997). Sustainability and technical progress. *The Scandinavian Journal of Economics, 99*(1), 1–13.

Racing to a Renewable Transition?

William R. Thompson and Leila Zakhirova

Abstract Historical patterns of energy transitions over the past 500 years suggest that systemic leadership transitions have become tied to transitions in the primary source of energy. A relatively inexpensive energy source is critical to literally fueling technological and economic growth that outpaces, for a time, the rest of the world. The Dutch had peat, the British had coal, and the United States had petroleum. As the world transitions away from petroleum due to dwindling supplies, and in an effort to save the planet from a climate crisis, what might be the next big source of energy that will power the world economy? The question is particularly compelling if we are in the early days of a political-economic transition with China possibly surpassing the United States at some point in this century. Which of the two leading economies, and the world's biggest carbon emitters, are leading the world in replacing petroleum with renewable alternatives? And is their pace to the new age of renewable energy fast enough to reverse climate change? We argue that both states conversion to non-fossil fuel are critical for the welfare of their own economies and the movement to respond to the threats emanating from climate change. However, neither state, albeit for different reasons, is embracing renewable or non-fossil fuel energy to the extent necessary to feel very confident that an energy transition will occur soon enough to do much good.

Keywords Energy transition • Renewables • System leadership • Energy • CO_2 emissions • Global warming

W.R. Thompson (✉)
Department of Political Science, Indiana University, 1100 East 7th St., Bloomington, IN, 47405, USA
e-mail: wthompso@indiana.edu

L. Zakhirova
Department of Political Science, Concordia College, 901 8th St S., Moorhead, MN, 56562, USA
e-mail: lzakhiro@cord.edu

T. Devezas et al. (eds.), *Industry 4.0*, Studies on Entrepreneurship, Structural Change and Industrial Dynamics, DOI 10.1007/978-3-319-49604-7_9

1 Introduction

> The choice we face is not between saving our environment and saving our economy—it's a choice between prosperity and decline. The nation that leads the world in creating new energy sources will be the nation that leads the twenty-first-century global economy (Barack Obama, April 22, 2009)

The above Obama citation accurately poses a central dilemma for the twenty-first century. Historical patterns in energy transitions over the past several hundred years suggest that systemic leadership transitions have become tied to transitions in the primary source of energy. A relatively inexpensive energy source is critical to literally fueling technological and economic growth that outpaces, for a time, the rest of the world. The Dutch had peat, the British had coal, and the United States had coal, electricity, and petroleum. As the world transitions away from coal and petroleum due to dwindling supplies of petroleum and their harmful side-effects, and in an effort to save the planet from a climate crisis, what might be the next big source of energy that will power the world economy and its lead economy(ies)? The question is particularly compelling if we are in the early days of a political-economic transition with China possibly surpassing the United States at some point in this century. Which of the two leading economies, and the world's biggest carbon emitters, are also leading the world in replacing fossil fuels with alternative energy sources? And is their pace to the next energy regime fast enough to address adequately the threat of fundamental climate change?

There is a complicated and disheartening paradox concerning the futures of systemic leadership and renewable energy. Systemic leadership has increasingly become reliant on an energy transition to make new technology relatively inexpensive to innovate and to move far ahead of the competition for a finite period of time. The incumbent systemic leadership appears to be eroding. Many observers nominate China as the likely successor. Given the finite supply of fossil fuels and the high costs of consuming these fuel sources (for the consumer and the environment), renewable energy is pitched as the best alternative and the likely focus of a new phase of dominant energy source. Yet if renewable energy is the next phase, most governmental and corporate forecasts do not foresee renewable energy becoming predominant in at least the next 50 years, if not much longer. Indeed, conservative forecasts target the turn of the next century as the transitional point for a renewable energy regime. If the forecasts prove to be accurate (and there are reasons to expect that they will not), there is less reason to anticipate a change in systemic leadership before that transition. If no economy can be expected to take the lead in (a) developing radical new technology that is (b) fueled by a new and relatively inexpensive source of energy, there is much less reason to expect the emergence of a powerful, new lead economy as has occurred repeatedly in the past millennia. Moreover, there is probably also less chance of an effective and timely global response to the climate crisis in the absence of systemic leadership. Such a response cannot be carried out completely from a top-down position but without leadership, bottom up participation in the major changes necessary are unlikely to carry the day.

If the forecasts prove to be inaccurate—perhaps because it will prove to be increasingly difficult to rely on fossil fuels through 2050 for different reasons—we may anticipate major energy shortfalls, global crisis, and no new systemic leadership.[1] Even if the shortfalls do not materialize, accelerated environmental deterioration in response to rising hydrocarbon consumption seems highly likely. Thus we still end up with a shortage of renewables, one form of global crisis or another and no new systemic leadership. Such a situation might facilitate a greater role for leadership by international institutions but one can hardly guarantee such an outcome. Indeed, the track record of international institutions suggests that they are unlikely to serve as a vanguard in the first half of the twenty-first century.

In this essay, we briefly scan the evidence for the fusion of energy transitions and technological leadership and the erosion of the incumbent regime. Turning to the question of systemic leadership transition, we examine the relative positioning of the United States and China in terms of renewables development. While it is clear which state is in the lead at the present, we then ask whether it matters all that much if the two states are competing over a small proportion of the total energy inventory. This question leads us to consider various forecasts for the relative consumption of different types of energy. If the technological leadership-energy transition argument and the forecasts are correct, there is little reason to anticipate a systemic leadership transition of any strength in this century. If the forecasts are wrong and overestimate the fossil fuels remaining to be extracted economically, we are likely to be confronted with major energy problems by roughly mid-century. Either way, we may be experiencing major energy problems with even more major environmental problems in about the same time frame.

2 Systemic Leadership Concerns

System leaders have emerged largely on the basis of possessing the world's lead economy for a finite period of time. Leads do not last forever. They are constructed gradually and often erode fairly gradually as well. The rate of erosion, however, may vary and the slower the rate of erosion, the more difficult it is discern relative decline. This tendency helps explain the ongoing disagreements about the current position of the United States. Not only was its concentration of leads more strongly established than those of its predecessors, its relative decline has also been slower. Nonetheless, other states and their economies have been catching up to the U.S. position. Figure 1, for instance, shows the fluctuations in the U.S. share of world knowledge and technologically intensive industries as a marker of

[1]Coal is relatively plentiful but deadly in its CO_2 emissions. Petroleum has its own emissions problems but is also expected to become less available in the future.

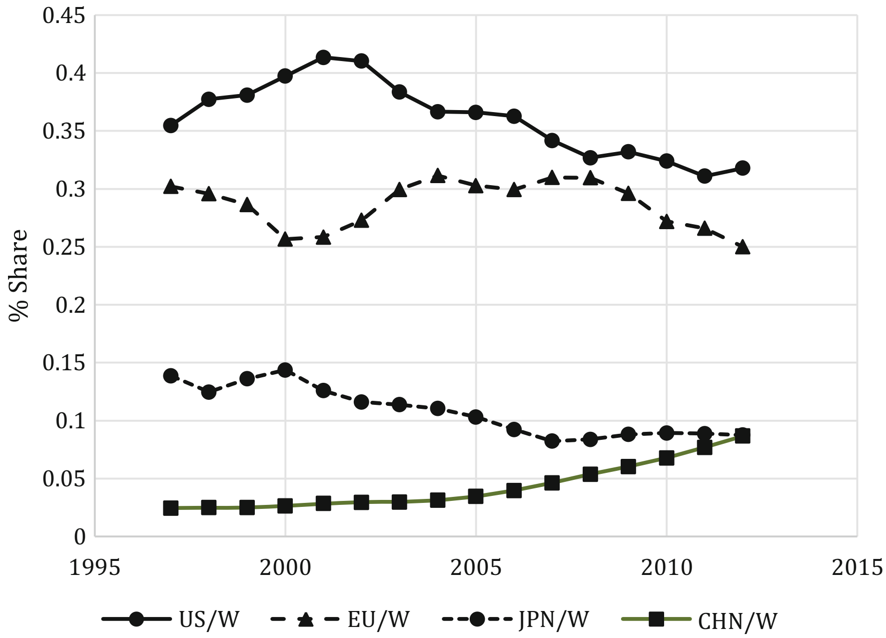


Fig. 1 Shares of world knowledge and technological intensive industries

technological leadership. The United States has maintained its lead but other states have narrowed the gap. European economies, in general, are closest, followed by Japan, both of which Fig. 1 shows falling back in the competition. China is moving up but still has some way to go. Thus, based on this indicator, the U.S. technological lead is slowly declining relative to what others have been able to achieve in narrowing the gap.[2]

A second dimension in systemic leadership is the ability to project force throughout the system in order to protect its predominance in the world economy. Global reach capabilities have varied over time but prior to the twentieth century, they were exclusively naval in nature. More technological change in the twentieth century permitted forays into the air and outer space. Consequently, any index of global reach must adjust for technological change. The power projection scores shown in Fig. 2 combine distributional information on blue water naval capability (ships-of-the-line, battleships, heavy aircraft carriers, nuclear submarines and their ballistic missiles) augmented by strategic bombers, land-based ballistic missiles, and military satellites).

[2]The relative decline data in Fig. 1 would be much more apparent if the knowledge/technologically intensive information also encompassed the second half of the twentieth century.

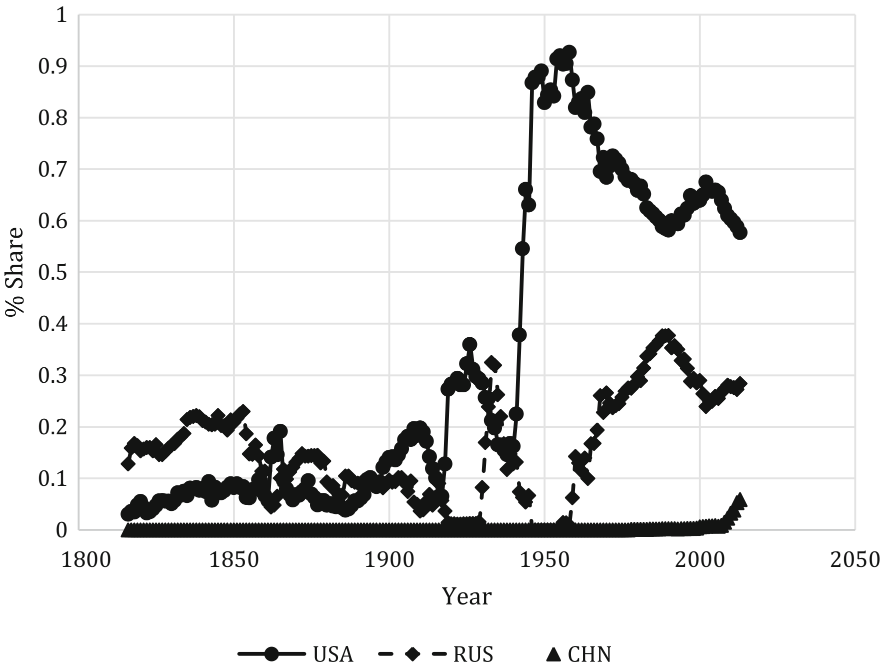


Fig. 2 Global power projection

While not demonstrated in Fig. 2, the U.S. lead in global reach capabilities was and still is much more impressive than its British predecessor's position. At the same time, though, the U.S. lead has hardly been static. Its strategic preeminence emerged abruptly in and immediately after World War II, was maintained into the 1960s, and has been declining slowly relative to attempts to narrow the military projection of force gap since then. The Soviet Union made some headway in closing the gap but stopped well short of catching up. Its successor, the Russian Federation, is trying once again, and so is China—although it has quite a ways to go to become competitive in this sphere.

The two figures demonstrate something similar: a still predominant United States that is gradually losing its relative technological and power projection leads. Given the impressive heights of concentrated power to which the United States rose, it may not be surprising that its decline trajectory is slow. At the same time, relative decline in military capability is even slower than in technology. One of the distinctive properties of our era is that the economies that have made the most headway in narrowing the technological gap have much less interest in narrowing the military capability gap. Alternatively, the economies that have shown the most interest in narrowing the military capability gap have struggled to narrow the technology gap.[3] Yet neither gap-narrowing process has stopped still and

[3]From a conflict perspective, both characteristics are beneficial for world peace.

presumably further inroads to the U.S. leads can be anticipated in the years to come. Continuing U.S. relative decline, albeit slowly, sets up an expectation of eventual succession as system leader.

The natural question is who is the most likely candidate and many observers have put China forward as the state with the best chance to become the next system leader. Much of this furor, however, is predicated on China's economy becoming the largest one in the world. Such an emphasis is misplaced and misunderstands the nature of the processes at risk. Possessing a large economy can be advantageous just as having the world economy's largest economy can be an asset, especially when it comes to creating markets for the consumption of commodities. More important, though, is the technological lead and that is not something that is conferred by possessing the world's largest economy. Moreover, the pattern is fairly clear that it is not just the technological lead that is critical. Increasingly in the past few centuries, the radical new technologies produced by possessing the technological lead require relatively inexpensive fuels that are capable of converting organic and non-organic materials into powerful sources of energy. Both the low price and the power of the energy have become critical to successful production and finding markets for the productions' consumption.

Therefore, it is not enough to ask who will have the largest economy or even the most technologically advanced economy. It is also required that we ask whether the most technologically advanced economy will be founded on a new, relatively inexpensive energy base. One thing is abundantly clear. Coal and petroleum, which have fueled the past two systemic leads, will not suffice. The petroleum supply is disappearing and, consequently, its price is rising generally.[4] Far more important, however, the costs of remaining dependent on fossil fuels appears to be far too great given the implications for escalating global warming. Thus, natural gas, often touted as a fossil fuel with less CO_2 emissions than coal or petroleum is a dubious replacement. The next system leader—should there be one—cannot depend on fossil fuels to provide cheap and reliable energy without great environmental costs. Inasmuch as dependence on coal and petroleum have led to unsustainable environmental costs already, the incentive to move to a different, non-carbon-based, source of energy seems all that more paramount.[5] Our questions then are twofold: (1) which of the likely candidates for technological pre-eminence

[4]At the time of this writing (early 2016), oil prices have fallen impressively due to declines in global demand, shale production, and Gulf decisions to maintain high production in a battle for market share. Presumably, this exception to the rising price of petroleum will prove to be short-lived.

[5]Another possibility is nuclear power but its history of accidents works against widespread development. As a consequence, economic forecasts rarely pay much attention to its potential and tend to predict that its future prominence will be no greater than its current status. Even if new technology that permits safer functioning reactors could overcome the reluctance in some parts of the world to expand this energy source, the production of uranium from mines is insufficient to fuel existing nuclear reactors. For example, in the period between 1995 and-2005, the gap between supply and demand for uranium was almost 50 %. While some of the gap is currently being filled by recovering uranium from military stocks and old nuclear warheads, these stocks are a finite

in the future are moving toward developing a non-carbon-based economy? and (2) is any candidate moving fast enough to make a difference?

2.1 Who Is on First Base?

As the world depletes the supply and desirability of fossil fuels what might be the next big source of energy that will power the world economy? More importantly, between the declining leader (the United States) and the aspiring heir apparent (China), which of the two (also the world's biggest carbon emitters) are leading the race in replacing fossil fuels with renewable alternatives? And is their pace to the new age of renewable energy fast enough to reverse or reduce the risks of climate change? To succeed the United States as the world's next leading global power, China's prospects would be greatly enhanced if it led in developing a new energy regime. To what extent is China doing that now given its current energy needs? This question is particularly compelling if we are in the early days of a political-economic transition with China expected to surpass the United States at some point in this century. Similarly, if the United States wishes to maintain and renew its lead, it will need to lead in the development of a new energy regime. To what extent is the United States doing that now given its own current energy and economic re-vitalization needs?

However, there is another possibility. Some observers think that natural gas will lay the energy foundation for the next phase of advanced economic development. Increasing supplies of gas are already negatively impacting investments in renewables. The value of increased natural gas is in part environmental—as gas is substituted for coal, the negative consequences for climate processes is less than it might otherwise be. If gas is set to become a half way medium towards genuinely greener energy sources, the next system leader will profit from ready access, preferably close to home. In this sense, the shale gas boom could contribute to the United States at least holding on to its lead status somewhat longer than it would have otherwise. If the natural gas windfall should become instrumental in pioneering the next phase of industrial revolution, it could also lead to the United States succeeding itself as system leader, much as Britain did in the nineteenth century. Whether it is a strong possibility, though, is debatable.[6] Hughes (2011) argues that gas wells are not very productive and a great deal of expensive investment will be required to maintain a large supply of gas. If the supply cannot be maintained, its price will increase. The price of natural gas will also rise if

resource and cannot be seen as a definitive solution to the problem of insufficient supply. For a greater discussion on the uranium supply-demand gap, see Bardi (2014).

[6]Helm (2012: 199–216) recommends a quick but temporary transition to natural gas in order to replace coal but he also realizes that this strategy can only work for a limited time. Smil (2015) anticipates that natural gas will become more important in coming decades but balks at agreeing with claims that natural gas will enjoy a golden age in the mid-twenty-first century.

substantial amounts are exported abroad. Moreover, moving to natural gas does not resolve the global warming problem. While overall CO_2 emissions might be reduced, no one has advanced an argument that they would be reduced sufficiently to cope with the widely perceived need to avoid warming $>2\ ^\circ C$. Natural gas usage also is accompanied by increased methane emissions that are considered even more damaging than CO_2 emissions. If the global warming persists or becomes even worse, what might it mean to be the world's lead economy in a global setting overwhelmed with environmental deterioration issues?

An expanded supply of natural gas, moreover, could divert the U.S. economy (or any other economy, including China's, for that matter) from its search for alternatives to fossil fuels.[7] If more gas is consumed in the coming decades because it is more readily available and less expensive, support for renewable development could languish in the absence of a situation approaching acute crisis. If the gas diversion discouraged the U.S. search for alternatives, any lead in developing sunrise energy resources could well reside elsewhere.

Consequently, we should not be surprised by increasing consumption of natural gas in the immediate future but it does not appear to be a long term winning strategy, particularly for aspiring system leaders. The next most overt candidate is renewables which usually translate primarily into hydro power, wind and solar technology. Dams, wind turbines, and solar panels can be set up widely, generate energy that is free of CO_2 emissions, and can be used to replace coal and gas for the generation of electricity. The technologies, of course, are not without major problems. Only so many new dams are likely to be built especially in an era facing water scarcity. Wind and sun as sources are intermittent and need to be stored and transmitted to consumers, thereby raising questions about their ultimate power potential. Yet renewables enthusiasts are convinced that this path is the best way to freeing the planet of its dependency on fossil fuels. Technically, various plans have been developed that suggest that it would be feasible to shift completely to renewables by 2050–2070. Unfortunately, technical feasibility is not the major obstacle to be overcome. The cultivation of renewables will require considerable investment and governmental subsidy to facilitate their adoption in a transitional phase. Considerable resistance to new energy sources must also be overcome. But these statements tend to be applicable to all new energy sources that are usually more expensive initially than are conventional supplies and hardly subject to unanimous endorsement as the next best thing.

[7] Smil (2015: 141) summarizes the consensus on Chinese shale prospects as characterized by a lack of technical experience, scattered and deep formations with many faults, poorly mapped, competing with scarce water supplies, and various types of bureaucratic resistance. None of this precludes Chinese shale exploitation; all of it makes it more likely to be developed slowly and less productive than sometimes forecasted.

2.2 Who Is Winning the Renewable Energy Race?

Given the expense and resistance associated with new energy sources, governments need to be more than mildly facilitative in encouraging adoption. Political commitment and strong endorsement are necessary. The natural questions, then, are which governments are inclined to support renewable development and, which economies possess greater renewable capabilities? Not surprisingly, small, affluent states located in western Europe tend to lead in this race. But these states will not be contenders for systemic leadership. Confining the question to the most obvious candidates leads to asking: is China or the United States in the lead?

For most energy analysts, it is something less than a mystery who is winning this race. Relative to their economies, China is investing twice as much in renewable energy as the United States and its support for the renewable industry is rapidly expanding (see Table 1). But China has not always been in the lead. In fact, the United States under the Carter administration was among the first to set a target for renewable energy. At a 1979 press conference, President Carter promised to the American people that the U.S. would obtain 20 % if its energy from solar power by 2000. The idea that the world needed to transition to what was then called solar energy (an umbrella term for most renewables) and that the U.S. needed to lead that race was largely a response to the oil crisis of the 1970s, which revealed the dangerous overreliance on Middle East oil by the world's advanced industrial economies. The desire to avoid the disastrous consequences of another cutoff of supply paved the way for the current boom in renewable energy. However, nothing like Carter's 20 % target actually happened. What followed instead were decades of stagnation. The extremely low oil prices in the 1980s and 1990s combined with the ongoing shale revolution of the early 2000s discouraged efforts to develop renewable energies. Consequently, even as late as 2015, renewables accounted for only about 9 % of the U.S. energy supply—not much different from the share it had in 1980. If we discount hydropower (which has been constant for many years) and biomass, renewables in 2015 constituted <3 % of the total U.S. energy supply (U.S. EIA 2015).

While the U.S. solar industry became what the *Economist* at the time described as "a commercial graveyard for ecologically minded dreamers," leadership in renewable energy was being established elsewhere in the world, first by Japan, then by Germany (Yergin 2011). With strong government support, both economies were motivated to revive the renewables industry for different reasons. Japan's economic growth in the 1960s and 1970s had been largely fueled by Middle Eastern oil making Japan extremely vulnerable to any disruptions in the supply of oil

Table 1 US–China investment in clean energy, 2012–2015 (in US$ billions)

Rank	Country	2012	2013	2014	2015
1	China	54.2	57.9	89.5	110.5
2	United States	36.7	40.3	51.8	56.0

Source: Bloomberg New Energy Finance (2013) imports. Motivated by the need to reduce Japan's dependence on Middle Eastern oil, the country launched an all-out national initiative in the 1970s to develop solar energy. Fueled by large government subsidies, the solar market took off in Japan.

Half way around the world, the leadership in renewables was passing to another economy. As a result of the opening of the Berlin Wall and in an effort to integrate East Germany's power system into West Germany's electric utilities grid, Germany became a global leader in renewable energy in the 1990s. Despite the oversupply of oil, which flooded the world market in the mid-1980s forcing oil prices to plummet from a high of $34/barrel to $10/barrel in 1986, Germany was more concerned with the devastating consequences of another event that took place the same year. The nuclear power plant accident at Chernobyl served as a mobilizing force for Germans who did not want to depend on nuclear energy, which up until then had been touted as the energy source of the future. By the early 1990s, wind turbines were proliferating across Germany. By the turn of the century, the government adopted a program to phase out all nuclear power, which at the time was providing over a quarter of Germany's total electricity supply, providing a further impetus for the development of renewables. In addition to wind, Germany became the world's biggest market for solar photovoltaics (PVs).[8] In 2009, renewables' share of Germany's electricity consumption reached 14 %, exceeding its 2010 goal. By 2012, Germany's solar PV capacity accounted for nearly one-third (32 %) of the global total, whereas the U.S. and China's share were very similar at around 7 % each (REN21 2013: 40–41). However, 2 years later, Germany's early commanding lead had slipped to about 21 % while China's share had improved to almost 16 % and the U.S. share had increased more moderately to a little over 10 % of global capacity (REN21 2015: 58–60).

China entered the renewables race in a significant way relatively late in the game. Even though it had laws in place that called for solar and wind energy as early as the early 1970s, renewables were largely considered "antipoverty measures for the benefit of the rural poor" (Yergin 2011: 543–544). But by the turn of the century, renewables had become a global business and China wanted to play a major role in shaping it. By the mid-2000s, rapid economic growth led to equally rapid growth in energy consumption forcing China to import energy in order to respond to its domestic energy crisis, with electricity demand outrunning the availability of coal and electricity. The decisive change came with the Renewable Energy Law of 2005, followed by the Medium-and Long-Term Development Plan for Renewable Energy of 2007, which not only jump-started renewables in China but also set out specific targets for renewables to reach 15 % of total energy by 2020 (Yergin 2011: 544). As a result of massive government stimulus spending during the 2008 global financial crisis, China's renewable energy moved into high gear. Within a decade, Chinese companies became the dominant market for solar panel

[8]The word "photovoltaics" comes from the words "photo" and "volta." It refers to the direct conversion of sunlight into electricity.

Table 2 Cumulative installed solar PV capacity in top 5 countries, 2000–2014 (megawatts)

Year	Germany	China	Italy	Japan	USA
2000	76	19	19	330	0
2001	186	30	20	453	0
2003	296	45	22	637	28
2004	435	55	26	860	73
2005	1105	64	31	1132	131
2006	2056	68	38	1422	172
2007	2899	80	50	1709	275
2008	4170	100	120	1919	427
2009	6120	140	458	2144	738
2010	10,566	300	1181	2627	1172
2011	17,554	800	3502	3618	2022
2012	25,039	3300	12,803	4914	3910
2013	32,643	7000	16,139	6743	7271
2014	38,200	28,199	18,460	23,300	18,280

Source: Figures are as published in BP, *Statistical Review of World Energy* (2014, 2015), modified from Roney (2014)

manufacturing and second only to Germany in terms of cumulative installed solar PV capacity (see Table 2).

To put things in perspective, in 1980, 85 % of all solar modules were being produced in the U.S. By 2005, this share had shrunk to <10 %. Japan, China and Germany replaced America as the leaders in photovoltaics (Quaschning 2010: 113). China has also made significant inroads into the wind-turbine industry, surpassing the U.S. and Germany as the global leader in cumulative wind power installed capacity in 2010 and maintaining its lead ever since (see Table 4). Concerned with energy security and obsessed with the need to control all aspects of producing energy, China put in place policies requiring that as much as 70 % of the parts for domestic wind installations had to be Chinese made. Although the requirement was short-lived, such policies "gave Chinese wind-turbine suppliers time to expand the scale and sophistication of their operations and building on China's comparative advantage in manufacturing costs, to be more competitive with foreign companies and abroad" (Yergin 2011: 545).

In 2010, China also became the global leader in renewable energy (in terms of annual investment), taking that title from Germany, which had held it for several years (REN21 2013: 47). According to many scenarios, China is expected to maintain its leadership in the coming decades. Just in terms of new installed wind power capacity worldwide in 2014, China led the race by accounting for 45 % of the total share as opposed to the United States which accounted for <10 %. In terms of the cumulative capacity wind power installed, China accounted for about 31 % of the global share, followed by the US at about 18 %. In wind energy capacity, Table 3 demonstrates that China is no doubt the global leader at present.

Table 3 Cumulative installed wind power capacity in top 5 countries 1980–2014 (megawatts)

Year	China	U.S.	Germany	Spain	India
1980	n.a.	8	0	0	0
1981	n.a.	18	0	0	0
1982	n.a.	84	0	0	0
1983	n.a.	254	0	0	0
1984	n.a.	653	0	0	0
1985	n.a.	945	0	0	0
1986	n.a.	1265	0	0	0
1987	n.a.	1333	5	0	0
1988	n.a.	1231	15	0	0
1989	n.a.	1332	27	0	0
1990	n.a.	1484	62	0	0
1991	n.a.	1709	112	5	39
1992	n.a.	1680	180	50	39
1993	n.a.	1635	335	60	79
1994	n.a.	1663	643	70	185
1995	38	1612	1130	140	576
1996	79	1614	1548	230	820
1997	170	1611	2080	512	940
1998	224	1837	2875	834	1015
1999	268	2472	4442	1812	1077
2000	346	2539	6113	2235	1220
2001	404	4275	8754	3337	1456
2002	470	4685	11,994	4825	1702
2003	568	6372	14,609	6203	2125
2004	765	6725	16,629	8263	3000
2005	1272	9149	18,415	10,027	4430
2006	2599	11,575	20,578	11,623	6270
2007	5910	16,824	22,194	15,145	7845
2008	12,020	25,076	23,826	16,689	9655
2009	25,805	35,086	25,673	19,160	10,926
2010	44,733	40,298	27,097	20,623	13,065
2011	62,364	46,929	29,071	21,674	16,084
2012	75,324	60,007	31,270	22,784	18,421
2013	91,412	61,091	34,250	22,959	20,150
2014	114,609	65,879	39,165	22,987	22,465

Source: Global Wind Energy Council (2013), www.gwec.net. Missing values modified from Roney (2014)

If we look across the board at the different types of renewable energy in Table 4, China led at the beginning of 2015 in 4 of 7 categories and placed in the top 5 in all but geothermal. If one were handicapping a race, the odds of China leading the pack are quite good. Still, the United States is ranked first or second in five categories and

Table 4 Total installed renewable capacity by source as of end of 2014 global rankings for top 5 countries

Source type	Renewable power (w/hydro)	Renewable power (no hydro)	Biopower	Geothermal	Solar PV	Wind	Hydro
1	**China**	**China**	**USA**	**USA**	Germany	**China**	**China**
2	**USA**	**USA**	Germany	Philippines	**China**	**USA**	Brazil
3	Brazil	Germany	**China**	Indonesia	Japan	Germany	**USA**
4	Germany	Spain/Italy	Brazil	Mexico	Italy	Spain	Canada
5	Canada	Japan/India	Japan	New Zealand	**USA**	India	Russia

USA and China positions designated in bold
Source: REN 21 (2015: 19)

places in the top 5 in all 7 categories. The race's outcome, therefore, is not entirely a foregone conclusion.

Despite China's lead in the renewable race, the country's push to boost renewables faces serious challenges. For instance, from 2000 to 2010, the increase in Chinese energy demand (just in terms of growth) was equivalent to two Latin Americas worth of energy consumption (Herberg 2013). In the span of a decade, the Chinese have created two continents worth of energy demand and they are expected to create another continent worth of demand every 5–6 years. This is detrimental to the global climate particularly since up to two-thirds of Chinese energy supply comes from coal. The biggest challenge for China is to acquire enough energy to feed the energy intensive juggernaut that its economy has become since the late 1990s.

Even as China's wind and solar power generating capacity rapidly expands, they may together account for just 5 % of China's total electricity generation in 2020. Even though China is moving fast on renewables, it still has been burning 10 % more coal each year just to keep up with the annual growth in energy demand. Air pollution from burning coal, however, has encouraged some retreat from this strong commitment to coal. Yet renewables cannot become China's primary source of energy any time in the near future since China's economy will continue to demand more energy than renewables can supply. Nuclear power and natural gas, both seen as cleaner alternatives to coal and oil, will be used but so will petroleum and coal. As one observer notes:

> Whether China can become a true energy innovator will determine the country's economic future. If it is successful, it will ease the country's rise to great-power status and aid the global fight on climate change. If it is not, the Chinese dream might see an unpleasant awakening, as China throws off the shackles of poverty only to see them replaced by the shackles of energy imports that have long plagued "Western foreign policy" (Huenteler 2014).

Of course, China is hardly alone in being far from developing a newfound reliance on renewables. If China has problems feeding its economic growth, the United States has major policy disagreements about how best to stimulate economic growth. New U.S. leads in petroleum and natural gas production thanks to shale

Table 5 Closeness to reaching 100 % renewables (at the end of 2014)

% Attained	States
50+	Norway, Paraguay
40–49	Iceland, Tajikistan
30–39	–
20–29	Portugal, Sweden
10–19	Costa Rica, Switzerland, New Zealand, Montenegro, Georgia, Spain, Kyrgyzstan, Albania, Uruguay, Canada, Denmark, Greece, Bosnia and Herzegovina
5–9	Armenia, Turkey, Ireland, Austria, Romania, El Salvador, Colombia, Slovenia, Brazil, Sri Lanka, Italy, Venezuela, Germany, Zambia, France, Latvia, Mozambique, Panama, United Kingdom, Bulgaria, Cambodia, Kosovo, Nicaragua, Peru, Chile
1–4.9	Croatia, Australia, Argentina, Vietnam, Japan, Sudan, Ecuador, Myanmar, Slovak Republic, Macedonia, North Korea, *USA*, Honduras, Congo, Guatemala, Philippines, Serbia, Syria, Ghana, Finland, *China*, Namibia, Russian Federation, Dominican Republic, Kenya, Azerbaijan, Estonia, Iraq, Morocco, Czech Republic, Mexico, Cyprus, Belgium, Cameroon, Poland, India, Bolivia, Congo Dem. Rep., Zimbabwe, Ethiopia, Angola, Netherlands, Pakistan, Cote d'Ivoire, Ukraine, Nepal, Malaysia, Iran, Lithuania, Gabon, Lebanon, Luxembourg, Israel
0–0.9	43 more states

USA and China positions designated in italics
Source: Based on rank orders reported in Jacobson et al. (2015: 14)

techniques are rather difficult to suppress or ignore in this context, even if the gains are unlikely to be long-term in nature. Most of the world is in a similar position. With the exception of a very few small countries, most states have much farther to go than they have come in making the transition to a fully renewable commitment. Table 5 singles out only four states (Norway, Paraguay, Iceland, and Tajikistan) that are deemed close to half way in making this transition. The rest of the world is far behind, including the two potential leaders, the United States and China—both of which are judged to be <5 % along the way towards the full transition to a renewable economy.

These trends are more than sobering and suggest that the renewable age (despite all investments and installed capacities) may well not emerge before the twenty-second century. China and the United States may be leading the race, but leading at a snail's pace is not good enough, given the impending climate crisis. So while we are asking "who is ahead in the renewable race"—the United States or China?—the more important question is "are they leading fast enough?" In 2013, renewables accounted for 10.1 % of global energy consumption, whereas fossil fuels covered around 78.3 % of the world's primary energy needs (REN21 2015: 27). It is not enough to lead the renewable race if the leading economies are not yet making systemic changes along the way. More importantly, if neither the United States nor China is transitioning away from fossil fuels to renewable energy-based economies fast enough—and we argue that the ascent to system leadership is predicated on riding a complex of new technology and new fuel—then new system leadership is less likely to happen in the next half century or so.

To help answer these questions, we have to know what the current map of U.S. and Chinese renewable energy consumption patterns look like, and what the map might look like by 2050. In the face of the climate crisis, the idea of an inexhaustible and environmentally friendly energy source is deeply appealing. Yet, currently, both leading economies heavily rely on fossil energy to fuel their economies. Since the 1990s, China's coal consumption has grown rapidly, providing more than two-thirds of China's total energy, and four-fifths of its electricity, and most of it is produced domestically. This rate of coal production growth is unsustainable—something that even Chinese decision-makers now acknowledge. Data from the World Energy Council shows that China's proven coal reserves will last 34 years given its annual production rates, based on figures for 2011. "That is down from about 100 years just a decade ago and means China will have exhausted its reserves by 2049, if it keeps going at the current rate" (Huenteler 2014).[9] But the longer term scarcity problem is overshadowed by the nearer-term costs in environmental and human welfare within China. The severity of Chinese public health problems due to its dependence on coal alone have been enough to give central decision-makers strong incentives for re-thinking their approach.

The fear of an unwavering commitment (at least to 2030) to coal on China's part may not reflect what is sometimes called the "new normal" growth model (Green and Stern 2015). An intensive effort to rapidly alter China's lack of economic development has been successful in some respects but it also has led to the previously noted and quite major environmental and public health problems. Less emphasis on quantitative growth and more emphasis on qualitative growth—that is, the new normal—may lead to a more rapid abandonment of coal than has been planned in the past. Such a move would address the environmental degradation problem at the same time that it prepares the Chinese economy for a switch to more high-tech economic production.[10] Time will tell how this new normal model plays out.

What we do know at this juncture is that the demand for fossil fuels remains fairly high. Our collective appetite for energy is so great that most of the known fossil deposits could be used up during the twenty-first century. For decades, peak oil pessimists have been warning about the imminent end to fossil energy reserves. Yet this end never quite seems to be in sight, and most people take no heed of the warnings as long as gas prices remain relatively low. That is because in the past constant technological advances in the exploitation of oil and natural gas have always resulted in a revision of the forecasts about how long reserves would last.

[9]Chinese energy statistics have been revised in recent years to reflect even greater consumption rates than were reported earlier.

[10]Chinese decision-makers appear to be aware that many high-growth LDCs have stalled mid-point in their development trajectories by failing to move away from low-tech to high-tech innovation and production.

The large coal reserves still available worldwide could enable us to use fossil energy sources for decades or even another century. But the number of new finds, especially of oil, has declined substantially during recent years, and new supplies cannot be exploited fast enough to meet rising demand. For instance, in 2013, total global oil consumption grew by nearly 1.4 million barrels per day, slightly above the historical average. The United States recorded the largest increment to global oil consumption in 2013 (400,000 more barrels per day), outpacing Chinese growth (39,000 more barrels per day) for the first time since 1999. In the next year, circumstances reversed themselves with the United States consuming little more petroleum in 2014 than in 2013 while China increased its consumption by 392,000 barrels daily. Global consumption rose for other fossil fuels as well. Table 6 shows world consumption of petroleum, coal, and natural gas all still on the rise. Chinese coal consumption may actually have peaked but its demand for natural gas continues to grow. The decline in American coal consumption is offset by the increased use of natural gas.

Oil remains the world's leading fuel, still accounting for about a third (32.6 %) of global energy consumption in 2014 (BP Statistical Review 2015). As long as fossil fuels remain attractively priced thanks to ample supply and continue to enjoy subsidies that are much more than subsidies to renewables, they are likely to dominate the global energy mix for some time to come. Non-hydro renewables now account for <3 % of total energy consumption.

We may well be in the midst or the early days of an energy transition but it will be slow and protracted. This makes it difficult to imagine that renewable energies will become the primary source of energy in the next half century. The production and use of renewable fuels has grown more quickly in recent years as a result of higher prices for oil and natural gas, and a number of state and federal government incentives for renewable energy. The use of renewable fuels is expected to continue

Table 6 Fossil fuel consumption, 2010–2014

	2010	2011	2012	2013	2014
Petroleum					
USA	19,180	18,882	18,490	18,961	19,035
China	9,266	9,791	10,231	10,664	11,056
World	87,867	88,974	89,846	91,243	92,086
Coal					
USA	525	495.4	437.9	454.6	453.4
China	1740.8	1896	1922.5	1961.2	1962.4
World	3451.9	3611.2	3777.4	3867	3881.5
Natural Gas					
USA	682.1	693.1	723.2	739.9	759.4
China	110.5	134.9	151.2	170.8	185.5
World	3193.7	3265.3	3345.8	3381	3393

Source: BP Statistical Review (2015); petroleum is measured in thousand barrels daily, the coal metric is million tonnes oil equivalent and natural gas is enumerated in billion cubic metres

Table 7 Forecasted world proportions of renewables in future energy mixes

	2020	2025	2030	2035	2040	2050	2060	2100
EIA					12 (US)			
IEA				12–23				
BP				7				
BP				17 (US)				
BP				10 (CHN)				
Exxon Mobil		15			15			
Shell Global a	7.1		11.5		16.3	22.8		
Shell Global b	9.5/ 9.6		9.8/ 11.5		12.2/ 18.2	16.7/ 18.2	19.6/ 32.7	
WEC						20–30		
Bloom-berg			16–22					
Matias and Devezas (2007)								100
Li (2014)								100
Hartley et al. (2014)								100

to grow over the next 30 years, although most forecasts project that we will still rely on non-renewable fuels to meet most of our energy needs.

Table 7 supports this generalization. Forecasts are based on scenarios and some forecasters like to develop conservative and liberal sets of assumptions (hence the range of outcomes separated by a slash in Table 7). Yet none of these forecasts can envision more than a third of the world's energy being supplied by renewable sources as late as 2060. Based on these types of calculations the Age of Renewables is unlikely to emerge before the advent of the twenty-second century.

Nonetheless, even the "liberal" entries in Table 7 are generally conservative. As REN21/ISEP notes "oil companies, some industry groups, the IEA, and the EIA" all tend to specialize in conservative estimates of energy mix change, usually restricting their renewable proportion to about 20 % of the total. Estimates that boost the renewable proportion to as much as 40 % tend to focus on climate mitigation concerns. Anything higher is reserved for renewable advocates who claim that renewables could become the main source of energy as early as 2050 if governments and populations were prepared to make the shift.[11]

We remain unconvinced that the conservative estimations will prove to be good guides to the future of fossil fuel consumption. For one thing, the predictions for lingering petroleum reliance are predicated on dubious Middle Eastern claims about reserves and the expectation of continuing oil discoveries. New oil discoveries may have peaked in 1965 according to Hughes (2012). So far, the potential for Arctic oil

[11]See, for instance, Williams et al. (2014, 2015), Jacobson et al. (2015), Teng et al. (2015), Ackerman et al. (2016), and Labor Network for Sustainability (2016).

exploitation seems limited. Even so, there are very good reasons to expect that any transition to renewables will entail an extraordinary investment and the overcoming of quite significant political resistance. It may be that the conservative estimates will prove reasonably accurate assuming that the supply of carbon-based fuels can somehow be elongated into the future. If so, the full availability of a new source of energy would seem to be unlikely during the next half-century and perhaps not until the twenty-second century.

At the same time, the current plans for an accelerated decarbonization by 2050 that could respond to the threats associated with global warming (as exemplified by Guerin et al. 2014; Jacobson et al. 2015) do not call for a full movement to renewables. The mix of fuel sources would vary from state to state depending in part on their development level, differential access to various energy sources and a focus on electrification of building heating/air conditioning and transportation. For instance, the United States energy mix (Williams et al. 2014) might be composed of renewables (39 % including hydro), nuclear (30 %), and gas and coal that have been ccs treated for carbon capture and storage (30 %). The Deep Decarbonization Project's Chinese plan (Teng et al. 2015) is less specific about its energy combinations but appears to have a 2050 China resembling a 2010 OECD country with much less emphasis on non-ccs coal. These 2050 plans are scenarios that allow for variable combinations of different energy sources[12]. We suspect that this more flexible approach is more likely to reflect 2050 conditions than either a continued petroleum or new renewable age. Yet some form of intensive decarbonization by 2050 presumes a widespread and serious response to global warming and climate change and, we recognize, that remains an iffy assumption.

If we are right about the historical pattern of new and inexpensive energy providing a base for systemic leadership, the future's economic context will not be favorable to new systemic leadership—either a renewed version led by the United States or a new version championed by China. Barring some type of unpredictable game changer, China may grow relatively richer than it has been and the US lead on the rest of the world will continue to diminish slowly—without any abruptly radical or fundamental shift in relative power distribution. This outcome will be unfortunate if systemic leadership is considered crucial to the likelihood that there will be serious and widespread response to climate change problems. At best, one might hope for U.S–Chinese cooperation on this question—something that may be emerging but will always be vulnerable to the vicissitudes of their ongoing rivalry.

Thus, the immediate future could very well be more of what we are currently experiencing—at a minimum, increasing verbal acknowledgement of a serious problem and maybe movement towards dealing with the issue in some meaningful way(s). The Paris 2015 meeting was more positive than earlier meetings in Copenhagen or Kyoto. Yet, it still lacked a concrete agreement to take action due in large part to the U.S inability to obtain legislative ratification for concrete measures. This

[12]As a consequence, 2050 decarbonization plans are often accompanied by multiple scenarios reflecting variable mixes of energy inputs to arrive at the same end.

secondary problem reflects less a lack of systemic leadership than it does political polarization within the United States.[13] Still, the two processes are not un-related and, more generally do reflect current weaknesses in the evolution of systemic leadership. Hence, our expectation of more of the same in the immediate future. One difference, however, is that the future is likely to be characterized by a world approaching exhaustion of its hitherto relatively inexpensive fossil fuels and continued and perhaps escalating problems with greenhouse gas emissions. More renewable energy will become available for sure. What remains unclear is how much is needed, by when, how long the transition away from fossil fuels will take, and whether systemic leadership—old or new—will be able to cobble together an effective response to global warming.

2.3 Conclusion

Brown (2012) has argued that "The great energy transition from fossil fuels to renewable sources of energy is under way….The old energy economy, fueled by oil, coal, and natural gas, is being replaced with an economy powered by wind, solar, and geothermal energy."

Most everyone concerned with the health of the environment would agree with Brown's assertion that the time to transition from fossil fuels toward renewable energy is past due. However, despite Brown's optimism about the future of renewable energy, a new world economy sustained by renewables is hardly emerging quickly. Nor is it likely to emerge quickly. At the current rate of global energy transition it may be the dawn of the twenty-second century before renewable alternatives constitute the primary source of global energy.

Over the past several hundred years, systemic leadership transitions (from the Dutch to the British to the American) have become tied to transitions in the primary source of energy (from peat to coal to petroleum, respectively). That is because historically leading states have had a commanding lead in energy and the associated infrastructure. System leaders need cheap energy to fuel their economic leads. The technology that they pioneer tends to change the type of energy (and vice versa) that is most critical to operating at the production frontier. Thus, energy transitions have profound and long-lasting implications for economic growth, which in turn exerts considerable influence on the international system's power hierarchy.

In the past, systemic leadership in the West emerged first in Europe and then in the United States at least partially due to the adoption of fossil fuels after failures to fully do so in China and the Netherlands. To go beyond agrarian economies, leading economies needed additional sources of energy. While wind and peat were important parts of the rise to systemic leadership, it took coal for the British and petroleum for the Americans to become predominant powers in the international

[13] The problem is complicated. For a cursory treatment of the serial fluctuations in U.S. and Chinese positions on climate change measures at COP meetings prior to Paris, see Darwall (2013).

system. Systemic leadership transitions ultimately hinged on the development of a combination of technological leadership with access to relatively inexpensive energy (and the energy had to fit the radical new technology). Assuming the radical new technology of the mid to latter twenty-first century will be IT based, that means whatever is useful for generating electricity at the least. Yet petroleum, if not coal, presumably will be much more scarce by 2050 and therefore more expensive. Of course, the major new wrinkle in this energy transition story is that the use of carbon-based fuels is causing major environmental problems, which will only become more severe if more carbon-based fuels continue to be utilized—as is anticipated. Hence, renewables are green for both the environment and system leadership. In this chapter we have argued that as the world approaches exhaustion of relatively inexpensive sources of fossil fuel, the ascendance of new systemic leadership will hinge on a transition towards non-fossil fuel energy sources. Thus it matters somewhat who is ahead in the renewables race but if they are not ahead that much, it suggests a low probability of a politico-economic transition in the twenty-first century. The absence of new or re-invigorated systemic leadership, in turn, bodes poorly for the construction of a global response to global warming. Not only do system leaders need to negotiate a global consensus to make changes, they must also lead in radically transforming their own economies. So far, it has proven easier to accomplish the former than the latter.

References

Ackerman, F., Cumings, T., & Fields, S. (2016). *The clean energy future appendix*. Cambridge, MA: Synapse Energy Economics. synapse-energy.com/sites/default/files/Clean-Energy-Future-Appendix-15-054.pdf

Bardi, U. (2014). *Extracted: How the quest for mineral wealth is plundering the planet*. White River Junction, VT: Chelsea Green Publishing.

Bloomberg New Energy Finance. (2013). *Global renewable energy market outlook 2013*. http://about.bnef.com/presentations/global-renewable-energy-market-outlook-2013-fact-pack/

British Petroleum. (2014). *BP statistical review of world energy*. http://www.bp.com/content/dam/bp/pdf/Energy-economics/statistical-review-2014/BP-statistical-review-of-world-energy-2014-full-report.pdf

British Petroleum. (2015). *Statistical review of world energy 2015*. www.bp.com/en/global/corporate/about-bp/energy-economics/statistical-review-of-world-energy.html

Brown, L. (2012). *The great transition: Part I: From fossil fuels to renewable energy*. Earth Policy Institute. http://www.earth-policy.org/plan_b_updates/2012/update107

Darwall, R. (2013). *The age of global warming: A history*. London: Quarter Books.

de Oliveira Matias, J. C., & Devezas, T. C. (2007). Consumption dynamics of primary energy sources: The century of alternative energies. *Applied Energy, 84*(7–8), 763–770.

Global Wind Energy Council. (2013). *Global wind report: Annual market updated*. http://www.gwec.net/wp-content/uploads/2014/04/GWEC-Global-Wind-Report_9-April-2014.pdf

Green, F., & Stern, N. (2015). *China's "new normal": Structural change, better growth, and peak emissions*. London: Grantham Research Institute on Climate Change and the Environment (London School of Economics and Political Science) and Centre for Climate Change Economics and Policy (University of Leeds and London School of Economics and Political Science).

Guerin, E., Mas, C., & Waisman, H. (Eds.). (2014). *Pathways to deep decarbonization—2014 report*. Unsdsn.org/wp-content/uploads/2014/09/DDPP_Digit.pdf
Hartley, P., Medlock III, K. B., Temzelides, T., & Zhang, X. (2014). *Energy sector innovation and growth*. Working paper 14-009. Houston, TX: Rice Initiave for the Study of Economics (RISE).
Helm, D. (2012). *The carbon crunch* (revised and updated ed.). New Haven, CT: Yale University Press.
Herberg, M. (2013). *China's energy future*. Los Angeles, CA: USC US-China Institute https://www.youtube.com/watch?v=NtMW_aa-r1I.
Huenteler, J. (2014, May 14). China's coal addiction a threat to its energy security. *The South China Morning Post*.
Hughes, J. D. (2011). *Will natural gas fuel America in the 21st century?* Post Carbon Institute. www.postcarbon.org/report/PCI-report-nat-gas-future.pdf
Hughes, J. D. (2012). The energy sustainability dilemma: Powering the future in a finite world. http://www.eeb.cornell.edu/howarth/HUGHES%20Cornell%Ithaca%20May%20%202012.pdf/
Jacobson, M. Z., Delucchi, M. A., Bauer, Z. A., Goodman, S. C., Chapman, W. E., Cameron, M. A., et al. (2015). *100 % clean and renewable wind, water, and sunlight (wws) all-sector energy roadmaps for 139 countries of the world*. Stanford, CA: Department of Civil and Environmental Engineering, Stanford University.
Labor Network for Sustainability. (2016). *The clean energy future*. www.labor4sustainability.org/wp-content/uploads/2015/10/cleanenergy_10212015_main.pdf
Li, M. (2014). *Peak oil, climate change, and the limits to China's economic growth*. London: Routledge.
Quaschning, V. (2010). *Renewable energy and climate change*. West Sussex: Wiley.
REN21. (2015). *Renewables 2015: Global status report*. Paris: REN21 Secretariat. www.ren21.net/wp-content/uploads/2015/07/REN12-GSR2015_Onlinebook_low1.pdf
REN21/ISEP. (2013). *Renewable energy policy network for the 21st century. Renewables 2013: Global status report*. Paris: Institute for Sustainable Energy Policies. www.res4me.org/uploads/iocs/13612002098REN21_GFR_2013_print.pdf
Roney, J. M. (2014). *World wind power poised to bounce back after slowing in 2013*. Earth Policy Institute. Available at http://www.earth-policy.org/indicators/C49/wind_power_2014
Smil, V. (2015). *Natural gas: Fuel for the 21st century*. Chichester: Wiley.
Teng, F., Alun, G., Yang, X., Wang, X., Liu, Q., Chen, Y., Tian, C., & Zheng, X. (2015). Pathways to deep decarbonization in China. SDSN-IDDRI. http://deepdecarbonization.org/wp-content/uploads/2015/09/DDPP
Teng, F., Qiang, L., Chen, Y., Tian, C., Zheng Z., Gu, A., et al. (2015). *Pathways to deep decarbonization in China*. Deepdecarbonization.org/wp-content/uploads/2015/09/DDPP_CHN.pdf
U.S. Energy Information Administration [EIA]. (2015). *Annual energy outlook 2015 with projections to 2040*. www.eia.gov/forecasts/aeo/pdf/0383(2015).pdf
William, J. H., Haley, B., Kahrl, F., Moore, J., Jones, A. D., Torn, M. S., & McJeon, H. (2014). Pathways to deep decarbonization in the United States. Deep Decarbonization Pathways Project of the Sustainable Development Solutions Network and the Institute for Sustainable Development and International Relations. https://ethree.com/publications/index US2050.pdf
Williams, J. H., Haley, B., & Jones, R. (2015). *Policy implications of deep decarbonization in the United States* (Vol. 2). Deep Decarbonization Pathways Project of the Sustainable Development Solutions Network and the Institute for Sustainable Development and International Relations.
Yergin, D. (2011). *The quest: Energy, security and the remaking of the modern world*. London: Penguin.

Metal Matrix/Nanocarbons Composites Based on Copper and Aluminum

Oleg V. Tolochko, Vesselin G. Michailov, and Andrey I. Rudskoi

Abstract Metal matrix composites possess high-temperature capability, high thermal conductivity, low thermal expansion coefficient, and high specific stiffness and strength. Recently discovered carbon nanostructures such as carbon nanofibers (CNFs), nanotubes (CNTs) and graphene are promising components for next-generation high-performance structural and multifunctional composite materials. Here, we utilized new approach to fabricate composite materials on the basis of aluminum, and copper matrix. In this chapter we summarize our knowledge and also present new results on the preparation of a good dispersion of CNTs, CNFs and graphene in a matrix, with the intention of improving the mechanical or electrical properties of metal-carbon composite nanomaterials. This study shows the way to solve one of the largest problem to creating strong, electrically or thermally conductive CNT/CNF or graphene composites: the difficulty of achieving a good dispersion of the carbon nanomaterials in a metal matrix. We show that discontinuously reinforcement can be successfully developed for industrial applications.

Keywords Metal matrix composites • Carbon nanostructures • Copper, aluminum • Mechanical properties • Thermal and electric conductivity

O.V. Tolochko (✉)
Materials Science Department, Peter the Great State Polytechnic University, 29 Polytechnicheskaya St., Saint Petersburg, Russia
e-mail: tolochko_ov@spbstu.ru

V.G. Michailov
Research Center of Lightweight Materials, Brandenburg University of Technology – BTU CS, Cottbus, Konrad-Wachsman-Allee 17, Senftenberg, Germany
e-mail: michailov@b-tu.de

A.I. Rudskoi
Peter the Great State Polytechnic University, 29 Polytechnicheskaya St, Saint Petersburg, Russia
e-mail: rector@spbstu.ru

T. Devezas et al. (eds.), *Industry 4.0*, Studies on Entrepreneurship, Structural Change and Industrial Dynamics, DOI 10.1007/978-3-319-49604-7_10

1 Introduction

Metal matrix composites (MMCs) possess high-temperature capability, high thermal conductivity, low thermal expansion coefficient, and high specific stiffness and strength. These unique physical properties generated vast optimism for the utilization of MMCs in critical space system applications in the late 1980s. Also enhanced ecological requirements to the transport vehicles in air, on the land and in water, especially for the reduction of harmful emissions, associated essentially with the reduction of dead weight. Along with the new design and construction methods in order to optimize the shape and limit the use of the materials properties, it is of special relevance the physical and mechanical properties of materials. Discontinuously reinforced MMCs such as graphite/aluminum, magnesium and copper-based composites were developed both for aerospace and commercial applications.

Carbon nanomaterials, such as carbon nanotubes (CNTs), nanofibers (CNFs) and graphene have recently achieved tremendous scientific interest due to their remarkable and useful properties, such as exceptional tensile strength, elastic modulus and electrical and thermal conductivity (Dresselhaus et al. 1996; Reich et al. 2004). These materials are promising candidates for the next generation high performance structural and multi-functional composite materials (Calvert 1999; Zhang et al. 2007; Riggs et al. 2000).

One of the largest obstacles to create strong, electrically or thermally conductive CNT/CNF composites is the difficulty of obtaining a good dispersion of the carbon nanomaterials in a matrix. Typically, time-consuming steps of the carbon nanomaterial purification, functionalization and mixing with matrix are required. Here, we utilize a new approach to grow CNTs/CNFs directly on the surface of the matrix, matrix precursor or filler particles. As precursor matrix and fillers we utilized cement (clinker), copper powder, fly ash particles, soil and sand. Carbon nanofibers were successfully grown on these materials without additional catalyst. Investigations of the physical properties of the composite materials based on these carbon-modified particles revealed enhancement in mechanical and electrical properties.

We have proposed a novel approach to create well-dispersed carbon nanomaterials in the matrix by their directly growth on the surface of matrix particles without any catalyst of additional treatments. This allowed us to get homogeneous introduction of carbon nanomaterials in the metal matrix. The synthesis of CNFs on the surface of metal powder was carried out in a simple one-step chemical vapor deposition (CVD) process.

Here, we utilized this approach to fabricate new composite materials on the basis of aluminum and copper matrix. In this paper we summarize our knowledge and also present new results on the preparation of a good dispersion of CNTs, CNFs and graphene in a matrix, with the intention of improving the mechanical and electrical properties of metal-carbon composite nanomaterials.

2 Experimental

2.1 Composite Powders Synthesis Method

We have proposed a novel approach to create well dispersed carbon nanomaterials by directly growing these on the surface of matrix particles, resulting in further improvements of the mechanical and electrical properties of the composite materials. Utilizing copper and aluminum particles and varying the carbon source from C_2H_2 to C_2H_4, we were able to alter the carbon product from nanofibers to few-layered graphene on the surface of micron-sized copper particles.

For the synthesis on the surface of copper particles we used the experimental setup described elsewhere (Nasibulina et al. 2010; Rudskoy et al. 2014). It consisted of a quartz tube heated in a resistive furnace in the temperature range from 700 to 940 °C. For the CNF synthesis, 30 cm^3/min of acetylene together with 260 cm^3/min of H_2 were introduced into the reactor. While for the graphene synthesis a gas mixture of C_2H_4/H_2 with a mole ratio of 1:2 was used. For the experiments, copper metal powder with purity of 99 % having an average particles size of 5–10 μm was purchased from Goodfellow. For the synthesis of the carbon materials the reactor was heated to the desired temperature in an N_2 atmosphere (100 cm^3/min). Then, a ceramic crucible with the copper powder (about 2 g) was placed in the middle of the reactor. After 5 min N_2 was replaced by 240 cm^3/min H_2 flow for 5 min for the surface copper oxide reduction. Then the carbon source was introduced together with hydrogen. After 10–20 min, a mixture of C_2H_4 and N_2 was switched to pure N_2 and the crucible was moved inside the quartz tube outside the hot zone to cool down to room temperature.

Also, we have shown the possibility of fabrication of a composite "aluminum–carbon nanofibers" material by gas phase synthesis in an acetylene-hydrogen mixture on the surface of micro-particles of aluminum containing up to 15 wt.% carbon nanofibers. In the method used we deposited a nickel or cobalt catalyst onto the surface of an aluminum powder from an aqueous solutions of corresponding nitrate.

2.2 Powders Compacting Methods

As the first step copper- and aluminum-based composite powders were compacted through uniaxial cold pressing in cylindrical molds of 8 mm in diameter with subsequent sintering. The compacting pressure was kept at about 500–600 MPa. The sintering was carried out at 950 °C during 2 h in an H_2 atmosphere for copper-based composites and 600 °C for Al-CNF composites. However, the syntheses of the graphene happen at temperatures about 940 °C, when significant sintering of the copper powder occurs, that makes impossible to apply the same technological route, so the partly sintered Cu-graphene powders were subjected to cold rolling.

The hot pressing was implemented with the help of a DA0040 hydraulic press developing a force of 106 kgf in a "truncated-hemisphere" high-pressure chamber at a pressure of 5 ± 0.5 GPa and varied temperature. Powders of various compositions were used in advance to make strong tablets 10 mm in diameter and about 4 mm high at a pressure of 600 MPa is a press mold produced from a steel tool with a hardness of 62–64 *HRC*. Then a tablet was placed into a high-pressure container between two graphite tablets into a cylindrical graphite heater. The high-pressure container was heated by passing an alternate current (0.3–0.8 kA) at a voltage of 3–6 V in the mode of stabilization of the electric power. The current and the voltage in the pressing process were detected for tracing the resistivity of the specimen in the heating process. The accuracy of the stabilization of the electric power was at least 5 %. The temperature inside the container was determined from the "power–temperature" calibration dependence and from the graphite-diamond phase transition in the presence of a nickel-manganese catalyst (Kidalov et al. 2008a, b). The accuracy of the determination of the temperature was ±50 °C.

2.3 Study of Structure and Properties

Microscopic investigations of the produced materials were carried out with the help of a high resolution analytical scanning electron microscope (SEM) equipped with field emission gun FE-SEM JEOL 7500FA and high resolution transmission electron microscope (TEM) JEOL JEM2200FS with corrector of spherical aberrations (Cs ≤ 0.005 mm). For TEM observation, the samples were dispersed and treated in ethanol by tip ultrasonicator (Branson 450D, microtip) at the power of 40 W (3 × 3 s pulses). The Raman spectra were recorded at room temperature; for the source of monochromatic radiation, the YAG laser (with a wavelength of 532.25 nm and a power of 30 mW) was used.

Brinell hardness was measured by a hardness tester Zwick/Roell ZHU by using 2.5 mm steel ball at the load of 62.5 kg. Elongation measurements of the sample were carried out using a tensile machine IR-5040-5 with a maximum load of 5 kN applying a three-point bending procedure. Samples for the test were prepared in the form of a plate with the size of 20 × 5 × 1 mm^3.

The friction tests were conducted on an end tribometer using a pin-on-disc scheme. The counter material was a copper (C11000) ring. Sliding was performed under dry friction at ambient conditions (room temperature 24 °C and humidity 50 %); linear velocity and constant load were set up as 0.8 m/s and 20 N correspondingly. Before each test the samples and counterpart surfaces were cleaned and polished. The wear (I) was calculated out from the loss of each specimen's volume and way of friction according to: $I = V/SL$, where V is the volume loss, S is the contact area, L is the track length. The coefficients of the friction were obtained by measuring the tangential force on the specimen using a strain gauge bridge.

3 Results and Discussion

3.1 Cu-Nanocarbons Composites

Figure 1 shows the dependence of the weight of the samples treated at different synthesis temperature and the SEM images of Cu-nanocarbons composite powder particles. The results of this study are compared with those obtained with acetylene as a carbon source reported in Nasibulin et al. (2013). As can be seen from the figure a notable ethylene decomposition occurred at temperatures above 750 °C. At lower temperatures neither the sample weight increase nor colour change were observed. The maximum of the mass increase of 6 % was found at the temperature of 890 °C. The decrease of the carbon deposited on the powder at the temperatures above 900 °C may be explained by gas phase ethylene decomposition and formation of aerosol carbon particles, which were deposited on the cold parts of the reactor.

So, for the CNF synthesis, a mixture of C_2H_2 and H_2 was utilized, while for the graphene synthesis, C_2H_2 was replaced by C_2H_4. Figure 2 shows SEM images of copper powders particles covered by carbon nanofibers, and multigraphene layers.

For mechanical tests, the prepared samples were mixed with a certain amount of pure copper powder. The concentration of CNFs in the composite was varied from 1 to 10 mass%. To prepare the composite, the mixture was compacted by the uniaxial cold pressing and sintered. HBW test showed significant increase in hardness up to 60 kg/mm^2 with increasing CNF concentration to 1–3 % and subsequent slow decrease (Fig. 3a). The decrease in hardness at higher carbon content may be explained by thickening of the grain boundary CNFs clusters while retaining the same matrix grain size. In contrast to the hardness, the specific

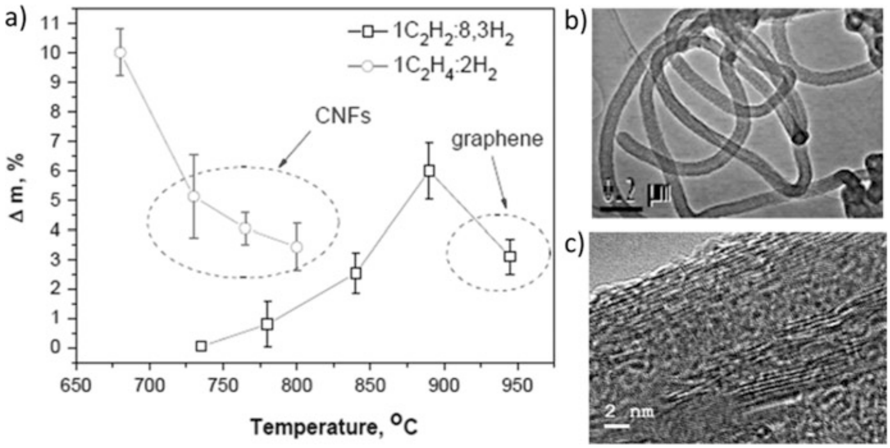


Fig. 1 (**a**) Temperature dependence of the sample weight after 20 min treatment with different carbon sources resulted in the formation of (**b**) CNFs in the presence of acetylene and (**c**) a few-layered graphene in the presence of ethylene

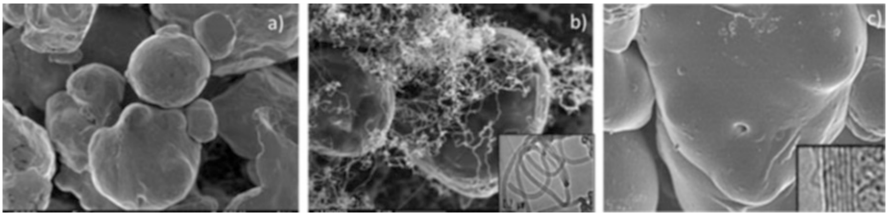


Fig. 2 SEM image of copper powders particles (a), particles covered by carbon nanofibers (**b**), and multigraphene layers (**c**)

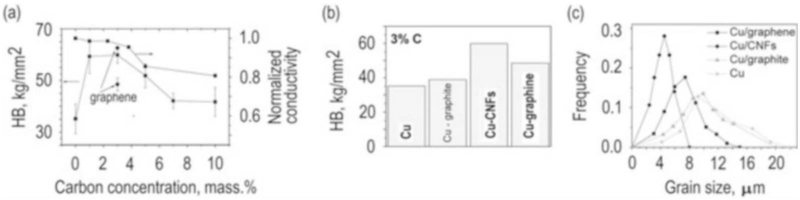


Fig. 3 (**a**) Mechanical and electrical properties of copper-based composites: hardness, and normalized conductivity of composite materials. (**b**) Mechanical properties of copper-based composites. (**c**) Grain size distribution in various composites

elongation naturally decreases, when the amount of CNF/Cu was increased in pure copper.

We have also carried out the comparative analysis with Cu/graphite prepared by traditional method and Cu/graphene composites (Fig. 3b). As expected any carbon additions to copper led to hardening of the composite. The best results were obtained for Cu/CNFs composite. Cu/CNF composite prepared by cold pressing and subsequent sintering showed 1.7 times increase in the hardness compared to pure copper. Copper hardened by graphene showed only 39 % increase in HB with significant ductility decrease.

In order to better understand the mechanical test results, we examined the microstructure of the produced composite materials. Figure 3c shows the grain size distribution in various composites measured on the basis of optical images. As one can see the hardness of the composites correlates perfectly with the grain size: the smaller the size the greater the hardness. As seen from the figure, Cu/CNFs composite has the narrowest grain size distribution with the smallest mean size of around 4 µm, that means it possess the most homogeneous and fine structure.

As known Cu-graphite composites are of great interest for tribological applications, so the tribological characteristics of the proposed Cu-CNFs material were investigated and compared to traditional Cu-graphite ones (Larionova et al. 2014). The composition 17–33 vol.% of carbon was chosen based on industrial Cu-graphite composite (Russian Standard—GOST 26719-85).

Coefficient of friction and wear loss of the tested materials are presented in Table 1. Wear in Cu-33 vol.% CNFs material reduced by more than 8 times

Table 1 Tribological properties of Cu-C composites

Material	Coefficient of friction	Wear ($\times 10^{-6}$)
Cu	0.9	22.0
Cu-33 vol.% graphite	0.16	6.0
Cu-17 vol.% CNFs	0.13	6.2
Cu-33 vol.% CNFs	0.08	2.5

compared to pure copper and by 2 times over Cu-33 vol.% graphite samples (Koltsova et al. 2015).

Images of the friction surfaces show that under wear conditions and the plastic deformation of copper matrix the graphite particles squeeze out and smear over contact surface. This graphite debris impedes adhesive wear and decrease of friction coefficient by reducing metal to metal contact points. As reported in (Moustafa et al. 2002; Sarmadi et al. 2013) the prevailing wear mechanisms of Cu-graphite composite under similar load is delamination: shear stress occurred under counterparts sliding causes nucleation and growth of the cracks in subsurface area resulting in detachment of the upper layers (Sarmadi et al. 2013).

It is clearly seen from the worn surface image, that the wear of the Cu-CNFs composites occurs through thin flakes exfoliation what is a characteristics of delamination wear as well. Considering the higher hardness of Cu-CNFs composite over Cu-graphite one (Fig. 3), plastic deformation of the copper matrix affects thinner subsurface area. Due to composite powder preparation method applied in this work, the copper grains are covered by CNFs layer, and, as was mentioned above, the copper matrix itself has finer grain structure, so the cracks caused by shear stress propagate along this CNFs reach boundary layer in the very upper subsurface area, resulting in the fine flakes exfoliation.

3.2 Al-Nanocarbons Composites

Also, we have shown the possibility of fabrication of a composite “aluminum–carbon nanofibers” material by gas phase synthesis in an acetylene-hydrogen mixture on the surface of micron particles of aluminum up to 15 wt.% carbon nanofibers.

In the method used we deposited a nickel catalyst onto the surface of an aluminum powder from an aqueous solution of by nickel nitrate spraying from an aqueous solution. Right before the synthesis, the specimens coated with the catalyst were annealed additionally in a hydrogen environment for 10 min at 550 °C to provide decomposition of the nickel nitrate and reduction of the oxide to metallic nickel that served a catalyst of growth of carbon nanofibers on the surface of the aluminum powder. Carbon nanostructures were synthesized at a temperature of 550 °C for 5–10 min.

Figure 4 presents the dependence of the variation of the mass of a specimen with respect to the initial weighed portion of aluminum on the time of the synthesis (Rudskoy et al. 2014). With increase of the time of the synthesis the rate of growth of the carbon nanostructures decreases due to deactivation of the catalyst, i.e., disappearance of the places of dominant nucleation of carbon nanostructures on the surface of the powder.

Uniform distribution of carbon nanofibers was obtained at a minimum content of nickel catalyst equal to 0.035 wt.% (Fig. 5). Transmission electron microscopy shows that the carbon product is represented by carbon nanofibers. The structure of the fibers at a high magnification looks like a "pile of coins" (Rakov 2006). The carbon products obtained at 0.035 % nickel were studied by the method of Raman spectroscopy. The Raman spectra show the presence of a peak at 1320 cm^{-1} (a D-peak) and tangential fluctuations (a G-peak) at 1590 cm^{-1}; the high parameter I_D/I_G

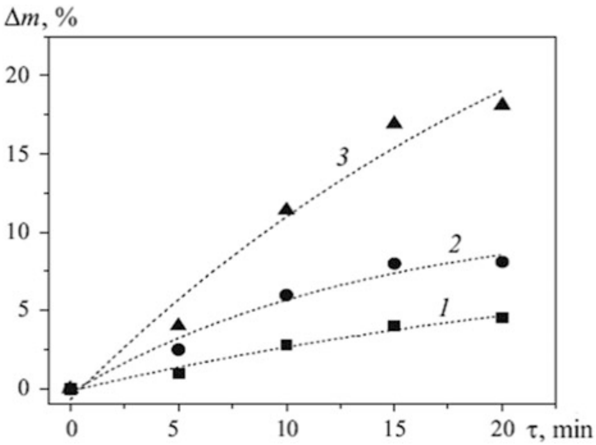


Fig. 4 Dependence of variation of the mass of a specimen on the time of synthesis at a temperature of 550 °C at different contents of nickel: (1) 0.015 % Ni; (2) 0.035 % Ni; (3) 0.07 % Ni

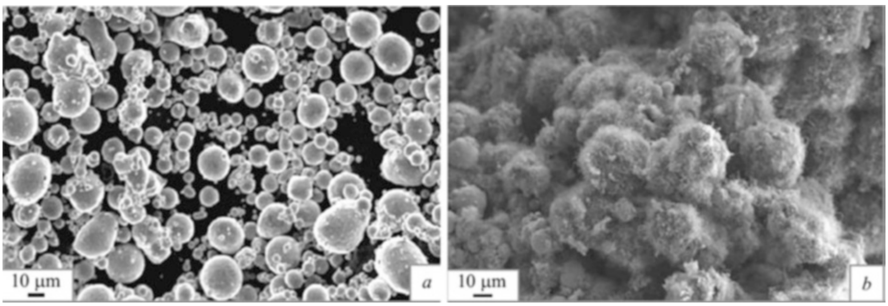


Fig. 5 SEM image of (a) aluminum powders particles, (**b**) and particles covered by carbon nanofibers

of the integral intensity confirms formation of multilayer graphene structures that correspond to the structure of carbon fibers.

Preliminary experiments on fabrication of compact materials containing 1 wt.% carbon nanostructures have shown that compact specimens can be obtained by uniaxial pressing at a pressure of 500–600 MPa and sintering at a temperature of 600 °C. The density of the sintered specimens exceeds 98 % of the theoretical value. As compared to the control specimens from aluminum powder with catalyst, the *HV* hardness increases by 20 % at satisfactory ductility. When the specimens obtained are deformed by rolling with reduction of up to 40 % at room temperature, visible defects are absent. These results allow us to expect that the properties of a powder material can be raised considerably by using hot pressing, extrusion and other modern methods of compactation.

The specimens were hot pressed at a temperature of 480, 720 and 980 °C at a pressure of 5 GPa. The dependence of the melting temperature of aluminum on the pressure is described by the formula $dT/dp = 6.41 * 10^{-2}$ [K/MPa] (Babichev et al. 1991), i.e., the melting temperature of aluminum at a pressure of 5 GPa is about 980 °C. The study of the density and of the microstructure of the specimens showed that in all the cases after hot pressing they had a density exceeding 98 % of the theoretical value.

During the hot pressing we detected the parameters of the hot-pressing process, i.e., the heating current, the heating voltage, the pressure in the oil system of the hydraulic press, and the displacement of the press plunger. Detection of the electric current in the pressing process allowed us to judge on the variation of the resistivity of the specimen during the heating (Rudskoy et al. 2015). For example, the resistivity of the specimens with carbon nanofibers changes the most at 980 °C. This result may mean that the carbon of the nanofibers dissolves at this temperature yielding an aluminum carbide and/or an aluminum oxide.

The microstructures of the samples after compacting process were analyzed; their images are presented in Fig. 6. Cu-CNF composite contains finer grains (around 4 μm) with carbon homogeneously dispersed between them. Apparently a good dispersion of CNFs in the matrix prevented the grain growth during sintering.

To determine the effect of the conditions of compactation on the chemical state of the material, we studied compact specimens of Al-CNF by the method of x-ray photoelectron spectroscopy (XPS). It turned out that aluminum was contained in them both in a metallic form and in the form of an Al_2O_3 oxide (Fig. 6b). The results of the XPS show the Al/Al_2O_3 ratio increases somewhat upon growth in the temperature. The XPS spectra of carbon (Fig. 6c) have a peak in the region of 284.8 eV, which indicates that carbon is in the state of *sp*2-hybridization typical for flat nets of graphite, graphene and nanotubes. At 720 and 980 °C we observe formation of an aluminum carbide, which is proved by the presence of a component with bonding energy 281.0–282.5 eV. When the pressing temperature is increased, the content of the carbide grows (Rudskoy et al. 2015).

The x-ray phase analysis has shown that the specimens compacted at 480 and 720 °C exhibit only aluminum peaks. At 980 °C we observe formation of an

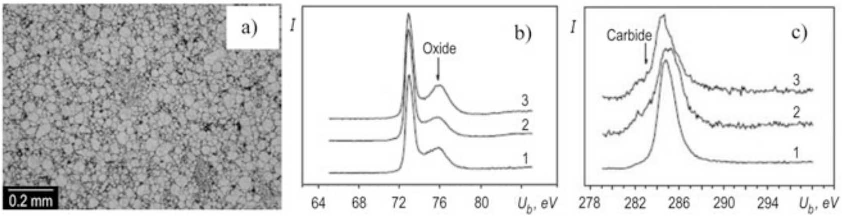


Fig. 6 Typical microstructures of Al-CNF (**a**) and X-ray photoelectron spectra of Al2p2 (**b**) and C1s (**c**) of specimens with 1 wt.% CNF compacted at a pressure of 5 GPa and a temperature of (*I* is the radiation intensity and U_b is the bonding energy): (1) 480 °C; (2) 720 °C; (3) 980 °C

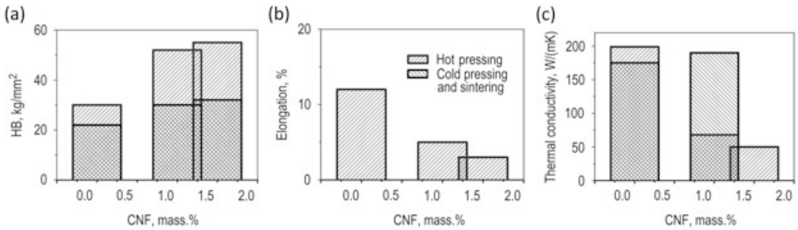


Fig. 7 (**a**) Mechanical and electrical properties of aluminum/CNF composites: hardness (**a**), elongation at the bending test (**b**), and thermal conductivity of composite materials (**c**). Hot pressing temperature is of 720 °C and pressure 5 GPa

aluminum alpha-oxide (Al_2O_3). The presence of crystalline carbide phases has not been detected by the x-ray diffraction analysis.

An aluminum composite material prepared using 1 wt.% CNFs obtained by uniaxial cold pressing and sintering showed 30 % increase in the hardness compared to pure aluminum, whereas the composites prepared by hot pressing showed 80 % increase in the hardness (Fig. 7). Composite materials have satisfactory ductility of about 5 % at the compacting temperature of 720 °C. Thus, the aluminum based material reinforced with carbon nanostructures should be appropriate for creating high-strength and light compacts for aerospace and automotive applications and power engineering.

In contrast to the aluminum-CNF specimens obtained by cold pressing and sintering, the hot-pressed specimens have a quite low ductility. It should be noted that all the specimens with maximum hardness pressed at 480 °C have a virtually zero ductility, which may be caused by virtual absence of bonding on the aluminum–carbon interfaces. The minimum values of the thermal conductivity of such specimens show the same. The elongation of the specimens with 1 wt.% CNF pressed at 720 °C and tested for bending attains 5 % at a hardness of 54 *HB*; that of the aluminum specimens without CNF attains 12 %. When the pressing temperature is increased to 980 °C and the carbon content in the material exceeds 1 wt.%, the ductility decreases. The study of fracture surfaces has shown that in all the cases the specimens fractured over grain boundaries, but at a pressing temperature of 720 °C the aluminum particles were deformed locally, which indicates formation of aluminum–carbon bonds.

The thermal conductivity of all the specimens independently of their composition grows with the pressing temperature. We have mentioned that the melting temperature of aluminum at 5 GPa is about 980 °C. For example, the thermal conductivity of the specimens fabricated by sintering of aluminum powder at 980 °C is 217 W/(m*K), i.e., it virtually corresponds to the thermal conductivity of high-purity aluminum [237 W/(m*K)] at 300 K (Babichev et al. 1991), i.e., at a temperature much lower than that of melting the specimens have a high thermal conductivity [200 W/(m*K)], which is primarily a result of the plastic deformation of aluminum under a high-pressure treatment and of the low residual porosity of the hot-pressed material.

The thermal conductivity of the aluminum-CNF composites obtained by hot pressing is substantially lower than that of the specimens of aluminum-CNF obtained by cold pressing and sintering, i.e., 150–190 W/(m*K). The low values of the thermal conductivity of the aluminum-CNF composites amounting to 25–100 W/(m*K) are connected with the thermal barrier on the filler/matrix interface. The heat resistance of the interfaces grows due to formation of aluminum carbide and aluminum oxide, because the thermal conductivity of most metal carbides is much lower than that of pure metals and ranges within 10–40 W/(m*K) and that of the aluminum oxide is 30 W/(m*K) (Babichev et al. 1991). The specimens fabricated by cold uniaxial pressing followed by sintering have a low hardness (up to 32 *HB*), but substantially better thermal conductivity than the hot-pressed specimens. On the whole, we should note that when the hot-pressing temperature is increased to 720 °C, the density of the specimens increases and so does the cohesion between the carbon nanostructures and the matrix. This raises the thermal conductivity and the ductility of the material due to formation of an aluminum carbide on the aluminum-CNF interface. When the temperature of the hot pressing is increased to 980 °C, the relative content of the Al_4C_3 aluminum carbide increases, and an alpha-Al_2O_3 crystalline phase forms. Depending on the carbon content in the range of 0–1.5 wt.% the hardness of the specimens increases from 30 to 57 *HB*, and the grain size in the structure decreases. The 'aluminum–carbon nanostructures' material has the highest ductility at 1 wt.% CNF and a pressing temperature of 720 °C.

4 Conclusions

The study of microstructures of the samples after compacting process shows that Cu-CNF composite contains finer grains (around 4 μm) with carbon homogeneously dispersed between them. Apparently a good dispersion of CNFs in the matrix prevented the grain growth during sintering. Experiments on fabrication of aluminum compact materials containing 1 wt.% carbon nanostructures by cold pressing and sintering shows that the density of the sintered specimens exceeds 98 % of the theoretical value.

Comparative analysis of the hardness of different Cu/C composite materials was performed. Cu/graphene composite showed a 39 % increase in hardness compared to pure copper, while 10 % and 70 % improvements were found for Cu/graphene and Cu/CNFs composites, respectively. An aluminum composite material prepared using 1 wt.% CNFs obtained by uniaxial cold pressing and sintering showed 30 % increase in the hardness compared to pure aluminum, whereas the composites prepared by hot pressing showed 80 % increase in the hardness. Composite materials have satisfactory ductility of about 5 % at the compacting temperature of 720 °C. Thus, the aluminum based material reinforced with carbon nanostructures should be appropriate for creating high-strength and light compacts for aerospace and automotive applications and power engineering.

Finally, we can conclude that carbon nanomaterials were successfully grown on the surface of copper and aluminum metallic powders. Investigations of the physical properties of the composite materials based on these carbon-modified particles revealed enhanced mechanical and electrical properties. This study shows the way to solve one of the largest problem to creating strong, electrically or thermally conductive CNT/CNF or graphene composites, which is an essential benefit as compared with the same materials that are developed in the world.

Acknowledgments This work was supported by Contract No. 14.Z50.31.0018 with the Ministry of Education and Science of the Russian Federation.

References

Babichev, A. P., Babushkina, N. A., & Bratkovskii, A. M. (1991). *Physical quantities. A handbook [in Russian]*. Moscow: Énergoatomizdat.

Calvert, P. (1999). Nanotube composites: A recipe for strength. *Nature, 399*, 210–211. doi:10.1038/20326.

Dresselhaus, M. S., Dresselhaus, G., & Eklund, P. C. (1996). *Science of fullerenes and nanotubes*. San Diego: Academic Press.

Kidalov, S. V., Shakhov, F. M., Davidenko, V. M., Yashin, V. A., Bogomazov, I. E., & Vul', A. Y. (2008a). Static synthesis of microdiamonds from a charge containing nanodiamonds. *Technical Physics Letters, 34*, 640–642. doi:10.1134/S106378500808004X.

Kidalov, S. V., Shakhov, F. M., Davidenko, V. M., Yashin, V. A., Bogomazov, I. E., & Vul', A. Y. (2008b). Effect of carbon materials on the graphite – diamond phase transition at high pressures and temperatures. *Physics of the Solid State, 50*, 981–985. doi:10.1134/S1063783408050302.

Koltsova, T., Larionova, T., Fadin, Y., & Tolochko, O. (2015). Copper-based composite materials reinforced with carbon nanostructures. *Materials Science (Medziagotyra), 21*, 364–368. doi:10.5755/j01.ms.21.3.7348.

Larionova, T. V., Koltsova, T. S., Fadin, Y., & Tolochko, O. V. (2014). Friction and wear of copper–carbon nanofibers compact composites prepared by chemical vapor deposition. *Wear, 319*, 118–122. doi:10.1016/j.wear.2014.07.020.

Moustafa, S. F., El-Badry, S. A., Sanad, A. M., & Kieback, B. (2002). Friction and wear of copper–graphite composites made with cu-coated and uncoated graphite powders. *Wear, 253*, 699–710. doi:10.1016/S0043-1648(02)00038-8.

Nasibulin, A. G., Koltsova, T. S., Nasibulina, L. I., Anoshkin, I. V., Semencha, A. V., & Tolochko, O. V. (2013). A novel approach to composite preparation by direct synthesis of carbon nanomaterial on matrix or filler particles. *Acta Materialia, 61*, 1862–1871. doi:10.1016/j.actamat.2012.12.007.

Nasibulina, L. I., Koltsova, T. S., Joentakanen, T., Nasibulin, A. G., Tolochko, O. V., Malm, J. E. M., et al. (2010). Direct synthesis of carbon nanofibers on the surface of copper powder. *Carbon, 48*, 4559–4562. doi:10.1016/j.carbon.2010.07.028.

Rakov, E. G. (2006). *Nanotubes and fullerenes [in Russian]*. Moscow: Logos.

Reich, S., Thomsen, C., & Maultzsch, J. (2004). *Carbon nanotubes*. Weinheim: Wiley-VCH.

Riggs, J. E., Guo, Z. X., Carroll, D. L., & Sun, Y. P. (2000). Strong luminescence of solubilized carbon nanotubes. *Journal of the American Chemical Society, 122*, 5879–5880. doi:10.1021/ja9942282.

Rudskoy, A. I., Koltsova, T. S., Shakhov, F. M., Tolochko, O. V., & Mikhailov, V. G. (2015). Effect of hot pressing modes on the structure and properties of an 'aluminum – carbon nanofibers' composite material. *Metal Science and Heat Treatment, 56*, 525–530. doi:10.1007/s11041-015-9793-6.

Rudskoy, A. I., Tolochko, O. V., Kol'tsova, T. S., & Nasibulin, A. G. (2014). Synthesis of carbon nanofibers on the surface of particles of aluminum powder. *Metal Science and Heat Treatment, 55*, 564–568. doi:10.1007/s11041-014-9670-8.

Sarmadi, H., Kokabi, A. H., & Seyed Reihani, S. M. (2013). Friction and wear performance of copper–graphite surface composites fabricated by Friction Stir Processing (FSP). *Wear, 304*, 1–12. doi:10.1007/s11041-015-9793-6.

Zhang, X., Li, Q., Holesinger, T. G., Arendt, P. N., Huang, J., Kirven, P. D., et al. (2007). Ultrastrong, stiff, and lightweight carbon-nanotube fibers. *Advanced Materials, 19*, 4198–4201. doi:10.1002/adma.200700776.

Additive Technologies - The Basis of Digital Custom Manufacturing

Anatoly A. Popovich, Vadim Sh. Sufiiarov, and Alexey V. Grigoriev

Abstract Additive manufacturing technologies, also known as 3D printing, have demonstrated a tremendous growth for the past 30 years, since the development of the first polymer machines to manufacturing functional metal parts with advanced characteristics and 3D-bioprinting. This study presents and describes the developed approach for the production of an individual designed hip prosthesis based on titanium alloy using additive manufacturing technologies. For the hip prosthesis manufacturing it was applied the selective laser melting technology, and titanium alloy powders containing 6 % aluminum and 4 % vanadium as selected raw materials produced by gas and plasma atomization technologies. It is presented the influence of selective laser melting modes on the quality of samples. Evaluation of the microstructure and phase composition of the compacted material before and after heat treatment have been made. It is showed the way, in accordance with the conception of fully digital production, for creation and manufacturing of customized products.

Keywords Additive manufacturing • Digital manufacturing • Powder metallurgy • Titanium alloy • 3D printing • Hip prosthesis • Implant

1 Introduction

Additive Manufacturing technologies (AM) have received a significant interest due to their ability to produce in short times complex geometries without tooling, using a wide range of materials. A huge interest in AM is demonstrated by academia as well as automotive, aeronautical, medicine and other industries. Manufacturing of parts is carried out directly from Computer-Aided Design data, creating cross-sectional layers of an object, and operating in fully digital environment, what

A.A. Popovich, Dr.Sc. • V.Sh. Sufiiarov, Ph.D. (✉)
Peter the Great Saint-Petersburg Polytechnic University, 29 Polytechnicheskaya St., Saint-Petersburg, Russia
e-mail: vadim.spbstu@yandex.ru

A.V. Grigoriev
JSC “Klimov”, 11 Kantemirovskaya St., Saint-Petersburg, Russia

T. Devezas et al. (eds.), *Industry 4.0*, Studies on Entrepreneurship, Structural Change and Industrial Dynamics, DOI 10.1007/978-3-319-49604-7_11

gives rise to a higher rate of automation, design freedom and shorter time-to-market. The process starts with a 3D model, which is sliced into layers by special software. Each layer is created by selective deposition of some material, which is bonded to other material (by fusing, gluing or other methods). Repeating of the treatment of layers leads to forming a solid 3D object. One of the promising AM technologies for manufacturing metal parts is Selective Laser Melting (SLM). SLM uses metal powder as a raw material, which is spread with a thin layer to form a powder bed, and laser beam for selective fusing of powder bed areas.

Selective laser melting has been already successful applied for producing samples from steels (Averyanova et al. 2012; Holzweissig et al. 2015), nickel alloys (Mumtaz and Hopkinson 2009; Sufiiarov et al. 2015), and aluminium alloys (Louvis et al. 2011).

One of the great advantage of additive manufacturing technologies, in particular selective laser melting, is the near absence of limitations on shape of produced parts. The ability to produce complex-shaped parts with such thin wall elements, as cellular and lattice structures, opens a new page in production technologies and makes it possible to design and manufacture products with the geometry that would be impossible to produce by conventional techniques. Owing to full melting of powder material, manufactured by SLM samples have a high relative density (close to 100 %), and also presence rapid melting and solidification of thin layers (20–100 μm), that is typical for this technology (Mercelis and Kruth 2006), leads to formation of fine-dispersed microstructure and high mechanical properties (Sun et al. 2013).

Titanium alloys, and Ti-6Al-4V in particular, are widely used in different industries. One of the application of Ti-6Al-4V alloy is manufacturing of medical implants due to its high biocompatibility and a combination of good mechanical properties (Niinomi 2003).

SLM technology allows manufacturing complex-shaped individual implants from titanium alloys for each specific patient with high mechanical properties (Yadroitsev et al. 2014; Sallica-Leva et al. 2016). Using the advantages of AM it is possible to make textured surfaces in implants for improving osseointegration.

AM technologies have the potential to significantly change traditional ways of production in terms of industrial mechanical equipment, assembly processes, supply chains and parts design. The possibility of mass digital manufacturing of customized parts is the main characteristic of AM; this fact, coupled with the reality of the Internet of Things and the Internet of Services, certainly will make it as one of the key technologies of “Industry 4.0” (Lasi et al. 2014; Rüßmann et al. 2015; Brettel et al. 2016).

2 Materials and Methods

Titanium Ti-6Al-4V alloy powders produced by gas atomization and plasma atomization were used as raw materials for investigation of selective laser melting process.

Particle size distribution measurements were determined by laser diffraction method on Analysette 22 NanoTec plus with a total measuring range of 0.01–2000 μ, equipped with three semiconductor lasers, fully meeting the requirements of ISO 13320 "Particle size analysis—Laser diffraction method". Ultrasonic radiation was used in the measurement process, allowing to split small particles conglomerates. The measurements were conducted three times for each powder, and the results of these measurements were averaged.

Investigations of particle shape and morphology of surfaces were performed on scanning electron microscope (SEM) JEOL 6060, operating at magnifications of $4–10^6$, accelerating voltage of 200 V–30 kV.

The study of selective laser melting process was conducted on SLM280HL system (SLM Solutions GmbH), with YLR-Faser-Laser power up to 400 W, scan speed up to 15 m/s.

X-ray diffractograms of samples were obtained with diffractometer Bruker D8 Advance on CuKα radiation.

Microstructure studies were performed using Leica DMI 5000 light microscope. Mechanical tests were carried out on Zwick/Roell—Z100 machine. The heat treatments were carried out using a vacuum furnace ALD MonoTherm in vacuum with pressure of $10^{-3}–10^{-4}$ mbar.

3 Results and Discussion

3.1 Powder Characterization

In the study two powders of Ti-6Al-4V alloy were investigated. The first one was produced by gas atomization, which is the common technology of powder production for additive manufacturing. The other one was produced by plasma atomization. Plasma atomization is a relatively new process, which was developed for production of high purity powders of reactive metals and alloys with high melting point such as titanium, zirconium, tantalum etc.

The particles of gas atomized powder have mostly spherical form (see Fig. 1).

At the side of some particles we can find small satellites, formed by collisions of particles with molten metal droplets during atomization process. The measurement of particle size distribution of gas atomized Ti-6Al-4V showed that this powder has the following parameters: $d_{10} = 27$ μm; $d_{50} = 47$ μm; $d_{90} = 76$ μm. The surfaces of the particles are smooth, but some particles have surface cavities caused by crystallization process.

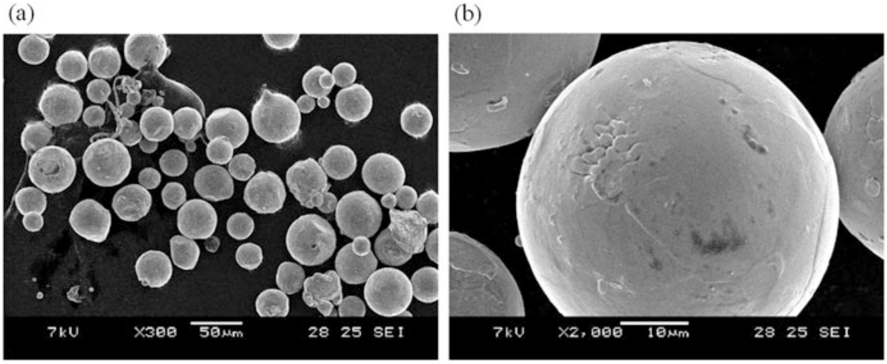


Fig. 1 SEM images of Ti-6Al-4V powder particles produced by gas atomization at (**a**) 300× and (**b**) 1000× magnifications

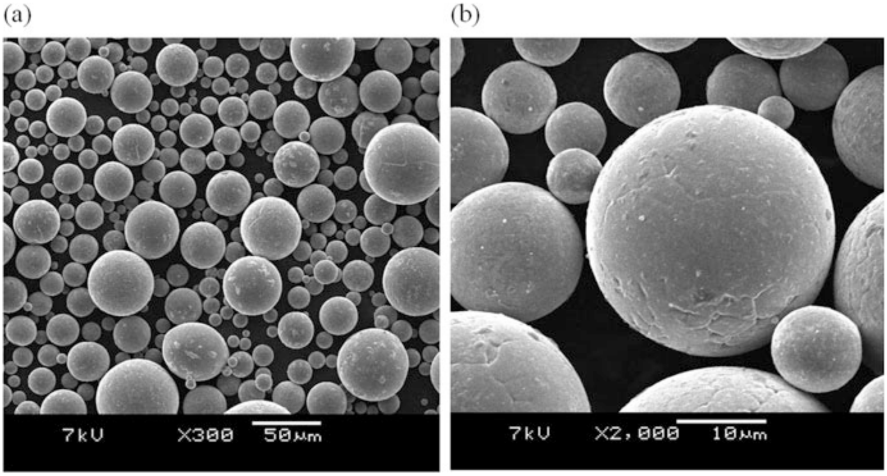


Fig. 2 SEM images of Ti-6Al-4V powder particles produced by plasma atomization at (**a**) 300× and (**b**) 1000× magnifications

Figure 2 shows SEM images of Ti-6Al-4V powder produced by plasma atomization. All observed particles have the form of ideal spheres, with smooth surfaces, and the few cavities are shallower. This powder presents the following particle size distribution: $d_{10} = 14$ μm; $d_{50} = 33$ μm; $d_{90} = 70$ μm.

Plasma atomized powder have more spherical particle form and more quantity of small particles (less 70 μm) in comparison with Ti-6Al-4V powder produced by gas atomization.

Measurements of flowability and apparent density of powders showed that powders have the same value of flowability 22.1 s/50g, but because of differences in particle size distribution gas atomized powder have lower apparent density 2.39 g/sm^3 compare to plasma atomized one 2.72 g/sm^3.

Table 1 Mechanical properties of Ti-6Al-4V after SLM

Specimen	Tensile strength (MPa)	Yield strength (MPa)	Elongation at break (%)
SLM with gas atomized powder	1110–1270	1080–1200	2.5–3.8
SLM with plasma atomized powder	1180–1210	1100–1170	2.6–3.5

3.2 *Compact Material Research*

The detailed description of selective laser melting process and the investigation of influence of main parameters (laser power, scanning speed, hatch distance) on the quality of produced parts are presented in a previous work (Sufiiarov et al. 2015). In this work, for the production of the specimens, as well as the hip prosthesis, we used the parameter set that provides the highest relative density of Ti-6Al-4V (~99.9 %), as described in work (Popovich et al. 2015).

To determine the mechanical properties of compact material after selective laser melting there were made blanks, which were then machined to the size of standard specimens for mechanical tests. The results of the mechanical tests of specimens from Ti-6Al-4V alloy powders produced by different technologies are presented in Table 1. The mechanical properties of specimens made of different powders have similar values, but specimens made from plasma atomized powder exhibit less variation in strength and ductility characteristics, i.e. present more stable results.

Further researches were made with Ti-6Al-4V plasma atomized powders. Two types of heat treatment (HT) in vacuum were applied to Ti-6Al-4V material after SLM using different temperatures: 950 °C for 1.5 h and 800 °C for 4 h. Mechanical tests were made with specimens after the heat treatments. The results of our tests with comparison with mechanical properties of specimens produced by other technologies (Niinomi 2003) and requirements of standards are presented in Table 2.

The tensile properties of SLM specimens after heat treatments meet the requirements of ASTM F2924 for additively manufactured Ti-6Al-4V alloy and ISO 5832-3 “Implants for surgery from Titanium 6-Aluminium 4-Vanadium alloy”. But it is important to point out that some specimens after heat treatment at 800 °C showed elongation at break of less than 10 %, whereas all specimens after heat treatment at 950 °C have elongation at break of more than 10 %.

The microstructures of the compacted materials before and after heat treatment are shown in Fig. 3.

Before heat treatment the microstructure consists of fine needle-like metastable martensitic α'-phase (see Fig. 3a, b), which is the result of rapid solidification and high cooling rates during the SLM. Martensitic α'-phase is hard but brittle, therefore

Table 2 Mechanical properties of Ti-6Al-4V after SLM and subsequent heat treatment

Specimen	Tensile strength (MPa)	Yield strength (MPa)	Elongation at break (%)
SLM HT 800 °C, 4 h	1070–1090	957–1005	9.8–10.3
SLM HT 950 °C, 1.5 h	1080–1088	945–1012	10.4–10.7
Casting	976	847	5.1
Superplastic forming	954	729	10
ASTM F2924	≥825	≥895	≥6–10
ISO 5832-3	≥860	≥780	≥8–10

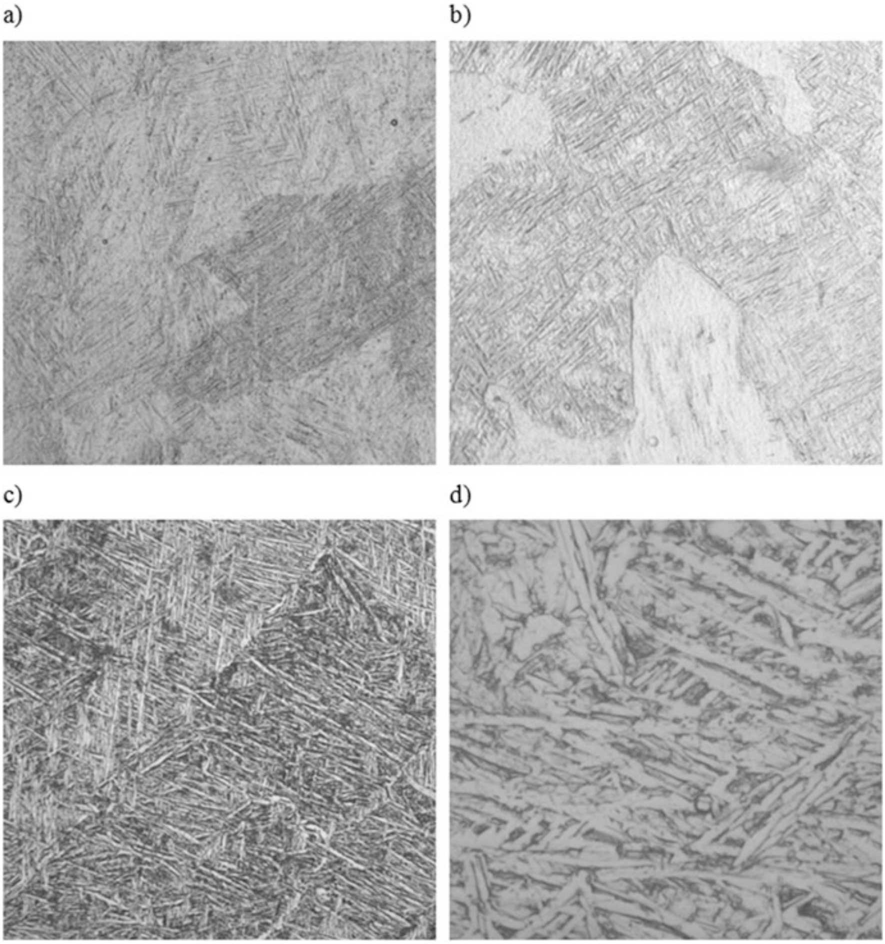


Fig. 3 Microstructure of Ti-6Al-4V alloy produced by SLM before (**a**, **b**) and after heat treatment (**c**, **d**)

compacted material after SLM shows high tensile strength, but low elongation at break (see Table 1).

After heat treatment occurs the partial decomposition of martensitic α'-phase into α- and β-phases, with enlargement of α'-phase needles and formation of β-phase on boundaries of grains and martensitic needles. The heat treatment increases the elongation at break and slightly decreases the tensile strength.

3.3 Manufacturing of Hip Prosthesis

After the investigations of the powder materials described above on selective laser melting processes, as well as the determination of microstructure and mechanical properties of specimens, the next stage of the work was manufacturing of a hip prosthesis.

Very recently it was developed at Saint Petersburg Polytechnic University a novel approach for production of individual designed hip prostheses from titanium alloys using additive technologies. The approach was carried out in accordance with the conception of a fully digital production and creation of customized products.

The basic points of the developed approach are:

- formation of a computer model of a pelvis, based on the patient's computer tomography (CT) data;
- designing of the future prosthesis taking in account structural features of the pelvis and possibilities of AM to produce complex shape elements for improving an accretion of the prosthesis with bone tissue;
- manufacturing with polymer additive technologies of pelvis models and developed hip prosthesis to provide surgeon possibility to simulate surgery;
- manufacturing of metal hip prosthesis by selective laser melting of titanium alloy powder.

CT-data of the patient's pelvis structure was used to make a physical model of the patient's pelvis, which was manufactured by Selective Laser Sintering of polyamide powder. After making a polymer model of the patient's bone it was developed the design of a prosthesis configuration. During the development of the future hip prosthesis it was took into account anatomical features of the patient and implemented partial texturing of the prosthesis surface.

The CAD-model of the prosthesis was positioned relative to the building platform of the SLM machine using 'Materialise Magics' software. Support structures were also created to preserve the prosthesis geometry during the manufacturing process (see Fig. 4a). The prosthesis orientation was chosen to minimize thermal stresses in the manufacturing process and to reduce the number of support structures. Then the custom-developed individual hip prosthesis was manufactured by SLM out of Ti-6Al-4V powder (see Fig. 4b).

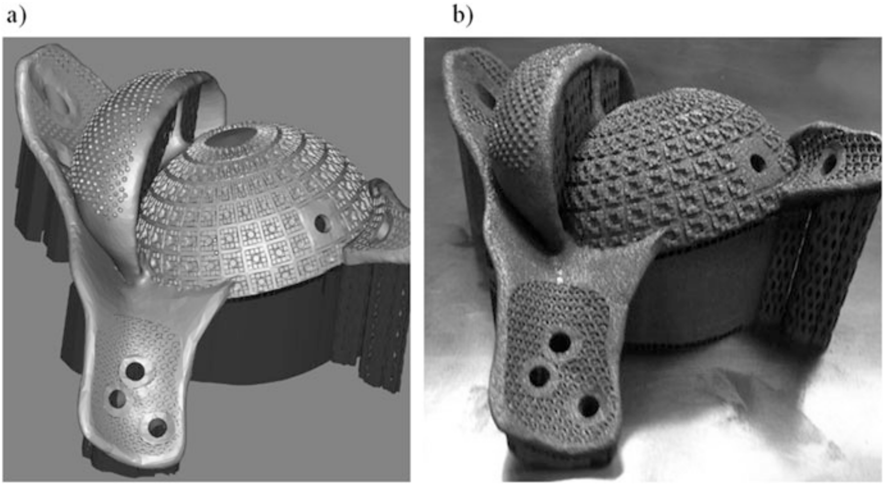


Fig. 4 The Ti-6Al-4V hip prosthesis: (**a**) computer model with support structures; (**b**) hip prosthesis manufactured by SLM from Ti-6Al-4V

After manufacturing by selective laser melting, the prosthesis with the building platform were heat treated in vacuum furnace, then the prosthesis was demounted from the building platform. Figure 5 shows pictures of the demounted prosthesis.

At Fig. 6 it is showed a close view of the special texturized surface of the produced prosthesis.

Additional preparation and cleaning of the prosthesis were made. This prosthesis was inserted in a patient during a surgery in autumn of 2015. Postoperative supervision has shown a good result; the patient can move with the installed prosthesis.

3.4 Outlook on Future of AM

The study of new approaches to the design of products that cannot be produced using conventional production methods has further increased the interest in AM. The most of traditional approaches of designing are performed based upon knowledge of known and conventional industrial production processes and their limitations. Possibilities of layer-based synthesis lead to the emergence of new approaches for designing, including design optimization of parts based on technological advantages of AM.

Intensive development of additive technologies challenges software developers in terms of capabilities to design new products, which take into account technological features of additive technologies. It is not a trivial challenge because additive technologies imposed severe demands on quality of computer models and a huge number of non-linear surfaces (such as in cellular structures) requires

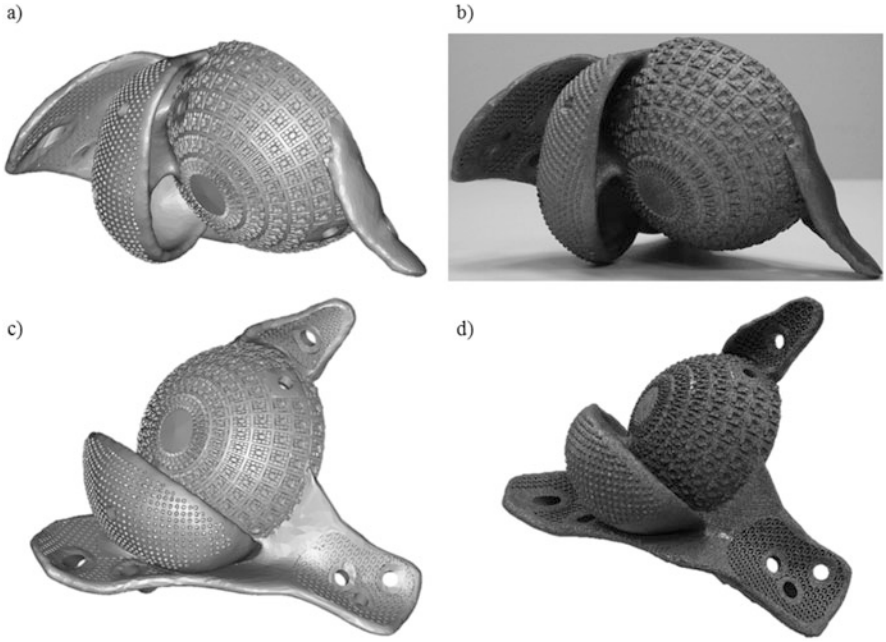


Fig. 5 Computer model of an individual hip prosthesis at different angles of view (**a**, **c**) and manufactured by SLM from Ti-6Al-4V metal hip prosthesis (**b**, **d**)

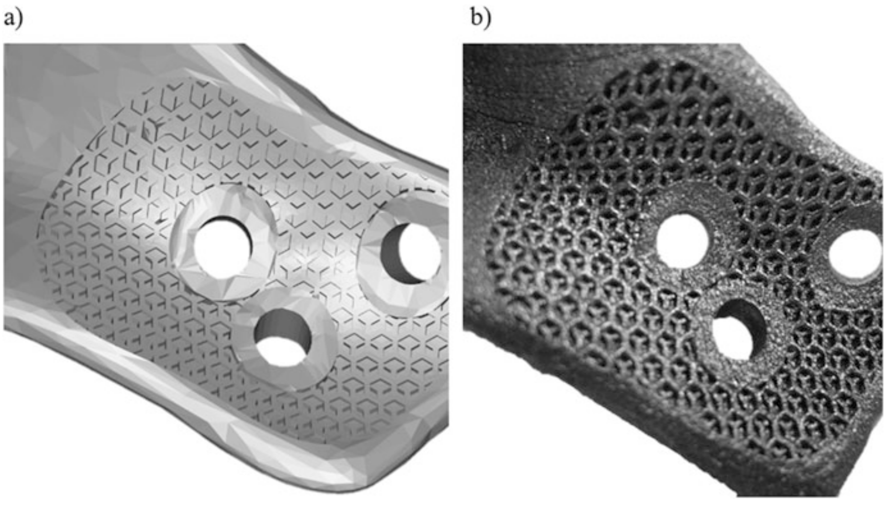


Fig. 6 Close view on texturized surface of the hip prosthesis: (**a**) computer model; (**b**) metal part

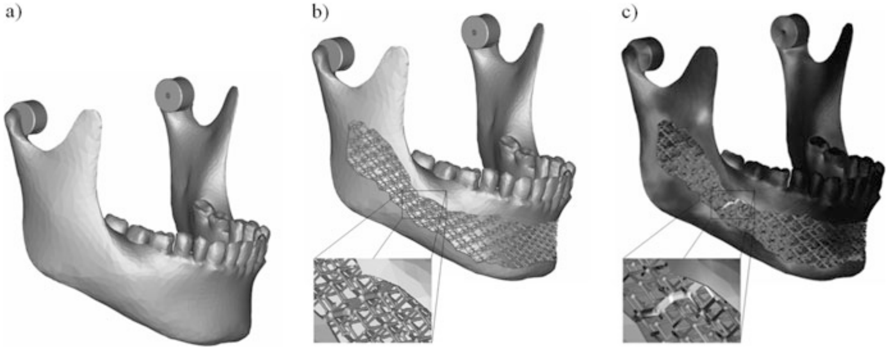


Fig. 7 Example of lightweight lower jaw in Autodesk Nastran, which provides the necessary strength and improved osseointegration process: (**a**) initial model; (**b**) lightweight model; (**c**) stress distribution calculation results of lightweight model

high computing power. The optimization and development of new software that provide the requirements of designing and analysis for additive technologies constitutes to date one of most challenging tasks of software developers. One of the world's leaders in CAx systems market, Autodesk Inc, has released recently one of its new product software (Autodesk Nastran). The main functionality of this software is to facilitate the design by creating lattice structures while maintaining the required strength characteristics. Figure 7 shows the computer model of a lower jaw designed with Autodesk Nastran considering the stress distribution calculation results.

As can be seen in Fig. 7b the lattice structures that fill the volume of bone have different sized lattices in different areas and are determined by the stress distribution results aiming to satisfy the needed strength and stiffness (Fig. 7c).

Today, we can safely say that additive technologies are a real breakthrough and, certainly, may and have to be one of the basis of "Industry 4.0" as element of digital custom manufacturing, because they have the needed flexibility, rapid manufacturing conversion of one type parts to another, high level of informatization, and ability to interact within the conception of Internet of Things and Internet of Services.

4 Conclusions

In this chapter are presented the results of a study of titanium alloy containing 6 % aluminum and 4 % vanadium powder materials produced by gas and plasma atomization technologies. It was investigated the mechanical properties and microstructure of specimens manufactured by selective laser melting of titanium alloy powder produced by gas and plasma atomization. Selective laser melting technology is based on melting of a thin layer of powdered material by brief exposure to laser radiation, a process that is characterized by rapid melting and solidification. Solidification and subsequent quenching in solid state formed a non-equilibrium highly dispersed microstructure mainly composed of titanium-martensite. Subsequent heat treatment leads to the decomposition of titanium-martensite to α- and β-phase. Accordingly, the mechanical properties are changed: before heat treatment the material shows high mechanical properties, but low ductility; after heat treatment the strength properties are reduced, but the ductility increases.

The chapter also describes the developed approach for manufacturing individual designed hip prosthesis by selective laser melting of plasma atomized Ti-6Al-4V powder. The hip prosthesis has special texturized surfaces for better osseointegration. This prosthesis was successfully inserted in a patient during a surgery in autumn of 2015. Postoperative supervision has shown a good result; the patient can move freely with the installed prosthesis.

At example of manufacturing hip prosthesis, it can be say that additive technologies may and have to be one of the basis of "Industry 4.0" as element of digital custom manufacturing and have great potential for further developing.

References

Averyanova, M., Cicala, E., Bertrand, P., & Grevey, D. (2012). Experimental design approach to optimize selective laser melting of martensitic 17-4 PH powder: Part I – Single laser tracks and first layer. *Rapid Prototyping Journal, 18*(1), 28–37.

Brettel, M., Klein, M., & Friederichsen, N. (2016). The relevance of manufacturing flexibility in the context of industrie 4.0. *Procedia CIRP, 41*, 105–110.

Holzweissig, M. J., Taube, A., Brenne, F., Schaper, M., & Niendorf, T. (2015). Microstructural characterization and mechanical performance of hot work tool steel processed by selective laser melting. *Metallurgical and Materials Transactions B, 46*(2), 545–549.

Lasi, H., Fettke, P., Kemper, H. G., Feld, T., & Hoffmann, M. (2014). Industry 4.0. *Business & Information Systems Engineering, 6*(4), 239–242.

Louvis, E., Fox, P., & Sutcliffe, C. J. (2011). Selective laser melting of aluminium components. *Journal of Materials Processing Technology, 211*(2), 275–284.

Mercelis, P., & Kruth, J.-P. (2006). Residual stresses in selective laser sintering and selective laser melting. *Rapid Prototyping Journal, 12*(5), 254–265.

Mumtaz, K., & Hopkinson, N. (2009). Top surface and side roughness of Inconel 625 parts processed using selective laser melting. *Rapid Prototyping Journal, 15*(2), 96–103.

Niinomi, M. (2003). Recent research and development in titanium alloys for biomedical applications and healthcare goods. *Science and Technology of Advanced Materials, 4*(5), 445–454.

Popovich, A., Sufiiarov, V., Polozov, I., & Borisov, E. (2015). Microstructure and mechanical properties of Ti-6AL-4V manufactured by SLM. *Key Engineering Materials, 651–653*, 677–682.

Rüßmann, M., Lorenz, M., Gerbert, P., Waldner, M., Justus, J., Engel, P., et al. (2015). Industry 4.0: The future of productivity and growth in manufacturing industries. Boston Consulting Group. http://www.inovasyon.org/pdf/bcg.perspectives_Industry.4.0_2015.pdf. Accessed 31 May 2016.

Sallica-Leva, E., Caram, R., Jardini, A. L., & Fogagnolo, J. B. (2016). Ductility improvement due to martensite α' decomposition in porous Ti–6Al–4V parts produced by selective laser melting for orthopedic implants. *Journal of the Mechanical Behavior of Biomedical Materials, 54*, 149–158.

Sufiiarov, V., Popovich, A., Borisov, E., & Polozov, I. (2015). Selective laser melting of heat-resistant Ni-based alloy. *Non-ferrous Metals, 1*, 32–35.

Sun, J., Yang, Y., & Wang, D. (2013). Mechanical properties of a Ti6Al4V porous structure produced by selective laser melting. *Materials & Design, 49*, 545–552.

Yadroitsev, I., Krakhmalev, P., & Yadroitsava, I. (2014). Selective laser melting of Ti6Al4V alloy for biomedical applications: Temperature monitoring and microstructural evolution. *Journal of Alloys and Compounds, 583*, 404–409.

Private Astronautics and Its Role in Space Exploration

Alexander B. Zheleznyakov and Vadim V. Korablev

Abstract The authors of this study refer to as “private astronautics” to those activities which are associated with non-state financing of the development of space capabilities that represent a complete system and not just parts of it, as for example, launch vehicles spacecraft, satellites for applied or scientific purposes, etc. These developments are carried out with the expectation of their further implementation in the commercial market, but are not initiated by government. Today the appearance of “private astronautics” is an accomplished fact. Its role should increase in the near future many times. If in 2015 the contribution of “private astronautics” into world space activities is still negligible and difficult to estimate, then by 2020 we can expect about 15–20 % of the world market, and by 2030—up to 40 %. Only in the case of a serious deterioration in the political international situation can result in the non-realization of this scenario. Any crisis in relations between the leading space powers will certainly affect the development of this emerging field of “private astronautics”. But continuing the present trend it is to expect the favorable development of this sector, which will make the entrepreneurs that are presently risking their capital as major players in the field of space activities.

Keywords Private astronautics • Commercial use • Space activities • Launch services • State funding • Private funding

A.B. Zheleznyakov
Rocket and Space Corporation “Energia”.194064, Tikhoretsky Prospect 21, St. Petersburg, Russia

V.V. Korablev (✉)
Peter the Great St. Petersburg Polytechnic University, 29 Politechnicheskaya St., St. Petersburg 195251, Russia
e-mail: korablev@spbstu.ru

T. Devezas et al. (eds.), *Industry 4.0*, Studies on Entrepreneurship, Structural Change and Industrial Dynamics, DOI 10.1007/978-3-319-49604-7_12

1 Introduction

"Private astronautics" is playing an increasing important role in space exploration in recent years. The appearance and development of "private astronautics" is perhaps one of the most interesting trends these days. The reason for its appearance and so rapid development is the changed international situation and the formation of new relations in the world economy, as well as the rapid information revolution and many other factors influencing directly or indirectly this process. The result of these changes was the substantial increase of investments in the space sector, not only from government but also from private business.

It is also important to note that investments are, mainly, not short term in nature, as it was before, but also with focus on the long term, with payback periods of 10 years or more. This relatively new trend is fundamentally changing the approach to the financing of space activities.

It is quite natural that private business is interested in projects that will quickly generate profits. Therefore, the main resources are currently focused on the large-scale projects, such as: "space tourism, suborbital and orbital; creation of cargo ships, which can provide traffic on the route "Earth–earth orbit–Earth" in the interests of public sector customers; the establishment of similar vehicles to ensure that passenger traffic on the same route in the interests of those same customers; launching services and more (Isaeva 2014; Kuybida 2012; Zheleznyakov 2014).

In the near future, the interests of the private investors in space activities will expand and may go beyond earth orbit, moving into the depths of the Solar system, as for example, to the Moon, Mars, Jupiter moons, and asteroids.

In addition, we should not forget "the role of personality". Mostly thanks to the appearance in this field of such persons as Elon Musk, Charles Branson, Richard Bigelow, and some other, "private astronautics" has received a powerful impulse for its further development. These individuals not only financially provide activities for their companies, organizing the development, manufacture and operation of space vehicles, but they are generators of ideas, often the best in their field.

The results of such symbiosis "ideas and opportunities" are very effective.

1.1 About Terminology

Still, there is no unambiguous interpretation of the concept of "private astronautics" and understanding what kind of activities can and should be attributed to this category. For example, the American space program since its inception to the present time has had also the participation of big private companies. The same can be said of the space programs of the European countries, Canada and Japan. However, unlike the American space program, the space activities in these countries has had and continue to have a significant proportion of the state participation.

At the same time, the Soviet space program was created as an element of state policy and remained so until the collapse of the Soviet Union. The actual Russian space program largely retains features of its predecessor, although some changes have emerged in it, even in spite of tight control by the state of all aspects of such activities (Zheleznyakov 2015, 2016).

A similar situation is characteristic for the Chinese space exploration activities, but we can observe some changes occurring, which may probably intensify in the near future, in relations between the public sector and private entities.

To understand the essence of the term "private astronautics" and give the definition of this concept, as well as to describe it as a new phenomenon, it is necessary to consider the changes occurred in its structure over time, as for example, considering some countries separately. Most appropriate for this purpose is to look at the US space program, where actually the phenomenon emerged in force. The American space program, in the 60 years of history of space exploration, was strongly influenced by political factors, but several changes that happened all along its unfolding was also the result of economic activities.

Space activity includes three main components:

1. Order and financing
2. Production of space technology.
3. Use of the results of space activities.

Analytically, these components can be represented by the following expression:

$$\text{Order} + \text{Funding} \rightarrow \text{Production} \rightarrow \text{Use} \tag{1}$$

In this expression, a lot depends on who can prepare an order, who carries out financing of this order, who is the manufacturer of technical means for its implementation, and who is the consumer of the results of space activities. The authors believe that this expression for space activities characterizes the space program of any country, either private or public.

1.2 Stages of Space Activities

For a more accurate evaluation of expression above and to detect the moment of actual emergence of "private astronautics", let us consider the several stages of space activities. Again, this will be done on the example of the American space program.

The first stage covers the period from mid-1950s to mid-1960s, when happened the most acute confrontation in the international scenario of the two superpowers, the USSR and the USA, representing two opposite socio-political systems: the socialist camp and the capitalist world. The confrontation was played also in the emerging arena of space activities.

This stage was characterized by the fact that the main customer and the main financier, and virtually the only consumer of the results of space activities, was the state. At the same time, the main manufacturer of spacecraft equipment was the private sector. Such distribution of functions is characteristic for all branches of the U.S. economy and looks natural, as a logical and the only possible option.

On the other side there were also exceptions, as for example, customers of a number of experiments aboard the spacecraft were made and financed by private companies, as well as the manufacture of a range of scientific apparatus engaged in government research centers, and technology developed in the manufacture of missiles and satellites. Most of the reached results found applications in other fields of human activities not related to space. However, such exceptions were isolated cases and did not play a special role in the unfolding of space exploration.

The expression (1) for the space activity at this stage can be written as[1]:

$$100\%(\text{state}) \rightarrow 100\%(\text{private sector}) \rightarrow 100\%(\text{state})$$

The second stage covers the period from mid-1960s to mid-1970s. In these years was reached the peak of confrontation of the USSR and the USA, and was also the period when the superpowers were able to overcome their global contradictions and to understand that cooperation in space can bring, at least, some political dividends.

This phase was characterized by the increasing role of the private sector as customer spacecraft. This was due to the fact that the government allowed some large companies to participate in space activities where they could act not only as producers but as customers of such products.

We are talking about the market of telecommunication services. The first commercial spacecraft (satellite)—Intelsat-1—was launched in 1965. Since then, the number of spacecraft owned by a private customer was constantly growing.

Allowing private traders to the market of space telecommunications services, the government, however, retained the strict technical control of any commercial activity in space, placing at the forefront the interests of national security.

In addition, the private sector has been actively expanding the list of products of the ground segment, based on the broad masses of consumers.

The expression (1) for space activities at this stage looks like this:

$$\begin{aligned} 90\%(\text{state})\ 10\%(\text{private sector}) &\rightarrow 100\%(\text{private sector}) \\ &\rightarrow 90\%(\text{state})\ 10\%(\text{private sector}) \end{aligned}$$

The third stage covers the period from mid-1970s to mid-1980s. These years, on the one hand, had a new aggravation of contradictions between the superpowers, but, on the other hand, during the same period, began to expand the space cooperation, involving mainly in-orbit activities.

[1] All numerical estimates were made by the authors.

This stage is characterized by the same factors as for the previous, with the exception of some weakening of technical control by the state of commercial activity in space.

Since changes occurred only in the technical field and did not touch the question of relations between the state and private owners, the expression for space activities has not changed:

$$\begin{aligned}90\%(\text{state})\ 10\%(\text{private sector}) &\rightarrow 100\%(\text{private sector})\\ &\rightarrow 90\%(\text{state})\ 10\%(\text{private sector})\end{aligned}$$

The fourth stage covers the period from mid-1980s to mid-1990s. In these years there were global changes in the international arena (the collapse of the Soviet Union and the socialist camp, the political changes in China, the end of the "Cold war"), reconstructed the economic relations between the two countries, started the information revolution (personal computers, Internet, mobile devices). All this became the impetus for a shift in space activities.

This stage is characterized by a significant expansion of the list of the areas of outer space activities, to which private traders were allowed. This concerns the use of satellites for remote sensing and navigation systems in the ground segment. In addition, private traders gained access to the means of launching payload into orbit. Access to launch services immediately spawned a new field of commercial space activities—the marketing of launch services.

New information technologies have increased the interest of private traders to the new traditional telecommunications systems. This fact immediately led to a significant increase in the number of orbiting commercial spacecraft.

Thus, the expression of the space activities at this stage has undergone significant change:

$$\begin{aligned}85\%(\text{state})\ 15\%(\text{private sector}) &\rightarrow 100\%(\text{private sector})\\ &\rightarrow 60\%(\text{state})\ 40\%(\text{private sector})\end{aligned}$$

The fifth stage covers the period from the mid-1990s to 2003. During these years, there was a certain stabilization of the situation in the international arena, at least with regard to international politics, resulting in a new alignment of tade relations between countries, and modern information technology has become an integral part of everyday life. Some of the difficulties were caused by the economic crisis of 1998; however, the collapse of the world economy was short and small in this period.

This fifth phase was characterized by the further strengthening of the position of private traders and the corresponding extension services related to the space activity, as for instance the navigation market.

First, it concerns the launch services. At the turn of the century private businesses began to invest for the first time in the development of new rockets. Moreover, it was not the case of participating in any government program. In those years, a new launch vehicle was not created, although funding for such work was already available.

Another feature of this period can be called as the birth of “space tourism”, in other words, the interest of the private sector in the organization and implementation of paid space flights for everyone. At the beginning it was, primarily, on sub-orbital flights over the border of atmosphere, for it is easier in technical terms, and cheaper than orbital tourism.

At this stage, the expression of the space activity can be written as follows:

$$\begin{aligned}80\%(\text{state})\ 20\%(\text{private sector}) &\rightarrow 100\%(\text{private sector})\\ &\rightarrow 50\%(\text{state})\ 50\%(\text{private sector})\end{aligned}$$

The sixth stage covers the period from 2003 to the present. In political and economic terms, these years can be considered unstable (“Arab spring”, the cooling of relations between Russia and Western countries, the fall in oil prices, the slowdown of the world economy).

This period is characterized by the failure of NASA regarding a number of activities and the transfer of these activities to private owners. In fact, the American Space Agency has set new “rules of the game” and gave to private owners those areas that were “uninteresting”. This list includes not only applied satellite systems (Earth remote sensing, communications, navigation), but manned flights into earth orbit and delivering cargo on board the International Space Station. NASA left to the private sector the realization of manned flights on earth orbit, conserving however some big projects to explore deep space and other scientific missions (Lindsey 2013; Higginbotham 2013).

As a result, in the United States was observed the active creation of carrier rockets and space ships, which soon was crowned with success—the launching of “Antares” by “Orbital Science Corporation” (now “Orbital ATK”) and “Falcon-9” from “Space X”. The same companies have created the cargo transport spacecraft “Cygnus” and “Dragon”, which began regular flights to the International Space Station. At the same time “Space X” began to promote to the world market its launching services.

Foreseen to 2017, the traders intend to introduce their first manned spacecraft to deliver crews into orbit. As of same year is scheduled to begin suborbital flights organized for “space tourists”.

As a result, by the end of 2015 the expression of space activity can be written as follows:

$$\begin{aligned}70\%(\text{government})\ 30\%(\text{private sector}) &\rightarrow 100\%(\text{private sector})\\ &\rightarrow 40\%(\text{state})\ 60\%(\text{private sector})\end{aligned}$$

It was during the sixth stage that we can really talk about “private astronautics” as such. Ceasing to be a “commercial component”, private space activity has become a separate subject of the whole space endeavor, becoming a strong thriving force and, importantly, constantly improving its own infrastructure.

Thus, the authors see as more correct the following definition of the term “private space flight”: *non-state financed activities for the development of space capabilities that represent a complete system and not just parts of it, and the commercial use of these developments*.

Important to stress that the main feature of this actual (6th) phase of space activities is the fact that costs and revenues are been “cleaned” from the influence of the state. The main objective of “private astronautics” is to reduce the cost of space launching systems.

Similar dynamics of change in the relationship between government and the private sector in the expression for space activities has been observed in Europe, Japan and India. However, in the European countries the contribution of the private sector in the commissioning and financing of works was slightly lower than in the United States. Even to a lesser extent private traders invest in Japan and very few in India. However, the trend for increasing the share of private investment in all these countries is the same.

As already noted, in the USSR, and later in Russia, the state determined, directed, and controlled space activities, and unilaterally dispose of its fruits. This situation has not changed much after the changes in the Russian economy occurred in the early 1990s. Even after becoming joint-stock companies, enterprises of rocket and space industry actually continued to be controlled by the state and owned by the Russian government. Until recently, approximately before 2010, the expression of the space activity for Russia can be written as follows:

$$100\%(\text{state}) \rightarrow 100\%(\text{state}) \rightarrow 100\%(\text{state})$$

Minor exceptions only confirm the general rule.

However, in more recent years there have appeared in Russia some companies that can be considered representatives of the “private astronautics”. Yet they are few, as the number of employees working in them. Consequently, their “contribution” to the Russian space program is still almost imperceptible. It is hoped that in the coming years they will show themselves as full actors in rocket and space activity.

The statements above can be applied to other space powers (China, North and South Korea, Iran). For the time being there is no “private astronautics” in these countries, but the private traders can increase their role in the near future, especially in China, where there are appropriate conditions.

1.3 The Present Status of “Private Astronautics”

According to various estimates, the global space market in 2015 was \$250–350 billion. As shown in Fig. 1 below, approximately 33 % of this volume is the

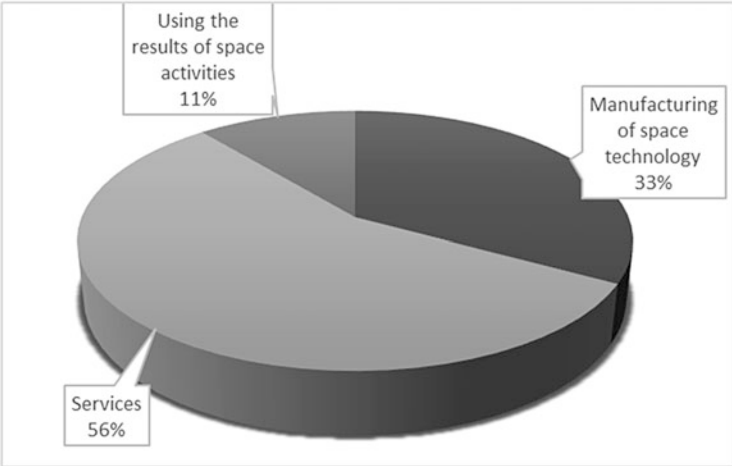


Fig. 1 Shares in 2015 of the global space activity

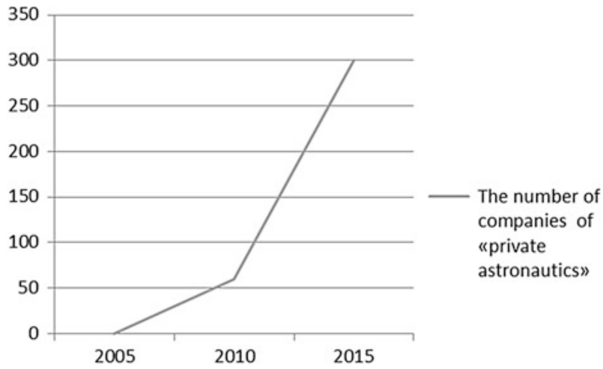
development and production of satellites and launching vehicles, as well as launching services, 56 %, and services (telecommunications, navigation and other) provided by space technology, 11 %—using the results of space activities (implementation of data obtained by satellite remote sensing) (Azarenko and Wokin 2011; Zhiganov et al. 2014; Losev 2015; Losev 2016).

As already mentioned, the authors of this chapter consider the timing of the birth of the phenomenon called "private astronautics" as being the first decade of the twenty-first century, although the prerequisites for its occurrence were recorded a decade earlier. For convenience, we will assess the current status in comparison with 2005–2006.

Ten years ago in the United States, the representatives of the "private astronautics" could include a few companies falling under this definition by the authors. These are "Space X", "Orbital Science Corporation", "Bigelow Aerospace", "Sierra Nevada", engaged in the development of space launching vehicles and spacecraft, and "BDU Origin", "Armadillo", "Virgin Galactic", involved in the development of suborbital rockets and spacecraft, and some few other. Important to stress that all the listed companies continue their activity today, while some other not named companies disappeared from the market. The total number of representatives of the "private astronautics" in those years can be counted in the range 10–15.

The first successes that were achieved by private businesses in the field of launching services, as well as the development of other means of space technology, attracted attention of the industry and initiated the emergence of new groups working on the same principles as major representatives of the "private astronautics". However, the question in this case was not about such a large-scale, science-intensive and costly projects, such as, for example, the development of the launching vehicle "Falcon-9" and "Antares", or "Spaceship Two" and "New

Fig. 2 Growth dynamics in the number of private companies involved in space activities



Shepherd". Their activity is fully covered under the definition of "private astronautics". By 2010–2011 the total number of such companies was estimated as 50–60.

During the last 5 years more and more companies came to the club of "private astronautics", after NASA has given up on the mercy of private traders such directions, as the traffic "Earth—ISS—Earth" and manned flights into Earth orbit. The private investment in these areas increased significantly and significantly increased the number of companies, which, in varying degrees, became engaged in the problems of the "private astronautics".

Currently, the number of such companies, according to some estimates, is ranging from 150 to 300. (Ribeiro et al. 2016). The dynamics of change in the number of representatives of the "private astronautics" can be seen in the chart given in Fig. 2.

Similar dynamics is observed in Europe. The only difference is that the number of companies engaged in the "private astronautics" is an order of magnitude less than in the US.

The number of Russian companies, which can be attributed to this category, is still very low. The most famous is "Dauria Aerospace" engaged in the manufacture of small satellites, "Kosmocar" developing a project on space tourism, and "Lin Industrial" designing ultra-light launching vehicles. Their contribution to the missile and space activities of Russia, as noted, is insignificant (Yakovlev and Farafontova 2013; Nadezhdin 2016).

1.4 Development Trends

After analyzing the current status of "private astronautics", we can assume the following scenarios for its further development.

The first scenario, which can be called "optimistic scenario", implies the normalization of relations between Russia and Western countries, the growth of the global economy, and the return of oil prices to the level of 2008.

In this case, we can predict significant strengthening of space activities and the inclusion of the new space faring states, the expansion of the list of services included in the results of this activity, the creation of new technologies used in space and on Earth, as well as the increase of interest in space by ordinary citizens. All this will inevitably lead to the strengthening of the position of "private astronautics".

It can be assumed that the contribution of "private astronautics" to world space activity in the "optimistic scenario" will increase from 2–3 %[2] in 2015 up to 15–20 % in 2020, and by 2030—up to 35–40 %.

The second scenario, which we call a "common sense option", assumes that current status quo in the international arena will be maintained, with Western countries' sanctions against Russia and Russian retaliatory sanctions against Western countries for at least the next 5 years, low oil price, and the slowdown of the world economy as well as of China.

In this case, space activities will remain at current levels, expanding the list of services from space activities and the creation of new technologies required in the production of missile and space technology. In this scenario can be assumed that some growth will be observed, but will be based only on current needs, with some increase in state structures as the main customer of the results of space activities.

In such a situation, the contribution of "private astronautics" into world space activities will increase, but not as fast as in the "optimistic scenario". It can be assumed that by 2020, its global market share will not exceed 10 %, and by 2030 will amount to no more than 25–30 %.

Finally, the third scenario, which should be called "pessimistic variant", suggests further cooling of the relationship between Russia and Western countries, the stagnation of the world economy, the continuing low price of oil for more than 5 years, and the militarization of near-earth space.

In this case, it is possible to predict the strengthening of the role of state structures, as the main customer and consumer of the results of space activities, and also the strengthening of state control over the "private astronautics" and, as a consequence, reducing the interest of private business in space exploration.

In the "pessimistic scenario", you can expect a slight increase of the role "private astronautics" in the global space activities: up to 10 % in 2020 and up to 15 % in 2030. Figure 3 pictures the unfolding of the three above described space scenarios.

To predict the development of "private astronautics" for the period after 2030 is far more difficult due to the presence of many uncertainties. One can only assume that in the long-term future there will be an increasing contribution of private enterprises to the global space industry, but to give specific metric values to is inappropriate.

[2]"Contribution" of private businesses in space activities varies across countries. Therefore, the assessments made by the authors about the contribution of "private astronautics" in the global space industry are integrated.

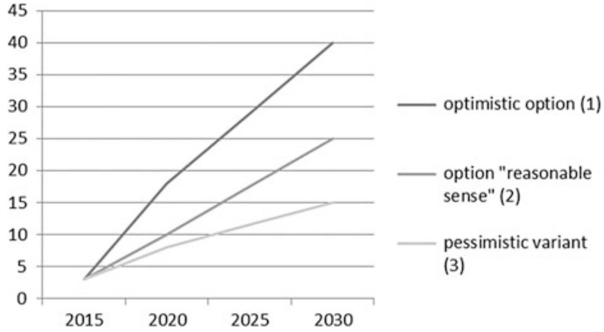


Fig. 3 Unfolding of three possible scenarios for the share (%) of "private astronautics" in the global market of space activities

2 Conclusion

No matter which scenario of the "private astronautics" will happen, optimistic or pessimistic, we can confidently say that its role will increase in the world's space industry in the coming years. The pace of development will depend entirely on the political situation in the world arena, as well as to the state of affairs in the world economy. In any case, "private astronautics" took place as a phenomenon and its role in the exploration of outer space will inevitably increase.

References

Azarenko, L. G., & Wokin, G. G. (2011). Space services: Features of formation of the market and infrastructural support. *Service in Russia and Abroad: Electronic Scientific Journal, 4*(23), 5–12.

Higginbotham, A. (2013). Robert bigelow plans a real estate empire in space. *Business Week*, 6–9.

Isaeva, N. (2014). Private space: What are the prospects of its development and influence on the Russian economy?. *School life.ru*. http://shkolazhizni.ru/world/articles/65996/

Kuybida, A. (2012). Private business requests to declassify the space. *Izvestia*. http://izvestia.ru/news/513491

Lindsey, C. (2013). NewSpace flights in 2013. *NewSpace Watch, 32*(14), 3–7.

Losev, A. (2015). Business is bursting to get into space. *Russian Business Newspaper*.

Losev, A. (2016). Commercial space market. *Electronics*. *Ecoruspace.me resource*. http://www.ecoruspace.me/%D0%9A%D0%BE%D0%BC%D0%BC%D0%B5%D1%80%D1%87%D0%B5%D1%81%D0%BA%D0%B8%D0%B9%20%D0%BA%D0%BE%D1%81%D0%BC%D0%B8%D1%87%D0%B5%D1%81%D0%BA%D0%B8%D0%B9%20%D1%80%D1%8B%D0%BD%D0%BE%D0%BA.html

Nadezhdin, V. (2016). Problems and prospects of the Russian Astronautics. Information-analytical portal of the Union State. http://www.soyuz.by/news/expert/25648.html

Ribeiro, J., Belderrain, M. C., & Devezas, T. (2016). Space exploration: Global analysis of a changing scenario, to be published in Acta Astronautica.

Yakovlev, M. S., & Farafontova, E. L. (2013). Private space companies in Russia. *Actual Problems of Aviation and Astronautics, 2*(9), 307–308.

Zheleznyakov, A. (2014). Space activities around the world in 2013. *Scientific and Technical News of St. Petersburg State Polytechnic University, 2*(195), 7–17.

Zheleznyakov, A. (2015). Space: The results of 2014. *Innovations, 2*, 5–11.
Zheleznyakov, A. (2016). Space: The results of 2015. *Innovations, 2*, 3–6.
Zhiganov, A. N., Zaichko, V. A., & Maksimov, A. V. (2014). Methodical approach to classification of space products and services. *Services in Russia and Abroad, 8*(4), 176–187.

The MANBRIC-Technologies in the Forthcoming Technological Revolution

Leonid Grinin, Anton Grinin, and Andrey Korotayev

Abstract In this chapter, we analyze the relationship between Kondratieff waves and major technological revolutions on the basis of the theory of production principles and production revolutions, and offer some forecasts about the features of the Sixth Kondratieff Wave/the Fourth Industrial Revolution. We show that the technological breakthrough of the Sixth Kondratieff Wave may be interpreted as both the Fourth Industrial Revolution and as the final phase of the Cybernetic Revolution. We assume that the sixth K-wave in the 2030s and 2040s will merge with the final phase of the Cybernetic Revolution (which we call a phase of self-regulating systems). This period will be characterized by the breakthrough in medical technologies which will be capable of combining a number of other technologies into a single system of new and innovative technologies (we denote this system as a system of MANBRIC-technologies—i.e. medical, additive, nano-, bio-, robo-, info-, and cogno-technologies).

Keywords Medical technologies • Additive technologies • Nanotechnologies • Biotechnologies • Robotics • IT • Cognitive technologies

1 Introduction: Cybernetic Revolution, Scientific-Cybernetic Production Principle, the Sixth Kondratieff Wave, and the Fourth Industrial Revolution

The production revolution which began in the 1950s and is still proceeding, has led to a powerful acceleration of scientific and technological progress. Taking into account expected changes in the next 50 years, this revolution deserves to be called

L. Grinin • A. Korotayev (✉)
National Research University Higher School of Economics, Myasnitskaya 20, Moscow 101000, Russia
e-mail: Leonid.Grinin@gmail.com; andrey.korotayev@gmail.com

A. Grinin
International Center for Education, Social Research and Humanities, Basovskaya 16/1, Moscow 109202, Russia
e-mail: antigrinin@yandex.ru

T. Devezas et al. (eds.), *Industry 4.0*, Studies on Entrepreneurship, Structural Change and Industrial Dynamics, DOI 10.1007/978-3-319-49604-7_13

'Cybernetic' (see our explanation below). The initial phase of this revolution (the 1950s—the 1990s) can be referred to as a scientific-information because it was characterized by the transition to scientific methods of planning, forecasting, marketing, logistics, production managements, distribution and circulation of resources, and communication. The most radical changes took place in the sphere of informatics and information technologies. The final phase will begin approximately in the 2030s or the 2040s and will last until the 2070s (note that many contributors to this volume denote this forthcoming technological breakthrough as the "Fourth Industrial Revolution"). We believe that it can be also interpreted as a phase of self-regulating systems of the Cybernetic Revolution (see below). We do not think that these interpretations are mutually exclusive; on the contrary, they are perfectly compatible. Now we are in the intermediate (modernization) phase of the Cybernetic Revolution which will last until the 2030s. It is characterized by powerful improvements and the diffusion of innovations made at the initial phase—in particular by a wide proliferation of easy-to-handle computers, means of communication, and the formation of a macro-sector of services among which information and financial services take center stage. At the same time the innovations necessary to start the final phase of the Cybernetic Revolution are being prepared.

Cybernetic Revolution is a great breakthrough from industrial production to production and services based on the operation of self-regulating systems.

Table 1 demonstrates the connection between three phases of the scientific-cybernetic production principle (which coincide with three phases of the Cybernetic Revolution) and three Kondratieff waves[1] (the fourth, fifth and sixth).

1.1 Peculiarities of the Fourth Kondratieff Wave in Connection with the Beginning of the Cybernetic Revolution

The fourth K-wave (the second half of the 1940s—1980s) fell on the initial phase of the Cybernetic Revolution. The beginning of a new production revolution is a special period which is connected with the fast transition to a more advanced technological component of economy. All accumulated innovations and a large number of new innovations generate a new system that has a real synergetic effect. It would appear reasonable that an upward phase of the K-wave coinciding with the beginning of a production revolution can appear more powerful than A-phases of other K-waves.

[1]See Kondratieff 1926, 1935, 1984, 1998, 2004 [1922]; Schumpeter 1939; Hirooka 2006; Devezas 2006, 2010, 2012; Korotayev et al. 2011; Grinin et al. 2012, 2014, 2016c; Korotayev and Grinin 2012; Grinin and Grinin 2015, 2016 for more detail on Kondratieff waves.

Table 1. The scientific-cybernetic production principle (initial phases) and Kondratieff waves

Phases of the scientific cybernetic production principle	The first phase (initial phase of the Cybernetic Revolution) 1955–1995 ≈ **40 years**	The second phase (middle phase of the Cybernetic Revolution)1995—the 2030s/40s ≈ **35–50 years**	The third phase (final phase of 'self-regulating systems' of the Cybernetic Revolution) the 2030s/40s–2055/70s ≈**25–40 years**	Total: ≈ **100–120** years
K-Waves and their Phases	The Fourth Wave, 1947–1982/1991 ≈ **35–45 years**	The Fifth Wave, 1982/1991–the 2020s. The beginning of the upward phase of the sixth wave (2020–2050s) ≈ **30–40 years**	The sixth wave, 2020–2060/70s. The end of the upward phase and downward phase (the latter ≈ 2050–2060/70s) ≈ **40–50 years**	About**110–120 years**
K-Waves and Their Phases	Upward phase, 1947–1969/1974s	Downward phase of the fifth wave, 2007–2020s	–	
K-Waves and Their Phases	Downward phase, 1969/1974–1982/1991	Upward phase of the sixth wave, 2020–2050s	–	
K-Waves and Their Phases	The fifth wave, 1982/1991–2020s, upward phase, 1982/1991–2007	–	–	

That was the feature of the upswing A-phase of the fourth K-wave (1947–1974) which coincided with the scientific-information phase of the Cybernetic Revolution. As a result, a denser than usual cluster of innovations (in comparison with the second, third and fifth waves) was formed during that period. All this also explains why in the 1950s and 1960s the economic growth rates of the World System were higher, than in the A-phases of the third and fifth K-waves. The downswing phase of the fourth K-wave (the 1970s–1980s) in its turn also fell on the last period of the initial phase of the Cybernetic Revolution. This explains in many respects why this downswing phase was shorter than those of the other K-waves.

1.2 The Fifth K-Wave and the Delay of the New Wave of Innovations

It was expected that the 1990s and the 2000s would bring a radically new wave of innovations, comparable in their revolutionary character with computer technologies, and therefore capable of creating a new technological paradigm. Those directions which had already appeared and those which are supposed to become the basis for the sixth K-wave were considered in position to make a breakthrough. However, it was the development and diversification of already existing digital electronic technologies and rapid development of financial technologies that became the basis for the fifth K-wave. Those innovations which were really created during the fifth K-wave as, for example, energy technologies, still have a small share in the general energy, and, above all, they have not developed properly. Some researchers believe that from 1970s up to the present is the time for the decelerating scientific and technological progress (see a discussion on this topic in Brener 2006; Khaltourina and Korotayev 2007; see also Maddison 2007). Polterovich (2009) also offers the notion of a technological pause. But, in general, the mentioned technological delay is, in our opinion, insufficiently explained. We believe that taking features of the intermediate modernization phase of a production revolution (i.e., the second phase of the production principle) into account can help explain this. Functionally it is less innovative; rather during this phase earlier innovations become more widely spread and are improved.

As regards the 1990s–2020s (the intermediate phase of the Cybernetic Revolution) the question is that the launch of a new innovative breakthrough demands that the developing countries reach the level of the developed ones, and the political component of the world catches up with the economic one (see Grinin and Korotayev 2010, 2014a, b; Korotayev and de Munck 2013, 2014; Korotayev and Zinkina 2014; Korotayev et al. 2011a, b, c, 2012, 2015).

Thus, the delayed introduction of innovations of the new generation is explained, first, by the fact that the center cannot endlessly surpass the periphery in development, that is the gap between developed and developing countries cannot increase all the time. Secondly, the economy cannot constantly surpass the political and other components, as this causes very strong disproportions and deformations. And the appearance of new general-purpose technologies, certainly, would accelerate economic development and increase disparities. Thirdly, introduction and distribution of the new basic technologies do not occur automatically, but only within an appropriate socio-political environment (see Grinin and Grinin 2016; Grinin and Korotayev 2014a, b; see also Perez 2002). In order for basic innovations to be suitable for business, structural changes in political and social spheres are necessary, eventually promoting their synergy and wide implementation in the world of business.

Thus, the delay is caused by difficulties of changing political and social institutions on the regional and even global scale, and also (or, perhaps, first of all) within the international economic institutions. The latter can change only as a result of the strong political will of the main players, which is difficult to execute in the framework of the modern political institutions. These institutions rather can change under the conditions of depressive development (and probable aggravation of the foreign relations) compelling them to reorganize and dismantle conventional institutions that are unlikely to be changed under ordinary conditions due to a lack of courage and opportunities (for our vision of the future of the world order see Grinin and Korotayev 2010, 2015a; Grinin et al. 2016a).

The above explains as well the reasons of different rates of development as regards the center and periphery of the World System during the fifth K-wave (for more detail see Grinin and Korotayev 2010, 2015a; Grinin et al. 2016a). The periphery was expected to catch up with the center due to the faster rates of its development and slowdown of the center development. However, one should not expect continuous crisis-free development of the periphery—a crisis will come later and probably in other forms. Without a slow-down of the development of the periphery and serious changes, full harmonization of the economic and political component will not happen. Consequently, it might be supposed that in the next decade (approximately by 2020–2025) the growth rates of the peripheral economies can also slow down, and internal problems will aggravate; this can stimulate structural changes in the peripheral countries, thus also increasing international tension.

The world order has already begun to change, and it will continue to change over the next 10–20 years and some visible results of this change may appear by the start of the new K-wave. We have called this change "the World System reconfiguration" (see Grinin and Korotayev 2012, 2015a: 159–166; Grinin et al. 2016a, b). Thus, we suppose that in the next 10–15 years the world will face serious and painful changes. The World System reconfiguration processes further explain the reasons for the very turbulent processes observed in the recent years.

2 Characteristics of the Cybernetic Revolution

2.1 What Are Self-regulating Systems and Why Are They So Important?

Self-regulating systems are systems that can regulate themselves, responding in a pre-programmed and intelligent way to the feed-back from the environment. These are the systems that operate either with a small input from human or completely without human intervention. Today there are many self-regulating systems, for example, pilotless electric cars, artificial Earth satellites, drones, navigation systems laying the route for a driver. Another good example is life-support systems (such as medical ventilation apparatuses or artificial hearts). They can control a number of parameters, choose the most suitable mode of operation and detect critical situations. There are also special programs that determine the value of stocks and other securities, react to price changes, buy and sell them, carry out thousands of operations in a day and fix a profit. A great number of self-regulating systems have been created but they are mostly technical and information systems (like robots or computer programs). During the final phase of the Cybernetic Revolution there will be a lot of self-regulating systems connected with biology and bionics, physiology and medicine, agriculture and environment. The number of such systems as well as their complexity and their autonomy will dramatically increase. These systems will also significantly reduce energy and resource consumption. Human life will become organized to a greater extent by such self-regulating systems (for example, by monitoring health, daily regimens, regulating or recommending levels of personal exertion, having control over the patients' condition, prevention of illegal actions, etc. [for more detail about self-regulating systems see Grinin and Grinin 2016]).

Thus, we designate the modern revolution 'Cybernetic,' because its main sense is the wide creation and distribution of self-regulating autonomous systems. Cybernetics, as is well-known, is a science of regulatory systems. Its main principles are quite suitable for the description of self-regulating systems (see, e.g., Wiener 1948; Ashby 1956; Foerster and Zopf 1962; Beer 1967, 1994; Umpleby and Dent 1999).

As a result, the opportunity to control various natural, social and production processes without direct human intervention (that is impossible or extremely limited now) will increase. In the fourth phase (of maturity and expansion) of the scientific cybernetic production principle (the 2070s and 2080s) the achievements of the Cybernetic Revolution will become quite systemic and wide-scale in its final phase (for more detail see Grinin 2006; Grinin and Grinin 2016).

Below we enumerate the most important characteristics and trends of the Cybernetic Revolution and its technologies. These features are closely interconnected and support each other.

2.2 *The Most Important Characteristics and Trends of the Cybernetic Revolution*

1. Increases in the amount of information and complexity in the analysis of the systems (including the ability of systems for independent communication and interaction).
2. Sustainable development of the system of regulation and self-regulation.
3. Mass use of artificial materials which previously lacked the appropriate architectural properties.
4. Qualitatively increasing controllability a) of systems and processes that vary in their constitution (including living material); and b) of new levels of managing the organization of matter (up to sub-atomic levels, as well as the use of tiny particles as building blocks).
5. Miniaturization as a trend of the constantly decreasing size of particles, mechanisms, electronic devices, implants, etc.
6. Resource and energy saving in every sphere.
7. Individualization as one of the most important technological trends.
8. Implementation of smart technologies and a trend toward humanization of their functions (use of a common language, voice, etc.).

2.3 The Most Important Characteristics and Trends of the Cybernetic Revolution

1. The transformation and analysis of information as an essential part of technologies.
2. The increasing connection between the technological systems and environment.
3. A trend toward automation of control is observed together with the increasing level of controllability and self-regulation of systems.
4. The capabilities of materials and technologies to adjust to different objectives and tasks (smart materials and technologies) as well as capabilities for *choosing optimum regimes in the context of certain goals and tasks*.
5. A large-scale synthesis of the materials and characteristics of the systems of different nature (e.g., of animate and inanimate nature).
6. The integration of machinery, equipment and hardware with technology (know-how and knowledge of the process) into a unified technical and technological system.[2]
7. Self-regulating systems (see below) will become the major component of technological processes. That is the reason why the final (forthcoming) phase of the Cybernetic Revolution is (or should) be called **the epoch of self-regulating systems** (see below).

2.4 Medicine as a Sphere of the Initial Technological Breakthrough and the Emergence of MANBRIC-Technology Complex

It is worth remembering that the (first) Industrial Revolution began in a rather narrow area of cotton textile manufacturing and was connected with the solution of quite concrete problems—at first, liquidation of the gap between spinning and weaving, and then, after increasing weavers' productivity, searching for ways to mechanize spinning. However, the solution of these narrow tasks caused an explosion of innovations conditioned by the existence of a large number of the major elements of machine production (including abundant mechanisms, primitive steam-engines, quite a high volume of coal production, etc.) which gave an impulse to the development of the Industrial Revolution. In a similar way, we assume that the Cybernetic Revolution will start first in a certain area.

Given the general vector of scientific achievements and technological development and taking into account that a future breakthrough area should be highly

[2]During the Industrial Epoch these elements existed separately: technologies were preserved on paper or in the engineer's minds. At present, thanks to IT and other technologies the technological constituent fulfils the managing function facilitating the path to the epoch of self-regulating systems.

commercially attractive and have a wide market, we forecast that the final phase of this revolution will begin somewhere at the intersection of medicine and a number of other technologies (we will provide reasons for this statement below). By the 2030s there can appear unique opportunities for a breakthrough in medicine (see below). However, when speaking about medicine, one should keep in mind that with respect to potential revolutionary transformations medicine is a very heterogeneous sphere. That is why the breakthrough will not occur in all spheres of medicine but in its one or two innovative fields. Perhaps, it has already formed (as biomedicine or nanomedicine) or it can form as a result of the introduction of some other innovative technologies into medicine. As for other branches of medicine, revolutionary transformations will begin there later. Moreover, some branches of medicine would be unable to transform due to their conservatism. Thus, more radical reforms will occur in these fields in the future.

In general, the breakthrough vector in medicine and associated branches can be defined as a rapid growth of opportunities for correction or even modification of the human biological nature. In other words, it will be possible to extend our opportunities to alter a human body, perhaps, to some extent, its genome; to widen sharply our opportunities of minimally invasive influence and operations instead of the modern surgical ones; to use extensively means of cultivating separate biological materials, bodies or their parts and elements for regeneration and rehabilitation of an organism, and also artificial analogues of biological material (bodies, receptors), etc.

This will make it possible to radically expand the opportunities to dramatically increase life expectancy and improve physiological abilities of people as well as health-related quality of life (HRQoL). It will be technologies intended for common use in the form of a mass market service. Certainly, it will take a rather long period (about two or three decades) from the first steps in that direction (in the 2030s and 2040s) to their common use (about some possible forthcoming medical technologies see Appendix).

On the whole, the drivers of the final phase of the Cybernetic Revolution will be medical technologies, additive manufacturing (3D printers), nano- and bio-technologies, robotics, IT, cognitive sciences, which will together form a sophisticated system of self-regulating production. We can denote this complex as MANBRIC-technologies (i.e. medical, additive, nano-, bio-, robo-, info-, and cognitive technologies). As is known, with respect to the sixth technological paradigm (known also as the sixth technological system or style) there is a widely used idea connected with the notion of NBIC-technologies (or NBIC-convergence) (see Lynch 2004; Bainbridge and Roco 2005; Dator 2006; Korotayev 2008; Akaev 2012). There are also some researchers (e.g., Jotterand 2008) who see in this role another set of technological directions—GRAIN (Genomics, Robotics, Artificial Intelligence, Nanotechnology). However, we believe that this set will be larger. And medical technologies will be its integrating part.

It should be noted that Leo Nefiodow has been writing about health as the leading technology of the sixth Kondratieff wave for a long time (Nefiodow 1996; Nefiodow and Nefiodow 2014a, b). He explains that health is much more than medicine and includes mental, psychosocial, environmental and spiritual aspects. He believes that medicine covers only a small part of the health problems we face today. We agree that health is more than medicine. However, we regard medicine as the most important business sphere connected with health care (note that the overwhelming majority of researchers in the health area work with medical technologies). We also agree with Nefiodow that business and profit far from always serve people. But we do not know any power beside medical business which has opportunities (in co-operation with such state agencies as the National Institutes of Health in the USA) to finance research and development in this area, to elaborate new ways to fight mortal diseases, to invest in prolongation of life expectancy. In Nefiodow's opinion, health area covers not only psychotherapeutic, psychological and psychiatric services, but also numerous measures of health improvement that, using his terms, will reduce social entropy. The problems with this argument, based on reducing social entropy (e.g., corruption, small and large crime, drug addiction, lack of moral guide, divorces, violence, etc.), is that social entropy (as Nefiodow himself points out) has always existed in society. Social changes can be really extremely important for the creation of starting conditions for a long-term upswing in reducing social entropy (see Grinin and Grinin 2013, 2015 for more detail). However, it is production and/or commercial technologies that represent the driving force of the K-Wave upswing phases. There is one more important point. The Nefiodows believe that it is biotechnologies that will become an integrating core of the new technological paradigm. However, we suppose that the leading role of biotechnologies will be, first of all, in their possibilities to solve major medical problems. That is why it makes sense to speak about medical technologies as the core of a new technological paradigm. Besides, we forecast a more important role of nanotechnologies than the the Nefiodows do (Nefiodow and Nefiodow 2014b: Chap. 3). Nanotechnologies will be of great importance in terms of the development of bio- and medical technologies (they are supposed to play a crucial role in the fight against cancer; at the same time nanotechnologies will play a crucial role in other spheres too, in particular in energy and resource saving).

Thus, we maintain the following:

1. Medicine will be the first sphere to start the final phase of the Cybernetic Revolution, but, later on, self-regulating systems development will cover the most diverse areas of production, services and life.
2. We treat medicine in a broad sense, because it will include (and already actively includes) for its purposes a great number of other scientific branches: e.g., the use of robots in surgery and care of patients, information technologies in remote medical treatment, neural interfaces for treatment of mental illness and brain research; gene therapy and engineering, nanotechnologies for creation of artificial immunity and biochips which monitor organisms; new materials for growing artificial organs and many other things to become a powerful sector of economy.
3. The medical sphere has unique opportunities to combine the above mentioned technologies into a single system.

4. There are also some demographic and economic reasons why the phase of self-regulating systems will start in medicine:

- Increase in average life expectancy and population ageing will favor not only the growth of medical opportunities to maintain health, but also allow the extension of working age, as population ageing will be accompanied by shortages in the working-age population.
- People, in general, are always ready to spend money on health and beauty. However, the growth of the world middle class and the cultural standard of people implies much greater willingness and solvency in this terms.
- Medical corporations usually do not impede technological progress, but, on the contrary, are interested in it.

Thus, today medicine is a very important sector of the economy, and tomorrow it will become even more powerful.

3 Global Demographic Factors

A number of global demographic factors explain why in particular in medicine the transition to the new technological paradigm should start.

This will be supported by a particularly advantageous situation developing by 2030 in economy, demography, culture, a standard of living, etc.—these will define a huge need for scientific and technological breakthrough. By advantageous situation we do not mean that everything will be perfectly good in the economy; just on the contrary, everything will be not as good as it could be. Advantageous conditions will be created because reserves and resources for prolonging previous trends will be exhausted, and at the same time the requirements of currently developed and developing societies will increase. Consequently, one will search for new developmental patterns.

Particular attention should be paid to the global population aging (see Figs. 1, 2, 3, and 4). As shown above, an especially rapid global increase in the number of age persons above retirement-age is expected to come in the next 20 years—their number will actually double during a short historical period, thus it will increase by almost 600 million and the total number will considerably exceed a billion.

However, a massive acceleration will be observed in particular of people aged 80 years or more. While by 2050 the number of persons of retirement age will approximately double, the number of elderly people aged 80 years or more will practically quadruple and, in comparison with 1950, their number by 2075 will increase almost by 50 times (see Fig. 4).

As we can see, over the next 35 years there will be a truly explosive growth of population over 80 years old. By the 2050s, the global population of this age group quadruple. A particularly rapid growth will occur in the 2030s and 2040s, when the number of very old (and hence having particular acute needs in medical care) people will grow from 200 to 400 million (by 200 million people) just within

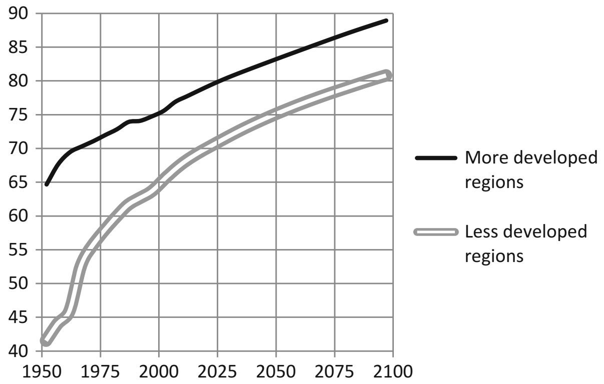


Fig. 1. Dynamics of the life expectancy at birth (years) in the World System core and global periphery, 1950–2015, the UN medium forecast to 2050 (*Source*: UN Population Division 2016)

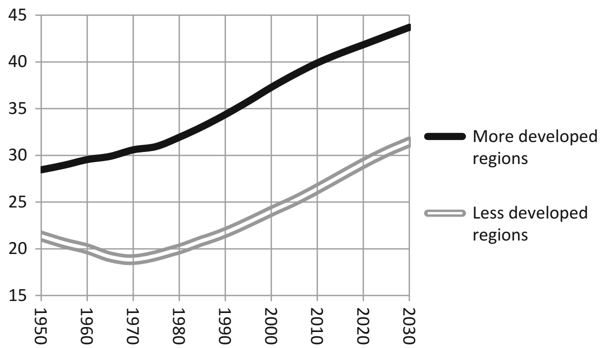


Fig. 2. Dynamics of the median age of population (years) in the World System core and global periphery, 1950–2015, with the medium forecast of the UN till 2030 (*Source*: UN Population Division 2016. We would like to remind the reader that if the median age of population of a given country equals, for example, 40 years, it means that half of the population of this country is younger than 40 years, and the other one is older)

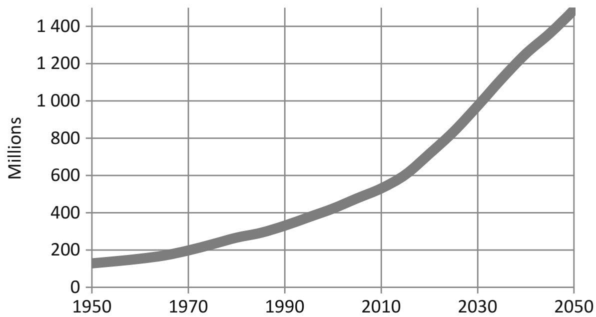
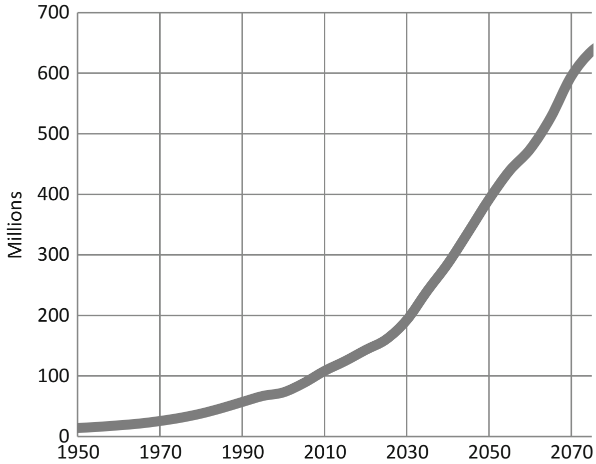


Fig. 3. Increasing number of persons of retirement age (over 65), 1950–2015, with the UN medium forecast till 2050 (*Source*: UN Population Division 2016)

Fig. 4. Increase of global number of elderly people (aged 80 years+), 1950–2015, with the UN average forecast till 2075 (*Source*: UN Population Division 2016)



20 years. And this is the very period when we expect a technological breakthrough in medicine. As we can see, the UN forecast shows that the demographic conditions for such a breakthrough in those years will be extremely favorable. An explosive growth of the number of very old people will mean an equally explosive increase in demand for medical services that will serve as a powerful incentive to increase even more research and development in this area. On the other hand, it would mean a very rapid growth of the market for such services—which implies that new developments in this area will have brilliant opportunities for commercialization that will stimulate further growth in life expectancy, forming a positive feedback loop that can act as a powerful driving force of this technological breakthrough.

So by the 2030s, the number of middle-aged and elderly people will increase; economy will desperately need additional labor resources while the state will be interested in increasing the working ability of elderly people, whereas the population of wealthy and educated people will grow in a rather significant way. In other words, the unique conditions for the stimulation of business, science and the state to make a breakthrough in the field of medicine will emerge, and just these unique conditions are necessary to start the innovative phase of production revolution!

It is extremely important to note that enormous financial resources will be accumulated for the technological breakthrough, such as: the pension money whose volume will increase at high rates; spending of governments on medical and social needs; growing expenses of the ageing population on health (related) services, and also on health services obtained by the growing world middle class. All this can provide initial large investments, an appeal of high investment of respective venture projects and long-term high demand for innovative products, thus a full set of favorable conditions for a powerful technological breakthrough will become available.

4 Self-regulating Systems Phase and the Sixth Kondratieff Wave

4.1 The 6th K-Wave A-Phase: Take-off Run Before the Start of the Final Phase of the Cybernetic Revolution

The sixth K-wave will probably begin approximately in the 2020s. Meanwhile the final phase of the Cybernetic Revolution has to begin later, at least, in the 2030s or 2040s. Thus, we suppose, that a new technological paradigm will not develop in a necessary form even by the 2020s (thus, the innovative pause will take longer than expected—see Grinin et al. 2016c: Chap. 6). However, it should be kept in mind that the beginning of the Kondratieff wave upswing phase is never directly caused by new technologies. This beginning is synchronized with the start of the medium-term business cycle's upswing. And the upswing takes place as a result of the levelling of proportions in economy, the accumulation of resources and other impulses that improve demand and conjuncture. One should remember, that the beginning of the second Kondratieff wave was connected with the discovery of gold deposits in California and Australia, the third wave was linked with an increase in prices for wheat, the fourth one with the post-war reconstruction, the fifth one with the economic reforms in the UK and the USA, as well as oil price shocks. And then, given an upswing, a new technological paradigm (which could not completely—if at all—realize its potential) facilitates overcoming of cyclic crises and allows further growth.

Consequently, some conjunctural events will also stimulate an upward impulse of the sixth K-wave. And, for example, the rapid growth of the underdeveloped world regions (such as Tropical Africa, the Islamic East, and some Latin American countries) or new financial and organizational technologies can become a primary impulse. Naturally, there will also appear some technical and technological innovations which, however, will not form a new paradigm yet. Besides, we suppose that financial technologies have not finished yet its expansion in the world. If we can modify and secure them somehow, they will be able to spread into various regions which underuse them now. One should not forget that large-scale application of such technologies demands essential changes in legal and other systems, which is absolutely necessary for developmental levelling in the world. Taking into account a delay of the new generation of technologies, the period of the 2020s may resemble the 1980s. In other words, it will be neither a recession, nor a real upswing, but rather somehow accelerated development (with stronger development in some regions and continuous depression in others—see Grinin et al. 2016c: Chap. 3).

Then, given the favorable conditions as they had been mentioned above, during this wave the final phase of the Cybernetic Revolution will begin. In such a situation it is possible to assume that the sixth K-wave's A-phase (the 2020s–2050s) will have much stronger manifestation and last longer than that of the fifth one due to more dense combination of technological generations. And since the Cybernetic

Revolution will evolve, the sixth K-wave's downward B-phase (the 2050s–the 2060/70s), is expected to be not so depressive, as those during the third or fifth waves. In general, during this K-wave (the 2020s–the 2060/70s) the Scientific and Information Revolution will come to an end, and the scientific and cybernetic production principle will acquire its mature shape.

4.2 There Is Another Scenario

The final phase of the Cybernetic Revolution can begin later—not in the 2030s, but in the 2040s. In this case the A-phase of the sixth wave will terminate before the beginning of the regulating systems revolution; therefore, it will not be based on fundamentally new technologies and will not become so powerful as is supposed in the previous scenario. The final phase of the Cybernetic Revolution in this case will coincide with the B-phase of the sixth wave (as it was the case with the zero wave during the First Industrial Revolution, 1760–1787—see Grinin et al. 2016c: Chaps. 2 and 6) and at the A-phase of the seventh wave. In this case the emergence of the seventh wave is highly possible. The B-phase of the sixth wave should be rather short due to the emergence of a new generation of technologies, and the A-phase of the seventh wave could be rather long and powerful.

4.3 The End of the Cybernetic Revolution and Possible Disappearance of K-Waves

The sixth K-wave (about 2020–the 2060/70s), like the first K-wave, will proceed generally during completion of the production revolution (see Grinin et al. 2016c: Chap. 6). However, there is an important difference. During the first K-wave the duration of the one phase of the industrial production principle significantly exceeded the duration of the whole K-wave. But now one phase of the K-wave will exceed the duration of one phase of production principle. This alone should essentially modify the course of the sixth K-wave; the seventh wave will be feebly expressed or will not occur at all (on the possibility of the other scenario see above). Such a forecast is based also on the fact that the end of the Cybernetic Revolution and diffusion of its results will promote integration of the World System and a considerable growth of influence of new universal regulation mechanisms. It is quite reasonable, taking into account the fact that the forthcoming final phase of the revolution will be the revolution in the regulation of systems. Thus, the management of the economy should reach a new level. K-waves appeared at a certain phase of global evolution and they are likely to disappear at its certain phase.

5 Conclusion: Some Possible Future Medical Technologies

5.1 Constant Health Monitoring as a Self-regulating Supersystem

Nowadays the boundary between medical diagnosis and treatment already becomes more and more imperceptible. Diagnostics is a constant necessary measure for disease controlling and drug dosage. During the final phase of the Cybernetic Revolution we expect breakthroughs in many fields of medical care. Thus, a very important direction of self-regulation can be associated with the development of the health monitoring system that will allow an early diagnosis and prevention of diseases. The key compounds of such devices are biosensors.

Biosensors are a good example of self-regulating systems and development of individualization (e.g., Cavalcanti et al. 2008). One can easily imagine that in the future biosensors will be able to become an integral part of human life fulfilling the function of a constant scanner of the organism or its certain organs and even transmitting the information about it to medical centers in case of potential threats or serious deterioration in the state of health. Built-in sensors will allow for controlling and regulating all vital processes, as well as prompting the time of drug intake and their dosage, time of physical activities and required exercises taking into account different circumstances, and recommending the most appropriate diet, etc.

Respectively, such mini-systems can be integrated into a large system which monitors a large number of people, for example in medical centers, therapeutic facilities, hotels, etc. We can forecast the decreasing number of hospitals, and such monitoring and remote online access can significantly relieve hospitals. One can imagine that such systems will be able to detect potentially dangerous situations and quickly respond to critical situations. That is a good example of prognostics and prevention of problems. We suppose that it will take much time to create such systems. Besides, there are complicated ethical and legal problems as regards to such monitoring as there always exists the danger that a watching ‘Big Brother’ will take advantage of this.

5.2 Artificial Antibodies and Growing Opportunities to Use the Immune System

There will never be any universal drug against all diseases. But strengthening the immune system is one of the universal directions which can transform this situation and help the struggle against different diseases. There is a special instrument of the human immune system—antibodies.

Antibodies are the molecules synthesized to fight against certain cells of foreign origin—antigens. The damage done by antigen usually leads to the destruction of foreign organisms and to recovery. Specific antibodies are produced for each antigen. They are produced by special immune cells—lymphocytes, which accumulate and circulate in the blood over the period of a lifetime. Thus, everyone has his own protective system based on the 'history of diseases'. It is one of the most important directions of development of individualization. Medicine is always connected with a patient's individuality. However, in the twentieth century there was a tendency towards mass medicine (connected with mass vaccination, preventive examinations, etc.). At present there are some signs of transition from mass medicine to personal/individual medicine, which is related to the general tendency of the Cybernetic Revolution towards individualization. But individualization to an even greater extent will be manifested when based on the unique characteristics of the organism, one of which is the immune system. Artificial antibodies can strengthen the tendency towards the individualization of medicine.

Scientists have repeatedly attempted to produce artificial antibodies. In 1970, Cesar Milstein and Georges Köhler found the method of producing the antibodies of a certain type, that is of monoclonal antibodies. Nowadays a focus of much medical research is into the production of antibodies by other means (Schirhagl et al. 2012) and also the creation of chemoreceptors (Dickert et al. 2001). Antibodies have already become widely used in pregnancy tests, in the diagnostics of many diseases, in laboratory experiments.

We suppose that during the final phase of the Cybernetic Revolution there will be a considerable progress in the creation of artificial antibodies and their acceptance by the organism. There is no doubt that progress in this field will lead to a breakthrough in medicine. The formation of artificial antibodies will play an important role in the prevention and treatment of many serious diseases, they will prevent the rejection of transplanted organs, etc. This will help make controlling the course of a disease easier and will help in suppressing the disease and defeating the disease if it is possible. Progress toward the creation of artificial antibodies will mean a significant growth of opportunities to control processes previously inaccessible for controllable interference and formation of self-regulating systems for regulation of such interference.

Control of programmed cell death (*apoptosis*) is one of the promising methods to defeat serious diseases including cancer. The researches into this field have been carried out since the 1960s. They show that some cells often die in strict compliance with the predetermined plan. Thus, the microscopic worm nematode's embryo consists of 1090 cells before hatching but later some of them die and there remain only 959 cells in the adult worm organism (Raff 1998; Ridley 1996). The mechanism of apoptosis is associated with the activity of signaling molecules and special receptors which receive the signal, launch the processes of morphological and biochemical changes, and as a result lead to the cell death.

An opportunity to trigger the self-destruction of pathogenic cells can make the struggle against diseases controllable. Besides, it would secure a rapid recovery without a long period of rehabilitation which is necessary after surgical intervention, chemotherapy or radiation treatment (it is an example of economy of energy and time for a patient). We suppose that during the final phase of the Cybernetic

Revolution medicine will be able to make progress in this direction and in the mature stages of the scientific-cybernetic principle of production to control it. In this case the movement toward creation of self-regulating systems will occur on the basis of the influence on the key elements of these subsystems of the organism in order to select the optimum regime in the context of certain goals and tasks. So in some cases it will be possible to cause the death of unwanted cells deliberately and in other cases to block the mechanism of death of necessary cells.

Also switching off the mechanism of cell self-destruction will help to save an organism from some diseases and, probably, to control the process of ageing. But it is only one of the possible ways to slow down the ageing process. On the processes determining ageing and opportunities to 'fight' ageing also see the monograph by Aubrey de Grey and Michael Rae (2007).

Acknowledgments The study was implemented in the framework of the Basic Research Program at the National Research University Higher School of Economics (HSE) in 2017. This research was supported by the Russian Science Foundation [Project # 14-11-00634].

References

Akaev, A. A. (2012). *Mathematical basis of Schumpeter—Kondratieff innovation-cyclical theory of economic development*. In A. A. Akaev, R. S. Grinberg, L. E. Grinberg, A. V. Korotayev, & S. Y. Malkov (Eds.), *Kondratieff waves. Aspects and perspectives*. Uchitel Publishers (in Russian).

Ashby, R. (1956). *An introduction to cybernetics*. London: Chapman and Hall.

Bainbridge, M. S., & Roco, M. C. (2005). *Managing nano-bio-info-cogno innovations: Converging technologies in society*. New York: Springer.

Beer, S. (1967). *Cybernetics and management*. London: English Universities Press.

Beer, S. (1994). *Decision and control: The meaning of operational research and management cybernetics*. London: Wiley.

Brener, R. (2006). *The economics of global turbulence. The advanced capitalist economies from long boom to long downturn, 1945–2005*. London – New York: Verso.

Cavalcanti, A., Shirinzadeh, B., Zhang, M., & Kretly, L. C. (2008). Nanorobot hardware architecture for medical defense. *Sensors, 8*(5), 2932–2958.

Dator, J. (2006). *Alternative Futures for K-Waves*. In T. C. Devezas (Ed.), *Kondratieff waves, warfare and world security*. Amsterdam: IOS Press.

de Grey, A. B., & Rae, M. (2007). *Ending aging: The rejuvenation breakthroughs that could reverse human aging in our lifetime*. New York: St. Martin's Press.

Devezas, T. C. (Ed.). (2006). *Kondratieff waves, warfare and world security*. Amsterdam: IOS Press.

Devezas, T. C. (2010). Crises, depressions, and expansions: Global analysis and secular trends. *Technological Forecasting and Social Change, 77*(5), 739–761.

Devezas, T. C. (2012). *The recent crisis under the light of the long wave theory*. In L. E. Grinin, T. C. Devezas, & A. V. Korotayev (Eds.), *Kondratieff waves*. Dimensions and prospects at the dawn of the 21st century, Uchitel Publishing.

Dickert, F. L., Hayden, O., & Halikias, K. P. (2001). Synthetic receptors as sensor coatings for molecules and living cells. *Analyst, 126*, 766–771.

Grinin, Л. E. (2006). *Productive forces and historical process*. Moscow: KomKniga/URSS.

Grinin, L. E., & Grinin, A. L. (2013). Macroevolution of technology. In L. E. Grinin, & A. V. Korotayev (Eds.), *Evolution: Development within Big History, Evolutionary and World-System Paradigms*. Volgograd: Uchitel Publishing.
Grinin, L. E., & Grinin, A. L. (2015). Global technological perspectives in the light of cybernetic revolution and theory of long cycles. *Journal of Globalization Studies, 6*(2), 119–142.
Grinin, L. E., & Grinin, A. L. (2016). *The cybernetic revolution and the forthcoming epoch of self-regulating systems*. Moscow: Moscow Branch of the Uchitel Publishing House.
Grinin, L. E., & Korotayev, A. V. (2010). Will the global crisis lead to global transformations. 1. The global financial system: *Pros* and *Cons*. *Journal of Globalization Studies, 1*(1), 70–89.
Grinin, L., & Korotayev, A. (2012). Does "Arab Spring" mean the beginning of world system reconfiguration? *World Futures, 68*(7), 471–505.
Grinin, L. E., & Korotayev, A. V. (2014a). *Globalization and the shifting of global economic-political balance*. In E. Kiss & A. Kiadó (Eds.), *The dialectics of modernity—Recognizing globalization. Studies on the theoretical perspectives of globalization*. Budapest: Arostotelész.
Grinin, L. E., & Korotayev, A. V. (2014b). Globalization shuffles cards of the world pack: In which direction is the global economic-political balance shifting? *World Futures, 70*(8), 515–545.
Grinin, L., & Korotayev, A. (2015a). *Great divergence and great convergence. A global perspective*. New York: Springer.
Grinin, L., & Korotayev, A. (2015b). Population ageing in the west and the global financial system. In J. Goldstone, L. Grinin, & A. Korotayev (Eds.), *History & mathematics: Political demography & global ageing*. Volgograd: Uchitel Publishing.
Grinin, L. E., Devezas, T. C., & Korotayev, A. V. (Eds.). (2012). *Kondratieff waves. Dimensions and prospects at the dawn of the 21st century*. Volgograd: Uchitel.
Grinin, L. E., Devezas, T., & Korotayev, A. V. (Eds.). (2014). *Kondratieff waves. Juglar—Kuznets—Kondratieff*. Volgograd: Uchitel.
Grinin, L. E., Ilyin, I. V., & Andreev, A. I. (2016a). World order in the past, present, and future. *Social Evolution & History, 15*(1), 60–87.
Grinin, L., Issaev, L., & Korotayev, A. V. (2016b). *Revolutions and instability in the Middle East*. Moscow: Uchitel (in Russian).
Grinin, L., Korotayev, A., & Tausch, A. (2016c). *Economic cycles, crises, and the global periphery*. New York: Springer.
Hirooka, M. (2006). *Innovation dynamism and economic growth. A nonlinear perspective*. Cheltenham: Edward Elgar.
Jotterand, F. (2008). *Emerging conceptual, ethical and policy issues in bionanotechnology*. Berlin: Springer.
Khaltourina, D. A., & Korotayev, A. V. (2007). *Secular cycles and millennial trends: A mathematical model*. In M. G. Dmitriev, A. P. Petrov, & N. P. Tretyakov (Eds.), *Mathematical modeling of social and economic dynamics*. Moscow: RUDN Publishing.
Kondratieff, N. D. (1926). Die langen Wellen der Konjunktur. *Archiv für Sozialwissenschaft und Sozialpolitik, 56*(3), 573–609.
Kondratieff, N. D. (1935). The long waves in economic life. *Review of Economic Statistics, 17*(6), 105–115.
Kondratieff, N. D. (1984). *The long wave cycle*. New York: Richardson & Snyder.
Kondratieff, N. D. (1998). *The Works of Nikolai D. Kondratiev. 4 vols*. London: Pickering and Chatto.
Kondratieff, N. D. (2004 [1922]). *The world economy and its conjunctures during and after the War*. Moscow: International Kondratieff Foundation.
Korotayev, A., & de Munck, V. (2013). Advances in development reverse inequality trends. *Journal of Globalization Studies, 4*(1), 105–124.
Korotayev, A., & de Munck, V. (2014). Advances in development reverse global inequality trends. *Globalistics and Globalization Studies, 3*, 164–183.

Korotayev, A. V., & Grinin, L. E. (2012). Kondratieff waves in the world system perspective. In L. E. Grinin, T. C. Devezas, & A. V. Korotayev (Eds.), *Kondratieff waves*. Dimensions and Prospects at the Dawn of the 21st Century, Uchitel Publishing.

Korotayev, A., & Zinkina, J. (2014). On the structure of the present-day convergence. *Campus-Wide Information Systems, 31*(2), 41–57.

Korotayev, A., Zinkina, J., & Bogevolnov, J. (2011a). Kondratieff waves in global invention activity (1900–2008). *Technological Forecasting & Social Change, 78*, 1280–1284.

Korotayev, A., Zinkina, J., Bogevolnov, J., & Malkov, A. (2011b). Global unconditional convergence among larger economies after 1998? *Journal of Globalization Studies, 2*(2), 25–62.

Korotayev, A., Zinkina, J., Bogevolnov, J., & Malkov, A. (2011c). Unconditional convergence among larger economies. In D. Liu (Ed.), *Great Powers, World Order and International Society: History and Future*. The Institute of International Studies, Jilin University.

Korotayev, A., Zinkina, J., Bogevolnov, J., & Malkov, A. (2012). Unconditional convergence among larger economies after 1998? *Globalistics and Globalization Studies, 1*, 246–280.

Korotayev, A., Goldstone, J., & Zinkina, J. (2015). Phases of global demographic transition correlate with phases of the great divergence and great convergence. *Technological Forecasting and Social Change, 95*, 163–169.

Lynch, Z. (2004). Neurotechnology and Society 2010–2060. *Annals of the New York Academy of Sciences, 1031*, 229–233.

Maddison, A. (2007). *Contours of the world economy, 1–2030 AD: Essays in macro-economic history*. Oxford – New York: Oxford University Press.

Nefiodow, L. (1996). *Der sechste Kondratieff. Wege zur Produktivität und Vollbeschäftigung im Zeitalter der Information*. Sankt Augustin: Rhein-Sieg-Verlag.

Nefiodow, L., & Nefiodow, S. (2014a). The sixth Kondratieff. The growth engine of the 21st century. In L. E. Grinin, T. C. Devezas, & A. V. Korotayev (Eds.), *Kondratieff waves*. Juglar—Kuznets—Kondratieff, Uchitel.

Nefiodow, L., & Nefiodow, S. (2014b). *The sixth Kondratieff. The new long wave of the world economy*. Sankt Augustin: Rhein-Sieg-Verlag.

Perez, C. (2002). *Technological revolutions and financial capital: The dynamics of bubbles and golden ages*. Cheltenham: Elgar.

Polterovich, V. (2009). Innovation pause hypothesis and modernization strategy. *Voprosy economici, 6*, 4–23 (in Russian).

Raff, M. (1998). Cell suicide for beginners. *Nature, 396*, 119–122.

Ridley, M. (1996). *The origin of virtue*. New York: Viking.

Schirhagl, R., Qian, J., & Dickert, F. L. (2012). Immunosensing with artificial antibodies in organic solvents or complex matrices. *Sensors & Actuators: B. Chemical, 173*, 585–590.

Schumpeter, J. A. (1939). *Business cycles. A theoretical, historical and statistical analysis of the capitalist process*. New York: McGraw-Hill.

Umpleby, S. A., & Dent, E. B. (1999). The origins and purposes of several traditions in systems theory and cybernetics. *Cybernetics and Systems: An International Journal, 30*, 79–103.

UN Population Division. (2016). *UN population division database*. Accessed January 17, 2016, from http://www.un.org/esa/population.

von Foerster, H., & Zopf, G. (Eds.). (1962). *Principles of self-organization*. New York: Pergamon Press.

Wiener, N. (1948). *Cybernetics, or control and communication in the animal and the machine*. Cambridge: MIT Press.

Global Pattern in Materials Consumption: An Empirical Study

Tessaleno C. Devezas, António M. Vaz, and Christopher L. Magee

Abstract This paper surveys the worldwide production of a basket of 98 important industrial materials between 1960 and 2013, divided into 9 main groups. The objective was to look for an answer to the question if it is possible to keep the global economy growing while simultaneously reducing the amount of materials resources necessary to maintain today's growth rate. Our results were analyzed within the framework of the dematerialization theory and the decoupling concept, checking the temporal unfolding in the past 50 years of the most widely used industrial materials, aiming to identify trends in dematerialization in the usage of this set of materials. One of the most striking patterns to be noted in our results is the spike-like aspect of the increasing consumption of some key materials in the first decade of the present century, a trend that seems to resume in most recent years, and is interpreted as a consequence of China's modernization in the past three decades. Our results do not support the conclusion that human society is experiencing dematerialization of the global economy, but some patterns found show that there are reasons for optimism.

Keywords Dematerialization • Decoupling • Jevon's paradox • Materials consumption • Rebound effect • Sustainability

1 Introduction

If we make a survey covering a reasonable fraction of the world population about which is the greatest challenge facing humanity at the dawn of this twenty-first century and that urge for solutions and/or short-term actions, certainly the answers would converge to how to maintain the healthy growth necessary to lift the world's poorest one billion people out of absolute poverty and manage the natural resources

T.C. Devezas (✉) • A.M. Vaz
Dept. of Electromechanics, University of Beira Interior, Covilhã, Portugal
e-mail: tessalen@ubi.pt

C.L. Magee
Massachusetts Institute of Technology, 77 Massachusetts Av., Cambridge, MA, USA
e-mail: cmagee@mit.edu

T. Devezas et al. (eds.), *Industry 4.0*, Studies on Entrepreneurship, Structural Change and Industrial Dynamics, DOI 10.1007/978-3-319-49604-7_14

required for the well-being of the entire Earth population in the near future—all while keeping environmental impact within acceptable limits and sustaining life's natural support system.

There is already a reasonable consensus that human use of materials is one of the main drivers of global environmental change. When we talk about 'environmental changes' we refer to all kinds of damage to the environment, i.e., depletion of natural resources, problems occurring during extraction, processing of raw materials, as well as emissions and wastes that are returned to the natural environment after processing and use of the materials. Undoubtedly the main agent of human-induced environmental change has been the almost exponential growth verified during the last century of the industrial metabolism, i.e., the inputs of materials and energy into the socioeconomic system. In despite of the fact that there are already many signals of decreasing rates of growth of the world population, of the global economic output (GDP) and of the per-capita materials consumption, there is still immense concern about the Earth's carrying capacity to support the continuing economic growth in a sustainable way.

This chapter intends to shed some new light on the interplay between economic activity and materials use, or in other words, on the key issue of '*decoupling*' economic growth and materials consumption, a theme of worldwide debate, also known in the technical literature as '*dematerialization*' of the economy.

2 Decoupling × Dematerialization

The words *decoupling* and *dematerialization* have been used by economists with the same meaning, but it's important in the context of the present article to point out some minor differences: they do not express exactly the same conceptual frame.

Decoupling at its simplest is reducing the amount of resources, such as water or fossil fuels or mineral ores, used to produce economic growth and delinking economic development from environmental deterioration. In other words, decoupling is a concept intrinsically coupled to wealth and the environmental aspects of the materials usage.

Decoupling applied to natural resources has grown out of the concept of 'ecoefficiency' developed by the World Business Council for Sustainable Development (WBCSD) in 1992, and the 2001 OECD definition of decoupling as breaking the link between 'environmental bads' and 'economic goods'. From a developing world context, the UN Economic Commission for Latin America and the Caribbean (ECLAC) in 2004 promoted the idea of 'non-material economic growth', essentially decoupling economic growth from resource consumption. As already shown by several authors, for instance Krausmann et al. (2009) and Steinberger et al. (2010), *decoupling is a consequence of material productivity*, measured as economic production per tonne of material consumed. Figure 1, constructed with the available data until 2013, shows the essence of the phenomenon of decoupling, being evident that in the last 50 years income per capita has

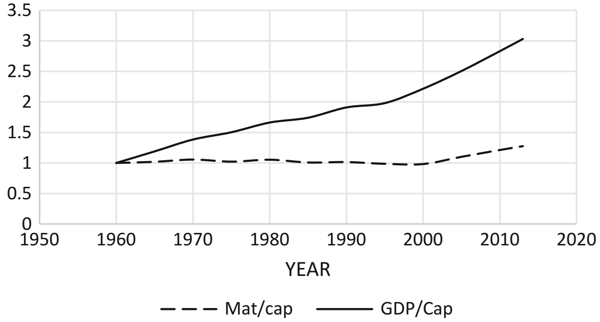


Fig. 1 Per capita evolution of the global GDP and global materials consumption in the last 50 years related to the values of both measures in 1960 (1960 = 1). In this timeframe, global GDP per capita has tripled and the global materials consumption per capita has increased about 30 % (see Sect. 4 for details)

grown more steeply than worldwide materials consumption per capita (details about these curves and the data used will be described in Sect. 4).

But it's very important to take in account that not only material productivity is responsible for this clear difference in rates of growth, but also other factors contribute to GDP growth, for instance in recent times the increasing role of the digital economy as already pointed out by Brian Arthur (2011) and several other economists. In fact, there are a lot of other economic activities that are nowadays increasingly contributing to affluence (GDP per capita) which do not require higher material consumption, a point that will be commented on later in this chapter.

Dematerialization, on the other hand, has perhaps a more restricted meaning, being usually defined in the technical literature as the reduction in the quantity of stuff (materials and energy) needed to produce something useful over time, which derives fundamentally from ongoing increases in technical capability. In other words, *dematerialization* is the phenomenon of reduction of the amount of material necessary for the same output which is a consequence of the continuing and unabated technological progress, and should be measured as the relation of the rate of material consumption with the rate of technological improvement for the production of a given good.

In a nutshell, economists often use both terms with the sense of *doing more with less*. But as just noted, both terms encompass much more, each in its respective context, than contained in this simple expression.

In 2007 the United Nations Environmental Programme (UNEP) established the International Resource Panel (IRP) with the mission of providing independent, coherent, authoritative and policy relevant scientific assessments on the management of natural resources and the environment for the highest net benefit of present and future generations. IRP has already delivered two extensive reports, the first one in 2011 under the title of *Decoupling—natural resource use and environmental impacts from economic growth* (IRP 2011), and the most recent one (IRP 2014) with the title of *Decoupling 2—Technologies, Opportunities, and Policy Options*.

Both reports called attention to the fact that overexploitation of resources, climate change, pollution, land-use change, and loss of biodiversity rose toward the top of the list of major international environmental concerns. One result is that 'sustainability' has become an overarching social and economic imperative among governments, international organizations, and businesses. Leaders in these sectors now understand that making progress towards a more sustainable economy requires an absolute reduction in resource use at a global level, while human wellbeing demands that economic activities should expand and environmental impacts diminish.

In the first report (2011), the Panel showed that breaking the link between human well-being and resource consumption is necessary and possible but in reality is hardly happening. In the follow-up report (2014) the Panel highlights existing technological possibilities and opportunities for both developing and developed countries to accelerate decoupling and reap the environmental and economic benefits of increased resource productivity. We can say that the second report has a slightly more optimistic vision in relation to the real *possibility* of getting 'decoupling' than the first one, for the authors show that many decoupling technologies and techniques that deliver significant resource productivity increases are already commercially available and used in both developing and developed economies. These technologies allow economic output to be achieved with fewer resource inputs, reducing waste and saving costs that can further expand the economy or reduce its exposure to resource risks.

But 'possibility' is different from 'reality'. The crude fact is that the twentieth century was characterized by an expressive explosion in demand of material resources that is set to accelerate as population growth and the increase in incomes continue to rise. More than 3 billion people are expected to enjoy "middle class" income levels in the next 20 years, compared to 1.8 billion today. A very recent report (GFN 2014) claims that '*if current population and consumption trends continue, by the 2030s, we will need the equivalent of two Earths to support us. And of course, we only have one*'.

Will it be possible to attain the goal of enabling about one third of current population within this time frame to newly enjoy the same living standard of to-day's most developed countries while reducing the amount of materials resources necessary (contrary then to the degrowth model discussed by some economists)?

The objective of this paper is to show our most recent statistics and empirical analysis on the worldwide consumption of materials, which we interpret as signaling that the answer to this question is probably positive. Considering the fact that our analysis will focus objectively on numbers and statistics about materials consumption and materials productivity, and not about the environmental impact of the industrial metabolism, we will refer to 'dematerialization' in our analysis throughout the paper.

3 Previous Work on Dematerialization

Very recently Magee and Devezas (2016) reviewed the existing published work on dematerialization theory and proposed an extension of it focusing on a fundamental question initially asked by Bernardini and Galli (1993) and pointed out by Kander (2005) as a possible important variable in the analysis of dematerialization: does improving technical capability offer substantial potential for continuing global economic growth accompanied by an *absolute* decrease in usage of materials from the Earth? Magee and Devezas argue that this is the critical question since environmental harm is difficult to avoid if the absolute amount of materials used continues to increase. In fact, browsing through all the work previously published on dematerialization one can find neither a satisfactory and careful analysis of this issue nor an answer to the question.

Magee and Devezas extended the theory by explicit consideration, for the first time in dematerialization analysis, of the ongoing technical progress on reducing the amount of the material needed for a given function coupled with the important issue of rebound [also called the Jevons' paradox (Jevons 1865), or the Khazzoom-Brookes postulate (Khazzoom 1980; Brookes 1984, 2000), or sometimes called backfire, and sometimes take back as well as rebound]. Whatever it is called, rebound involves the demand increase associated with technical improvement that cancels some of the decrease in demand due to fulfilling a given need with less material. The authors did not treat substitution among technologies in this simple extension nor structural change in the economy but only the direct effect of technological change and rebound over long periods of time.

The quantitative theory developed by Magee and Devezas identifies regions of possible dematerialization when the rate of technical improvement is sufficiently high and the demand elasticity sufficiently low. They also develop a methodology that is able to estimate the rate of improvement and demand elasticity when one knows both the technical performance and production as a function of time. Using the only data of this kind available (Nagy et al. 2013), Magee and Devezas conclude that no evidence yet exists for long-term dematerialization. In the present paper, however, performance data is not available so we cannot add to or amend the Magee and Devezas work. Instead, we look more closely at some material classes and call attention to potentially interesting recent usage trends and patterns.

Closing this section, it is important to comment on a new book by Vaclav Smil (Smil 2014)—*Making the Modern World: Materials and Dematerialization*. In this book Smil describe in detail the history of human use of materials and displays various statistics of the most important materials used, as well as details the complexity of the flow of materials in the modern world. In one of the chapters Smil posits the question: "Are we dematerializing?" In his attempt to answer this complex issue Smil seeks to make a distinction between *apparent dematerialization* and *relative dematerialization*.

Regarding the former Smil mentions the case of a *complex form of materials substitution* (using his words), in which a technology disappears, implying in the disappearance of a given set of materials, being substituted by another technology far more complex involving the usage of a larger infrastructure and of course of a larger amount of materials. The author mentions some examples, but perhaps the more interesting in this context is the case of the *universal dematerialization of design* (again using his words) which occurred after the introduction of CAD. Drafting boards, large amounts of blueprints and storage steel cabinets were substituted with electronic graphics displayed on screens and stored on magnetic devices or another media. The problem is that, following Smil's line of reasoning, such new technology requires extensive computational infrastructures, redundant mass data storage, flat screens, and specialized software (written by other computers), not to mention the increased demand for electricity.

As to the latter, *relative dematerialization*, Smil considers the reduction of material inputs in production enabled by (1) gradual improvements that do not involve new materials, (2) substitutions of constituent materials with lighter or more durable alternatives, (3) intensified recycling, and (4) introduction of entirely new devices that perform desired functions with a fraction of mass needed for their predecessors. After analyzing a series of examples that occurred during the second half of the twentieth century, Smil concludes that "*in an overwhelming majorities of cases the complex dynamic interactions of cheaper energy, less expensive and lighter materials, as well as cheaper manufacture, have resulted in such ubiquitous ownership of an increased range of products and more frequent use of a widening array of services that even the most impressive relative weight reductions accompanying these consumption increases could not be translated into any absolute cuts in the overall use of materials*". Claiming then the action of the 'rebound effect' or 'Jevons paradox' already mentioned above when commenting on the contribution of Magee-Devezas, Smil states that *relative dematerialization has been a key factor promoting a massive expansion of total material consumption.*

Smil's arguments, although based on sparse data on the materials and technologies involved in his analysis, are exposed in a discursive and possibly biased way intending to demonstrate that "*less has been an enabling agent of more*" (his words). Although Smil's general view qualitatively agrees with the Magee and Devezas results, the latter framework is more amenable to extensive testing over time. Given these largely negative predictions regarding the hypothesis of dematerialization of the economy, we think it important to scrutinize thoroughly and quantitatively the evolution of the use of a given basket of 98 important industrial materials over the last 50 years.

4 Methods and Data

In order to perform the statistical survey mentioned above, we examined the worldwide production of 98 industrial materials between 1960 and 2013, divided in 9 main groups: cellulose and derivatives (10 materials), wood (10 materials), fibers (4 materials), semimetals (6 materials), metals (31 materials), nonmetals

(4 materials), minerals (26 materials), rocks and stones (6 materials), and plastics (all plastics as a single material).

Data on the production of these materials were taken from the following databases: fibers: USDA (US Dept. of Agriculture), World Bank, CIRFS (European Man-made Fibers Association) and UNTAD Org; Metals, nonmetals, semimetals and minerals: USGS (US Geological Survey); Cellulose/Wood: FAO (Food and Agriculture Organization, UN); plastics: PEMRG (Plastics Europe Market Research Group), European Plastics Converters (EuPC).

It is important to point out that the present study does not represent an attempt to survey the complete worldwide material flow account (MFA), for we have not considered materials extracted from the biomass (exception of wood) and also not the materials from fossil energy carriers. The aim of the present work was limited to check the temporal unfolding in the past 50 years of the most widely used industrial materials, with the purpose to attempt to identify any trends in dematerialization in the usage of this set of materials.

Our reckoning indicates that the total tonnage of production of this selected set of materials amounted in 2013 to about 20.6 billion metric tons (Gt) (an increase to about 3 times, or 200 %, the production of the same set of materials in 1960). According to modern estimates (Krausmann et al. 2009; Steinberger et al. 2010) the global materials extraction at the beginning of the twenty-first century was in the range between 50 and 60 Gt, what means that the basket of materials chosen for this study are about one third of the total of materials used by humans. Also according to the same authors, the per capita global materials consumption in 2000 amounted to 8.1 tonnes per person per year; our calculation for 2013, considering the set of 98 materials, indicate a consumption of 2.87 tones per person per year (an increase of circa 27 % compared with 1960)—again about one third of the total of materials usage.

From the comparisons made above we can draw two important conclusions: firstly, since the objective of this work is identifying any trend towards dematerialization, the basket of the chosen industrial materials is a significant one related to the total materials used by modern man; secondly, the increase of per capita consumption (27 % compared with 1960) is less than the increase in overall consumption (200 % compared with 1960), which implies that population growth has been a greater driver of materials use than economic growth—a conclusion that will be further discussed later in this paper (in the same period population has grown 136 %).

The bulk of our analysis considers the aggregate global level of production of the selected materials. This observation is important because in this case we do not need to consider the net trade of materials imports and exports, which would be necessary for the analysis of production and consumption of individual nations. At the aggregate level, total net trade is zero, as also pointed out by Steinberger et al. (2010), for the total amount of resources extracted (usually coined as Direct Extraction, or DE) equals total amount of resources used (Direct Material Consumption, or DMC). In this paper the only case for which we made considerations about imports and exports was the comparison (for some most consumed materials) between the USA and China (Sect. 6).

Among the 98 analyzed materials, the most significant (in relation to the total tonnage) were by far cement (share 19.8 %) and iron ore (share 15.1 %). Steel followed with a share of 7.8 %. But considering that cement, in addition to being the most consumed material, has the unique feature among all the investigated materials, to be used exclusively in construction, and not be dispersed among numerous other applications and transformations, we conducted a detailed study of the same not only globally, but also for the largest world consumers/producers: the USA and China. For all other materials we discuss global changes in the tonnage of materials used per capita and per unit of GDP.

Data for the world population and world output (GDP) were retrieved from The World Bank. The data on GDP were considered in Purchasing Power Parity (at 2011 prices), which as we know eliminates inter-countries differences in price levels, so that differences in the volume of economic activity can be compared across countries, allowing a coherent set of space–time comparisons. Data from China's balance between imports and exports were taken from the National Bureau of Statistics of China.

5 Results

In Table 1 an overview is given of the changes between 1960 and 2013 of the main indicators analyzed in this study, namely GDP PPP, world population, total consumption of 98 materials, and per capita consumption of 98 materials. Table 2 exhibits the progression in brute consumption of each of the 9 materials groups investigated and their respective shares as well.

From the numbers presented in Table 1, we emphasize two main aspects:

(1) The global output (GDP) grew 7 times in the last half century, while in the same period global population grew 2.36 times and the global consumption of the selected 98 materials grew 3 times. In other words, the human capacity to generate wealth grew faster than population, and not necessarily consuming materials in the same proportion. Global growth of material consumption accompanied to some extent the population growth.

Table 1 Overview of the change in 50 years of the main global indicators

	GDP PPP (2011) US$	World Population	GDP PPP per capita	Consumption 98 materials (metric tons)	Per capita consumption 98 materials
1960	14.01×10^{12}	3.03×10^{9}	4.6×10^{3}	6.83×10^{9}	2.25 t/year
2013	100.40×10^{12}	7.17×10^{9}	13.89×10^{3}	20.59×10^{9}	2.87 t/year
Variation	616 %	136 %	202 %	201 %	27.5 %

Table 2 Global consumption per group of materials and their respective shares

	Cellul.& Deriv.	Wood	Fibers	Metals	Semi metals	Non metals	Miner.	Rocks & Stones	Plastics
1960 (10^6 Ton.)	200	4800	14.93	640	1.48	19.7	1090	39.81	6.65
1960 share	2.94 %	70.40 %	0.24 %	9.43 %	0.02 %	0.29 %	16.08 %	0.59 %	0,01 %
2013 (10^6 Ton.)	1035	7370	88.83	2860	11.67	70.84	8406	453.9	299
2013 share	5 %	35.8 %	0.45 %	13.9 %	0.05 %	0.35 %	40.8 %	2.2 %	1.45 %
Variation 1960–2013	415 %	53 %	495 %	346 %	689 %	260 %	670 %	1040 %	4396 %

(2) Reinforcing the conclusion above we can see that individual affluence (GDP/capita) grew 3 times, while individual material consumption increased less than 30 %, as already pointed out in Fig. 1 of the previous section.

By examining Table 2 we can see that among the 9 groups of materials, the groups with the highest growth were plastics (~45 times) and rocks & stones (~11.5 times), followed by semimetals (7.9 times) and minerals (~7.7 times). All other groups grew less than the growth recorded for GDP PPP. But the most striking feature to be noted in Table 2 is the radical transformation observed in the shares during the observed time span. While by the middle of the past century the (by far) most consumed material was the wood family (share ~70 %), its share was reduced by about half (share ~36 %) at the dawn of the present century, an important factor to be considered later in our discussion about the sustainability of continuing economic growth. The role of most consumed group changed to minerals, which jump from 16 % to 41 %, mainly due to the enormous growth for cement and feldspar, as will be analyzed later.

But we must not forget that these measures relate to the absolute tonnage of the respective groups, and the materials considered in our analysis have very disparate specific weights, and are also used in quite different quantities in their respective industrial applications. This is another important aspect to be considered ahead in our discussion about trends in dematerialization. See for instance the case of plastics that changed its share from a mere 0.01 % to a significant 1.45 %, considering that is the lightest of the materials considered.

Before going to the specifics of the evolution in consumption of each of the 98 materials analyzed, we present the per capita consumption unfolding of each of the 9 material groups since 1960, and the complete set of 98 materials as well, pictured in the next 10 graphs (Figs. 2, 3, 4, 5, 6, 7, 8, 9, 10 and 11, measured in 5-year intervals).

Examining this set of figures, we can observe that:

1. The global per capita consumption of the complete set of 98 materials (Fig. 2) diminished slightly (0.02 %) between 1960 (2.25 t/cap) and 2000 (2.21 t/cap), but increased about 30 % after this date (2.87 t/cap in 2013). In other words, the increase pointed out in Table 1 and commented on in point 1 above occurred only after the dawn of the new century.
2. Per capita consumption of wood (Fig. 4) is the only case where we can observe a steady decline over the entire observed period, translating visually the share reduction from ~70 % to ~36 % pointed out when commenting on Table 2. The global consumption of all types of wood diminished from 1.58 t/cap (1960) to 1.03 t/cap (2013).
3. Per capita consumption of cellulose/derivatives (Fig. 3) increased about 74 % between 1975 and 1995, stabilized after this date, and decreased 2 % between 2010 (153 kg/cap) and 2013 (144 kg/cap), indicating an important decreasing trend in the consumption of paper.
4. Per capita consumption of fibers (Fig. 5) increased evenly in the period 1960–1995, and evidenced an uptrend after 2000 mainly due to an increasing

Fig. 2 Per capita consumption unfolding of the complete set of 98 materials

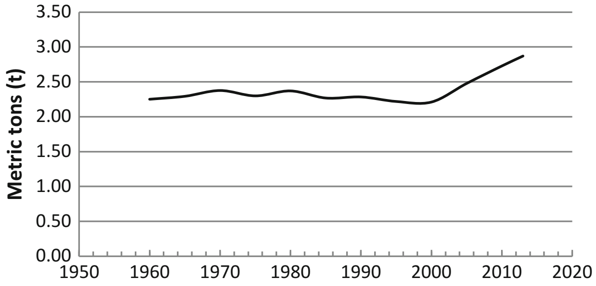


Fig. 3 Per capita consumption unfolding of the cellulose and derivatives group

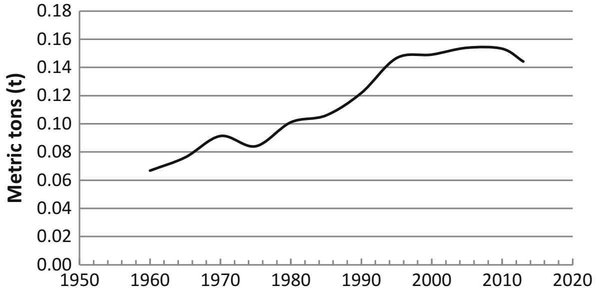


Fig. 4 Per capita consumption unfolding of the wood group

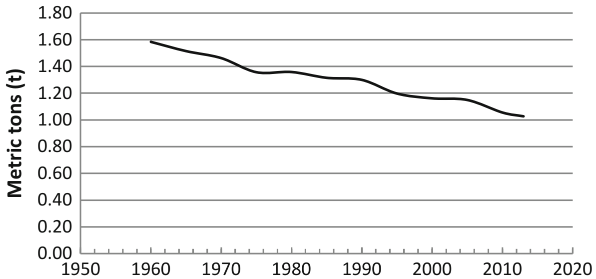


Fig. 5 Per capita consumption unfolding of the fibers group

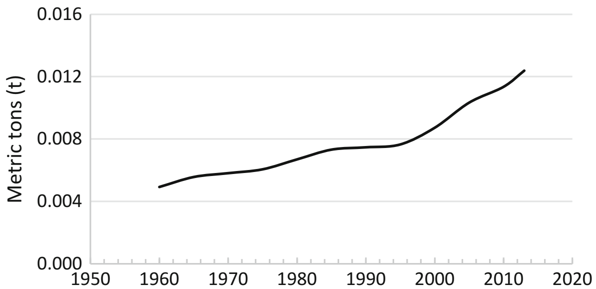


Fig. 6 Per capita consumption unfolding of the metals group

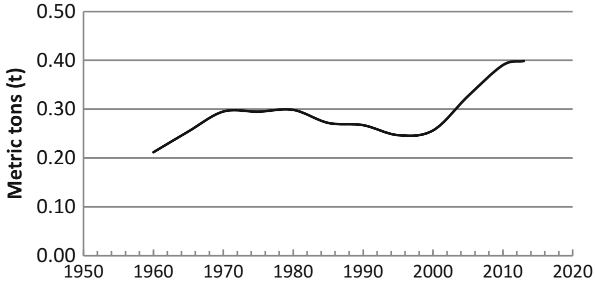


Fig. 7 Per capita consumption unfolding of the nonmetals group

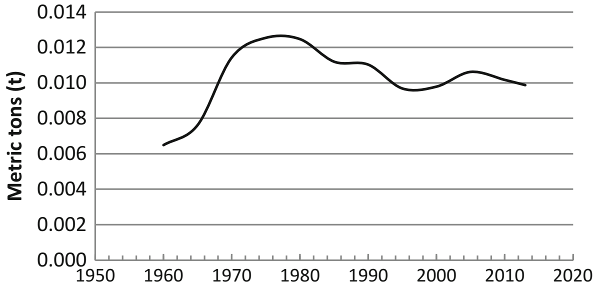


Fig. 8 Per capita consumption unfolding of the semimetals group

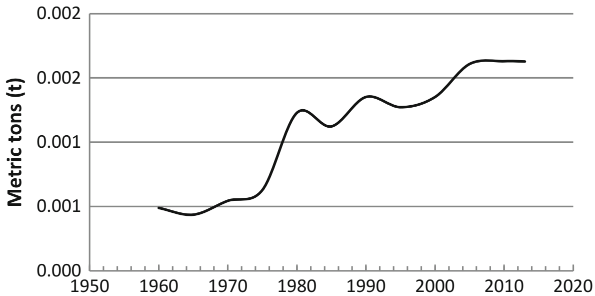
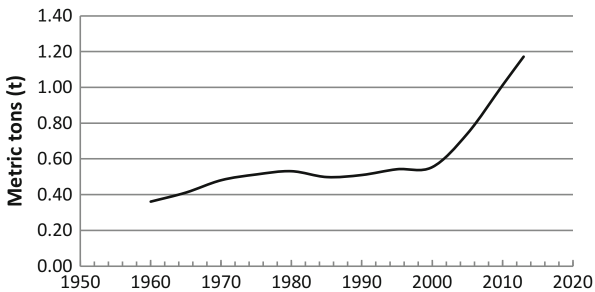


Fig. 9 Per capita consumption unfolding of the minerals group



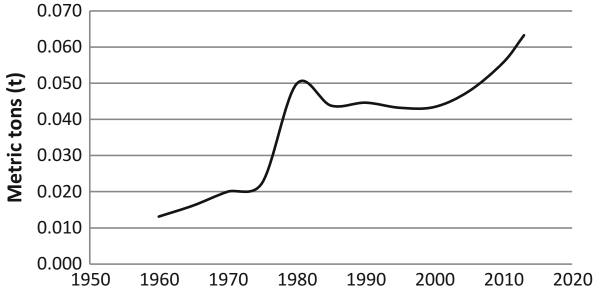


Fig. 10 Per capita consumption unfolding of the rock and stones group

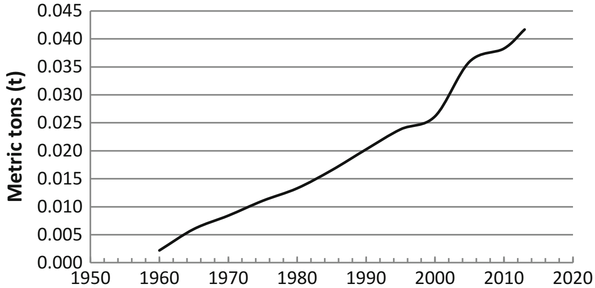


Fig. 11 Per capita consumption unfolding of the plastics group

consumption of synthetic fibers. This class of material represented only about 5 % of the total within the group in 1960, but jumped to 65 % in 2013. An inverse trend was presented by cotton, which decreased from 68 % in 1960 to 28 % in 2013.

5. Per capita consumption of metals (Fig. 6) decreased about 13 % in the period 1970–2000, but increased significantly after this date until 2010. In the 4 most recent years the consumption of the set of 31 metals seems to have stabilized.
6. Per capita consumption of nonmetals (Fig. 7) presents a curious behavior, increasing from 1960 until 1975, but decreasing after this date, and seems to have stabilized with a net consumption of about 1 kg/cap in the last years.
7. Per capita consumption of semimetals (Fig. 8) increased from 1960 until 2005, but apparently stabilized after this date with about 1.6 kg/cap.
8. Per capita consumption of minerals (Fig. 9) remained practically constant from 1975 until 2000, but grew steeply (112 %) after this date, changing from 513 kg/cap to about 1170 kg/cap. This growth translates visually the change in share already pointed out when discussing the numbers presented in Table 2, and is related with the enormous growth in cement consumption after 2000 (cement accounts to 48.5 % of total production within the group of minerals in 2013—a detailed discussion about this point is presented in Sect. 6.2).
9. Per capita consumption of rocks & stones (Fig. 10) and plastics (Fig. 11) increased over the observed period, but with the difference that plastics grew evenly while rocks and stones picked in the period 1975–1985, then decreased and increased again after 2005.

Let’s now scrutinize the development of each of the 98 selected materials by presenting in the five Tables 4, 5, 6, 7 and 8 their percentage variation, decade by decade since 1960. For the construction of these tables we have used the following criterion for defining boundaries of dematerialization within the logic of sustainability:

1. The *materialization* → *weak dematerialization* boundary is when the usage per GDP declines from its previous value (materialization is when more material per GDP is used in later than earlier periods). This boundary is in accord with previous definitions but in a growing economy still means significantly more materials usage over time.
2. The *weak dematerialization* → *moderate dematerialization* boundary is when the usage per capita decreases. Once this boundary is passed, growth in materials usage only occurs due to population growth, as already pointed out by Ziolkovski and Ziolkovska (2011).
3. The *moderate dematerialization* → *strong dematerialization* boundary is passed when there is an absolute decrease in usage of materials over time—despite economic or even population growth.

The last transition is the boundary fully supporting achievement of sustainability. We will look at this boundary for the various material types selected and attempt to judge from all the data whether the path of human civilization is headed in this direction.

In the boxes of these tables we use the following criteria of shadows and letters to differentiate *materialization* (Mat)—*weak dematerialization* (WD)—*moderate dematerialization* (MD)—*strong (absolute) dematerialization* (SD) considering that the percentage variation of GDP PPP (2011) and world population was as presented in Table 3. For instance, as shown in the grid below, black box with white letters means that the percent variation of the materials has exceeded the percent variation of GDP PPP in that decade, and so on.

Mat	WD	MD	SD
Xx %	**Xx%**	Xx%	Xx%

A quick look at this set of tables gives us the feeling of a very disparate behavior among the different materials: some have passed boundary (3), some other have not even passed boundary (1), and several materials exhibit oscillations, increasing and/or decreasing consumption in the successive decades.

Looking at the last column of these tables (total percentage variation 1960–2013) we observe that:

1. 23 materials have strongly materialized (consumption change higher than GDP change), but some of these cases are on balance good for the environment, as for instance the case of recovered paper and synthetic fibers.

Table 3 Percentage variation of GDP PPP and world population decade by decade since 1960

	1960–1970	1970–1980	1980–1990	1990–2000	2000–2010	2010–2013	1960–2013
GDP-PPP (2011)	68 %	45 %	37 %	34 %	45 %	11 %	617 %
World population	21 %	20 %	19 %	16 %	13 %	4 %	136 %

2. From the set of 31 metals observed, only 8 show significant increase in consumption and some of these are very light materials (as for instance aluminum and magnesium) or are consumed in very small quantities (as for instance niobium, strontium or vanadium).
3. Some of the materials of the mineral group exhibit strong growth in consumption, even though as pointed out above, some of them are consumed in very insignificant quantities, as for instance the case of rare earths. The most important material in this group—cement, considering its specific weight and its share within the group, deserves a specific analysis that is presented in the next section.
4. Plastics is one the materials exhibiting the highest growth, but among other reasons for its increase in consumption, this is a very likely case of replacement, whereby a lightweight material is being extensively used (for instance in composite for the automotive and aerospace industry) to substitute for heavier materials. Again, within the context of sustainability, this is probably a positive case.
5. Perhaps disappointing is the fact that only 10 materials present an absolute decline in consumption, this is not encouraging under the point of view of a trend toward dematerialization. Add to that the fact that some are cases of replacement (paper for packaging and wool for instance), and at least four cases are clearly not examples of technological improvement overcoming rebound leading to dematerialization but instead the consumption decline for asbestos, beryllium, mercury and thallium has occurred because of legal restrictions on their use due to toxicity issues.

But if, instead of looking only to the last column, we observe the behavior of consumption decade after decade for the entire set of materials, we find a dramatically different perspective. Table 9 below summarizes the results for the 98 materials, and Fig. 12 shows graphically the progress decade by decade of the extremes—number of materials materializing and dematerializing respectively.

Tables 4, 5, 6, 7 and 8 and Fig. 12 show that- on a decadal basis, the number of individual materials surpassing boundary (3) has increased significantly (in the most recent 4 years more than the half, 51 materials), while the number of materials confined in boundary (1) reduced about 30 %. Undoubtedly the most striking aspect observed in both (Fig. 12 and Table 9) is the fact that something anomalous seem to have happened in the period 2000–2010, with a significant rise in materialization and a corresponding decline in dematerialization. The most recent 4 years indicate that perhaps recently the trend is returning to the previous dematerializing trend.[1] This phenomenon is also very clear in the graphs presented in the sequence of Figs. 2, 3, 4, 5, 6, 7, 8, 9, 10 and 11 with a sudden change upwards in the curves after 2000. This alerts us to the fact that decade by decade results may not be indicative of long term change. This phenomenon is discussed in detail in the next section.

[1]The reader should keep in mind that, despite the fact that the last period is smaller (2010–2013, 4 years) the comparisons are made with the corresponding variations of GDP and population in the same short period.

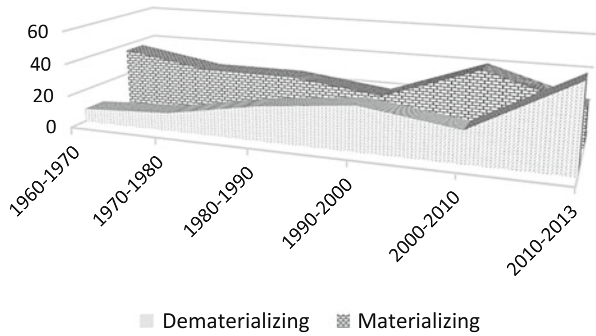


Fig. 12 Decade by decade progress of the extremes—number of materials materializing and dematerializing respectively

Table 4 Percentage variation of the materials of the groups cellulose and derivatives, wood, and fibers

Cellulose and derivatives (t)	1960-1970	1970-1980	1980-1990	1990-2000	2000-2010	2010-2013	1960-2013
Coated Papers	NA	NA	NA	NA	9%	**-8%**	**0%**
Mechanical pulp	37%	12%	39%	**0%**	**-17%**	**-13%**	55%
Other Fibre Pulp	48%	24%	111%	**-1%**	21%	**-26%**	244%
Other Papers Packaging	NA	NA	NA	NA	**-16%**	**-15%**	**-29%**
Paper and Paperboard	69%	35%	41%	36%	23%	**-1%**	435%
Paper for News Print	50%	18%	29%	20%	**-17%**	**-11%**	103%
Printig and Writing Paper	70%	55%	69%	42%	11%	**-4%**	572%
Pulp for paper	66%	25%	29%	11%	1%	**-2%**	193%
Recovered Paper	100%	64%	67%	70%	48%	1%	1293%
Wraping Papers	NA	NA	NA	NA	12%	**-2%**	9%

1960=1961

Wood (t)	1960-1970	1970-1980	1980-1990	1990-2000	2000-2010	2010-2013	1960-2013
Chips and particles	NA	NA	NA	NA	12%	**-11%**	0%
Fibre Furnish	73%	34%	40%	31%	22%	**0%**	412%
Plywood (m^3)*	102%	18%	22%	21%	46%	**0%**	415%
Roundwood (m^3)*	12%	11%	13%	**-2%**	1%	3%	42%
Sawnwood	20%	8%	10%	**-17%**	7%	1%	28%
Uncoated Mechanical	NA	NA	NA	NA	56%	**-8%**	43%
Uncoated Woodfree	NA	NA	NA	NA	**-6%**	**-4%**	**-9%**
Veneer (m^3)*	18%	22%	30%	4%	**-11%**	**-1%**	73%
Wood fuel	3%	9%	9%	**-1%**	4%	**0%**	26%
Wood Pulp	65%	24%	23%	11%	0%	1%	180%

*1960=1961 *(t)=(m^3)*

Fibers (t)	1960-1970	1970-1980	1980-1990	1990-2000	2000-2010	2010-2013	1960-2013
Cellulosic	28%	5%	**-12%**	**-11%**	53%	35%	117%
Cotton	16%	19	37%	**-1%**	29%	0%	143%
Synthetic	570%	126%	45%	97%	61%	18%	8087%
Wool	9%	0%	21%	**-30%**	**-20%**	**-1%**	**-27%**

Table 5 Percentage variation of the materials of the group metals

Metals (t)	1960-1970	1970-1980	1980-1990	1990-2000	2000-2010	2010-2013	1960-2013
Aluminum	115%	60%	25%	26%	70%	16%	960%
Beryllium	**-44%**	**50%**	**-24%**	**-20%**	**-10%**	27%	**-42%**
Bismuth	55%	**-3%**	**-5%**	9%	137%	**-6%**	250%
Cadmium	49%	10%	11%	0%	15%	**-6%**	98%
Chromium	53%	48%	40%	20%	59%	18%	610%
Cobalt (1970=1973)	220%	35%	13%	8%	148%	**0%**	309%
Copper	50%	22%	28%	43%	23%	12%	364%
Direct reduced Iron (1980=1984)	NA	NA	97%	133%	63%	**-3%**	624%
Gold	24%	**-18%**	79%	19%	**0%**	8%	135%
Indium (1970=1972)	NA	**-27%**	137%	184%	98%	20%	1068%
Lead	42%	4%	**-4%**	**-5%**	27%	34%	129%
Lithium	**-16%**	27%	76%	25%	140%	27%	613%
Manganese	34%	18%	**-6%**	**-23%**	113%	14%	176%
Magnesium	137%	44%	12%	19%	77%	18%	844%
Magnesium compounds	28%	32%	**-9%**	21%	115%	**-9%**	264%
Mercury	17%	**-30%**	**-40%**	**-67%**	58%	**-13%**	**-77%**
Nickel	96%	24%	25%	32%	28%	4%	438%
Niobium (1960=1964)	241%	78%	1%	62%	151%	**-5%**	2275%
Pig iron	66%	19%	3%	8%	82%	**-4%**	286%
Platinum	232%	61%	37%	25%	30%	**-5%**	1031%
Potash	101%	53%	**-1%**	**-2%**	33%	**-4%**	280%
Rhenium (1970/2013)	NA	45%	254%	7%	38%	**-1%**	649%
Silver	28%	14%	55%	9%	32%	8%	254%
Steel	72%	21%	8%	10%	68%	13%	368%
Strontium	421%	59%	153%	65%	**-7%**	**-10%**	2796%
Tantalum (1960=1969)	**-18%**	71%	**-27%**	170%	**-11%**	**-18%**	101%
Thallium	NA	NA	15%	0%	**-33%**	**0%**	**-23%**
Tin	27%	6%	**-10%**	26%	**0%**	6%	60%
Tungsten	4%	60%	**0%**	**-15%**	55%	19%	161%
Vanadium	196%	141%	**-8%**	23%	75%	10%	1467%
Zinc	77%	9%	20%	23%	40%	9%	334%

Table 6 Percentage variation of the materials of the groups semimetals and nonmetals

Semimetals(t)	1960-1970	1970-1980	1980-1990	1990-2000	2000-2010	2010-2013	1960-2013
Antimony	31%	**-4%**	**-10%**	95%	56%	**-16%**	245%
Arsenic	**-5%**	**-37%**	71%	18%	10%	**-14%**	33%
Boron	**-1%**	916%	11%	58%	**-10%**	**-13%**	1485%
Germanium	86%	37%	**-34%**	**-8%**	69%	31%	160%
Silicon (1960=1964)	45%	68%	50%	**-15%**	98%	14%	380%
Tellurium (2013=2012)	**-6%**	**-34%**	**-39%**	63%	**-13%**	**0%**	**-46%**

Nonmetals (t)	1960-1970	1970-1980	1980-1990	1990-2000	2000-2010	2010-2013	1960-2013
Bromine (1960=1961)	128%	40%	49%	23%	**-13%**	**-14%**	337%
Iodine	173%	40%	38%	22%	47%	9%	933%
Selenium	73%	**-2%**	38%	**-18%**	47%	1%	186%
Sulfur	114%	31%	5%	3%	18%	1%	259%

Table 7 Percentage variation of the materials of the group minerals

Minerals(t)	1960-1970	1970-1980	1980-1990	1990-2000	2000-2010	2010-2013	1960-2013
Asbestos	58%	35%	-15%	-47%	-5%	0%	-9%
Barite	45%	93%	-23%	12%	33%	6%	241%
Cement	81%	54%	18%	59%	98%	24%	1189%
Feldspar	61%	26%	87%	59%	116%	3%	1250%
Fluorspar	107%	20%	2%	-13%	56%	-2%	235%
Graphite	-10%	52%	58%	-11%	37%	-4%	155%
Gypsum	29%	52%	33%	4%	115%	6%	513%
Ilmenite	72%	32%	9%	21%	52%	5%	378%
Iron ore	47%	17%	11%	10%	140%	20%	505%
Iron oxide (1970=1976)	NA	-11%	2%	98%	47%	-31%	82%
Lime (1960=1963)	43%	24%	13%	-11%	155%	14%	419%
Mica (flake)	28%	44%	-3%	58%	238%	-1%	848%
Mica (sheet)	-32%	6%	-22%	-29%	0%	0%	-60%
Molybdenum	104%	35%	14%	6%	81%	5%	539%
Perlite (1970=1975)	NA	20%	3%	23%	10%	-8%	29%
Phosphate rock	128%	55%	10%	-19%	39%	22%	438%
Rare Earths	600%	72%	94%	72%	39%	-13%	4746%
Rutile	301%	5%	5%	-10%	55%	17%	612%
Salt	72%	16%	8%	7%	38%	-3%	209%
Soda Ash	33%	48%	13%	8%	44%	1%	249%
Sodium sulfate (1970=1972)	NA	20%	10%	15%	39%	0%	113%
Talc and Pyrophyllite	91%	56%	24%	-7%	-17%	-4%	174%
Titanium mineral	91%	34%	16%	18%	44%	2%	414%
Vermiculite	60%	38%	6%	-9%	-19%	-16%	47%
Wollastonite	99%	54%	140%	94%	9%	11%	1612%
Zirconium	209%	70%	25%	-14%	122%	-8%	1055%

Table 8 Percentage variation of the materials of the groups rocks and stones and plastics

Rocks and Stones (t)	1960-1970	1970-1980	1980-1990	1990-2000	2000-2010	2010-2013	1960-2013
Bauxite	109%	54%	27%	20%	76%	18%	925%
Diatomite	13%	-4%	11%	13%	1%	19%	61%
Gemstones	99%	-21%	402%	28%	10%	-5%	952%
Industrial Diamond	39%	11%	1090%	25%	815%	0%	21090%
Pumice and Pumicite	32%	7%	-28%	25%	32%	-8%	54%
Sand and gravel Ind.	NA	NA	-4%	4%	11%	20%	32%

Plastics (t)	1960-1970	1970-1980	1980-1990	1990-2000	2000-2010	2010-2013	1960-2013
Plastics	366%	90%	81%	50%	66%	13%	4396%

Table 9 Number of individual materials within boundaries decade by decade

Year	Boundary			Absolute dematerialization	Not available	Sum
	❶	❷	❸			
	Materialization	Weak dematerialization	Moderate dematerialization			
1960–1970	36	31	9	8	14	98
1970–1980	27	26	25	11	9	98
1980–1990	27	16	28	21	6	98
1990–2000	20	24	21	27	6	98
2000–2010	41	23	15	19	0	98
2010–2013	24	11	12	51	0	98
1960–2013	23	42	23	10	0	98

6 Discussion

6.1 *What Matters Most in the Results?*

As stated previously the objective of this work is the empirical examination of the global materials consumption in order to answer the question if it is possible to keep the global economy growing while simultaneously reducing the amount of materials resources necessary. Or to put the question in a more radical point of view: will it be necessary to stimulate degrowth to obtain an effective dematerialization of the socioeconomic system?

For the purpose of the empirical examination, we chose a basket of 98 important industrial materials that amounts to about one third of the total of materials used by humans. The large set of results presented in the graphs and tables of the previous section do not enable us to state that we are actually dematerializing nor to state that there is a strong tendency towards absolute dematerialization of industrial metabolism, nor to state that we will have significant reduction of our current environmental aggression in a short time span. However, the matter of fact is that some of our results are encouraging in suggesting that degrowth may not be necessary.

In order to see why we note that we have some encouraging results, it is necessary to analyze some details of the observed patterns. A fact to consider is that, as shown in Fig. 1, we are experiencing decoupling, which to some extent is due *to the more efficient usage of natural resources*. This phenomenon is still more evident when we look at the graph plotted in Fig. 13 exhibiting the Intensity-of-use (total metric tons/global GDP-PPP) of our 98-materials basket in the last 50 years—a clear declining trend—from 0.48 kg per dollar produced in 1960 to 0.20 kg per dollar produced in 2013. In other words nowadays each human being is producing wealth using 58 % less material resources than their parents did in the 1960s.

Regarding efficiency improvements in resource consumption, the recent results obtained by Dahmus (2014) are interesting. In his doctoral thesis at MIT, this author examines over large periods of time the historical effectiveness of efficiency improvements in reducing consumption for ten different technologies: pig iron

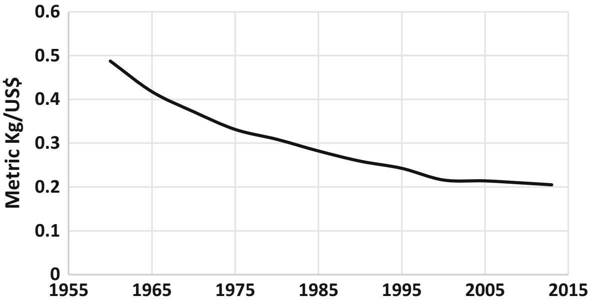


Fig. 13 Intensity-of-use of the basket of 98 materials in the period 1960–2013

production (1805–1990), aluminum production (1905–2005), nitrogen fertilizer production (1925–2000), electricity generation from coal (1925–2009), electricity generation from oil (1925–2009), electricity generation from natural gas (1925–2009), freight rail travel (1955–2009), passenger air travel (1955–2009), motor vehicle travel (1940–2009), and residential refrigeration (1960–2009). Dahmus claims that the historical evolution of reducing consumption is the result of the complex interplay between technological innovation, market forces, and government policy.

Dahmus results resembles our results to some extent when considered as the mean value over the large timespan observed—he concludes that none of the technologies investigated presented efficiency increases outpacing the quantity of goods and services provided. However, when examined decade by decade (as we did), there do exist examples of decade-long time periods in which efficiency improvements were able to match or outpace increases in quantity—noteworthy, these positive periods have happened mostly in recent times (2000–2009,[2] with the exception of motor vehicle travel), what again resembles our results when we look at the penultimate column of Tables 4, 5, 6, 7 and 8. The author concludes optimistically stating that with appropriate incentives, including for instance efficiency mandates and price mechanisms, future resource consumption, and its associated environmental impacts, could be stabilized and even reduced.

The most important features to be inferred from the set of results presented in Sect. 5 can be summarized as follows (not necessarily in order of importance):

1. Global growth (1960–2013) in materials consumption has accompanied vis-a-vis global output per capita (about 200 %, Table 1) and was mainly due to population increase in the same period, and can't be attributed to GDP growth, for as seen in Fig. 12a, per capita global material consumption remained practically constant from 1960 to 2000, when increased about only 27 %;
2. Looking at the individual materials groups (Table 2) we see that the most marked growths occurred for plastics (4396 %) and rocks & stones (1040 %).

[2]His decade of outpacing however matches our materializing decade and decade by decade comparisons can obviously be affected by economic disruption such as occurred globally starting in late 2007.

Plastics are treated in point 6. Regarding the latter, we see in Table 6 that industrial diamonds is the material exhibiting the largest growth rate (~21,000 %) among all materials investigated. Industrial diamonds, invented in the 1950s, has been intensively valued for their hardness and thermal conductivity, and has found innumerous applications as the ideal material for cutting and grinding tools, and after the 1980s has been increasingly used as heat sink for integrated electronic circuits, what has triggered its strong growth in consumption. In addition, the volume of production (about 900 tonnes in 2014) is insignificant compared with the total volume we are dealing with in this paper, as well as the fact that 90 % of this production happened in China. Another material exhibiting a large percent variation (~925 %) is bauxite, which as we know is the raw material for aluminum production, to be handled in the next paragraph.

3. Among the nine materials groups investigated, three of them account together for 90 % of the total tonnage reported for 2013 in our study, namely minerals (share 40.8 %), wood (35.8 %), and metals (13.9 %), as demonstrated in Table 2. Wood, as seen in Fig. 4, is the group with the steepest decline since 1960 and seems to continue this trend.[3] Minerals will be the subject of analysis in point 5 below. Metals is the largest set of materials (31 individual materials) studied, being to note that among them circa 6 metals exhibit percentage variations between ~1000 % and ~2800 %, namely aluminum (960 %), indium (1068 %) , niobium (2275 %), platinum (1031 %), strontium (2796 %), and vanadium (1467 %). But the most important aspect to be pointed out for this small set of materials with the largest percentage variations is the fact that their variations have fallen to one or two digits in the period 2010–2013, or even 'dematerialized' (as for instance niobium and platinum, both with ~5 %). This decline may have several reasons, from changing technologies to materials substitution and/or displacement. As an example of the former we can mention the case of strontium (the largest percentage variation between 1960 and 2013 among the metals): at the peak of the production of television cathode ray tubes (CRTs) in the 1970s up to 75 % of its use was directed to make faceplate glasses, but with the displacement of CRTs in TV sets its consumption has declined dramatically. As for the latter, the fact that aluminum is displacing steel in many applications, in automobile and aerospace industries for instance, is important in reducing the usage of several alloying elements, as niobium and vanadium among others. Aluminum is the typical example of a welcome substitution, for besides being a lightweight material, it is easily recycled at a cost of about one tenth of the cost of primary aluminum. In despite of its almost ubiquitous application in modern engineering and increasing consumption, aluminum production in 2013 represented only 0.23 % of the total tonnage of the basket of 98 materials, while steel still appears with a share of 7.9 % demonstrating that even at equal volume, steel is much more widely utilized than the more expensive aluminum.

[3]We also note that wood is being substituted at least partly as a fuel source by petroleum products which are not covered in the 98 materials basket.

4. The trends in materials usage in the group of cellulose & derivatives is interesting to look at sequentially. Figure 12b evidences the fact that the global per capita consumption of this group increased in the period 1960–1995, but stabilized and even decreased in the period 1995–2013. Looking at Table 4 we can clearly observe a predominance of black boxes and/or dark gray in the first four decades, switching to gray light or white boxes in the last two columns, reflecting a clear transition to boundary (3) moderate to strong dematerialization in more recent periods. As in other cases already pointed out as a positive trend in our analysis, the material presenting the larger percentage variation is recovered paper, which substitutes for new paper and thus is a force for dematerialization not materialization. The second largest variation is for printing and writing paper and the third for paper and paperboard (packaging), which accounts for a share of about 38 % among this group in 2013. Among the trends that have contributed to the decline of paper consumption we can mention two strong cases of technological displacements: firstly the demise of the market for photographic paper because of the shift in disseminating and storing photographic images in electronic form, and secondly the fact that online publication of newspapers, magazines, and e-books are now rapidly proliferating worldwide, bookstores are being closed, many newspapers have ceased to publish paper editions (or reduced significantly circulation of printed editions), etc., resulting in a reduction of the global consumption of about 26 % between 2000 and 2013. In addition, a myriad of blogs and discussion forums have emerged everywhere and are functioning as information sources for everyday life. Enormous facilities now exist to access information through mobile phones, tablets, notebooks, etc. A potentially telling example is the recent fact that after 244 years, the Encyclopedia Britannica has gone out of print. Before the dawn of the new century some authors (Ausubel and Sladovich 1990; Ausubel and Waggoner 2008) have pointed out that total printing paper consumption has soared despite claims that the electronic information revolution would create a 'paperless office'. In fact, amidst the information revolution brought by computational and information storage capability, paper remained a significant carrier of information, for new information storage technologies seem to supplement the range and augment the amount of information stored, rather than reduce the use of paper. However, the numbers concerning consumption of writing paper and paper for newsprint, as well as qualitative observation of recent events, suggest that a significant positive change may now be underway.
5. The second largest group investigated, minerals (26 materials), also deserves a close look. In this group we have the large percentage variation (4746 %) for rare earths (not properly rare on earth surface, but instead the word rare here means 'difficult' and costly to refine and process), which is really a subset of materials consisting of the 15 lanthanides plus scandium and yttrium. This family of materials has found a very large spectrum of applications in a myriad of new and advanced technologies that would be out of place to describe them in detail in this work. The interesting indicated result is that in despite of the broad range of technological applications their consumption has also decreased (overcoming

the boundary (3)) and their tonnage is insignificant (about 110×10^3 metric tons) compared with the totals considered in this research. Three other materials in this group have evidenced a marked percentage increase: cement (1189 %), feldspar (1250 %) and wollastonite (1612 %). Cement is a special case that will be analyzed in detail in Sect. 6.2, for it is a dominant construction material, represents the highest tonnage share (~20 %) among the 98 materials studied (48.5 % among the minerals), and exhibit the highest (24 %) increasing variation in the group over the recent period (2010–2013). Wollastonite is a mineral with several applications in the ceramics industry (about 40 % of the market), polymers (35 %, plastics and rubber), paint (15 %), and metallurgy (<10 %), but also its tonnage is not significant (about 500×10^3 metric tons in 2013) in our context. Feldspar is an aluminum silicate mineral with a broad range of applications in glassmaking, ceramics, and to some minor extent is used as filler in paint and in the polymer industry, but as can be seen in Table 6 is a material exhibiting a moderate dematerialization (3 %).

6. As seen in Fig. 11 and Table 2 plastics is the group evidencing the highest and steepest growth rate in the last 50 years. Smil (2014) observes that "plastics have been widely seen as the quintessential twentieth century material, with a particularly rapid post-World-War II diffusion as they replaced wood, metals, and glass in many household, industrial, and transportation products". We can add that plastics, perhaps better to say polymers, is a ubiquitous presence in our everyday life, being present all around us in our homes, clothes, cell phones, computers, packaging, cars, and even in airplanes. Not forgetting that the now also ubiquitous LCD screens and touch screens have polymers as important components. Their production increased six-fold in the decade 1950–1960, and 50-fold in the following 50 years, and seem to be maintaining this growth trend. Plastics can be divided into two large groups: thermoplastics (which can be softened with moderate heating, are easy to process and most of them can be recycled) and thermosets (which cannot be softened with moderate heating and are difficult to recycle). The former account today for nearly 80 % of the global production, further divided into several products, of which the most important are: polyethylene (PE, about 30 % of world's aggregate plastic output), polypropylene (PP, ~20 %), and polyvinyl chloride (PVC, ~12 %). These three dominant polymers represent today three-fifths of the global production of plastics. The output of thermosets is dominated by two important materials: polyurethanes (largely used for thermal insulation, among several other applications) and epoxy resins (a paramount material for the production of high strength composites). There is today a debate about the advantages and disadvantages on the usage of plastics, strongly dependent on the perception of consumers, producers or environmentalists. Environmentalists note at least three important environmental menaces: synthesis of plastics need the use of fossil fuels (mainly natural gas) as well as the use of hazardous chemicals (the case of PVC) and their disposal in land or in the oceans after use is a very critical and yet unsolved problem. Last but not least plastics have a limited lifespan in terms of functional integrity, usually ranging from 3 to 10 years (PVC can last

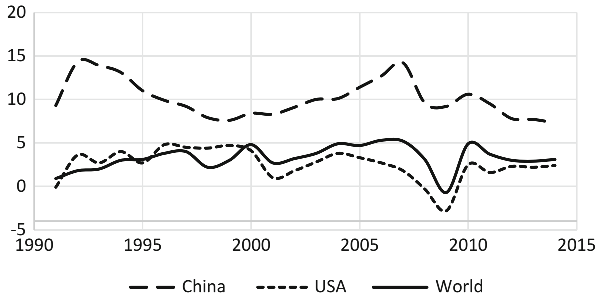
longer, about 2 or 3 decades), a negative aspect regarding increasing consumption, for if short-lived they promote the production of more goods for substitution (Lokensgard 2010). However, substantial progress has been made regarding alternative synthesis methods (the so called Green Plastics, Stevens 2002), biodegradability, and recycling (Goodship 2007, 2010), and will certainly help to mitigate these environmental threats at least to some extent. Despite their large range of applications, plastics account for a share of only 1.45 % (corresponding to a global per capita consumption of circa 42 kg) of the total tonnage analyzed in this work, and this fact points toward a positive aspect regarding their consumption: being a very lightweight material, which has replaced a broad spectrum of much heavier materials, their intensive usage induces a cascade-reducing effect in respect to transport costs, fuel consumption, and handling facilities.

6.2 The 2000–2010 Spike

As noted in the last paragraph of Sect. 5, the most striking aspect of our results regards the striking increase in consumption at the dawn of the new century, against the slightly decreasing trend in the previous decades. This increase seems to reverse again in the four most recent years, a fact clearly translated in the behavior observed in the curves of Fig. 12. Calling this sudden increase a a 'spike' is perhaps overly strong, but in any case we have in the period 2000–2010 indications of a new driving force that has markedly pushed up global material consumption.

It can be shown that the new driving force was China's modernization, a process that has started slowly in the 1980s, but accelerated enormously in the 1990s—in 1992 China's annual GDP growth reached the incredible mark of 14,3 %. After a slight decline in the second half of the 1990s, China's output continued to growth unabated in the first decade of the twenty-first century maintaining the two digits mark during most of this period. Importantly, during this period, China's global economic significance of course increased and in 2010 China became the world's second largest economy, and now more than twice as large as that of Japan. Figure 14 below shows comparatively the annual percentage growth rate for

Fig. 14 Annual percentage growth rate of GDP at market prices based on constant local currency (World Bank 2016)



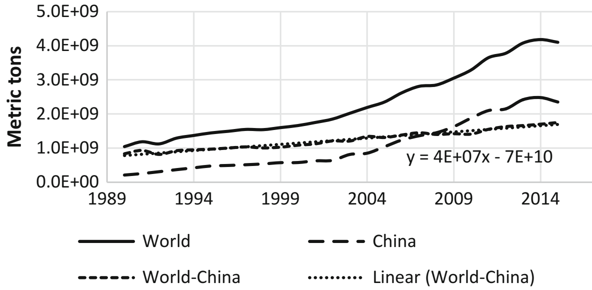


Fig. 15 Comparative cement production—world's total, China's total, and world without China.

China, USA and the World; even during the worldwide economic crisis in 2008–2009 China's economic growth continued at rates over 9 %. Looking back such high level economic growth was higher than the levels achieved during the previous peak-growth spells of Japan or South Korea.[4]

Between 1980 and 2010 China experienced four successive doublings of its economy and this evidently implied in an identical increase of all material flows, and consequently also to the necessity to resort to massive materials imports. This unprecedented growth also involved the establishment of an immense infrastructure network and logically also in a huge expansion of civil construction—no doubt the world's largest expansion of housing, commercial buildings, industries, roads, etc., which carried with it a huge demand for building materials. China became not only the second world's largest economy, but also the world's leading manufacturer and importer of commodities.

This huge growth has impacted the world economy and diminishes any dematerialization trend if one were happening in the second half of the twentieth century. As mentioned in Sect. 4 we choose the case of cement to demonstrate this impact, comparing China's, USA's and world's demand for cement, the material contributing with a share of almost 20 % to our basket of 98 materials. Cement is the dominant construction material, necessary to make concrete, and so has a unique application, unlike other materials of our basket, whose industrial applications are dispersed into a very large number of products. This fact makes its analysis regarding the consumption/production at the level of a nation or region much simpler, for it is far easier to survey data on import and export, which is for example very difficult in the case of plastics, used in a myriad of products.

Smil (2014) cites a very interesting comparison that illustrates the scale of China's 'concretization': "consumption of cement in the USA totaled about 4.56 Gt during the entire twentieth century—while China emplaced more cement (4.9 Gt) in new construction in just 3 years between 2008 and 2010, and in the 3 years between 2009 and 2011 it used even more, 5.5 Gt". Figure 15 shows the comparison between China's and worldwide cement production with that of the world without China in the period 1990–2015. China's production in 2014–2015 corresponds to

[4]We recognize that some observers are now questioning whether growth in recent years is in fact as high as reported by China.

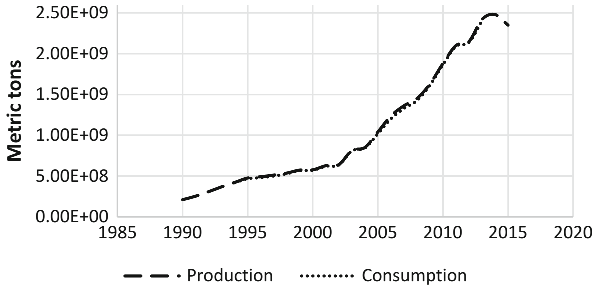


Fig. 16 Residual difference between cement production and consumption in China

about 57 % of the world production; it's evident that China's production has significantly pulled up the world total production. Global cement production minus China's production has grown almost linearly at a growth rate of only about 40 Mt annually.

In Fig. 16, a comparison between China's total cement production and China's own consumption between 1990 and 2015, obtained subtracting the exports from the total production; we can see that the difference is approximately nil, that is, practically all production was consumed in building China's infrastructures. Figure 17 shows, for the period 1990–2013, the cement per capita consumption in the USA, China, and in the world (without China and USA). For the determination of the per capita consumption in the USA was considered the net consumption, that is, the balance between total production, imports and exports. As can be observed China's per capita consumption quintupled between 1994 and 2013, while per capita consumption in the USA and the world has not changed significantly -consumption in the USA has even decreased slightly. In 2015 every Chinese was consuming about 1.71 tonnes of cement, while every American was consuming about 0.25 tonnes.

Naturally, such a large boom in construction boosted the consumption and imports of a bandwagon of other materials used in building, as lime, steel, iron, plate glass, aluminum, etc., as well as the corresponding minerals and rocks necessary for their production. For these other materials we have found similar results, that is, an accentuated increasing in consumption in China in the first decade of this century.

Concluding, China's modernization indeed provoked a very impressive imbalance in the global struggle of materialization vs. dematerialization. The results show that this country is using a lot of more materials per capita as its citizens increasingly can afford and demand better housing and standards of living. But much of this is a one-time building boom that will subside as Chinese housing stocks and infrastructures reach modern standards—there is no reason to anticipate that China will not follow the pattern shown in the rest of the world as that boom subsides. The last segment of Fig. 17 (2014–2015) suggest that this may be happening. However, the extent they will be replaced by other fast growing economies (India etc.) now emerging towards higher standards of living is going to be important to future dematerialization results.

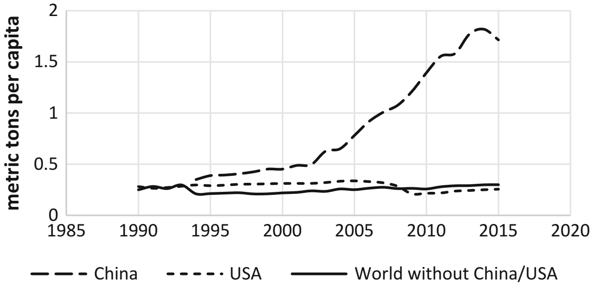


Fig. 17 Per capita cement consumption in the USA, China and the rest of the world (without China and USA)

7 Concluding Remarks

The objective of this paper was to survey the worldwide production and consumption of a basket of 98 materials most used in modern engineering applications in order to find patterns of dematerialization and or materialization during the last half century. The results do not enable us to state that human society is 'dematerializing', but we have pointed out that we have some positive trends which allow us to be optimistic regarding a reduced consumption of materials in the future and less harm to the environment as well. In a nutshell, the patterns indicate that business is not simply proceeding as usual, for the ever increasing technological capability is contributing to an ever improving efficiency in the usage of materials. Figures 12 and 13 translate very well our hopeful vision of a future scenario that, if not showing absolute dematerialization, at least evidencing a stabilized and sustainable path allowing economic growth without continuing increased material consumption. As shown in a previous paper by two of the authors, achieving this will take some decrease in demand elasticity as time goes on as well as increasing technological capability. The evidence for the importance of China in cement production may well be the kind of mechanism for obtaining long run reduction of demand elasticity.

Some may argue that the result presented in Fig. 12 (the 51 materials evidencing dematerialization) is mainly due to a global picture of economic degrowth or recession in the last couple of years. But regarding this point it is important to consider that many economists are suggesting that we are experiencing neither a global recession, nor a normal economic crisis, but instead that the global economy will not return to the roaring economic growth verified in several periods of the last 60 years. Many books (Gordon 2016; Heinberg 2011; Rubin 2012; Galbraith 2014) and technical reports (Summers 2016; Lang et al. 2016; Buchanan 2016) appeared in recent years claiming that economic growth is not happening as it used to be, in other words, wealth is being created without necessarily more material consumption—our analysis demonstrated that, from the basket of 98 materials studied, humans needed to consume 480 g to produce one dollar output in 1960, and now can produce the same dollar with only 200 g.

Acknowledgement Authors T. Devezas and A. Vaz aknowledge the financial support of the Portuguese Foundation for Science and Technology (FCT) through the Research Centre for Mechanical and Aerospace Science and Technologies (C-MAST, UID/EMS/00151/2013) and of the European Regional Development Fund and of the Commission for the Coordination and Development of the Central Region of Portugal (CCDRC) via project EMaDeS (CENTRO-01-0145-FEDER-000017).

References

Ausubel, J. H., & Sladovich, H. E. (1990). Dematerialization. *Technological Forecasting & Social Change, 37*, 333–348.

Ausubel, J. H., & Waggoner, P. E. (2008). Dematerialization: Variety, caution, and persistence. *Proceedings of the National Academy of Sciences (PNAS), 105*, 12774–12779.

Bernardini, O., & Galli, R. (1993, May). Dematerialization: Long-term trends in the intensity of use of materials and energy. *Futures, 25*, 432–448.

Buchanan, M. (2016). *Humanity's race against itself*. BloombergView (February 16).

Brookes, L. G. (1984). Long-term equilibrium effects of constraints in energy supply. In L. Brookes & H. Motamen (Eds.), *The Economics of Nuclear Energy*. London: Chapman and Hall.

Brookes, L. G. (2000). Energy efficiency fallacies revisited. *Energy Policy, 28*, 6–7.

Dahmus, J. (2014). Can efficiency improvements reduce resource consumption? *Journal of Industrial Ecology, 18*(6), 883–897.

Galbraith, J. K. (2014). *The end of normal: The great crisis and the future of growth*. New York: Simon & Schuster.

GFN. (2014). *Global Footprint Network (on line)*. Retrieved April 4, 2014, from http://www.footprintnetwork.org.

Goodship, V. (2007). *Introduction to plastics recycling* (2nd ed.). London: Smithers Rapra.

Goodship, V. (2010). *The instant expert: Plastics, processing, and properties*. Toronto: ChemTec Publishing.

Gordon, R. (2016). *The rise and fall of American growth*. Princeton: Princeton University Press.

Heinberg, R. (2011). *The end of growth: Adapting to our new economic reality*. Gabriola Island, Canada: New Society Pub.

IRP. (2011). International Resource Panel (UNEP). Decoupling—natural resource use and environmental impacts from economic growth. United Nations Environmental Programme.

IRP. (2014). International Resource Panel (UNEP). Decoupling 2—Technologies, Opportunities, and Policy Options. United Nations Environmental Programme.

Jevons, W. S. (1865). The coal question: Can Britain survive? In A. W. Flux (Ed.), *The Coal question: An inquiry concerning the progress of the nation, and the probable exhaustion of our coal-mines*. New York: Augustus M. Kelley.

Kander, A. (2005). Baumol's disease and dematerialization of the economy. *Ecological Economics, 55*, 119–130.

Khazzoom, J. D. (1980). Economic implications of mandated efficiency in standards for household appliances. *Energy Journal, 1*(4), 21–40.

Krausmann, F., et al. (2009). Growth in global materials use, GDP and population during the 20th century. *Ecological Economics, 68*(10), 2696–2705.

Lang, S., Putz, P., & Kopp, T. (2016). *Do Maure economies grow exponentially*?. arXiv: 1601.04028.

Lokensgard, E. (2010). *Industrial plastics—Theory and applications* (5th ed.). New York: Delmar, Congage Learning.

Magee, C. L., & Devezas, T. (2016). *A simple extension of dematerialization theory: Incorporation of technical progress and the rebound effect*. arXiv: 1602.00090, 26 pgs (to be published in Technological Forecasting & Social Change)

Nagy, B., Farmer, J. D., Bui, Q. M., & Trancik, J. E. (2013). Statistical basis for predicting technological progresso. *PLoS One, 8*, 1–7.

Rubin, J. (2012). *The end of growth*. Canada: Random House.

Smil, V. (2014). *Making the Modern World: Materials and dematerialization*. New Jersey: Wiley.

Summers, L. (2016). The age of secular stagnation: What it is and what to do about it. *Foreign Affairs,* February 15.

Steinberger, J. K., Krausman, F., & Eisenmenger, N. (2010). Global patterns of materials use: A socioeconomic and geophysical analysis. *Ecological Economics, 69*, 1148–1158.

Ziolkowska, J. R., & Ziolkowski, B. (2011). Product generational dematerialization indicator: A case of crude oil in the global economy. *Energy, 36*, 5925–5934.

Part III
Entrepreneurship Development

Challenges in Technology Entrepreneurship: Managing Inter-professional Conflicts in Biopharmaceutical Enterprises

Călin Gurău

Abstract Biopharmaceutical enterprises represent classical examples of knowledge organisations that integrate multiple professional cultures. The two main professional groups that should cooperate for the success of the new biopharmaceutical venture—business people and scientists—have, however, different sets of values, which can determine intra-organisational tensions and conflicts. The present study attempts to identify the existing conflicts between the professional sub-cultures that coexist in a biopharmaceutical SME, and to analyse the conflict management procedures employed by the manager-entrepreneur to reconcile the interests, requirements and visions of various groups that work in the organisation.

Keywords Professional groups • Biopharmaceutical SMEs • Inter-group tensions and conflicts • Conflict management • Managerial procedures

1 Introduction

Organizations are often multi-cultural systems (Lumineau et al. 2015). Sociologists who studied cultural issues in the organizational context found evidences of heterogeneous systems, comparable to the social heterogeneity of different categories of workers (Gregory 1983; Herkenhoff and Heydenfeldt 2011; Martin and Siehl 1983; Martin et al. 1985; Van Maanen and Barley 1985). Individuals embody and express professional (occupational) cultures, and sometimes the culture of specific social classes (e.g. working class). Therefore, organizational culture does not exist independently of the company employees, but is constructed by them, as the result of cultural exchanges and co-existence between different professional and social groups within the company (Irrmann 2002; Mudambi and Swift 2009).

An extensive study conducted between 1986 and 1994, by Francfort et al. (1995), on a sample of 262 organizations, evidenced the existence of multiple cultural identities. Only 34 % of their sample could be identified as having a

C. Gurău (✉)
Montpellier Business School, Montpellier Research in Management, 2300 Avenue des Moulins, 34185 Montpellier, France
e-mail: c.gurau@montpellier-bs.com

T. Devezas et al. (eds.), *Industry 4.0*, Studies on Entrepreneurship, Structural Change and Industrial Dynamics, DOI 10.1007/978-3-319-49604-7_15

unifying organizational culture (which does not means homogeneous culture), generally in firms having a limited professional identity differentiation, such as SMEs exposed to strong market uncertainties, or multinationals promoting a discourse of collective challenge rewarded by promotion. In 42 % of the cases there were several cultural groups within the same organization, often strongly defined but with no symbolic link between them. Finally, the case of cultural disintegration was found in 24 % of the investigated organizations, corresponding to an old and very cohesive culture that was progressively disintegrating due to important changes in the socio-economic context of the firm.

Biopharmaceutical enterprises represent classical examples of knowledge organizations (Hodgson 2001) that integrate multiple professional cultures. Mehta (2004) argues that the creation of a new biopharmaceutical firm can be the initiative of two types of entrepreneurs: the inventor of technology and the market perceiver (see Fig. 1). In reality, in most cases a successful biotech venture is the result of the collaboration between a scientist and a marketer. This team of entrepreneurs has to combine and apply their complementary competencies to lead the newly-created firm through its development stages (see Table 1).

The two main professional groups that should cooperate for the success of the new biopharmaceutical venture—business people and scientists—have, however, different sets of values. At the heart of the challenges facing the manager-

Fig. 1 New venture creation process in biopharmaceutical industry (Mehta 2004)

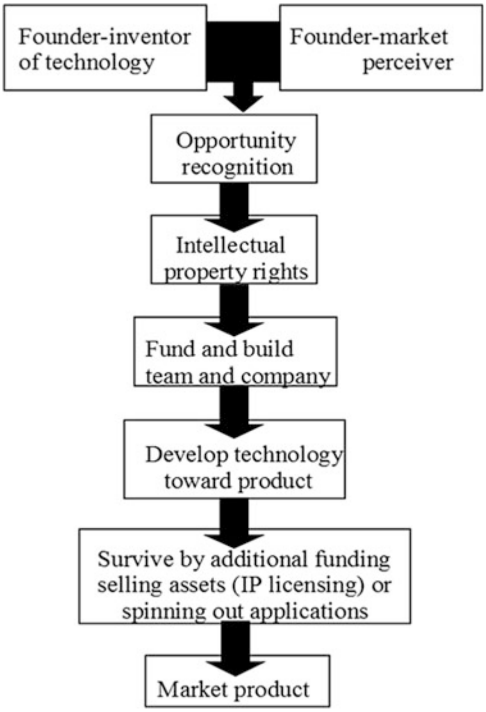


Table 1 Strengths and weaknesses of market perceivers and technopreneurs (adapted from Mehta 2004)

Venture stage	Market perceiver	Technopreneur
Recognize opportunity	**Has:** A comprehensive knowledge of markets. Clear definition of product characteristics. **Needs:** To find appropriate technology to meet the market's specific needs.	**Has:** A solid understanding of and expertise in specific, well-characterized technology. Established credibility with peers, investors and customers (academic researchers, biotechs and big pharma). **Needs:** To confirm that here is a market for the product. To define product characteristics.
Secure Intellectual Property rights	**Has:** A clear understanding of the market applications, so IP claims can be formulated easily. **Needs:** To comprehensively harvest IP portfolio.	**Has:** A strong position to easily license his/her own invention from the university into the start-up. **Needs:** To erase any perceived conflict of interest that may arise by being on both sides of license negotiation, as academic inventor and employee and company executive. An understanding of future IP needs.
Fund team and build company	**Has:** Business credibility in creating sound commercialization and business plan with clear market needs. **Needs:** To carefully evaluate timeline for technology development, balancing the attraction of large markets with a plan for growth as technology matures. Strong history of experience to overcome credibility gap if technology is licensed in without inventor participation.	**Has:** Strength in early phases of company, where main efforts are on research and most of the personnel are technically oriented. Credibility with investors due to technical expertise. **Needs:** To manage investors' questions in business and commercial areas. To learn how to manage non-scientists.
Develop technology to product	**Has:** An external perspective that brings a strong product-focus to the development process. Focus on scaling up production to commercial manufacturing levels, market acceptance and regulatory acceptance. **Needs:** Understanding of the transition from R&D to commercial manufacture.	**Needs:** Experience of commercial product development, particularly issues in scaling up. Unbiased perspective to evaluate the technology's realistic potential versus its elegance.
Survive	**Has:** Sensitivity to the needs of the business and finances so as to strategize and manage IP assets astutely.	**Needs:** To understand that his or her appropriate position within growing company may not be at the helm, but in a specific technical leadership position, such as CTO, CSO or on the Scientific Advisory Board.
Market	**Has:** Ability to deliver the market potential message to multiple stakeholders on their terms, put a commercial team in place and strike appropriate business partnerships.	**Needs:** To shift focus from developing technology to building a strong commercial team speedily and efficiently.

entrepreneur is a clash between two different rationales concerning the purpose of science (Ernø-Kjølhede et al. 2001).

The present study attempts to identify the existing conflicts between the professional sub-cultures that coexist in a biopharmaceutical SME, and to discuss the conflict management procedures employed by the manager-entrepreneur to reconcile the interests and the vision of various groups within the organization. After a succinct description of the business environment of the biopharmaceutical venture—in the context of a particular knowledge-intensive sector, we develop a series of research hypotheses that link the conflict situations and the conflict resolution procedures with the specific profile (business orientation) and evolution (lifecycle stages and growth) of the biopharmaceutical venture. The next section discusses the research methodology applied to collect and analyze secondary and primary data, followed by a presentation of results in relation to the formulated research hypotheses. The study concludes with a summary of the main findings and research limitations, and with propositions for future research (see Table 2).

2 The Business Environment of Biopharmaceutical Organizations

The biopharmaceutical firm represents a particular working environment, in which the scientific skills should be closely integrated with business decisions at every stage of the product development cycle.

For most innovative therapeutic drugs, the development process is extremely long, risky and expensive (Phrma 2015). The development of new medicines requires the screening of a huge number of initial substance candidates (the success ratio is usually one in 4000); it can take 10 years or longer until is launched on the market; and costs more than $500 million (DiMasi and Grabowski 2007). The complete value-added chain of activities between the generation of a new idea and the commercialization of the final product can be divided into three main stages: product innovation, product development and product commercialization (Evans

Table 2 The two perspectives on science (adapted from Ernø-Kjølhede et al. 2001)

The individualist perspective	The high-skilled employee perspective
The researcher is a 'self-employed' person who motivates himself	The researcher is an employee who sometimes needs to be motivated
The researcher must be autonomous and free to set his own agenda for research: free thinking is the basis for creativity and originality	The researcher must integrate his research agenda with the desires of stakeholders: free and institutional thinking
Research is a personal calling for the few; it is a highly elitist and unique activity	Research is a professional calling; it is a craft that can be taught
Researchers are individualists and loners	Researches are individualists and team players

and Varaiya 2003; Walsh 1993). Each of these stages comprises a series of essential activities and processes (Chandrasekar et al. 1999; Phrma 2015):

Stage I. Innovation:

- generating the seed-idea;
- testing the scientific originality and validity of the seed-idea;
- gathering the necessary human, technological and financial resources for the development of the original idea;
- creating the product prototype;
- patenting the product prototype and/or the production process.

Stage II. Development:

- organizing and starting a pilot-scale manufacturing;
- testing the product;

In the case of therapeutic drugs, these trials are organized in a series of successive sub-stages:

- preclinical study: before testing the product on humans, laboratory and animal studies are carried out to determine safety and biological activity;
- clinical evaluation: on the basis of the preclinical tests' results, the company may obtain an authorization to carry on tests on humans:
- phase one: safety trial—a small-scale safety trial is realized on healthy volunteers;
- phase two: safety and efficacy trial—after the initial trial is finished, the company begins a larger trial to assess whether the biopharmaceutical appears to be effective in treating the targeted disease or medical condition; unlike the phase one trial, this is conducted on patients.
- phase three: controlled safety and efficacy—products that display efficacy in the phase two trial are moved into a larger clinical evaluation setting to verify the results.

A phase three trial may require hundreds or even thousands of patients, depending on the disease being studied and the requirements for the statistical evaluation.

- collecting the results of trials and submitting an application for the product approval to the specialized governmental authorities;
- obtaining the approval of the specialized governmental authority for the mass commercialization of the product. The company must continue to monitor the product of any adverse reactions and to report these to the specialized governmental authorities.

Stage III. Commercialization:

- initiating the market research process.

Theoretically, it can be argued that market research should take place before the innovation stage. However, in most cases, biotechnology discoveries are made on the basis of a potential demand (Walsh 1993), having to create their own market in the commercialization stage.

- identifying the needs of potential customers and defining the main customer segment;
- designing an appropriate marketing-mix strategy (product, price, distribution and communication), to transform customer needs in a manifest market demand;
- applying the marketing-mix strategy for the targeted market segment;
- developing the market and improving the offer.

Most of the biopharmaceutical SMEs are concentrating their activities on the first two stages of the product development process (Saviotti 1998; Senker 1998). The reasons of this focus are related with the specific structure of the biopharmaceutical industry. The pharmaceutical sector has a well-developed infrastructure for product commercialization, controlled and managed by specialized traditional organizations (pharmacies, retail stores, hospitals, national health agencies). The newly-developed biopharmaceutical ventures provided an alternative development cycle of new drugs, to the already established channels controlled by large pharmaceutical corporations (Azzone and Pozza 2003). Relaying on a specialized knowledge of genetic engineering techniques and on high levels of innovativeness, these firms have fragmented the traditional value-added chain of the pharmaceutical industry (Pisano 2006).

There are also biopharmaceutical products for which the R&D process is simpler and less expensive: such as diagnostic kits or vaccines. On the other hand, the need for supporting services, reagents and laboratory instruments, has determined the development of many specialized service providers for the dedicated biopharmaceutical companies (the so-called platform technology enterprises).

The main competitive advantages of entrepreneurial biopharmaceutical firms are flexibility and innovativeness. Many research workers from biopharmaceutical firms have moved into business from academia, attracted by the potential rewards of commercial activities (Briggs 1994). These scientists have created within entrepreneurial organizations a highly stimulating intellectual environment. The R&D structure in many biopharmaceutical SMEs is similar to university research teams—instead of being focused around specific products, research teams are based around techniques or special areas of research (Pisano 2006), and scientists enjoy a high level of freedom and independence.

Studies have shown that the high innovativeness of biopharmaceutical SMEs is significantly depreciated if the firm is acquired and integrated in a larger, more bureaucratic organization (Briggs 1994). This situation explains the increased popularity of strategic alliances and research partnerships between small firms and large pharmaceutical corporations (Pisano 2006).

In this highly complex business environment, the entrepreneur-manager of a biopharmaceutical venture faces two conflicting challenges: on one hand, he has to integrate the business and the scientific cultures existent within the organization, solving the emerging professional conflicts and creating a common ground for dialogue and co-operation; on the other hand, he has to preserve and enhance the innovative culture of the organization, in order to insure the competitive advantage of the firm.

Considering this particular organizational context of the biopharmaceutical firm, this study focuses on the following research objectives:

1. To identify the existence of professional sub-cultures in an entrepreneurial biopharmaceutical firm, and to define their specific profile.
2. To investigate the conflicts that can develop between the interests, values and working styles of various professional groups.
3. To describe the conflict management procedures used by the entrepreneurial team to reconcile the professional sub-cultures existing within a biopharmaceutical firm.

In the next section, we develop a series of research hypotheses, in direct relationship with the research objectives presented above.

3 Research Hypotheses

In the biopharmaceutical sector, knowledge represents an essential asset for business competitiveness. However, biopharmaceutical knowledge is multi-dimensional and complex: on one hand it represents an organic integration of scientific and business knowledge; on the other hand it is a combination of codified knowledge (patents, research and clinical trial procedures), and human capital.

The critical moments for knowledge management are linked with the stages of strategic re-structuring of the entrepreneurial organization. Usually, a biopharmaceutical firm passes through a number of stages in its lifecycle: in the first instance, the entrepreneurial team will attempt to build its intellectual property (IP) portfolio, and to create a good professional team; in the second stage, the management team will establish strategic objectives and will initiate the process of product or service development; finally, the successful development of a final product/service will enable the firm to initiate the commercialization stage. The professional roles and cultures within the biopharmaceutical enterprise, as well as the relationships between various professional groups, will evolve and change during these stages. In this context, the following research hypothesis can be formulated:

H1 Conflict situations between professional groups will evolve during the lifecycle stages of the biopharmaceutical firm.

Another important change in the organizational structure is determined by the growth of the company. The transition from a micro-firm (with up to 10 employees), to a small firm (with 11–50 employees), and then to a medium-sized firm (51–250 employees), will transform the entrepreneurial culture, creating new relations among employees and determining specific tensions between professional groups. Therefore, a second research hypothesis can be formulated:

H2 Conflict situations between professional groups will evolve with the growth of the biopharmaceutical firm.

The business orientation of the firm is also capable to create specific reasons for conflict—in the case of biopharmaceutical firms, there might be important differences between the drug-development organizations—that are confronted with a long and risky R&D process, and platform technology firms—that get integrated quicker in the value-added chain of biopharmaceutical activities, providing specialized—and sometimes standardized—services for other pharmaceutical organizations. Taking into account these differences, we formulate the following research hypothesis:

H3 Conflict situations between professional groups will be influenced by the business orientation of the biopharmaceutical firm.

As a logic consequence of the relationship between inter-professional conflicts and biopharmaceutical firm's profile, it can also be expected a variation in the solutions provided by the manager in these situations—in other words:

H1a The conflict management procedures will evolve during the lifecycle stages of the biopharmaceutical firm.

H2a The conflict management procedures will evolve with the growth of the biopharmaceutical firm.

H3a The conflict management procedures will be influenced by the business orientation of the biopharmaceutical firm.

4 Methodology

In order to verify the validity of the research hypotheses formulated above, both secondary and primary data have been collected and analyzed. In the first stage, an extensive desk research has provided information about the specific characteristics of the biotechnology sector, and the main categories of competencies required in a biopharmaceutical firm. In the second stage of the research, the web sites of 82 UK biopharmaceutical SMEs have been accessed, using the information provided in the online database Biotechanalytics (http://www.biotechanalytics.com). The data published on these corporate web sites provided an insight into the main professional categories that coexist in a biopharmaceutical SME, the specific knowledge of each professional category, and their place in the organizational structure. Finally, in the third stage of the research project, these 82 organizations have been contacted by email or by telephone in order to organize a series of interviews with representatives of various professional groups and with the manager or the CEO of the firm. 28 companies have answered affirmatively to this demand, giving a response rate of 34.14 %—which is compares favorably with similar research projects.

In each of the investigated organizations, 3–5 semi-structured interviews have been conducted, with at least a representative of the scientific team, at least a member of the business team, and with the entrepreneur-manager of the company. The interviews lasted between 45 and 90 min, and comprised questions related to the professional cultures within the organization, the areas of conflict between various professional groups and the conflict management procedures used by the entrepreneur-manager to reconcile the differences between these professional groups.

The data obtained has been processed using both qualitative and qualitative techniques. The sources of intra-cultural conflicts and types of conflict management procedures used within organizations have been counted and cross-tabulated with the organizational characteristics of these companies (size, business orientation, and stage reached in the organizational lifecycle), in order to highlight the specific conditions in which these particular events take place, and to differentiate between different types of organizations. On the other hand, the main themes developed during the interviews have been analyzed in detail, in order to achieve a deeper understanding of the inter-cultural processes developed in these firms, the management of conflict between various types of knowledge, and the success of these strategies within the investigated organizations.

5 Presentation and Analysis of Data

5.1 The Organizational Profile of the Investigated Biopharmaceutical SMEs

Five of the investigated biopharmaceutical firms were defined as micro-enterprises (with less than 10 employees), 16 as small firms (with a personal between 11 and 50 employees), and 7 as medium-sized firms (with 51 to 250 employees). The majority were classified as drug-development firms—22, and six as platform technology/service providers. Finally, six were considered to be in the first stage of the firm development: building their intellectual property portfolio, 17 in the stage of product/service development, and the remaining five in the commercialization stage (already obtaining direct revenues from the commercialization of their products or services).

5.2 Professional Sub-cultures in Biopharmaceutical Firms

Both the data collected online and the information obtained through interviews indicated the existence of two major professional groups in biopharmaceutical firms:

(a) scientific/academic—comprising specialists and researchers in biology, biochemistry, medicine or genetics, and
(b) business professionals—people who have business/management studies and/or experience (obtained in large corporations, entrepreneurial firms, or as business consultants).

The distinction between these two sub-groups is not absolute—most of the scientists that are members of the management team have also business studies or previous a working experience in an enterprise. On the other hand, the scientists-founders of the firm are driven by an entrepreneurial spirit, which modifies their traditional professional perspective. In fact, a complete separation of company researchers and the scientific community is not beneficial for the firm. The scientists are encouraged to maintain their links with universities and research institutes, in order to obtain relevant information about recent scientific discoveries (Lam 2004). Therefore, the relation between these two main competencies—science and business—is complex, and in many cases they are in a relation of complementarity rather than conflict (Cohen et al. 1999). Well identified and applied, conflict management procedures can enhance this complementarity, increasing the competitive advantage of the biopharmaceutical firm.

5.3 The Level and the Type of Intra-organizational Professional Conflict

The information provided by interviewees indicates that, sometimes, the two main professional groups are in conflict. The main reasons for conflict are:

(a) the level and the distribution of resources—many scientists feel that the resources allocated for R&D are insufficient; on the other hand, business people often consider the demands of the research team as unreasonable and not justified by the expected results of their work;
(b) the strategic objectives—there is a heated debate between scientists and business specialists concerning the feasibility of certain strategic objectives;
(c) the timescale of projects and of their implementation—the scientists often feel pressured by marketers to obtain quick results, which is they consider unrealistic taking into account the existence of specific time requirements for biological experiments and analysis;
(d) internal communication represents a dual problem at the level of process and content—the interviewees consider that better procedures should be implemented to facilitate communication between the two professional groups, but also an adaptation of the content to the level of understanding of each party involved;
(e) the dissatisfaction with external communication is linked to the problem of strategic objectives—many scientists feel that the press releases published by

the management team are overly optimistic, neglecting or hiding the risks and uncertainties related with biopharmaceutical products;

(f) the professional responsibilities allocated to some members of the scientific teams are considered too heavy, limiting the research potential of these people;

(g) scientific autonomy—some scientists feel that their activity is too restricted by the commercial objectives of the firm, sometimes this problem is linked to the incapacity or the delay to publish the results of an important scientific discovery for intellectual property reasons; on the other hand, some managers feel that researchers should be more aware of their role in the enterprise, and respect a stricter work discipline.

At an abstract level, these reasons for conflict can be classified into two categories: conflict of values (strategic objectives, professional responsibilities, and scientific autonomy), and conflict of work style (level and distribution of resources, timescale of projects and of their implementation, internal and external communication). This categorization is relative in value, and further studies are necessary to define more clearly the main reasons for inter-professional conflicts. On the other hand, the two types of conflicts are closely related, since it can be argued that the differences in work style are manifestations of value systems' conflicts.

5.4 The Influence of Firm's Profile on the Type of Inter-professional Conflicts

Tables 3, 4 and 5 present the reasons of professional conflicts in relation to various characteristics of company's profile. Since in an organization it can be found more than one type of conflict, the data presented shows the percentage of firms in which this conflict was noticed and reported during the interview.

The stage of development influences the type of inter-professional conflicts that take place in the firm. In the first development stage—building the IP portfolio, the proportion of firms experiencing conflicts is relatively low, in comparison with the other two stages. This situation can be explained by the common aims of the two main professional groups—to codify the scientific knowledge into exclusive patents, which might create a specific synergy. However, even in this stage, half of the investigated firms have experienced conflicts related to the strategic objectives of the firm—in all the cases the debate was determined by the future business orientation of the firm—drug development or platform technology.

In the second stage of development, all firms experience multiple forms of conflict. The most frequent are again related to the strategic objectives of the firm (88.2 %), with the distribution of company's resources, timescale of projects, and eternal communication (each with 82.3 %).

Finally, for the firms that reach the commercialization stage, the frequency of some types of conflicts decline (strategic objectives and internal communication),

Table 3 The reasons of professional conflict considered in relation to the lifecycle stage of the investigated organizations

Reason of conflict/Stage of development	Building the IP portfolio		Product/ service development		Commercialization	
	N	%	N	%	N	%
Level and distribution of resources	0	0	14	82.3	3	60
Timescale of projects and of their implementation	1	16.7	14	82.3	3	60
Internal communication	1	16.7	13	76.5	1	20
External communication	1	16.7	14	82.3	2	40
Strategic objectives	3	50	15	88.2	1	20
Professional responsibilities	2	33.4	8	47.1	3	60
Scientific autonomy	1	16.7	9	52.9	4	80
Total	6	100	17	100	5	100

Table 4 The reasons of professional conflict considered in relation to the size of organizations

Reason of conflict/Firm size	Micro-enterprise		Small firm		Medium-sized firm	
	N	%	N	%	N	%
Level and distribution of resources	1	20	12	75	4	57.1
Timescale of projects and of their implementation	0	0	15	93.7	3	42.8
Internal communication	0	0	13	81.2	2	28.6
External communication	0	0	15	93.7	2	28.6
Strategic objectives	2	40	14	87.5	3	42.8
Professional responsibilities	1	20	11	68.7	1	14.3
Lack of scientific autonomy	0	0	9	56.2	5	71.4
Total of enterprises	5	100	16	100	7	100

Table 5 The reasons of professional conflict considered in relation to the business orientation of biopharmaceutical firms

Reason of conflict/Business orientation	Drug development		Platform technology	
	N	%	N	%
Level and distribution of resources	15	68.2	2	33.3
Strategic objectives	16	72.7	3	50
Timescale of projects and of their implementation	18	81.8	0	0
Internal communication	12	54.5	3	50
External communication	16	72.7	1	16.7
Professional responsibilities	13	59.1	0	0
Lack of scientific autonomy	9	40.9	5	83.3
Total	22	100	6	100

but others remain high or even grow in frequency (e.g. scientific autonomy). Considering the data presented in Table 3, the research hypothesis:

H1 Conflict situations between professional groups will evolve during the development stages of the biopharmaceutical firm.

, is validated.

The size of organizations influences as well the type of conflicts between professional groups (see Table 4). The micro-organizations experience the lower frequency and variety of conflicts. This phenomenon was already identified by other researchers (Francfort et al. 1995), being explained by the synergy of all organizational capabilities, in order to deal with the competitive challenges of an unpredictable business environment. The interviews have also identified as a possible cause of this situation the good communication among a small number of people, and the entrepreneurial drive of the founder(s).

In the small firm, the number and variety of conflicts increase significantly. Often, at this stage, the company is still confronted with competitive challenges, but the range of strategic options becomes more diversified, and the firm grows, becoming professionally polarized. Communication procedures become more complex, and a simple hierarchical system takes shape, usually with a business manager as Chief Executive Director, and with the founder-scientist as Scientific Officer, or Product Development Manager.

As the firm grows further, reaching a medium size, the organizational procedures become more codified, increasing the rigidity of the internal structure. The two professional groups are now clearly defined, but often coexist as a result of departmentalization and bureaucratization. The conflicts regarding scientific autonomy become frequent, as a reaction to the clear delimitation of competencies. These data confirm that:

H2 Conflict situations between professional groups will evolve with the growth of the biopharmaceutical firm.

In terms of business orientation, the drug development firms experience far more conflicts than platform technology organizations. The interviews have outlined an interesting interpretation of this situation—often, the high frequency of conflicts is considered as a sign of dynamism, flexibility, and innovativeness. This interpretation is however influenced by the outcome of these conflicts, and the efficiency of conflict management procedures implemented within the organization. From this perspective, the fewer conflicts manifested in platform technology firms can be a consequence of the increased departmentalization of the two professional groups, and of hierarchical organizational procedures. The high proportion of firms experiencing conflicts related with the scientific autonomy supports this interpretation. The research hypothesis:

H3 Conflict situations between professional groups will be influenced by the business orientation of the biopharmaceutical firm.

, is therefore supported by the collected data.

6 Conflict Management Procedures

Conflict management procedures implemented by the management team should represent an efficient response to the conflict situations developed within the organization. Paradoxically, many managers have emphasized during the interview that these procedures should not aim for the complete elimination of inter-professional tensions, but rather achieve a dynamic equilibrium between the two main competency poles of the organization. This state of managed tension is considered by many as far more beneficial than the absence of conflict, which can be a sign of organizational fossilization and bureaucratization.

The data collected through interviews has indicated the existence of two main categories of conflict management procedures:

(a) integrative procedures, which attempt to facilitate the communication within the organization, in order to increase the cohesiveness and the synergy of the personnel: outlining common challenges; cross-professional working teams; rewards/promotion linked to overall corporate results; and participatory procedures for establishing and implementing corporate strategy;
(b) regulatory procedures, which aim to concentrate the organizational capabilities in specialized groups (departmentalization), and then to establish clear interaction procedures among departments (regulation of inter-professional relationships and standardization of professional tasks), within a clearly-structured hierarchical system (parallel hierarchical structures).

In the initial stage of firm's development, the management attempts to increase the synergy of the organization, using mostly integrative procedures (see Table 6). The situation is completely reversed in the commercialization stage, in which the

Table 6 The conflict management procedures considered in relation to the lifecycle stage of the investigated organizations

	Building the IP portfolio		Product/ service development		Commercialization	
Conflict management/Stage of development	N	%	N	%	N	%
Outlining common challenges	5	83.3	16	94.1	2	40
Cross-professional working teams	5	83.3	16	94.1	1	20
Rewards/promotion linked to overall corporate results	6	100	15	88.2	3	60
Participatory procedures for establishing and implementing corporate strategy	6	100	8	47.1	2	40
Departmentalization	0	0	7	41.2	5	100
Regulation of inter-professional relationships	1	16.6	8	47.1	5	100
Standardization of professional tasks	1	16.6	9	52.9	5	100
Parallel hierarchical systems	0	0	5	29.4	5	100
Total of enterprises	6	100	17	100	5	100

firm is already mature, and needs a clearer specialization of professional roles and tasks. The phase of product/service development is probably the most complex in terms of strategic possibilities: the manager attempts to reduce the internal conflicts through a fine tune between integrative and regulatory procedures. The results presented in Table 6 demonstrate that:

H1a The conflict management procedures will evolve during the development stages of the biopharmaceutical firm.

The first category of conflict management procedures is used in all investigated micro-enterprises, enhancing the natural effect of organizational synergy as a response to adverse business conditions (see Table 7). In small firms, the use of integrative procedures still predominates, but the regulatory procedures start to be used on an important scale. Finally, the medium-sized firms are characterized by a preponderance of regulatory procedures, although many integrative procedures are still preserved. However, the growth of the firms makes more difficult the application of participatory procedures (only one medium-sized firm is still using it because its number of employees is quite small: 55). This data validates the research hypothesis:

H2a The conflict management procedures will evolve with the growth of the biopharmaceutical firm.

Table 8 shows that there is a certain variation in the conflict management procedures applied in relation to the business orientation of the firm. Although both categories of procedures are used in the investigated firms, drug development companies seem to predominantly use integrative management procedures, while in platform technology organizations, the regulatory procedures are more frequent. The research hypothesis:

Table 7 The conflict management procedures considered in relation to the size of organizations

	Micro-enterprise		Small firm		Medium-sized firm	
Conflict management procedures/Firm size	N	%	N	%	N	%
Outlining common challenges	5	100	14	87.5	4	57.1
Cross-professional working teams	5	100	15	93.7	2	28.6
Rewards/promotion linked to overall corporate results	5	100	14	87.5	5	71.4
Participatory procedures for establishing and implementing corporate strategy	5	100	10	62.5	1	14.3
Departmentalization	0	0	5	31.2	7	100
Regulation of inter-professional relationships	2	40	6	37.5	6	85.7
Standardization of professional tasks	0	0	9	56.2	6	85.7
Parallel hierarchical systems	0	0	5	31.2	5	71.4
Total of enterprises	5	100	16	100	7	100

Table 8 The conflict management procedures considered in relation to the business orientation of biopharmaceutical firms

Conflict management procedures/Business orientation	Drug development		Platform technology	
	N	%	N	%
Outlining common challenges	20	90.9	3	50
Cross-professional working teams	19	86.4	3	50
Rewards/promotion linked to overall corporate results	22	100	2	33.3
Participatory procedures for establishing and implementing corporate strategy	15	68.2	1	16.6
Departmentalization	10	45.4	2	33.3
Regulation of inter-professional relationships	11	50	3	50
Standardization of professional tasks	10	45.4	5	83.3
Parallel hierarchical systems	5	22.7	5	83.3
Total	22	100	6	100

H3a The conflict management procedures will be influenced by the business orientation of the biopharmaceutical firm.

, is therefore partially supported.

7 Concluding Remarks

Following the methodology used in the previous studies on organizational sub-cultures, this project attempted to identify the existence of professional cultures in biopharmaceutical firms, and to investigate the relation between different professional groups, using as field of research the UK biotechnology sector.

Considering the relation between the inter-professional conflicts and the profile of biopharmaceutical SMEs, three research hypotheses have been formulated:

H1 Conflict situations between professional groups will evolve during the development stages of the biopharmaceutical firm.

H2 Conflict situations between professional groups will evolve with the growth of the biopharmaceutical firm.

H3 Conflict situations between professional groups will be influenced by the business orientation of the biopharmaceutical firm.

On the other hand, the application of conflict management procedures was also linked to the specific characteristics of the firm, determining three additional research hypotheses:

H1a The conflict management procedures will evolve during the development stages of the biopharmaceutical firm.

H2a The conflict management procedures will evolve with the growth of the biopharmaceutical firm.

H3a The conflict management procedures will be influenced by the business orientation of the biopharmaceutical firm.

The data collected through interviews with representatives of the two main professional groups within biopharmaceutical enterprises—scientists and business specialists—have validated the formulated hypotheses.

Considering the conflict management procedures as a direct response to the conflict situations manifested within organizations, it is important to investigate the relationship between these two elements. The graphical representation of the identified categories of inter-professional conflicts and conflict management procedures, provides a two axes matrix (see Fig. 2), in which four distinct types of organizational situations can be identified:

(a) Co-operation—when the successful application of integrative procedures creates a dynamic synergy of different working styles—it is the most positive situation, which preserves the balance of competencies and the flexibility of the firm, enhancing its innovativeness and competitive advantage;
(b) Forced co-habitation—results from the application of integrative conflict management procedures to conflicts determined by different value systems—represents an instable situation which is a rich source of recurrent negative tensions and conflicts;
(c) Parallel systems—represents the combination of strong regulatory procedures with conflicts of professional values—the manager does not attempt to achieve a false reconciliation of different value systems, but separates drastically the two professional groups creating an organization with parallel hierarchical systems which minimizes the interaction opportunities;
(d) Specialization—the differences in work styles are solved using regulatory procedures, which implement clear professional roles and formalized interactions between the two main professional groups.

The explicative value of this model is relative, since in reality, the research findings show that most organizations present hybrid profiles. However, despite its

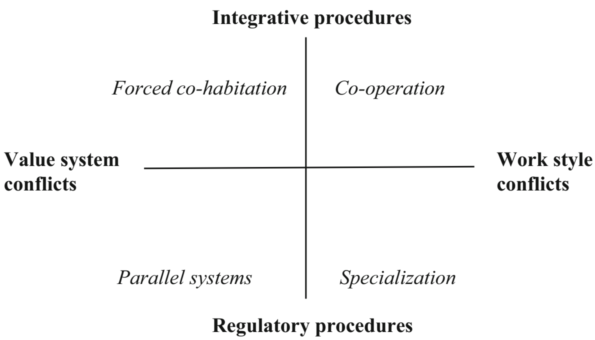


Fig. 2 The conflict management matrix

theoretical limitations, the Conflict Management Matrix can be successfully used by practitioners to diagnose their type of organization and to identify the best procedures to reconcile the existing internal conflicts.

The main limitations of this study were determined by the research methodology adopted. Despite a good response rate, the number of investigated organizations is quite small, which might raise questions regarding the generalization of findings to the entire population of study. Considering these limitations, the results of this study can still be used by academics and practitioners:

- on the basis of the presented findings, researchers can refine the model of inter-professional conflict and conflict management procedures in high-technology firms, developing similar research in other industrial sectors (informatics, electronics, agricultural biotech, etc.), or in other countries;
- the managers-entrepreneurs of biopharmaceutical SMEs can apply the existing results to better understand and improve the functioning of their organizations, applying more targeted conflict management procedures to the specific situations manifested in their company.

Future research into this subject should be developed using two alternative research methodologies. On one hand, more quantitative studies can provide a clearer typology of conflict management practices in various categories of organizations. On the other hand, a case study approach can offer a better understanding of the cultural environment of each investigated organization, and of the specific effects of various conflict resolution practices. Ultimately, these two approaches should be integrated into a synthetic model of inter-professional cultural conflict that must propose concrete measures to enhance the dynamic cohesiveness of highly innovating organizations.

References

Azzone, G., & Pozza, I. D. (2003). An integrated strategy for launching a new product in the biotech industry. *Management Decision, 41*(9), 832–843.

Briggs, P. S. (1994). *Biotechnology: A commercial evaluation*. Richmond: Script Reports.

Chandrasekar, C. R., Helliar, C. V., Lonie, A. A., Nixon, W. A., & Power, D. M. (1999). Strategic management of innovation risk in the biopharmaceutical industry: A UK perspective. *International Journal of Healthcare Technology and Management, 1*(1–2), 62–76.

Cohen, L., Duberley, J., & McAuley, J. (1999). Fuelling discovery or monitoring productivity: Research scientists' changing perceptions of management. *Organization, 6*(3), 473–497.

DiMasi, J. A., & Grabowski, H. G. (2007). The cost of biopharmaceutical R&D: Is biotech different? *Managerial and Decision Economics, 28*(4–5), 469–479.

Dorabjee, S., Lumley, C. E., & Cartwright, S. (1998). Culture, innovation and successful development of new medicines—An exploratory study of the pharmaceutical industry. *Leadership & Organization Development Journal, 19*(4), 199–210.

Ernø-Kjølhede, E., Husted, K., Mønsted, M., & Wenneberg, S. B. (2001). Managing university research in the triple helix. *Science and Public Policy, 28*(1), 49–55.

Evans, A. G., & Varaiya, N. P. (2003). Anne Evans: Assessment of a biotechnology market opportunity. *Entrepreneurship: Theory and Practice, 28*(1), 87–106.
Francfort, I., Osty, F., Sainsaulieu, R., & Uhalde, M. (1995). *Les Mondes Sociaux de l'Entreprise*. Paris: Desclée de Brouwer.
Gregory, K. (1983). Native-view paradigms: Multiple cultures and culture conflicts in organizations. *Administrative Science Quarterly, 28*(3), 359–376.
Guérin, G., Wils, T., & Lemire, L. (1996). Le malaise professionnel: Nature et mesure du concept. *Relations Industrielles/Industrial Relations, 51*(1), 62–96.
Herkenhoff, L. M., & Heydenfeldt, J. A. (2011). A correlational study of professional culture and intraorganizational conflict. *The International Journal of Management and Business, 2*(1), 61–76.
Hodgson, J. (2001). The headache of knowledge management. *Nature Biotechnology, 19*(6), 44–46.
Irrmann, O. (2002). Organizational culture and identity strategies in international management: An interdisciplinary review. 28th EIBA Conference Proceedings, 8–10 December, Athens. www.aueb.gr/deos/EIBA2002.files/PAPERS/C63.pdf. Accessed January 2007.
Lam, A. (2004). Work roles and careers of R&D scientists in network organisations. Brunel University, Working Paper No. 6. www.brunel.ac.uk/research/brese/pub/docs/lam_wp6.PDF. Accessed January 2007.
Lumineau, F., Eckerd, S., & Handley, S. (2015). Inter-organizational conflicts: Research overview, challenges, and opportunities. *Journal of Strategic Contracting and Negotiation, 1*(1), 42–64.
Martin, J., & Siehl, C. (1983). Organizational culture and counter-culture: An uneasy symbiosis. *Organizational Dynamics, 12*(2), 52–64.
Martin, J., Sitkin, S. B., & Boehm, M. (1985). Founders and the elusiveness of a cultural legacy. In P. J. Frost, L. F. Moore, M. R. Louis, C. C. Lundberg, & J. Martin (Eds.), *Organizational culture*. Beverly Hills: Sage.
Mehta, S (2004). Paths to entrepreneurship in the life sciences. *Bioentrepreneur*, vol. 26, October. http://www.nature.com/cgi-taf/gateway.taf?g=6&file=/bioent/building/entre/102004/full/bioent831.html. Accessed December 2004.
Mudambi, R., & Swift, T. (2009). Professional guilds, tension and knowledge management. *Research Policy, 38*(5), 736–745.
Persidis, A. (1996). Building molecular value. *Journal of Business Strategy, 17*(2), 18–21.
Phrma. (2015). Biopharmaceutical research & development: The process behind new medicines. www.phrma.org/sites/default/files/pdf/rd_brochure_022307.pdf. Accessed January 2016.
Pisano, G. P. (2006). Can science be a business? Lessons from Biotech. *Harvard Business Review, 84*(10), 114–124.
Saviotti, P. P. (1998). Industrial structure and the dynamics of knowledge generation in biotechnology. In J. Senker (Ed.), *Biotechnology and competitive advantage*. Cheltenham: Edward Elgar.
Senker, J. (1998). Biotechnology: The external environment. In J. Senker (Ed.), *Biotechnology and competitive advantage*. Cheltenham: Edward Elgar.
Van Maanen, J., & Barley, S. (1985). Cultural organization: Fragments of a theory. In P. J. Frost, L. F. Moore, M. R. Louis, C. C. Lundberg, & J. Martin (Eds.), *Organizational culture*. Beverly Hills: Sage.
Walsh, V. (1993). Demand, public markets and innovation in biotechnology. *Science and Public Policy, 20*(3), 138–156.

Perception Gaps in International Corporate Entrepreneurship: The Role of Knowledge Transfer Tools

Mara Brumana, Lucio Cassia, Davide Gamba, and Tommaso Minola

Abstract This chapter focuses on knowledge transfer in multinational corporations and analyses to what extent and under what conditions perception gaps associated to knowledge sharing have an impact on performance. We build and test a moderated mediation model which shows that perception gap on knowledge sharing is an antecedent of capability perception gap which, in turn, negatively influences subsidiary performance. Results obtained highlight that knowledge transfer tools—and particularly technology-based coordination mechanisms—play a crucial role in creating the best environment to share knowledge within multinational corporations.

Keywords Perception gap • International corporate entrepreneurship • Knowledge transfer • Knowledge transfer tool • Multinational corporation

1 Introduction

In a globalized society influenced by liberalism and free market system, the current economy is driven by an accelerated rate of change and innovation. We are living in the heat of the 4th industrial revolution (Schwab 2016)—better known as Industry 4.0. Leveraging on new technological trajectories such as those triggered by Internet of things and cloud computing (Kagermann et al. 2013), this revolution encompasses automation and knowledge exchange (Hermann et al. 2016) within an economic context characterized by volatility, uncertainty, complexity and ambiguity (Abidi and Joshi 2015).

In this new scenario, knowledge sharing relationships within multinational companies (MNCs) are nevertheless critical and one of the sources of the internationalization strategy's failure (Reus et al. 2015) due to geographical, cultural and institutional distance (Kostova 1999; Berry et al. 2010). In other words, coordination within the inter-organizational networks is one of MNCs' main issue

---

Authors are listed in alphabetical order.

M. Brumana • L. Cassia • D. Gamba • T. Minola (✉)
Center for Young and Family Enterprise and Department of Management, Information and Production Engineering, University of Bergamo, Via Pasubio 7b, 24044 Dalmine (BG), Italy
e-mail: mara.brumana@unibg.it

T. Devezas et al. (eds.), *Industry 4.0*, Studies on Entrepreneurship, Structural Change and Industrial Dynamics, DOI 10.1007/978-3-319-49604-7_16

(Bartlett and Ghoshal 1988). Relationships in MNCs can be modelled as a mixed motive dyad (Birkinshaw et al. 2000) in which interests of the two parties are frequently not perfectly aligned due to different perceptions of the whole. For example, where one of them hopes for autonomy, the other requires control. Again, where one sees entrepreneurial opportunity, the other could think about opportunism. In these situations, a series of differences in perception may emerge (Birkinshaw et al. 2000; Chini et al. 2005).

Perception gap is a common phenomenon in international business (Ghoshal and Nohria 1989; Nohria and Ghoshal 1994), which can create dysfunctional tensions in the MNC (Birkinshaw et al. 2000) and organizational conflicts, both associated with dissatisfaction and negative emotions (see for instance, Schotter and Beamish 2011). Perception gaps can damage the economic and social structure of an organization and hence its performance (Chini et al. 2005; Monteiro et al. 2008).

This chapter investigates the knowledge sharing process within MNCs and in particular: (1) the impact of perception gaps related to both knowledge sharing and subsidiary capability on subsidiary performance; (2) the role of knowledge transfer tools.

2 Theoretical Background and Hypotheses Development

2.1 Knowledge Transfer in MNCs

The research on international corporate entrepreneurship—namely, the definition, evaluation and exploitation of business opportunities in the international market—has put attention on MNCs (Zahra et al. 2004). The latter are defined, indeed, as enterprises that look for "the discovery, enactment, evaluation, and exploitation of opportunities across national borders to address future goods and services" (Oviatt and McDougall 2005: 540).

Within MNCs, two organizational actors are usually identified: headquarter (HQ) and subsidiary. The HQ is defined as the place where a company's executive offices are located. This is not necessarily the location where the majority of employees work. A subsidiary is defined as "an affiliate of a MNC located in a foreign country and of which the parent company holds at least 51 % of ownership" (Bouquet and Birkinshaw 2008: 585). The standard organizational configuration model of HQ and subsidiaries gives power to the first one assuming that it is located in a geographically central area with respect to subsidiaries (Freeman 1979) and it controls critical resources (Pfeffer and Salancik 1978): this is the HQ-centred model. However, when subsidiaries have a high degree of freedom and independence from the HQ in their local area, literature talks about subsidiary-centred models. In any case, the network originating from the interactions between HQ and subsidiaries is a powerful tool in terms of knowledge transfer (Oviatt and McDougall 1994).

Knowledge is defined as "accumulated practical skill or expertise that allows one to do something smoothly and efficiently" (Kogut and Zander 1992: 386). A well-known classification of knowledge is based on its nature (Campisi and Passiante 2007). Accordingly, scholars identify two categories of knowledge: explicit knowledge and tacit knowledge (Polanyi 1966). Explicit knowledge is the one that can be easily articulated, codified, accessed and verbalized (Hélie and Sun 2010) using words and numbers in the form of data, scientific formulas, and specifications (Nonaka et al. 1998). Conversely, tacit knowledge is highly embedded into the individual (Ravetz 1971), context specific (Sternberg 1994) and hard to formalize and share (Nonaka 1991).

Knowledge management corresponds to "a process or performance to create, acquire, share and apply knowledge which leads into learning increase and performance improvement" (Armstrong 2010: 22). Knowledge management involves processes related to identifying, sharing, and generating knowledge. Knowledge management requires a system to generate and keep knowledge stocks as well as knowledge transfer and organizational learning. The act of transferring knowledge is defined by the literature as "the process through which one unit (e.g., group, department, or division) is affected by the experience of another" (Argote and Ingram 2000: 151). Knowledge management projects' returns are directly proportionated to quality and adequacy of tools that support them (Quagini 2004). Example of tools used in order to manage knowledge within organizations are thematic forum and chat, dynamic FAQ (e.g., Yahoo! Answer), and content management systems (e.g., MediaWiki platforms). More generally, one of the classifications proposed by the literature divides knowledge transfer tools in two categories, which are at the same time substitutes and complements: personal coordination mechanisms (PCM) and technology-based coordination mechanisms (TCM) (Wang and Noe 2010; Ambos and Ambos 2009; Haas and Hansen 2005; Schulz and Jobe 2001).

PCM are associated with an intensive collaboration between recipients and senders. This relationship provides the social basis for knowledge exchange (Nonaka et al. 2000) and supports the knowledge sharing process within the organization (Szulanski 1996; Dyer and Nobeoka 2000; Håkanson and Nobel 2001; Almeida and Phene 2004). On the one hand, a strong relationship between sender and recipient will favour actors' understanding. On the other, it will also make the identification of relevant knowledge more plausible. The following tools are included among the PCM category.

- Coaching. The employee receives feedback and guidance from the supervisor while practicing new skills.
- Mentoring. An organizationally sponsored program that pairs an employee with a more experienced employee or manager in order to receive guidance and advice on career development. Mentors offer advice on what to do in a given situation, how to do it and why it is worth doing.

- Dual incumbency. Two employees temporarily occupy the same position. This tool is useful when sharing of tacit information is fundamental in order to achieve budget results.
- Job shadowing. A less experienced employee is paired with another employee owning the skills to be transferred. Knowledge is shared while dealing with everyday problems.
- On the job training (OJT). Instruction that takes place on the actual job site with task accomplishment as part of the learning process. It involves learning skills and applying knowledge in a hands-on and on-the-job manner.
- Retirees. Past employees return to work to provide the organization with knowledge and expertise, train less experienced staff, or share specialized knowledge.
- Job rotation. A formal program in which a person exchanges job responsibilities with another employee for the purpose of cross-training and gaining more knowledge of the organization. It is directly connected to Japanese Kaizen's concepts of poly-competence and poly-valence.
- Communities of practices (CoP). A group of individuals that, without being part of a formal work team, shares a common work practice over a period, getting together to share information and knowledge.

PCM are characterized by vis-à-vis contact among the involved actors and can consist in liason personnel (e.g. mentoring), temporary task forces (e.g. dual incumbency) or permanent groups (e.g. communities of practices).

TCM, instead, aim to maximize the exploitation of resources that are embedded within a network of separated units. Information technology tools represent the heart of this category of knowledge transfer tools. Firms that demonstrate ability to use information technology infrastructures for knowledge transfer, tend to exhibit higher organizational effectiveness (Irma Becerra-Fernandez 2001; Gold and Arvind Malhotra 2001). The following tools are included in this category:

- On-line learning. It is the combination of artificial intelligence, real-time learning methods, and computers, in order to transfer knowledge and information. As a user encounters a problem, he or she can access all organizational policies and procedures through the IT system, and even learn in real time using the training component.
- Wikies. Wikies are collaborative websites that allow individuals to easily edit, delete, and modify contents related to a specific topic.
- Work process map. Systematic documentation of a process, created using pictures and symbols to tell a story. It is a written and graphical representation of what to do, when to do it, and how to do operations that compose a critical process.
- Chat group and web-based discussion group. Transmission of knowledge from sender(s) to receiver(s) through chats or forums.

2.2 Perception Gaps in MNCs

When knowledge transfer is far from perfect, some distortions of reality—i.e., perception gaps—can emerge. With specific reference to MNCs, literature defines perception gap (PG) as "the difference in perception between the HQ and a subsidiary concerning the management processes of the MNC" (Chini et al. 2005: 146). According to social psychology (e.g., March and Simon 1958; Cyert and March 1963), theories of perception explain how individuals interpret, evaluate, and act based on their perceptions of others.

Two main perception gaps are studied in this chapter: knowledge sharing perception gap (KSPG) and capability perception gap (CPG) (Asakawa 2001; Monteiro et al. 2008). The first one sees the multinational company as an open system made of different organizations, which has to face uncertainty (Thompson 1967). Since international operations include a great degree of uncertainty, the number of information shared between HQ and subsidiary is essentially high. Problems arise when both sides perceive the level of uncertainty surrounding the local environment differently. In such a situation, knowledge sharing perception gap emerges (Asakawa 2001; Chini et al. 2005). The second one is related to gaps about subsidiary capabilities (e.g., ability to collect and share information about local environment's market) and influences knowledge transfer frequency within MNCs. Capability perception gap brings thus to subsidiary isolation—the branch does not experience any knowledge transfer flows due to reciprocal underestimation of branch's capabilities—and eventually to lower subsidiary performance (Monteiro et al. 2008).

2.3 Research Question and Hypothesis

Assuming that knowledge transfer is an antecedent of capabilities development, we argue that the higher is the gap associated with knowledge sharing, the higher will be the gap related to the capabilities which should stem from the knowledge sharing process. In other words, our research aims to investigate KSPG as a possible antecedent of CPG. In addition, we suggest that the adoption of TCM and PCM will influence this relationship. Finally, following Monteiro et al. (2008), we hypothesise that capability perception gap, by influencing the amount of knowledge flows (i.e., frequency) exchanged within MNCs, has an effect on subsidiary isolation and hence on subsidiary performance.

Accordingly, our research question is the following: to what extent and under what conditions KSPG influences CPG within a MNC and hence subsidiary performance? In order to address this question, we propose four sets of hypotheses and we build and test a moderated mediation model in which KSPG (both downstream and upstream flows of knowledge) influences CPG and this relationship is moderated by the typology of knowledge transfer tool. Finally, CPG influences subsidiary performance.

The first hypothesis is related to the downstream flow of knowledge and we claim that the higher the level of knowledge sharing perception gap related to HQ information sharing process, the higher the level of capability perception gap.

- *Hypothesis 1A*. KSPG—associated with knowledge sharing processes from the HQ to the subsidiary—is positively related to CPG.
- *Hypothesis 1B*. Knowledge transfer tools adopted by the HQ negatively moderate the relationship of KSPG and capability perception gap, in such a way that the positive relationship between HQ KSPG and CPG capability perception gap becomes weaker, at greater level of TCM and PCM adoption by the HQ.

Similarly, with reference to the upstream flow of knowledge transfer, we propose that the higher the level of knowledge sharing perception gap related to subsidiary information sharing process, the higher the level of capability perception gap. Accordingly:

- *Hypothesis 2A*. KSPG—associated with knowledge sharing processes from the subsidiary to the HQ—is positively related to CPG.
- *Hypothesis 2B*. Knowledge transfer tools adopted by the subsidiary negatively moderate the relationship of KSPG and capability perception gap, in such a way that the positive relationship between branch's KSPG and CPG capability perception gap becomes weaker, at greater level of TCM and PCM adoption by the subsidiary.

Moreover, in line with Monteiro et al. (2008), we hypothesise that CPG negatively influence subsidiary performance. Thus:

- *Hypothesis 3*. Capability perception gap is negatively associated with subsidiary performance.

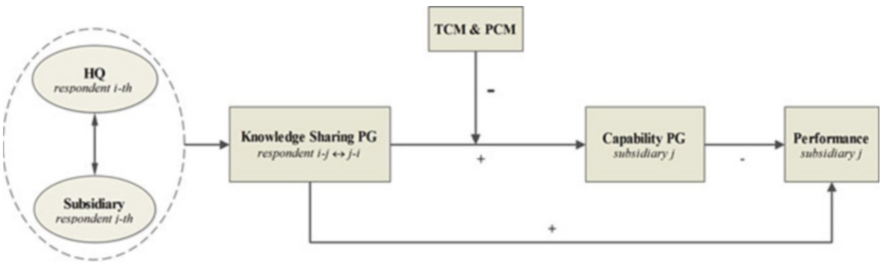
An additional hypothesis was set to evaluate the mediating role of CPG in the overall model.

- *Hypothesis 4*. CPG mediates the relationship between HQ KSPG (subsidiary KSPG) and subsidiary performance.

Figure 1 summarises the overall model and the expected relationships among variables.

3 Methods

In line with previous research (Chini et al. 2005; Monteiro et al. 2008), data have been collected in an Italian MNC through a structured questionnaire (see Appendix 1 for a brief description of the case). The survey was sent to 51 HQ employees and to 13 managers from the nine branches. The overall response rate was 56 %



Note. Dyadic pair (dashed oval) is composed by the "i-th" respondent's evaluation from HQ and the "j-th" subsidiary. Downstream KSPG is identified by the subscript "i-j", while upstream KSPG is identified by the subscript "j-i". CPG and performance' subscript "j" identifies the j-th subsidiary object of evaluation.

Fig. 1 Proposed theoretical model. Note. Dyadic pair (dashed oval) is composed by the "i-th" respondent's evaluation from HQ and the "j-th" subsidiary. Downstream KSPG is identified by the subscript "i-j", while upstream KSPG is identified by the subscript "j-i". CPG and performance' subscript "j" identifies the j-th subsidiary object of evaluation.

(34 respondents). In particular, 29 responses (about 57 % of 51) came back from HQ and 5 (about 59 % of 9) from branches. Received surveys were combined together as dyads (dyadic pairs). As we tried to measure mutual information perception gap between HQ and subsidiaries, following Asakawa's approach (2001), we considered a questionnaire to be valid only when both sides (HQ and related branch) had responded. In so doing, 78 dyads were collected. Respondents assessed the extent to which they agreed with each question on a 1–5 or 1–7 Likert scale. From the proposed survey, the following items were taken into account (see Appendix 2 for a complete list of the items, and Appendix 3 for variables' operationalization):

Capability Perception Gap We operationalized this construct by asking both HQ and subsidiary respondents about branches' capabilities. In specific, using the multi-item proposed by Monteiro et al. (2008), the rating of the j-th subsidiary capability was measured through a three item-scale, using a 1–7 Likert scale (1 = much below the average, 4 = average, and 7 = much above average). According to Birkinshaw et al. (2000) approach, we first subtracted the HQ response from the subsidiary response, computing bipolar perception gap reaching a 13-degrees ordinal scale. Using mathematical notation:

$$\mathrm{Cpg}_k = \mathrm{C\_h}_{i \to j} - \mathrm{C\_s}_j$$

Knowledge Sharing Perception Gap from HQ to Subsidiary The work of Asakawa (2001) inspired the operationalization of this variable which represents information flow from the HQ to the subsidiary. Respondents evaluated the knowledge flow from HQ to subsidiary based on a 1–5 Likert scale, where 1 means strongly disagree and 5 means strongly agree. As for the capability perception gap item, for each evaluation, we first subtracted the HQ response from the subsidiary response, obtaining a 9-degrees ordinal scale. Using mathematical notation:

$$\mathrm{KSHpgk} = \mathrm{KS\_h_i} - \mathrm{KS\_s_j}$$

Knowledge Sharing Perception Gap from Subsidiary to HQ It represents information flow from the subsidiary to the HQ (Asakawa, 2001). Respondents evaluated this knowledge flow based on a 1–5 Likert scale, where 1 means strongly disagree and 5 means strongly agree. As for the capability perception gap item, we first subtracted the subsidiary response from the HQ response, obtaining a 9-degrees ordinal scale. In other words, using mathematical notation:

$$\mathrm{KSSpg_k} = \mathrm{KS\_s_j} - \mathrm{KS\_h_i} \rightarrow {}_{\mathrm{j}}$$

Subsidiary Financial Performance Following the operationalization suggested by Monteiro et al. (2008), we asked to subsidiary respondents to self-evaluate their branch's financial performances. Evaluations was based on a 1–7 Likert scale (1 = much below the average, 4 = average, and 7 = much above average). Using mathematical notation:

$$\mathrm{FO_j} = \{1, 2, 3, 4, 5, 6, 7\}$$

Knowledge Transfer Tool Two different approaches were used for personal coordination mechanisms (PCMs) and technology-based coordination mechanisms (TCMs). Operationalization of this multiple item was adopted from Ambos and Ambos (2009), while introducing the dyad concept (Asakawa 2001). In particular, the PCM variable is composed by three evaluations on personal coordination mechanisms. Respondents were asked to indicate how much they personally relied and used on liaison personnel, temporary task forces and permanent teams. Respondents evaluated the three knowledge sharing's personal coordination mechanisms based on a 1–7 Likert scale, where 1 = strongly disagree and 7 = strongly agree, obtaining a 7-degrees ordinal scale. In addition, we introduced the PCM use gap, which represents the "distance" of utilization of these tools between the HQ and the subsidiary, obtaining a 13-degress ordinal scale. In other words, using mathematical notation:

$$\mathrm{PCMpg_k} = \mathrm{PCM\_s_j} - \mathrm{PCM\_h_i}$$

TCM is a multiple item that tries to assess the respondent's unit ability to use technological tools for inter-unit knowledge sharing. Items refer to different tools, which are typically part of knowledge management systems, including technological instruments to map knowledge, team collaboration tools, pointers to expertise, repositories of best practices and lessons learnt. Respondents evaluated the four knowledge sharing's technology-based coordination mechanisms based on a 1–7 Likert scale, where 1 means strongly disagree and 7 means strongly agree, obtaining a 7-degrees ordinal scale. As made for PCM, we introduced the TCM use gap, which represents the "distance" of utilization of these tools between the HQ and the subsidiary, obtaining a 13-degress ordinal scale. In other words, using mathematical notation:

$$\mathrm{TCMpg_k} = \mathrm{TCM\_s_j} - \mathrm{TCM\_h_i}$$

In our empirical model, as control variables we account for the effect of respondent's tenure, the nature of knowledge (by measuring its tacitness), the geographical distance between the HQ and the subsidiary and the level of economic development of the host country.

Tenure The operationalization of this variable was made by asking the HQ respondents to quantify the duration of their working relationship with the company.

Tacitness of Knowledge The operationalization of this variable was retrieved from Monteiro et al. (2008). Respondents evaluated four sentences related to their perception of their organization's knowledge based on a 1–7 Likert scale, where 1 means strongly disagree and 7 means strongly agree. We computed the absolute value of the measure, in order to have a general view of the tacitness level's gap (the direction of the flow does not matter in this case) reaching a 7-degrees ordinal scale. Using mathematical notation:

$$\mathrm{TACpg_k} = \left|\mathrm{TAC\_s_j} - \mathrm{TCM\_h_i}\right|$$

Geographical Distance The operationalization of this variable was retrieved from Ambos and Ambos (2009). Following this paper's approach, we measured geographic distance as the absolute number of air miles[1] between the location of the two units—the sender and the receiver. The natural logarithm of these values was used in our analysis (Monteiro et al., 2008).

Host Country Economic Level According to existing literature, the level of economic development of the host country may affect knowledge sharing flows (Monteiro et al. 2008; Gupta and Govindarajan 2000). Thus, 2014 per capita incomes in U.S. dollars were collected from the World Bank report.[2]

Before starting regression analysis, we performed three preparatory tests that entail checking for data: reliability of scales used, validity of constructs and check for potential multi-collinearity of independent variables. The first preparatory test is

[1]Air-miles distance computed using *www.distancefromto.net*

[2]Data referred to 2014 retrieved from *data.worldbank.org*

projected to evaluate the reliability of the scales used in the survey, specifically in all multi-item scales. We have checked the reliability of the scale using Cronbach's coefficient alpha (1951), which represents a statistical measure for internal consistency (Gliem and Gliem 2003; Litwin 1995; Churchill 1979). According to results, constructs are reliable and characterized by a value of alpha higher than 0.8. The only exception is represented by tacit knowledge, which has a Cronbach's alpha of 0.51. This construct must thus be used carefully. The second preparatory test consists in the assessment of the validity of constructs using factor analysis. Factor analysis is a statistical technique used to obtain a reduction of the number of factors that explain a phenomenon. The main goal is to determine a number of "latent" variables smaller than the initial one. Since the reliability and the internal consistency of the constructs selected has already been proven, we employ PCA (Principal Axes Factoring) as extraction method. In detail, only one eigenvalue greater than 1 emerged for each construct (tacitness is the only exception, which has two eigenvalues greater than 1). The results obtained for tacitness suggest its decomposition in two control variables:

$$\text{Factor 1} \rightarrow \text{knowledge external exploitation} = \frac{\text{item 3} + \text{item 4}}{2} \tag{3.1}$$

$$\text{Factor 2} \rightarrow \text{knowledge internal exploitation} = \frac{\text{item 1} + \text{item 2}}{2} \tag{3.2}$$

Finally, the multi-collinearity of the independent variables was examined using Variance Inflation Factor (VIF). The latter is an indicator of collinearity, where large values indicate a strong relationship between the independent variables studied within the model. According to Hair et al.'s rule of thumb (2006), VIF values greater than 10 indicate multi-collinearity. No multi-collinearity was found in our sample.

4 Results

Concerning Hypothesis 1—in order to clearly explain the relationships between variables—four models have been implemented.

Table 1 tests Hypothesis 1 and shows that subsidiary capability perception gap is higher when knowledge perception gap is high. The effect of knowledge sharing perception gap on capability perception gap is reduced by TCM. The final model is synthesized in the following equation:

Table 1 Results from regression analysis—CPG and TCM for downstream knowledge flows

Variables	Model 1	Model 2	Model 3	Model 4
Firm tenure	−0.142	−0.182**	−0.202***	−0.215***
	(0.107)	(0.0784)	(0.0775)	(0.0762)
Distance	0.116	−0.0690	−0.132	−0.0739
	(0.0999)	(0.0767)	(0.0824)	(0.0865)
Tacitness internal	−0.503***	−0.252***	−0.0898	−0.0501
	(0.107)	(0.0839)	(0.120)	(0.119)
Tacitness external	0.0203	0.0559	−0.0103	−0.0444
	(0.112)	(0.0822)	(0.0880)	(0.0880)
KSPG HQ		0.659***	0.574***	0.519***
		(0.0805)	(0.0912)	(0.0938)
TCM use gap			−0.246*	−0.283**
			(0.132)	(0.131)
TCM · KSPG				−0.178*
				(0.0957)
Constant	−0.159**	−0.469***	−0.491***	−0.513***
	(0.0801)	(0.0801)	(0.0801)	(0.0801)
Adj R-squared	0.2123	0.5703	0.5827	0.5947
Observations	78	78	78	78

Standard errors in parentheses
***p<0.01, **p<0.05, *p<0.1

$$\begin{aligned}\text{Capability PG}_{\text{TCM}} = \quad & -0.513\\ & +0.519^{*}\ \text{Knowledge sharing PG}_{\text{HQ}\rightarrow\text{subsidiary}}\\ & -0.215^{*}\text{Firm tenure}\\ & -0.283^{*}\text{TCM}\\ & -0.178^{*}\text{TCM Knowledge sharing PG}\end{aligned} \tag{3.3}$$

When focusing on PCM, results confirm that subsidiary capability perception gap is higher when knowledge perception gap is high (see Table 2). The use of PCM, however, does not affect the relationship between knowledge sharing perception gap and capability perception gap.

The final model for PCM is synthesized in the following equation:

$$\begin{aligned}\text{Capability PG}_{\text{PCM}} = \quad & -0.520\\ & +0.496^{*}\ \text{Knowledge sharing PG}_{\text{HQ}\rightarrow\text{subsidiary}}\\ & -0.215^{*}\text{Firm tenure}\\ & -0.332^{*}\text{PCM}\end{aligned} \tag{3.4}$$

As made for hypothesis 1, four models were implemented for hypothesis 2. We found that a gap in TCM use negatively moderates the effect of knowledge sharing perception gap on capability perception gap. The moderating effect of PCM use gap on the upstream of knowledge, instead, is not statistically significant. The final

Table 2 Results from regression analysis—CPG and PCM for downstream knowledge flows

Variables	Model 1	Model 2	Model 3	Model 4
Firm tenure	−0.142	−0.182**	−0.213***	−0.215***
	(0.107)	(0.0784)	(0.0758)	(0.0754)
Distance	0.116	−0.0690	−0.107	−0.0764
	(0.0999)	(0.0767)	(0.0745)	(0.0811)
Tacitness internal	−0.503***	−0.252***	−0.0740	−0.0582
	(0.107)	(0.0839)	(0.103)	(0.104)
Tacitness external	0.0203	0.0559	−0.0109	−0.0219
	(0.112)	(0.0822)	(0.0823)	(0.0827)
KSPG HQ		0.659***	0.510***	0.496***
		(0.0805)	(0.0943)	(0.0951)
PCM use gap			−0.327***	−0.332***
			(0.120)	(0.119)
PCM KSPG				−0.0899
				(0.0976)
Constant	−0.159**	−0.469***	−0.515***	−0.520***
	(0.0801)	(0.0801)	(0.0801)	(0.0801)
Observations	78	78	78	78

Standard errors in parentheses
***$p<0.01$, **$p<0.05$, *$p<0.1$

model related to the upstream flows of knowledge for TCM is synthetized in the following equation:

$$\begin{aligned} \mathit{Capability\ PG}_{TCM} = \quad & -0.513 \\ & -0.215 * \text{Firm tenure} \\ & +0.519 * \text{KSPG} \\ & -0.283 * \text{TCM} \\ & -0.178 * \text{TCM KSPG} \end{aligned} \tag{3.5}$$

On the other hand, the final model for PCM is synthesized in the following equation:

$$\begin{aligned} \text{Capability PG}_{\text{PCM}} = \quad & -0.360 \\ & -0.213 * \text{Firm tenure} \\ & -0.686 * \text{PCM} \end{aligned} \tag{3.6}$$

Finally, with reference to hypotheses 3 and 4, we implemented two models (see Table 3).

Results show that subsidiary financial performance is higher when capability perception gap is low. In conclusion, the multiple regression model has produced results that validate the third hypothesis proposed and confirm Monteiro et al. (2008) findings: the effect of capability perception gap on subsidiary financial

Table 3 Results from regression analysis—Subsidiary performance

Variables	Model 1	Model 2
Host country economic level	-0.0183^{***}	-0.0146^{***}
	(0.00514)	(0.00477)
Capability perception gap		-0.301^{***}
		(0.0734)
Constant	4.986^{***}	4.667^{***}
	(0.222)	(0.216)
Observations	78	78
R-squared	0.142	0.229

Standard errors in parentheses
$^{***}p<0.01$, $^{**}p<0.05$, $^{*}p<0.1$

Table 4 Results from mediation analysis—KSPG and performance for downstream knowledge flows

Effect	Value
Indirect	−.16225201
Direct	−.32631083
Total	−.48856284

performance is significant and negative. The final model is synthesized in the following equation:

$$\begin{aligned}\text{Subsidiary performance} = \; & +4.667 \\ & -0.0146^{*}\text{Host country economic level} \\ & -0.301^{*}\text{Capability perception gap}\end{aligned} \tag{3.7}$$

In order to enrich our analysis, in hypothesis 4 we focused on the possible direct relationship between knowledge sharing perception gap and subsidiary financial performance. The results show that the proportion of the total effect mediated by capability perception gap on the direct relationship between HQ knowledge sharing perception gap and subsidiary financial performance—obtained as a ratio between the indirect effect and the total one—accounts for about 33 % (see Table 4).

Since the effect of knowledge perception gap from subsidiary to HQ is not statistically significant, we didn't perform this analysis for the upstream flow of knowledge. In order to test the aforementioned mediation in detail, we relied on Baron and Kenny (1986) who state that four conditions have to be fulfilled. In specific, the independent variable has to affect the dependent variable, the independent variable has to affect the mediator, the mediator has to affect the dependent variable, and—when the mediator is added—the effect of the independent variable on the dependent variable have to become weaker. Being the various realtionships statistically significant (at 10 %) with respect to the four conditions above mentioned, we conclude that mediation is valid.

5 Conclusions

The present study contributes to the debate on knowledge transfer in MNCs. Our analysis demonstrates that a perception gap on the knowledge sharing process by the MNC's actors is an antecedent of capability perception gap which, in turn, negatively influences subsidiary performance. Results obtained also highlight that knowledge transfer tools—and particularly TCM—play a crucial role in creating the best environment to share knowledge inside the MNC's network: HQ should introduce technology-based coordination mechanisms rather than personal ones, both on downstream and upstream knowledge flows. In fact, the positive effect of knowledge sharing perception gap on capability perception gap is reduced by the introduction of technology based sharing tools. Investments in content management tools such as Wikis or e-learning platforms can mitigate the emergence of capabilities perception gap thus enhancing the subsidiary financial performance. This is in line with the "death of distance" brought by the advent of the Industry 4.0's information technology tools (Ambos and Ambos 2009). The present study has thus direct and interesting implications especially with reference to the role of knowledge sharing tools. At the same time, our analysis suffers from certain limitations related to the fact that data collection has been limited to a single firm. Moreover, concerning flows of knowledge within MNCs, we only focused on "vertical" flows—knowledge shared between HQ and subsidiary—while neglecting "horizontal" flows—knowledge shared among peer subsidiaries. Future research can build on this chapter's quantitative approach in order to further analyse the introduced relationships. For example, studies related to other industries (e.g., services) or to different HQ geographical location could be implemented in order to compare achieved results.

Appendix 1: Case Study[3]

PACK is an Italian manufacturing firm operating in the packaging sector and established in the '1970s, when Mr. Jack Neri started a mechanical workshop in the North of Italy. As a birthday gift for their age of majority, each of Jack's four sons received from the father 25 % of the workshop's shares. When the last son reached the majority, the whole ownership of the firm had been transferred from the first to the second generation. This is the starting point of the "Brothers Neri" mechanical workshop, which few years later, was renamed PACK. Through the years—especially when web communications were not diffused—packaging industry's customers asked to their suppliers to be closely located to them in order to enhance the efficiency of their production plants. The Neri brothers thus

[3]In order to ensure anonymity, both the company and the family members are identified with pseudonyms.

started to establish their subsidiaries around the world: the internationalization process started in 1996, when PACK opened the first representative greenfield office in Malaysia, followed the next year by the Mexican one.

Currently, with about 700 employees and a total turnover of around 130 million euros, PACK operates all around the world with ten subsidiaries and three representative offices. The relationships between the Italian HQ and its subsidiaries have been strengthened by the use of a customized Enterprise Resource Planning (ERP) system used to manage and control communication flows within the MNC's network.

Appendix 2: Survey's Items

Variable	Item
Capability perception gap	Based on your working experience, evaluate subsidiaries according to the following dimensions (1 = strongly under the average; 7 = strongly up the average): collecting information, distributing information, analysing and acting on information.
Knowledge sharing perception gap $_{HQ \rightarrow sub}$	Based on your working experience, evaluate HQ with a score from 1 (completely disagree) to 5 (strongly agree). – The parent sends enough information to local branches. – The parent's expectations with respect to the local branches are clearly defined.
Firm tenure HQ_i	How long have you been working for the MNC?
Tacitness	Based on your working experience, indicate how much do you agree with the following claims (1 = strongly disagree; 7 = strongly agree) – A manual describing how our activities are executed could be written. – New staff can easily learn how to perform the services that our local company offers by talking to skilled employees. – Training new personnel is typically a quick and easy job for us. – New personnel with a university education can perform the services that our local company offers.
Knowledge sharing perception gap $_{sub \rightarrow HQ}$	Based on your working experience, evaluate the subsidiary with a score from 1 (completely disagree) to 5 (strongly agree). – The parent receives enough information from the local branches.
TCM & PCM	Indicate how frequently each of the following knowledge management processes and tools are used (1 = strongly under the average; 7 = strongly up the average) TCM: group learning facilities (from multiple sources or at multiple points at time), mapping specific types of knowledge (i.e. an individual, specific system, or database), chat groups/web-based discussion group, pointers to expertise (skills "yellow pages" within the company) PCM: liaison personnel, temporary task forces, permanent teams

(continued)

Variable	Item
Subsidiary financial performance	Based on your working experience, evaluate subsidiaries according to the following dimensions (1 = strongly under the average; 7 = strongly up the average): overall sales revenue, overall market share, operating profit (EBIT ≈ Revenues—Costs)

Appendix 3: Variables Operationalization

	Typology of variable	Variable	Source
Hypothesis 1	Dependent	Capability perception gap	Monteiro et al. (2008)
	Independent	Knowledge sharing perception gap $_{\text{HQ} \rightarrow \text{sub}}$	Asakawa (2001)
	Moderator	TCM & PCM	Ambos and Ambos (2009)
	Control	Firm tenure HQ_i	/
		Tacitness_i	Monteiro et al. (2008)
		Geographical distance	
Hypothesis 2	Dependent	Capability perception gap	Monteiro et al. (2008)
	Independent	Knowledge sharing perception gap $_{\text{sub} \rightarrow \text{HQ}}$	Asakawa (2001)
	Moderator	TCM & PCM	Ambos and Ambos (2009)
	Control	Firm tenure HQ_i	/
		Tacitness_j	Monteiro et al. (2008)
		Geographical distance	
Hypothesis 3	Dependent	Subsidiary financial performance	Monteiro et al. (2008)
	Independent	Capability perception gap	Monteiro et al. (2008)
	Control	Host country economic level	Monteiro et al. (2008)

References

Abidi, S., & Joshi, M. (2015). *The VUCA company*. Mumbai: Jaico.

Almeida, P., & Phene, A. (2004). Subsidiaries and knowledge creation: The influence of the MNC and host country on innovation. *Strategic Management Journal, 25*(8–9), 847–864.

Ambos, T., & Ambos, B. (2009). The impact of distance on knowledge transfer effectiveness in multinational corporations. *Journal of International Management, 15*(1), 1–14.
Argote, L., & Ingram, P. (2000). Knowledge transfer: A basis for competitive advantage in firms. *Organizational Behavior and Human Decision Processes, 82*(1), 150–169.
Armstrong, M. (2010). HR strategic management. *Tehran, Cultural Researches Office Publications.*
Asakawa, K. (2001). Organizational tension in international R&D management: The case of Japanese firms. *Research Policy, 30*(5), 735–757.
Bartlett, C., & Ghoshal, S. (1988). Tap your subsidiaries for global reach. *Harvard Business Review, Nov–Dec 1986*, 87–94.
Berry, H., Guillén, M., & Zhou, N. (2010). An institutional approach to cross-national distance. *Journal of International Business Studies, 41*(9), 1460–1480.
Birkinshaw, J., Holm, U., Thilenius, P., & Arvidsson, N. (2000). Consequences of perception gaps in the headquarters–subsidiary relationship. *International Business Review, 9*(3), 321–344.
Bouquet, C., & Birkinshaw, J. (2008). Weight versus voice: How foreign subsidiaries gain attention from corporate headquarters. *Academy of Management Journal, 51*(3), 577–601.
Campisi, D., & Passiante, G. (2007). *Fondamenti di knowledge management: Conoscenza e vantaggio competitivo*. Rome: Aracne.
Chini, T., Ambos, B., & Wehle, K. (2005). The headquarters–subsidiaries Trench: Tracing perception gaps within the multinational corporation. *European Management Journal, 23*(2), 145–153.
Cyert, R. M., & March, J. G. (1963). *A behavioral theory of the firm*. Englewood Cliffs, NJ: Prentice Hall.
Dyer, J., & Nobeoka, K. (2000). Creating and managing a high-performance knowledge-sharing network: The Toyota case. *Strategic Management Journal, 21*(3), 345–367.
Freeman, L. (1979). Centrality in social networks conceptual clarification. *Social Networks, 1*(3), 215–239.
Ghoshal, S., & Nohria, N. (1989). Internal differentiation within multinational corporations. *Strategic Management Journal, 10*(4), 323–337.
Gold, A., & Arvind Malhotra, A. (2001). Knowledge management: An organizational capabilities perspective. *Journal of Management Information Systems, 18*(1), 185–214.
Gottschalk-Mazouz, N. (2002). Diskursethische Varianten. *Deutsche Zeitschrift für Philosophie, 50*(1), 87–104.
Haas, M., & Hansen, M. (2005). When using knowledge can hurt performance: The value of organizational capabilities in a management consulting company. *Strategic Management Journal, 26*(1), 1–24.
Håkanson, L., & Nobel, R. (2001). Organizational characteristics and reverse technology transfer. *MIR: Management International Review*, 395–420.
Hélie, S., & Sun, R. (2010). Incubation, insight, and creative problem solving: A unified theory and a connectionist model. *Psychological Review, 117*(3), 994.
Hermann, M., Pentek, T., & Otto, B. (2016, January). Design Principles for Industrie 4.0 Scenarios. In *2016 49th Hawaii International Conference on System Sciences* (HICSS) (pp. 3928–3937). IEEE.
Kagermann, H., Helbig, J., Hellinger, A., & Wahlster, W. (2013). *Recommendations for implementing the strategic initiative INDUSTRIE 4.0: Securing the future of German manufacturing industry; final report of the Industrie 4.0 Working Group*. Forschungsunion.
Kogut, B., & Zander, U. (1992). Knowledge of the firm, combinative capabilities, and the replication of technology. *Organization Science, 3*(3), 383–397.
Kostova, T. (1999). Transnational transfer of strategic organizational practices: A contextual perspective. *Academy of Management Review, 24*(2), 308–324.
Irma Becerra-Fernandez, R. (2001). Organizational knowledge management: A contingency perspective. *Journal of Management Information Systems, 18*(1), 23–55.
March, J. G., & Simon, H. A. (1958). *Organizations*. Oxford: Wiley.

Monteiro, L., Arvidsson, N., & Birkinshaw, J. (2008). Knowledge flows within multinational corporations: Explaining subsidiary isolation and its performance implications. *Organization Science, 19*(1), 90–107.

Nonaka, I. (1991). The knowledge-creating company. *Harvard Business Review, 69*(6), 96–104.

Nonaka, I., Konno, N., & Toyama, R. (1998). Leading knowledge creation: A new framework for dynamic knowledge management. In *Second Annual Knowledge Management Conference, Haas School of Business*, University of California, Berkeley, CA.

Nonaka, I., Toyama, R., & Konno, N. (2000). SECI, Ba and leadership: A unified model of dynamic knowledge creation. *Long Range Planning, 33*(1), 5–34.

Nohria, N., & Ghoshal, S. (1994). Differentiated fit and shared values: Alternatives for managing headquarters-subsidiary relations. *Strategic Management Journal, 15*(6), 491.

Oviatt, B., & McDougall, P. (1994). Toward a theory of international new ventures. *Journal of International Business Studies, 25*(1), 45–64.

Oviatt, B., & McDougall, P. (2005). Defining international entrepreneurship and modeling the speed of internationalization. *Entrepreneurship Theory and Practice, 29*(5), 537–554.

Pfeffer, J., & Salancik, G. (1978). *The external control of organisations* (p. 175). New York: Harper & Row.

Polanyi, M. (1966). *The logic of tacit inference*. *Philosophy, 41*(155), 1–18.

Quagini, L. (2004). *Business intelligence e knowledge management. Gestione delle informazioni e delle performances nell'era digitale*. Milan: Franco Angeli.

Ravetz, J. (1971). *Scientific knowledge and its social problems*. Oxford: Oxford University Press.

Reus, T., Lamont, B., & Ellis, K. (2015). A darker side of knowledge transfer following international acquisitions. *Strategic Management Journal, 37*, 932–944.

Schotter, A., & Beamish, P. (2011). Performance effects of MNC headquarters–subsidiary conflict and the role of boundary spanners: The case of headquarter initiative rejection. *Journal of International Management, 17*(3), 243–259.

Schulz, M., & Jobe, L. (2001). Codification and tacitness as knowledge management strategies: An empirical exploration. *The Journal of High Technology Management Research, 12*(1), 139–165.

Schwab, K. (2016). *The fourth industrial revolution*. Geneva: World Economic Forum.

Sternberg, R. (1994). *Tacit knowledge and job success. Assessment and selection in organizations: Methods and practice for recruitment and appraisal* (pp. 27–39). London: John Wiley.

Szulanski, G. (1996). Exploring internal stickiness: Impediments to the transfer of best practice within the firm. *Strategic Management Journal, 17*(S2), 27–43.

Wang, S., & Noe, R. (2010). Knowledge sharing: A review and directions for future research. *Human Resource Management Review, 20*(2), 115–131.

Zahra, S. A., Neck, H. M., & Kelley, D. J. (2004). International corporate entrepreneurship and the evolution of organizational competence: A knowledge-based perspective. In *Advances in entrepreneurship, firm emergence and growth* (Vol. 7, pp. 145–171). Bingley: Emerald Group Publishing.

Family Business and Entrepreneurship: Competencies and Organizational Behavior

Serena Cubico, Giuseppe Favretto, Piermatteo Ardolino, Stefano Noventa, Diego Bellini, Giovanna Gianesini, and João Leitão

Abstract Family-owned enterprises go through various stages of growth and development over time once the second and subsequent generations enter the business. They present a unique set of challenges and problems, mostly due to generational transition, role confusion, different vision of the business, qualification and skills. This quantitative study aims at increasing the understanding of which competencies are relevant to family business success and how they significantly influence both organizational behavior and family dynamics. Findings highlight the importance of family business as a unique context to study organizational behavior and confirmed how family enterprises are influenced by family culture and kinship ties.

Keywords Competences • Succession • Aptitude • Skills • Family Business Management

1 Introduction

1.1 Family Business in Italian Economic System

Family businesses represent the large majority of enterprises (Astrachan and Shanker 2006; Beckhard and Dyer 1983; Corbetta 1995; International Family Enterprise Research Academy [IFERA] 2003): *"in terms of contributions, and especially numbers, family businesses represent a dominant form of economic organization throughout the world"* (Chrisman et al. 2003 p. 441).

S. Cubico (✉) • G. Favretto
Department of Business Administration, University of Verona, Via Cantarane 24, Verona, Italy
e-mail: serena.cubico@univr.it

P. Ardolino • S. Noventa • D. Bellini • G. Gianesini
Centre for Assessment, University of Verona, Via S. Francesco 22, Verona, Italy

J. Leitão
Management and Economics Department, University of Beira Interior and CEG-IST, Covilhã, Portugal

University of Lisbon, Lisbon, Portugal

T. Devezas et al. (eds.), *Industry 4.0*, Studies on Entrepreneurship, Structural Change and Industrial Dynamics, DOI 10.1007/978-3-319-49604-7_17

In Italy, the number of family-owned companies is a percentage of 70 % as emerged in the most recent national census of industry and services (Istat—Italian National Institute of Statistics 2013). Family businesses represent the core of the economy; capitalistic family ventures are largely represented by small companies that have family members in managerial roles (Cnel 2005; Colli 1998).

The crucial role of family in creating social and economic prosperity is confirmed in different ways (Habbershon and Pistrui 2002). The family is seen as a controlling subject in the economy and job creation (Shanker and Astrachan 1996), as a major source of start-up capital (Steiner 2001), and as the most enduring organizational type for entrepreneurial activity in a developing economy (Pistrui et al. 1997).

Other data (Olivari 2009) show that family businesses proved to be stronger than others during the economic crisis and the factors influencing success appear to be: a stable leadership, low use of financial support and sustainable growth. The same author also highlights that the low number of young and women in strategic roles represents a risk factor.

In this scenario, in order to have a working definition of family businesses, we use a traditional one: *"organizations where two or more extended family members influence the direction of the business through the exercise of kinship ties, management roles, or ownership rights"* (Tagiuri and Davis 1982, reprint 1996 p. 199). This definition allows us to reflect on the specific dynamic that characterizes the enterprise in which there are combined factors of family, ownership, management, and employment.

1.2 Intergenerational Coexistence in Family Business

"*The feeling that in recent decades the distance between the present and past and the different modes of thinking and acting of the younger generation and the elderly have increased progressively in our and other societies is shared by many observers of social reality...*" (Calvi 2005 p. 115). We have chosen this statement to show that contact between generations is a crucial element. Life expectancy has increased, entrepreneurs/founders remain longer at work with positions of responsibility, and following generations share most of their working lives with their parents. In family businesses, family, job, property and business converge so that family members are involved in the enterprise and share a much more intense part of life than others who work in different environments (Gersick et al. 1997).

> *Family and context, family and business, tend to lose their original boundaries and become a 'relational field' where the family and social actors interchange* (Scabini et al. 2008 p. 23).

In the family business, people of different ages, or better yet, people of different generations of the same family live together (Davis and Tagiuri 1989). This coexistence creates the conditions for the development of specific dynamics between younger and older (Togni 2008). Juniors pass through three professional stages: inclusion, autonomy and success; seniors live those of: maturity, continuance and integrity. For both generations, life cycles are characterized by different motivations, defences, relational dynamics, group dynamics, conflict management, time structure, conflict management and change. For the same author (p. 83) *"the challenge is to get the best out of the time of this relationship and to establish effective daily generational balance in decisions, risks, self-realization and company management"*.

1.3 Entrepreneurial Competences: Variables and Models

In our attempt to define competences, a subject that involves different disciplines, we have tried to synthesize by highlighting selected approaches that have an interesting relationship to entrepreneurial competences. The first important identification of this theme comes from White (1959) and its subsequent application by McClelland (1973, 1976); the American tradition defines competences as those personal characteristics that are related to a superior performance and motivation (this model represents the behavioural approach, starting from an observation of performance in order to highlight individual differences). After McClelland, Spencer and Spencer (1993 p. 6) identify "*motives, traits, self-concepts, attitudes or values, content knowledge, or cognitive or behavioral skills—any individual characteristic that can be measured or counted reliably and that can be shown to differentiate significantly between superior and average performers, or between effective and ineffective performers*" as competences.

Many other models are present in the literature, for example: the British tradition of a *functional approach* (a framework concerning occupational standards of competences, Mansfield and Mitchell 1996); Cheetham and Chivers' *holistic model* (1996, 1998) with the five dimensions of cognitive, functional, personal, ethical, and meta-competencies; the *multi-dimensional approach* developed in mainland Europe (Winterton et al. 2005); the *holistic model of competences* that integrates cognitive, functional and social competences with the meta-competences at its core (Delamare-Le Deist and Winterton 2005).

This brief summary of competence models (which does not claim to be complete) allows us to use as our proposal the recommendation of European Parliament Council to foster "*key-competences for lifelong learning*" and that defines entrepreneurship as *"an individual's ability to turn ideas into action. It includes creativity, innovation and risk-taking, as well as the ability to plan and manage projects in order to achieve objectives"* (European Parliament and Council 2006, pp. 17–18) and "***Skills** [that] relate to proactive project management (involving, for example the ability to plan, organise, manage, lead and delegate, analyse, communicate, debrief, evaluate and record), effective representation and negotiation, and the ability to work both as an individual and collaboratively in teams. The ability to judge and identify one's strengths and weaknesses, and to assess and take risks as and when warranted, is essential*" (European Commission 2007 p. 11).

This institutional position invites scholars interested in generational coexistence and transition in family businesses to study in depth the different variables that are related to entrepreneurial competences, described as: skills, aptitudes, attitudes, motivations, relationships.

2 Family Business and Competence Dynamics: A Research

The composition of the board in family firms is an expression of family characteristics and objectives (Voordeckers et al. 2007; Corbetta and Salvato 2004) and *"family businesses are highly dependent on a single decision-maker, the owner"* (Feltham et al. 2005 p. 13). Strategic daily decision-making in small and medium sized family businesses is therefore strictly bound up with the figure of the entrepreneur (Garcia-Alvarez and Lopez-Sintas 2001) and coexistence with the new generation (junior) requires sharing spaces for decisions and roles.

With the scenarios described, it is important to define and to know the dynamic of the competences in these two generations in order to understand where and how it is possible to intervene in facilitating the crucial moment of generational transition.

2.1 Objectives, Instruments and Subjects

Our research main **Objectives** are:

- to describe the competences that characterize different generations (work motivations, skill levels in different organizational tasks, entrepreneurial aptitude)
- to illustrate the quality of the relationship during work and family life;
- to describe different attitudes toward the generational transition process;
- to detect links between competencies and relationship among generations in family businesses.

The research **Hypotheses** are:

H1 Generations of the same family involved in family business present differences in entrepreneurial motivations;

H2 Senior and junior describe their own and other professional skills in different ways;

H3 Senior and junior perceive the quality of relationship in different measures;

H4 Entrepreneurial aptitude is present in different measures in two generations;

H5 Senior and junior have different attitudes toward the generational transition process;

H6 Competences are related to relationships between generations.

In Table 1 a summary of variables and hypothesis.

Table 1 Construct, variables and hypothesis of the research

Competences model	Variables	Research hypothesis
Motivations	Needs: – Earning a daily living – Ensuring the future for oneself and one's family – Feeling a part of the community – Feeling appreciated and esteemed – Being responsible for one's own success	*Generations of the same family involved in family business present differences in entrepreneurial motivations*
Skills	Own Expertise / Other Expertise: – Personal updating of Technology and data processing – Innovation of products / services – Thinking about the future of the company – Appreciation of Collaborators – Circumventing fiscal constraints – Administration and Accounting Management – Attention to issues and safety rules – Cash management and investments – Proper use of banking services – Anticipation and interpretation of customer needs – Choosing the most profitable suppliers – Sales skills – Being competitive with competitors – Carving out an exclusive market	*Senior and junior describe their own and other professional skills in different ways*
Quality of intergenerational relationships	Work Relationships / Family Relationships: – Completely in harmony – Parallel and independent – Conflictual Mutual Positive Comments – Expressed – Received	*Senior and junior perceive the quality of relationship in different measures*
Entrepreneurial aptitudes	Factors: – Goal Orientation – Leadership – Adaptability – Need for Achievement – Need for Self-Empowerment – Innovation – Flexibility – Autonomy	*Entrepreneurial aptitude is present in different measures in two generations*

(continued)

Table 1 (continued)

Competences model	Variables	Research hypothesis
Attitude toward generational transition	Generational Transition: – is a problem – we are prepared	*Senior and junior have different attitudes toward the generational transition process*
	– Own Expertise/Other Expertise – Work Relationships/Family Relationships – Mutual Positive Comments	*Competences are related to relationships between generations*

The **Instruments** used are an ad hoc *Questionnaire on Generational Transition [QGT]* (Favretto et al. 2003) and the *TAI—Entrepreneurial Aptitude Test* (Cubico et al. 2010).

The *QGT*[1] is made up of 19 items, both in the "Senior form" and "Junior form". The items are equal or comparable to each other so that a statistical comparison between the two figures with respect to the investigated variables under investigation can be made. The senior's questionnaire differs only in that it also contains some items that are designed to collect data about the enterprise (those listed in the analysis of company characteristics).

The 19 items of both forms of the questionnaire can be grouped into three main areas: conception of role, skills and vision of the entrepreneur; psycho-social components of senior-junior relationship; image of generational transition.

The *TAI* "*Test di Attitudine Imprenditoriale*" (Cubico et al. 2010; Sartori et al. 2007) describes "*the potential toward creating and developing enterprise and self-employment*" with regards to eight factors (Cubico et al. 2010 p. 427–428):

– *Goal Orientation*—tendencies toward creativity and innovation, degree of determination in reaching goals, and personal perception as to overall handling of work situations;
– *Leadership*—aptitudes toward management and leadership;
– *Adaptability*—ability to perceive environmental change and adaptability;
– *Need for Achievement*—the desire for fame, success and social affirmation and respect from others;
– *Need for Self-Empowerment*—the desire to realize oneself through one's job which, apart from any economic goals, must be enjoyable, satisfying and interesting,
– *Innovation*—curiosity for what is new;
– *Flexibility*—tendency to reorient one's goals according to external situation;
– *Autonomy*—necessity of having one's own independent space to make decisions and choices.

[1]Cronbach's Alpha= .858

Both instruments have been applied separately to each senior and junior (in all cases there was an interviewer).

The research **Subjects** are a group of 57 northern Italian micro, small and medium sized family businesses in different sectors (commerce, tourism, food farming, electro- and metal-mechanic, transports, wood and furniture, services for companies and people).

In particular:

– 117 entrepreneurs, *fathers* and *sons*

  • 57 seniors; average age 58.51; s.d. 7.63; 81 % male
  • 60 juniors; average age 30.12; s.d. 6.45; 73 % male.

2.2 Results

The results are divided as follows: motivations, skills, quality of the intergenerational relationships, aptitude, attitude toward generational transition, and relationship within these variables.

Motivations The motivations toward the entrepreneurial job are an interesting way to understand which expectations senior and junior put into the enterprise; in Table 2 we can see that "To ensure personal and family future" is the most selected motivation in both generations (although a low level), and that juniors are more interested in the goals of success than the seniors.

The desire to be responsible for personal success characterizes the new generation, and this makes us think that it is the *sons* who feel less significant in the life of the enterprise. We accept H1 in part.

Table 2 What needs of the entrepreneur are satisfied in the enterprise?

Needs...	Mean s.d. SENIOR	Mean s.d. JUNIOR	p level
Earning a daily living	1.91 .96	1.70 .79	n.s.
Ensuring the future for oneself and one's family	1.98 .98	1.89 .96	n.s.
Feeling a part of the community	1.04 .19	1.09 .45	n.s.
Feeling appreciated and esteemed	1.11 .32	1.13 .48	n.s.
Being responsible for one's own success	1.13 .39	**1.39** .59	**.009**

(ANOVA; N = 108; Likert Scale = 1 insufficient−5 excellent)

Skills It is interesting to reflect on the skills that seniors and juniors perceived themselves and others as having. The first generation is characterized by the capacity to think of the enterprise's future and to choose the best suppliers. In particular, seniors recognize, more significantly than juniors, that they possess skills for management and accounting, financial investment, and for dealing with bank and suppliers. juniors perceived themselves as being skilled in the understanding customer needs and personal updating of technology; however, it was this last area of expertise that significantly characterized the younger generation's self-definition (Table 3).

When we ask them to reflect on the skills that characterize the other generation, the seniors highlight juniors' capacities in technological and human resources task, while the juniors recognize seniors' skills in fiscal, accounting, investment and banking services (Table 4).

Table 3 What is your skill level in each of the following areas?

Own area of skill	Mean s.d. SENIOR	Mean s.d. JUNIOR	p level
Personal updating of technology and data processing	2.22 1.16	**3.39** 1.09	**.000**
Innovation of products/services	3.54 1.09	3.44 1.00	n.s.
Thinking about the future of the company	4.09 1.15	3.76 .97	n.s.
Appreciation of collaborators	3.57 1.08	3.37 .96	n.s.
Circumventing fiscal constraints	2.07 1.21	2.02 1.16	n.s.
Administration and accounting management	**3.30** 1.11	2.85 1.19	**.047**
Attention to issues and safety rules	3.30 1.09	3.22 1.14	n.s.
Cash management and investments	**3.57** 1.13	2.61 1.05	**.000**
Proper use of banking services	**3.39** 1.12	2.65 1.07	.001
Anticipation and interpretation of customer needs	3.91 .99	3.89 1.02	n.s.
Choosing the most profitable suppliers	**3.85** .86	3.37 1.09	**.012**
Sales skills	3.80 1.16	3.44 1.18	n.s.
Being competitive with competitors	3.70 .90	3.37 .94	n.s.
Carving out an exclusive market	3.37 1.19	3.31 1.28	n.s.

(Anova; N = 108; Likert Scale = 1 insufficient−5 excellent)

Table 4 What is the skill level of the other generation in each of the following areas?

Area of skill of other	Mean s.d. SENIOR	Mean s.d. JUNIOR	p level
Personal updating of technology and data processing	**4.07** .93	2.15 1.16	**.000**
Innovation of products/services	3.74 1.01	3.46 1.02	n.s.
Thinking about the future of the company	3.67 1.23	3.96 1.81	n.s.
Appreciation of collaborators	**3.78** 1.02	3.28 1.17	**.020**
Circumventing fiscal constraints	2.24 1.24	**2.80** 1.28	**.024**
Administration and accounting management	2.81 1.30	**3.46** 1.19	**.008**
Attention to issues and safety rules	3.48 1.19	3.26 1.17	n.s.
Cash management and investments	2.83 1.24	**3.61** 1.14	**.001**
Proper use of banking services	2.78 1.22	**3.43** 1.24	**.007**
Anticipation and interpretation of customer needs	3.83 1.01	3.81 1.17	n.s.
Choosing the most profitable suppliers	3.70 1.13	3.67 1.20	n.s.
Sales skills	3.85 1.17	3.76 1.24	n.s.
Being competitive with competitors	3.63 .99	3.65 .94	n.s.
Carving out an exclusive market	3.30 1.27	3.63 1.19	n.s.

(ANOVA; N = 108; Likert Scale = 1 insufficient−5 excellent)

If we compare Tables 3 and 4, we see that the junior tends to see in the senior a higher number of skills than in himself, and that there is agreement in the two generations regarding some specific skills.

H2 is acceptable for the large part: senior and junior describe most of their own and other professional skills in different ways.

Quality of Intergenerational Relationships In Tables 5 and 6, the perception of quality in the work and family relationships is described. In particular, the most positive relationship is in family life but we can see that the junior is significantly less satisfied than the senior in both environments.

The quality of intergenerational relationships is described in the comments that seniors and juniors express and perceived receiving too; in Table 7 we can see the level of low satisfaction in the new generation.

Table 5 How do you define your work relationships?

Work relationships	Mean s.d. SENIOR	Mean s.d. JUNIOR	p level
Completely in harmony	**3.78** 0.92	3.39 1.10	**.050**
Parallel and independent	3.44 1.25	3.20 1.17	n.s.
Conflictual	2.19 1.20	2.28 1.17	n.s.

(ANOVA; N = 108; Likert Scale = 1 insufficient–5 excellent)

Table 6 How do you define family relationships?

Family relationships	Mean s.d. SENIOR	Mean s.d. JUNIOR	p level
Completely in harmony	**4.15** .96	3.67 1.17	**.021**
Parallel and independent	3.83 1.13	3.61 1.16	n.s.
Conflictual	1.89 1.08	2.04 1.17	n.s.

(ANOVA; N = 108; Likert Scale = 1 insufficient–5 excellent)

Table 7 How often do you express/receive positive comments?

Mutual positive comments	Mean s.d. SENIOR	Mean s.d. JUNIOR	p level
Expressed	**3.33** 1.10	2.83 1.14	**.023**
Received	3.09 1.10	2.67 1.17	n.s.

(ANOVA; N = 108; Likert Scale = 1 never–5 always)

Entrepreneurial Aptitudes The results on the eight factors of the entrepreneurial aptitude indicate slight (but not significant) differences between the two generations (Table 8).

H4 is rejected; it is encouraging to see that entrepreneurial aptitude is present in the same measure in the two generations in the family business.

Attitude toward Generational Transition The image of the generational transition is positive and the entrepreneurs feel to be prepared for it, but juniors are more pessimistic about the vision of the future (Table 9) and see generational transition as a significant problem, more so than for the first generations.

Relationship with Competences We consider it interesting to understand the links between skills, relationships, quality and image of generational transition. These bonds could lead to a better understanding of competence dynamics in the family business.

Table 8 Factor results in senior and junior—entrepreneurial aptitude

Entrepreneurial aptitude factor	Mean s.d. SENIOR	Mean s.d. JUNIOR	p level
F1 Goal orientation	64.67 10.89	66.07 9.81	n.s
F2 Leadership	60.79 14.30	59.67 14.59	n.s
F3 Adaptability	64.49 7.44	66.78 8.29	n.s
F4 Need for achievement	58.39 15.84	57.33 13.25	n.s
F5 Need for self-empowerment	72.86 9.91	73.25 11.04	n.s
F6 Innovation	71.68 13.61	74.10 14.79	n.s
F7 Flexibility	68.04 11.28	68.50 12.26	n.s
F8 Autonomy	60.26 12.73	60.25 14.03	n.s

(ANOVA; N = 117)

Table 9 How do you define generational transition?

Generational transition...	Mean s.d. SENIOR	Mean s.d. JUNIOR	p level
...is a problem	2.69 1.43	**3.33** 1.29	**.015**
... we are prepared	3.06 1.27	3.15 1.22	n.s.

(ANOVA; N = 108; Likert Scale = 1 insufficient−5 excellent)

In order to describe the skill profile synthetically, we have created[2] two variables and have labelled them as follows: Own Expertise and Other Expertise (respectively the score of the skills attributed to self and to others). The Own/Other Expertises present many correlations with the relationships' perceptions and image of generational transition, as in Table 10.

It is possible to synthesize the most interesting correlations (rho>.300**)

– Own Expertise is positively correlated with: work/family harmony and family parallel relationships;
– Other Expertise is positively correlated with: work/family harmony, expressed/received positive comments.

[2]The procedure verifies scale reliability in order to input the answer with a unique score.

Table 10 Correlations between skills and relationships variables (Spearman's Rho; N = 108)

	1	2	3	4	5	6	7	8	9	10	11	12
1.Work harmony	–											
2.Rel. parallel in work	.264**	–										
3.Work conflict	−.568**	−.274**	–									
4.Family harmony	.666**	.246*	−.317**	–								
5.Rel. parallel in family	.274**	.517**	−.114	.445**	–							
6.Family conflict	−.569**	−.152	.553**	−.712**	−.401**	–						
7.Expressed comments	.507**	.259**	−.293**	.507**	.220*	−.344*	–					
8.Received comments	.451**	.189	−.297**	.420**	.213*	−.371**	.560**	–				
9.Generational transition: problem	−.273**	.024	.215*	−.176	−.064	.225*	−.168	−.164	–			
10.Generational transition: we are ready	.164	.150	−.151	.203*	.083	−.108	.116	.144	−.124	–		
11.Own expertise	.316**	.228*	−.101	.350**	.435**	−.229*	.223*	.233*	−.075	.165	–	
12. Other expertise	.383**	.201*	−.136	.457**	.280*	−.290**	.433**	.344**	−.219*	.130	.289**	–

These data highlight the importance of the awareness of personal and other competences in the relationships between generations (in family and work environment).

Table 11 presents finding of the Regression Model Analysis. In the work and family environments, Model 2, which includes the Others' Expertise, is more explicative of the quality in the relationship. We can underline the importance of considering the other generation's competences as an important variables in creating positive relationships between generations in the family business.

H6 is partially confirmed: competencies are related to the relationship between generations.

3 Discussion

We can see that some differences emerge and that the resulting data reveal that the crucial moment of generational transition can be influenced by different personal and relational variables also. The question can also be raised as to whether these differences can be an element of positive value or not.

It is important to understand the differences discerned between seniors and juniors and whether they can help the company in the future or whether they present a risk as an obstacle in generational transition, thereby inflicting damage on the company itself.

The research presented reveals that recognizing mutual competences and the different components that characterize in intergenerational coexistence is a possible way of helping the generational transition process.

The dynamics of the competences influence relationships and the subsequent possibility of having a rewarding work life.

The limits of the work is recognizable in the lack of a single definition and model for competences, however we think that this limitation will lead to the future definition of a synthetic model to include variables that appear correlated.

Table 11 Regression model: competences on relationship variables (N = 108)

Dependent variables	Relationships			
	Harmony in Work		Harmony in Family	
	Model 1 *ß*	Model 2 *ß*	Model 1 *ß*	Model 2 *ß*
Constant	2.467	1.189	2.361	1.031
Own expertise	.171	.34	.209	.076
Other expertise		.486**		.472**
Adj R^2	.014	.223	.029	*.225*
Model F	*1.969*	*10.485*	*2.974*	*10.600*
Significance	*.165*	*.000*	*.089*	*.000*

(*p<.05; **p<.01)

Acknowledgment João Leitão acknowledges Fundação para a Ciência e a Tecnologia (FCT) for support under the Sabbatical Scholarship Program (Ref.: SFRH/BSAB/127908/2016).

References

Astrachan, J. H., & Shanker, C. M. (2006). Family businesses's contribution to US economy: A closer look. In P. Z. Poutziouris, K. X. Smirnyos, & B. Klein (Eds.), *Handbook of research in family business* (pp. 56–64). Cheltenham: Edward Elgar Publishing.

Beckhard, R., & Dyer Jr., W. G. (1983). Managing continuity on the family owner business. *Organizational Dynamics, 12*(1), 5–12.

Calvi, G. (2005). Le percezioni esistenziali, un indice della distanza tra generazioni. In G. Calvi (Ed.), *Generazioni a confronto. Materiali per uno studio* (pp. 115–152). Venezia, IT: Marsilio.

Cheetham, G., & Chivers, G. (1996). Towards a holistic model of professional competence. *Journal of European Industrial Training, 20*(5), 20–30.

Cheetham, G., & Chivers, G. (1998). The reflective (and Competent) practitioner: A model of professional competence which seeks to harmonise the reflective practitioner and competence-based approaches. *Journal of European Industrial Training, 22*(7), 267–276.

Chrisman, J. J., Chua, J. H., & Steier, L. P. (2003). An introduction to theory of family business. *Journal of Business Venturing, 18*(4), 441–448.

Cnel—Consiglio nazionale dell'economia e del lavoro (2005). *Piccole Imprese: concorrenza e competitività nel mercato globale*. Roma, IT: Pronounce 76. http://www.cnel.it/53?shadow_documenti=10526.

Colli, A. (1998). Networking the market: Evidence and conjectures from the history of the Italian industrial districts. *European Yearbook of Business History, 1*, 75–92.

Corbetta, G. (1995). Patterns of development of Family businesses in Italy. *Family Business Review, 8*(4), 255–265.

Corbetta, G., & Salvato, C. A. (2004). The board of directors in family firms: One Size Fits All? *Family Business Review, 17*(2), 119–134.

Cubico, S., Bortolani, E., Favretto, G., & Sartori, R. (2010). Describing the Entrepreneurial Profile: The Entrepreneurial Aptitude Test (TAI). *International Journal of Entrepreneurship and Small Business, 11*(4), 424–435.

Davis, J. A., & Tagiuri, R. (1989). The influence of life stage on Father-Son work relationships in family companies. *Family Business Review, 2*(1), 47–74.

Delamare-Le Deist, F., & Winterton, J. (2005). What is competence? *Human Resource Development International, 8*(1), 27–46.

European Commission (2007). Key competences for lifelong learning: European reference framework. https://www.britishcouncil.org/sites/default/files/youth-in-action-keycomp-en.pdf.

European Parliament and Council. (2006). Recommendation of the European Parliament and of the Council of 18 December 2006 on Key Competences for Lifelong Learning. Official Journal of the European Union (2006/962/EC). http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2006:394:0010:0018:EN:PDF.

Favretto, G., Sartori, R., & Bortolani, E. (2003). Il passaggio generazionale nella PMI a gestione familiare. *Risorsa Uomo, 9*(3-4), 361–374.

Feltham, T. S., Feltham, G., & Barnett, J. J. (2005). The dependence of family business on a single decision-maker. *Journal of Small Business Management, 43*(1), 1–15.

Garcia-Alvarez, E., & Lopez-Sintas, J. (2001). A taxonomy of founders based on values: The root of family business heterogeneity. *Family Business Review, 14*(3), 209–230.

Gersick, K. E., Davis, J. A., Hampton, M., & Lansberg, I. (1997). *Generation to generation: Life cycles of the family business*. Boston, MA: Harvard Business School Press.

Habbershon, T. G., & Pistrui, J. (2002). Enterprising families domain: Family-influenced ownership groups in pursuit of transgenerational wealth. *Family Business Review, 15*(3), 223–237.

IFERA—International Family Enterprise Research Academy. (2003). Family business dominate. *Family Business Review, 16*(4), 235–239.

Istat—Italian National Institute of Statistics. (2013). *Rapporto Annuale 2013. La situazione del paese*. Roma, IT: Istat http://www.istat.it/it/files/2013/05/Rapporto_annuale_2013.pdf.

Mansfield, B., & Mitchell, L. (1996). *Towards a competent workforce*. London, UK: Gower.

McClelland, D. (1973). Testing for competence rather than for 'intelligence'. *American Psychologist, 28*(1), 1–14.

McClelland, D. (1976). *A guide to job competency assessment*. Boston, MA: McBer.

Olivari, L. (2009). Più forti di fronte alla crisi. *L'impresa, 1*, 48–51.

Pistrui, D., Welsch, H., & Roberts, J. (1997). The [re]-emergence of family businesses in the transforming Soviet Bloc: Family contributions to entrepreneurship development in Romania. *Family Business Review, 10*(3), 221–238.

Sartori, R., Favretto, G., & Sartori, A. (2007). Aspetti quantitativi dell'attitudine imprenditoriale. Il TAI (Test di Attitudine Imprenditoriale). In G. Favretto & R. Sartori (Eds.), *Le età dell'impresa* (pp. 101–122). Milano, IT: Franco Angeli.

Scabini, E., Tamanza, G., & Pozzi, M. (2008). Dinamiche familiari e trasmissione generazionale nelle imprese di famiglia. In G. Trentini & M. Togni (Eds.), *Continuità generazionale d'impresa. Dimensioni psicologiche e relazionali* (pp. 19–38). Franco Angeli: Milano, IT.

Shanker, M. C., & Astrachan, J. H. (1996). Myths and Realities: Family Business's Contributions to the US Economy. *Family Business Review, 9*(2), 107–123.

Spencer, L., & Spencer, S. (1993). *Competence at Work: A Model for Superior Performance*. New York: Wiley.

Steiner, L. (2001). Family Firms, plural Forms of Governance, and the Evolving Role in Trust. *Family Business Review, 14*(4), 353–368.

Tagiuri, R., & Davis, J. A. (1982). Bivalent Attributes of the Family Firm. Cambridge, MA: Working Paper Harvard Business School [reprint (1996), *Family Business Review*, *9*(2), 199–208].

Togni, M. (2008). La relazione tra generazioni nell'impresa. Fasi e percorsi di sviluppo. In G. Trentini & M. Togni (Eds.), *Continuità generazionale d'impresa. Dimensioni psicologiche e relazionali* (pp. 83–106). Franco Angeli: Milano, IT.

Voordeckers, W., Van Gils, A., & Van den Heuvel, J. (2007). Board composition in small and medium-sized family firms. *Journal of Small Business Management, 45*(1), 137–156.

White, R. (1959). Motivation reconsidered: The concept of competence. *Psychological Review, 66*, 279–333.

Winterton, J., Delamare-Le Deist, F., & Stringfellow, E. (2005). *Prototype typology of knowledge, skills and competences*. Thessaloniki, GR: CEDEFOP http://www.cedefop.europa.eu/files/3048_en.pdf.

Metrics for Innovation and Entrepreneurial Networks

Luís Farinha and João J. Ferreira

Abstract This study aims to find a model to measure the regional dynamics of the Triple and Quadruple Helixes (Academia-Firms-Local Governments and Civil Society), based on innovativeness pillars. In accordance with this purpose, different strands of literature are identified according to their focus on specific regional competitiveness governance mechanisms. We put forward an overview of the state-of-the-art of research and is duly assessed in order to develop and propose a framework of analysis that enables an integrated approach in the context of collaborative dynamics. We conclude by presenting the Regional Helix Scoreboard (RHS) for the study of regional collaborative dynamics, which integrates and associates different backgrounds and identifies a number of key performance indicators for research challenges.

Keywords Entrepreneurship • KPIs • Innovation • BSC • Regional Competitiveness • Regional Helix Scoreboard

1 Introduction

Sustainability has to be approached within a perspective combining the environmental, economic and social dimensions (Gopalakrishnan et al. 2012). From the perspective of regional development, the competitive production of companies determines the levels of earnings and employability at the level of the regional business environment while demand is determined according to relative costs (Budd and Hirmisf 2004).

L. Farinha (✉)
Instituto Politécnico de Castelo Branco and NECE—Research Unit in Business Sciences, Av. Pedro Álvares Cabral, no 12, 6000-084 Castelo Branco, Portugal
e-mail: luis.farinha@ipcb.pt

J.J. Ferreira
Department of Management and Economics and NECE—Research Unit in Business Sciences, University of Beira Interior, Estrada do Sineiro, 6201-209 Covilhã, Portugal
e-mail: jjmf@ubi.pt

T. Devezas et al. (eds.), *Industry 4.0*, Studies on Entrepreneurship, Structural Change and Industrial Dynamics, DOI 10.1007/978-3-319-49604-7_18

According to Budd and Hirmisf (2004), competitiveness is productivity, the main determinant in the long run of an economy's standard of living. However, Jiang and Shen (2010) argue that competitiveness needs to be viewed in a balanced way, allowing a focus on the sustainable development orientation. Acs and Amorós define competitiveness as a result weighted by the behaviour of the different variables making up the Global Competitiveness Index (GCI) published annually by the World Economic Forum (WEF) (Acs and Amorós 2015), where the pillars of innovation and business sophistication are included. In a more territorial approach, progressive regions have a competitive advantage in attracting opportunities for development, capturing high-tech companies and talent, ensuring greater wealth creation and employability (Audretsch et al. 2012).

According to the Regional Innovation Scoreboard (RIS) published by the EU, innovation is a key factor determining productivity and economic growth (EC 2014a, b). In parallel, the interest in regional innovation and regional innovation systems as a source of competitive advantage has grown significantly over the past three decades (Asheim et al. 2009, 2011).

The concepts of creativity and innovation are often used as synonyms in the literature, although some authors emphasize the distinction: creativity results in the creation of new ideas; innovation requires its implementation in practice (Osterwalter and Pigneur 2010). From the perspective of Porter and Stern, there is a set of factors transversal to the economy that support innovation and including: the human and financial resources allocated to scientific and technological advances, the level of technological sophistication, the public policies affecting innovation related activities, intellectual property protection, fiscal incentives for innovation and enacting and effectively implementing antitrust and abuse of power legislation (Porter and Stern 2001).

Several experts have been advocating innovation and entrepreneurship as determinants of competitiveness and regional development (Porter and Stern 2001; Wong et al. 2005; Nordqvist and Melin 2010).

Entrepreneurship in GEM perspective presents a number of advantages associated with the creation of new businesses to the local economy, some of which are innovating, creating new jobs and wealth (Kelley et al. 2011).

Overall, there is a significant relationship between business activities, networking and productivity, not forgetting the importance of entrepreneurship and innovation capacity in the context of competitive aggressiveness (Huggins et al. 2014). Nowadays, strategic alliances, cooperation and networking of work play an important role in terms of regional development. These collaborative networks translate into a competitive advantage in the global market (Semlinger 2008). We should therefore ascertain the actual capacities of local companies to sell their products in external markets, the value of such sales and their productive levels of efficiency while also taking into account the incorporation of local resources and including both human and capital dimensions (Turok 2004).

Networks of cooperation and business are key to the competitiveness of regions by facilitating the implementation of new investment projects, allowing the increase in supply of new local suppliers or the formation of industry clusters, thereby making it possible to compete in the most demanding markets (Porter and Stern 2001; Budd and Hirmis 2004). According to the logics underpinning regional development, the predominance of the relationships between (A) Academia—(I) Industry—(G) Government (state or municipal) and specific local activities (for example, local technology transfers, the development of human capital and networking), in conjunction, determine better overall results (Smith and Bagchi-Sen 2012).

As a tool for measuring performance, Kaplan and Norton developed the Balanced Scorecard (Chytas et al. 2011); now used worldwide as a strategic management tool (Chytas et al. 2011; Wu and Chang 2012). However, some limitations are recognized to the model itself including not being able to respond effectively to all situations under analysis. To meet the changing demands to measuring performance resulting from alliances between institutions and projects management for regional development, new models of performance measurement have now been developed from the original BSC model (Philbin 2008; Al-Ashaab et al. 2011; Ioppolo et al. 2012). Unfortunately, the traditional BSC and its upgrades are neither totally appropriate nor useful to measure the performance of the Triple Helix regional interactions (Academia—Industry—Government), in the regional context of innovation, entrepreneurship and competitiveness. Therefore, this paper proposes a Regional Helix Scoreboard model (RHS) to measure the A-I-G interactions and thereby enriching the literature in this area.

Our purpose here is therefore to address four research questions: Question 1. In regional networking, are knowledge and technology transfer and R&D significant for competitiveness? Question 2. Do the A-I-G collaborative networks play an important role in innovation and entrepreneurship? Question 3. What is the role of the government in A-I-G networks? Question 4. How can we measure the impact of A-I-G collaborative networks in regional competitiveness?

This study aims to develop an integrative conceptual model aimed at measuring the dynamic interactions of the triple helix, balancing innovation and entrepreneurship initiatives as pillars of the competitiveness of regions—the “Regional Helix Scoreboard” (RHS).

The chapter is structured as follows: firstly, we carry out a literature review on innovation, entrepreneurship, competitiveness and the emergence of the triple helix system and its dynamics. Secondly, we set out a Regional Helix Scoreboard model for regional competitiveness. Finally, we put forward our concluding remarks.

2 Cooperation and Collaborative Networking: The Triple Helix Approach

In the current regional policy, business and cooperation networking are increasingly seen as the key to success (Semlinger 2008). Networks of R&D cooperation are assumed to be real organizational and economic contexts where companies join other institutions (companies, research centres, universities or others), creating umbrella networks to various locations in order to develop technological projects that can positively affect competitiveness, also here inserting public institutions aimed at promoting the development of their technology policies, sometimes supported by public framework programs to promote the establishment of networks for the development of R&D projects (Arranz and Arroyabe 2008). Organizations need to establish networks with external entities in order to acquire or have access to resources not otherwise available, especially the acquisition of technological resources, access to infrastructure and technological know-how or the establishment of agreements to comply with financial, economic and legal issues or at the level of knowledge transfer (Awazu 2006).

For many people, the terms “cooperation” and “collaboration” are indistinguishable. Cooperation involves communication and information exchanges; the complimentary goals and aligning activities; the compatibility of goals, individual identities and working apart (Verdecho et al. 2012c; Al-Ashaab et al. 2011). Meanwhile, collaboration adds joint goals, joint identities, creating together and joint responsibility—corresponding to a higher level of integration and maturity.

Backing up this perspective on how regional competitiveness and development determine the productive capacity of companies and regional levels of income and employability (Budd and Hirmisf 2004); other authors highlight the predominance of relationships between academia—industry—government (state, regional or local) and specific local activities in determining the best business results and outcomes (Smith and Bagchi-Sen 2012). Etzkowitz argues that the interactions of the triple helix are the key to innovation in societies increasingly based on knowledge, helping students, researchers and policy makers to respond to certain questions (Etzkowitz 2008): How do we strengthen the role of academia in economic and social development at the regional level?, How can governments encourage citizens to take an active role in promoting innovation?, How can firms collaborate with academia and government? The metaphoric model of the Triple Helix focuses on A-I-G interaction dynamics, including the transfer of knowledge through the development of Research, Development and Innovation (RD&I) between the network partners, facilitating access to new forms of financing, still managing an increase of production and market access synergies (Etzkowitz 2003). Various evolutionary stages need accounting for in terms of the many interactions between the triple helix spheres (Etzkowitz 2003, 2008). The evolution of innovation systems and the current dispute over which path is most appropriate for university—industry relationships effects the different institutional agreements in terms of the overall A-I-G relationships (Etzkowitz and Leydesdorff 2000).

The Triple Helix emerges from regional areas of knowledge, innovation and consensus and thus can play an important role in regional development and competitiveness, through the interaction between the different institutional spheres in a networking logic (Etzkowitz 2008; Hira 2013). Ozman further underlines that networks indisputably play an important innovation role (Ozman 2009). Consistent to this interpretation, a series of academic studies has recognised that cooperation between the three institutional spheres (A-I-G) is fundamental to improving regional and national innovation systems (Etzkowitz 2003, 2008; Hira 2013; Ozman 2009; Leydesdorff and Meyer 2006; Nissan et al. 2011).

The stability of interactions A-I-G takes on an important role in RD & I capacity and also in production competitiveness (Etzkowitz 2003; Etzkowitz and Leydesdorff 2000); Etzkowitz, and proving fundamental to boosting regional innovation systems (Smith and Bagchi-Sen 2012; Hira 2013; Farinha et al. 2016; Smith and Bagchi-Sen 2006).

Contemporary relationships deriving from interactions ongoing between the spheres of university and industry are resulting in a third hybrid current whether out of common interests in basic research, partnership projects between industry and higher education institutions as well as through the joint establishment of research and development programs making recourse to multiple sources of Etzkowitz (2008). Regional policies of the last decades have been fostering the emergence of spin-off companies, resulting from the knowledge transfer mechanisms and A-I technology. The science and technology park represent an organizational innovation, capable of supporting these spin-offs and the business community in general, tend to be organized in sectoral clusters. Incubators and accelerators companies have come even to facilitate reinforced cooperation network between companies, and between them and the Academy and the political decision (Salvador 2010).

Apply the conceptual model of the Triple Helix to the problem of competitiveness of territories, highlighting the need to increase the levels of innovation and development of local businesses, proves the starting point for a better theoretical understanding (Kim et al. 2012).

Given the changes in societies that have shaken off domination by a central instance, some authors have felt the case for presenting possible new alternative model scales with four or more helixes based on new variables (Leydesdorff 2012; MacGregor et al. 2010) fostering regional competitiveness and development (Audretsch et al. 2015, 2012). Innovation is assumed by Porter and Stern, as the decisive challenge able to ensure global competitiveness (Porter and Stern 2001). Referred to by some researchers within the scope of a fourth helix-pillar are independent organisations, without any profit motive and combining public and private financing. The dynamics of interaction A-I-G aims facilitating the capture public and private investment, while seeking time to strengthen the RD&I, otherwise the micro and small businesses do not have access. The sharing of infrastructure and technical knowledge, or strengthening of synergies in terms of product offering on the market, are also a priority (MacGregor et al. 2010; Audretsch and Keilbach 2005).

3 Finding a Tool for Measuring A-I-G Network Performance Levels

As proposed by Kitson in 2004, regional competitive advantage furthermore inherently requires articulated involvement and action across a multi-level scenario, within which feature the different variants of capital (Kitson et al. 2004). By strengthening the RD&I; of human and creative capital stimulus; improvement of productive capital; capture of finance capital; or reorientation of policy options to support innovation and entrepreneurship, it becomes possible to increase the business sophistication and increased competitive position in the global market (Kelley et al. 2011).

Lawton Smith focus on entrepreneurship and the geography of talent directly linked to economic performance and also constituting a strong contribution to the sustainable development of the regions (Smith et al. 2005). Other authors argue that entrepreneurial activity is an important mechanism for regional development through job creation and creating local wealth, whether it comes from the transfer of knowledge and technology from academia or simply through the creation of new businesses (Salvador 2010; Nordqvist and Melin 2010).

According to Van Looy, the logic of "university ventures" is tightly bound up with the existence of shortcomings in the innovation market (Van Looy et al. 2011). There are two trends shaping the framework of contemporary developments in relationships between university and industry: interests in basic research financed by research entities and councils and industrial projects, which universities are invited to participate in, with a third hybrid current emerging from the formulation of joint research programs making recourse to multiple sources of financing (Etzkowitz 2008).

According Carayannis sustainability should also be perceived within a three-dimensional approach: environmental, economic-financial and social, thereby boosting the competitive advantage of regions (Carayannis et al. 2012). Harris in 2009 point out how ethics and entrepreneurship remain inherently bound up and of particular relevance within the framework of entrepreneurial activities and regional development.

The triple helix spheres, while set out contextualised within their external environment (the political, economic, social, cultural and technological contexts), describe the dynamic and interactive movements of partnerships, supported by and in the format of cooperative networks striving to boost competitiveness (Bernasconi 2005).

The increasing levels of local intellectual capital and institutional support enable the development of an interactive group of private and public interests, acting through a network of organisational and institutional agreements and fostering the dissemination of knowledge, technologies and regionally located innovation skills and capacities (Hira 2013).

Aiming to answer the basic research questions formulated, and based on the literature review conducted, we found the need to measure the performance of the

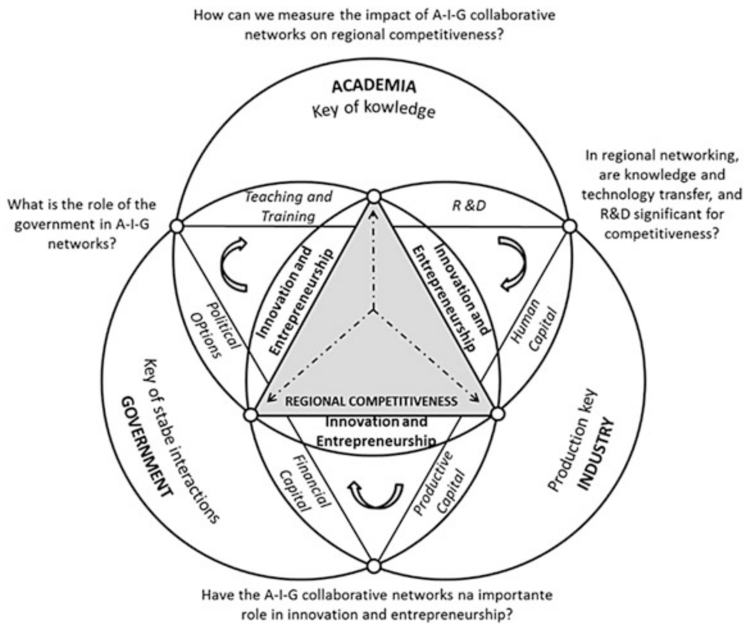


Fig. 1 The triple helix dynamics. Source: Own elaboration

resulting A-I-G network interactions in order to make it possible to measure their impact on regional competitiveness (see Fig. 1).

The balanced scorecard (BSC) was developed by Kaplan and Norton (1992) to give managers a balanced view of organizations working with other important strategic factors—from continuous improvement and partnerships to teamwork and global scale, besides the classical financial measurements. The BSC seeks to provide answers to four basic questions:

- How do customers see us? (Customer perspective)
- What must we excel at? (Internal business perspective)
- Can we continue to improve and create value? (Innovation and learning perspective)
- How do we look to shareholder? (Financial perspective).

Targeting a new impulse to the implementation of the BSC, an article published in the Harvard Business Review set out three case studies applying the scorecard to measure performance and strategy (Kaplan and Norton 2001a; Kaplan and Norton 2001b) [22], [68].

The BSC developed by Kaplan and Norton (2001a), is today used globally as a strategic management tool for measuring performance (Chytas et al. 2011; Dror 2008; Kanji and Sá 2002; Lazzarotti et al. 2011; Philbin 2008; Sundin et al. 2009; Taylor and Baines 2012; Tseng 2010; Verdecho et al. 2012a, b, c; Wu and Chang 2012; Al-Ashaab et al. 2011; Ioppolo et al. 2012). However, some limitations are recognized as to the model itself and identified as unable to respond effectively to all situations under analysis (Kanji and Sá 2002; Philbin 2008; Verdecho et al. 2012c; Al-Ashaab et al. 2011; Ioppolo et al. 2012).

4 Data Collection and Method

Having defined the general framework of innovation, entrepreneurship, competitiveness and the dynamics of the triple helix, we performed a systematic review of the literature about BSC and BSC for networking performance measurement. Such a literature review establishes the state of the art in a specific field (D'Este and Perkmann 2010).

Our objective is to analyse the state of the art on BSC to measure the performance of networks of cooperation and collaboration between different agents. To this end, we applied the following procedure. We first identified all the relevant research published on BSC from 1990 to 2013. We carried out an extensive search through the titles and abstracts of published, peer-reviewed articles held in the bibliographical database Thomson Reuters (ISI) Web of Knowledge, using a predetermined series of keywords (BSC; Cooperation BSC; Collaborative BSC; and Networking performance measuring). We subsequently performed a manual search of the journals with the highest article citations over the past 21 years (1992–2013), filtering the records of the 50 most cited articles in that time period from the around 1,000 articles identified.

Analyzing the trend towards a greater increase in publications about BSC, we encountered a boom between 2010 and 2013. For performing the literature review search, we followed a constructivist methodological approach consisting of the identification of a problem of practical relevance, its theoretical connections and the acquisition of its main postulates (Alfaro et al. 2009; Verdecho et al. 2012c). Thus, we decided to apply a filter adjusted to this time period under analysis—we filtered the records with the most cited articles from 2010 to 2013 achieving a synthesis with the 25 most cited articles in ISI journals. This composed a summary table providing a comparison between the articles within the historical perspective of the 50 most cited and the most recent trend of the 25 most cited (see Table 1).

Having completed this stage, and better understanding trends in the usage of the BSC performance measurement within networks of cooperation and collaboration, we again applied a filter to our database.

This time, restricted to the keywords "cooperation BSC; collaborative BSC; and networking performance measurement", applied to the period prescribed between 1992 and 2013. After content analysis, we attained a list with the 10 most cited articles, which establish the basis for the development of our model (see Table 2).

In the range of selected articles, we find works applied to different areas or sectors of activity, from health, management or science of computing, and applying different methodological approaches.

Table 1 Synthesis of articles most cited by time series

Journal	N° of articles	Citations	Journal	N° of articles	Citations
1992–2013: Top 50—times cited (historical perspective)			2010–2013: Top 25—times cited (most recent trend)		
Accounting Organizations and Society	6	382	Omega—International Journal of Management	3	16
Harvard Business Review	5	2072	Accounting Review	2	6
Accounting Review	4	351	Expert Systems with Applications	2	9
Long Range Planning	3	201	Environmental Monitoring and Assessment	1	21
Expert Systems with Applications	3	157	Technological and Economic Development of Economy	1	20
California Management Review	2	171	Journal of The Royal Society Medicine	1	11
International Journal of Operations & Production Management	2	171	International Journal of Hospitality Management	1	9
Computers & Industrial Engineering	2	148	Evaluation and Program Planning	1	7
European Journal of Operational Research	2	114	Plos Medicine	1	6
Journal of Operational Research	2	114	Clinical Psychology-Science and Practice	1	5
Computers in Industry	2	101	Journal of Operations Management	1	4
Others	17	1301	Others	10	26
Sum	*50*	*5283*	*Sum*	*25*	*140*

Source: Own elaboration

5 Proposed Regional Helix Scoreboard (RHS)

Collaboration amongst enterprises is a common strategy deployed to increase competitiveness. Therefore, this requires the measuring of the performance of these business processes under a strategic approach from an inter-enterprise perspective, defining and using performance measurement/management frameworks composed of performance related factors (objectives, performance indicators, etc.) that facilitate the management of activities as well as monitoring strategy and processes (Verdecho et al. 2012a, b, c; Grubic and Fan 2010; Alfaro et al. 2009; Stanley et al. 2010; Herath et al. 2010).

Promoting high quality research networks inherently requires the establishment of evaluation tools for measuring performance and the corresponding definition of metrics and performance indicators (Singhal et al. 2012). While companies

Table 2 BSC for networking performance measurement: top 10—times cited

Author(s)	Journal	Title	Methodology/ Method	Year	Citations
Alfaro et al. (2009)	International Journal of Computer Integrated Manufacturing	Business process interoperability and collaborative performance measurement	Conceptual	2009	8
Stanley et al. (2010)	Contemporary clinical trials	Development and implementation of a performance measure tool in an academic pediatric research network	Case study	2010	2
Herath et al. (2010)	Journal of Accounting and Public Policy	Joint selection of balanced scorecard targets and weights in a collaborative setting	Mathematical programming models/Simplex Method	2010	2
Ioppolo et al. (2012)	Land Use Policy	Developing a Territory Balanced Scorecard approach to manage projects for local development: Two case studies	Two case studies	2012	2
Perkmann et al. (2011)	R&D Management	How should firms evaluate success in university-industry alliances? A performance measurement system	Conceptual	2011	2
Verdecho et al. (2012a)	Omega—International Journal of Management	A multi-criteria approach for managing inter-enterprise collaborative relationships	Case study/ Analytic Network Process (ANP)	2012	1
Verdecho al. (2012b)	Decision Support Systems	Prioritization and management of inter-enterprise collaborative performance	Quantitative/ Analytic Network Process (ANP)	2012	1
Wu and Chang (2012)	Decision Support Systems	Using the balanced scorecard in assessing the performance of e-SCM diffusion: A multi-stage perspective	Quantit./ struct. equation modelling (SEM)	2012	1
Chytas et al. (2011)	International Journal of Information Management	A proactive balanced scorecard	Fuzzy logic/ Fuzzy Cognitive Maps (FCMs)	2011	1
Al-Ashaab et al. (2011)	Production, Planning & Control	A balanced scorecard for measuring the impact of industry-university collaboration	Two case studies	2011	1

Source: Own elaboration

increasingly engage in formal alliances with universities, there is a lack of tools for evaluating the results of these collaborations (Perkmann et al. 2011).

The BSC is considered such a strategic measurement tool. Various companies have applied it to measure four key perspectives of their organisation's performance: financial, customer, internal business processes and learning and growth. However, this original model was not developed to measure the impact of collaborative research projects ongoing under an open innovation strategy (Al-Ashaab et al. 2011).

In order to meet these new measuring performance requirements resulting from collaborative alliances between institutions, new performance measurement models were developed out of the original balanced scorecard (Philbin 2008; Al-Ashaab et al. 2011; Ioppolo et al. 2012). Al—Ashaab put forward a balanced scorecard for measuring the impact of industry–university collaboration—the collaborative BSC, and Ioppolo developed the Territory BSC to manage local development projects (Al-Ashaab et al. 2011; Ioppolo et al. 2012) (see Table 3).

The KTForce is a project supported by the INTERREG IVC Capitalisation Programme, co-financed by the European Regional Development Fund (ERDF), and its aim is to improve the effectiveness of regional development policies in the fields of innovation and the knowledge economy. This involves eleven partners from six regions, covering "modest and moderate innovator" regions (Lithuania, Portugal and Romania) and "innovation follower and leader" regions (France, Germany and Ireland) (EC 2014a, b; Sternberg 2012; Farinha et al. 2014). For the measuring of innovative performance, the KTForce project comprises a set of indicators, distributed by pillars "Technology licensing", "Spin-offs creation and entrepreneurship" and "University-Industry relations".

The quality and abilities of the labour force (human capital); the extension, depth and focus of social and institutional networks (social/institutional capital), the range and quality of infrastructure as well as cultural assets (cultural capital), the presence of a creative and innovative class (knowledge/creative capital) and the quality of infrastructural policies and results (infrastructural capital) are all deemed to be critical factors in supporting and determining regional economic outcomes (Huggins et al. 2014; Čučković et al. 2013; Balkyte and Tvaronavičiene 2010).

Recent developments saw concerted efforts by emerging countries to transform their industrial-based economy into post-industrial knowledge-based economy. The growth of science and technology is necessary to support this economic transformation strategy.

Seeking to fill some of the gaps in the literature on a global model for A-I-G interaction performance measurement, we now proceed with setting out a new conceptual model, based upon the Triple Helix model, as defended by a vast range of authors (Etzkowitz 2008; Etzkowitz et al. 2005; Etzkowitz and Leydesdorff 2000; Viale and Etzkowitz 2010), focused on innovation and entrepreneurship as critical factors to regional competitiveness through their capacities to stimulate new investment and job creation, thus driving economies to attain new standards of competition (Kelley et al. 2011).

In this context, we present the Regional Helix Scoreboard (RHS) (see Fig. 2).

Table 3 Characterization of different performance measurement models

Metrics									
	Networking view								
	Inter-enterprises		*Academic research network*				*A-I Collaboration*		*Intern. Cooper. Projects*
Author	Inter-enterprises business process interoperability	Collaborative enterprise network	General contributions	Intellectual development	Project implementation	Timely project completion	Inputs; In-process activities; Outputs; Impact	Competitiveness; Sustainable Development; Innovation; Strategic Knowledge Partnerships; Human Capital; Internal Business Processes	Internal Processes; New Public Governance; Learning & Innovation; Territory Strategic Development
Alfaro et al. (2009)	X								
Stanley et al. (2010)			X	X	X	X			
Verdecho et al. (2012c)		X							
Verdecho et al. (2012b)	X								
Perkmann et al. (2011)							X		
Al-ashaab et al. (2011)								X	
Ioppolo et al. (2012)									X

Source: Own elaboration

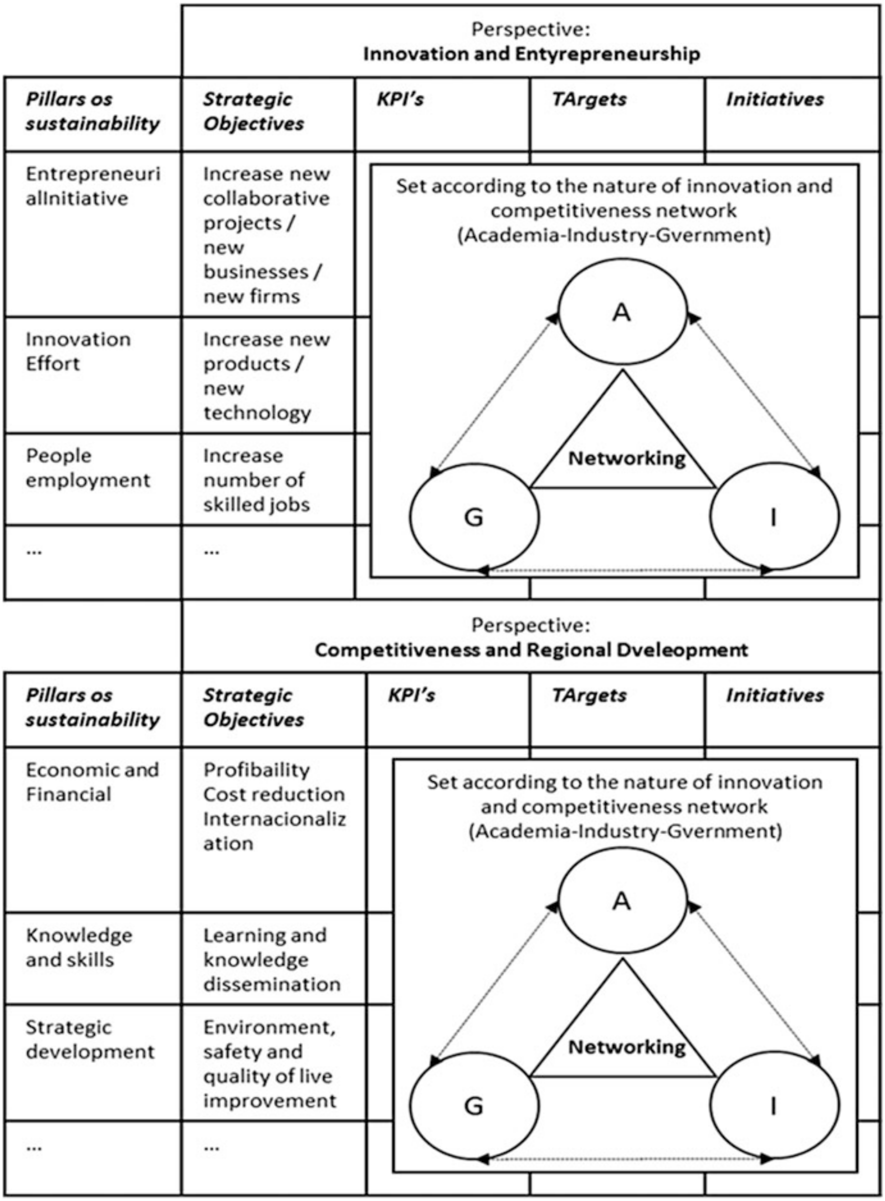


Fig. 2 Regional Helix Scoreboard model (RHS). Source: Own elaboration

Adjusted to the dynamics of the Triple Helix, and designed from the various inputs identified in the literature review, the THS derives from the initial BSC model from Kaplan and Norton, focuses on perspectives about "Innovation and entrepreneurship" and "Competitiveness and regional development" in order to measure the performance of the A-I-G interactions (Kaplan and Norton 2001a). Thus, for each of the perspectives, the model proposes a set of pillars of sustainability, which are the defined strategic objectives, KPIs, targets and initiatives and collectively aiming to answer the central research question: how do innovation and entrepreneurship linked to the dynamics of the triple helix contribute to increasing regional competitiveness and development?

In this perspective, "innovation and entrepreneurship" are identified through three main pillars of sustainability: "entrepreneurial initiative", "innovation effort", and "people employment". Regarding the perspective "regional competitiveness and development", the following pillars of sustainability were selected: "economic and financial", "knowledge and skills", and "strategic development". For each perspective and for each pillar of sustainability, strategic objectives and KPIs are defined and subject to adjustment in accordance with the nature of the respective innovation and competitiveness network.

Some of the most relevant **"Innovation and entrepreneurship"** strategic objectives and KPIs are:

- Strategic objectives/KPIs:
  - Increase in new collaborative projects/new business/new companies
    - Number of new companies created
    - Number of technology based companies created (spin-off)
    - Number of companies created > 250 employees
    - Number of grants funding start ups
    - Total number of value propositions developed
    - Total number of new business plans developed
    - Number of successful proposals developed collaboratively to obtain public funding.
  - Increase in new products/new technology
    - Number of patent requests and patents
    - Number of industrial property licenses
    - Number of intangibles resulting from collaborative projects in the form of patents, licenses, copyright or trademarks.
  - Increase in jobs
    - Number of jobs created
    - Number of skilled jobs created.

For **"Competitiveness and regional development":**

- Strategic objectives/KPIs:
  - Profitability
    - Turnover
    - Sales

– Cost reduction

  - Percentage of cost savings thanks to alliances
  - Percentage of cost savings thanks to university-based research.

– Internationalization

  - Export volume percentage
  - Linkages between international cluster networks.

– Learning and knowledge dissemination

  - R&D spending
  - Number of joint publications in scientific journals or conferences.

– Environment, safety and quality of life improvement

  - Number of projects developing new models and/or methods to improve sustainability practices: health and safety, recycling methods, sustainable construction, etcetera
  - Percentage of component reutilisation
  - Number of collaborative projects that environmentally or socially improved any region or facility.

Finally, so that evaluative conclusions may be drawn about performance, the definition of future target values and Initiatives with targets to meet are proposed.

We would note that depending on the nature of the A-I-G interaction (A-I collaborative projects; regional clusters, science parks and technology business incubators, etcetera), attention should be paid to appropriately adjusting the strategic objectives and KPIs.

6 Conclusion Remarks

This study puts forward an integrative conceptual model displaying a dynamic and interactive triple helix model able to clarify the role of innovation and entrepreneurship as factors of regional competitiveness. Entrepreneurship is defined in the literature as a high risk dynamic and with an especially high binomial level of effort-reward. Companies need to be able to innovate in the global marketplace, designing, producing and commercialising new products and evolving faster than their rivals.

The development of regions may correspondingly be segregated into exogenous development and endogenous development (Hira 2013). The triple helix model focuses on interactions ongoing between universities—industry—government as the key to improving the conditions necessary to innovation, based on changing the paradigm from industrial societies to knowledge based societies. Strengthening this perspective on regional competitiveness and development, the productive private sector capacity determines the prevailing levels of regional earnings and employability [20]. From the Global Entrepreneurship Monitor perspective, the launch of new companies results in investment inflows, new jobs and driving overall competitiveness and development (Kelley et al. 2011).

The TH relational model reflects the interaction of relationships ongoing between three institutional spheres (university—industry—government) designed to secure regional competitive advantage within the framework of actions interrelated across a multi-level scenario. The TH model thereby serves as the point of departure for designing and implementing empirically based studies, susceptible to providing responses to the questions raised relative to the interactions taking place in the different spheres. This is, in turn, based on the assumption of a positive relationship between the dynamics of innovation and entrepreneurship for regional competitiveness and development that needs empirical validation with recourse to the appropriate research methodologies (quantitative and/or qualitative). Our model, in the context of the dynamics of regional areas of the triple helix, seeks to answer the research questions presented in the introduction to this chapter and thus provides a mechanism for measuring the impact of these networks on regional competitiveness.

Considering the pertinence of developing this theme in future research, and irrespective of the prevailing economic conjuncture—with recessionary pressures at the global level and reflecting in the rescaling and postponement of new investment projects despite the corresponding need for job creation within the framework of a globalised and competitive economy in which innovation stands out as a key factor for competitiveness, all combine to ensure the priority attributed to regional development and its associated competitiveness.

This inherently requires the dissemination of knowledge and technology through a sustainable inter-organisational network. Based on this assumption, as future lines of research, we would suggest the empirical testing of the RHS as a tool for measuring the performance of triple/quadruple helix dynamics created from the Balanced Scorecard model and its developments in the fields of territorial and inter-organization collaborative relationships, now adjusted to the specific interactions of the triple helix.

Finally, we suggest studies which combine quantitative and qualitative research, with the creation and validation of instruments for collecting data through observation and document analysis, field notes, interviews and questionnaires in order to most fully test the Regional Helix Scoreboard here developed.

References

Acs, Z., & Amorós, J. (2015). Entrepreneurship and development: The role of clusters. *Small Business Economics, 39*(3), 561–574.

Al-Ashaab, A., Flores, M., Doultsinou, A., & Magyar, A. (2011). A balanced scorecard for measuring the impact of industry–university collaboration. *Production Planning and Control, 22*(5–6), 554–570.

Alfaro, J. J., Rodriguez-Rodriguez, R., Verdecho, M. J., & Ortiz, A. (2009). Business process interoperability and collaborative performance measurement. *International Journal of Computer Integrated Manufacturing, 22*(9), 877–889.

Arranz, N., & Arroyabe, J. C. (2008). The choice of partners in R&D cooperation: An empirical analysis of Spanish firms. *Technovation, 28*(1–2), 88–100.

Asheim, B. T., & Coenen, L. (2005). Knowledge bases and regional innovation systems: Comparing Nordic clusters. *Research Policy, 34*(8), 1173–1190.

Asheim, B. T., Ejermo, O., & Rickne, A. (2009). When is regional 'Beautiful'? Implications for knowledge flows, entrepreneurship and innovation. *Industry and Innovation, 16*(1), 1–9.

Asheim, B. T., Smith, H. L., & Oughton, C. (2011). Regional innovation systems: Theory, empirics and policy. *Regional Studies, 45*(7), 875–891.

Audretsch, D. B., & Keilbach, M. (2004). Entrepreneurship and regional growth: An evolutionary interpretation. *Journal of Evolutionary Economics, 14*(5), 605–616.

Audretsch, D. B., & Keilbach, M. (2005). Entrepreneurship capital and regional growth. *The Annals of Regional Science, 39*(3), 457–469.

Audretsch, D. B., Link, A. N., & Peña, I. (2012). Academic entrepreneurship and economic competitiveness: Introduction to the special issue. *Economics of Innovation and New Technology, 21*(5–6), 427–428.

Audretsch, D. B., Belitski, M., & Desai, S. (2015). Entrepreneurship and economic development in cities. *The Annals of Regional Science, 55*(1), 33–60.

Awazu, Y. (2006). Managing technology alliances: The case for knowledge management. *International Journal of Information Management, 26*(6), 484–493.

Balkyte, A., & Tvaronavičiene, M. (2010). Perception of competitiveness in the context of sustainable development: Facets of sustainable competitiveness. *Journal of Business Economics and Management, 11*(2), 341–365.

Bernasconi, A. (2005). University entrepreneurship in a developing country: The case of the P. Universidad Católica de Chile, 1985–2000. *Higher Education, 50*(2), 247–274.

Budd, L., & Hirmis, A. (2004). Conceptual framework for regional competitiveness. *Regional Studies, 38*(9), 1015–1028.

Budd, L., & Hirmisf, K. (2004, December). Conceptual framework for regional competitiveness. *University Business, 3*, 1015–1028.

Carayannis, E. G., Barth, T. D., & Campbell, D. F. (2012). The Quintuple Helix innovation model: Global warming as a challenge and driver for innovation. *Journal of Innovation and Entrepreneurship, 1*(1), 2.

Chytas, P., Glykas, M., & Valiris, G. (2011). A proactive balanced scorecard. *International Journal of Information Management, 31*(5), 460–468.

Čučković, N., Jurlin, K., & Vučković, V. (2013). Measuring regional competitiveness: The case of Croatia. *Southeast European and Black Sea Studies, 13*(4), 503–523.

D'Este, P., & Perkmann, M. (2010). Why do academics engage with industry? The entrepreneurial university and individual motivations. *The Journal of Technology Transfer, 36*(3), 316–339.

Dror, S. (2008). The balanced scorecard versus quality award models as strategic frameworks. *Total Quality Management and Business Excellence, 19*(6), 583–593.

EC. (2014a). European competitiveness report 2014: Helping firms grow. Belgium.

EC. (2014b). Innovation union scoreboard. Brussels.

Etzkowitz, H. (2003). Innovation in innovation: The triple helix of University—Industry—Government relations. *Social Science Information, 42*(3), 293–337.

Etzkowitz, H. (2008). *The triple helix: University–Industry–Government innovation in action* (1st ed.). Oxon: Routledge.

Etzkowitz, H., & Leydesdorff, L. (2000). The dynamics of innovation: From national systems and "Mode 2" to a triple helix of University—Industry—Government relations. *Science and Technology*, 109–123.

Etzkowitz, H., Mello, J. M. C., & Almeida, M. (2005). Towards 'meta-innovation' in Brazil: The evolution of the incubator and the emergence of a triple helix. *Research Policy, 34*(4), 411–424.

Farinha, L., Ferreira, J., & Gouveia, J. B. (2014). Innovation and competitiveness: A high-tech cluster approach. *The Romanian Review Precision Mechanics, Optics & Mechatronics, 45*, 41–48.

Farinha, L., Ferreira, J., & Gouveia, B. (2016). Networks of innovation and competitiveness: A triple helix case study. *Journal of the Knowledge Economy, 7*(1), 259–275.

Gopalakrishnan, K., Yusuf, Y. Y., Musa, A., Abubakar, T., & Ambursa, H. M. (2012). Sustainable supply chain management: A case study of British Aerospace (BAe) systems. *International Journal of Production Economics, 140*(1), 193–203.

Grubic, T., & Fan, I.-S. (2010). Supply chain ontology: Review, analysis and synthesis. *Computers in Industry, 61*(8), 776–786.

Herath, H. S. B., Bremser, W. G., & Birnberg, J. G. (2010). Joint selection of balanced scorecard targets and weights in a collaborative setting. *Journal of Accounting and Public Policy, 29*(1), 45–59.

Hira, A. (2013). Mapping out the triple helix: How institutional coordination for competitiveness is achieved in the global wine industry. *Prometheus, 31*(4), 271–303.

Huggins, R., Izushi, H., Prokop, D., & Thompson, P. (2014). Regional competitiveness, economic growth and stages of development. *Zbornik radova Ekonomskog fakulteta u Rijeci, časopis za ekonomsku teoriju i praksu, 32*(2), 255–283.

Ioppolo, G., Saija, G., & Salomone, R. (2012). Developing a territory balanced scorecard approach to manage projects for local development: Two case studies. *Land Use Policy, 29*(3), 629–640.

Jiang, Y., & Shen, J. (2010). Measuring the urban competitiveness of Chinese cities in 2000. *Cities, 27*(5), 307–314.

Kanji, G. K., & Sá, P. M. (2002). Kanji's business scorecard. *Total Quality Management, 13*(1), 13–27.

Kaplan, R. S., & Norton, D. P. (2001a). *The strategy-focused organization: How balanced scorecard*. Boston: Harvard Business School.

Kaplan, R. S., & Norton, D. P. (2001b). Transforming the balanced scorecard from performance measurement to strategic management: Part I. *Accounting Horizons, 15*(1), 87–104.

Kelley, D. J., Singer, S., Herrington, M. (2011). The global entrepreneurship monitor 2011 global report.

Kim, Y., Kim, W., & Yang, T. (2012). The effect of the triple helix system and habitat on regional entrepreneurship: Empirical evidence from the U.S. *Research Policy, 41*(1), 154–166.

Kitson, M., Martin, R., & Tyler, P. (2004). Regional competitiveness: An elusive yet key concept? *Regional Studies, 38*(9), 991–999.

Lazzarotti, V., Manzini, R., & Mari, L. (2011). A model for R&D performance measurement. *International Journal of Production Economics, 134*(1), 212–223.

Lengyel, I., & Rechnitzer, J. (2013). Drivers of regional competitiveness in the Central European Countries. *Transition Studies Review, 20*(3), 421–435.

Leydesdorff, L. (2012). The triple helix, quadruple helix, . . ., and an N-Tuple of helices: Explanatory models for analyzing the knowledge-based economy? *Journal of the Knowledge Economy, 3*(1), 25–35.

Leydesdorff, L., & Fritsch, M. (2006). Measuring the knowledge base of regional innovation systems in Germany in terms of a triple helix dynamics. *Research Policy, 35*(10), 1538–1553.

Leydesdorff, L., & Meyer, M. (2006). Triple helix indicators of knowledge-based innovation systems Introduction to the special issue. *Research Policy, 35*(10), 1441–1449.

Leydesdorff, L., Dolfsma, W., & Vanderpanne, G. (2006). Measuring the knowledge base of an economy in terms of triple-helix relations among 'technology, organization, and territory'. *Research Policy, 35*(2), 181–199.

MacGregor, S. P., Marques-Gou, P., & Simon-Villar, A. (2010). Gauging readiness for the quadruple helix: A study of 16 European organizations. *Journal of the Knowledge Economy, 1*(3), 173–190.

Mendes, P., Santos, A., Perna, F., & Ribau Teixeira, M. R. (2012). The balanced scorecard as an integrated model applied to the Portuguese public service: A case study in the waste sector. *Journal of Cleaner Production, 24*, 20–29.

Nissan, E., Galindo Martín, M.-Á., & Méndez Picazo, M.-T. (2011). Relationship between organizations, institutions, entrepreneurship and economic growth process. *The International Entrepreneurship and Management Journal, 7*(3), 311–324.

Nordqvist, M., & Melin, L. (2010). Entrepreneurial families and family firms. *Entrepreneurship and Regional Development, 22*(3–4), 211–239.

Osterwalder, A., & Pigneur, Y. (2010). *Business model generation: A handbook for visonairies, game changers and challengers*. Hoboken, NJ: Wiley.

Ozman, M. (2009). Inter-firm networks and innovation: A survey of literature. *Economics of Innovation and New Technology, 18*(1), 39–67.

Perkmann, M., Neely, A., & Walsh, K. (2011). How should firms evaluate success in university-industry alliances? A performance measurement system. *R&D Management, 41*(2), 202–216.

Philbin, S. (2008). Process model for university-industry research collaboration. *European Journal of Innovation Management, 11*(4), 488–521.

Porter, M. E., & Stern, S. (2001). Innovation: Location matters. *MIT Sloan Management Review, 42*(4), 28–36.

Salvador, E. (2010). Are science parks and incubators good 'brand names' for spin-offs? The case study of Turin. *The Journal of Technology Transfer, 36*(2), 203–232.

Semlinger, K. (2008). Cooperation and competition in network governance: Regional networks in a globalised economy. *Entrepreneurship and Regional Development, 20*(6), 547–560.

Singhal, S., McGreal, S., & Berry, J. (2012, June). Application of a hierarchical model for city competitiveness in cities of India. *Cities, 31*, 114–122.

Smith, H. L., & Bagchi-Sen, S. (2006). University–Industry interactions: The case of the UK biotech industry. *Industry and Innovation, 13*(4), 371–392.

Smith, H. L., & Bagchi-Sen, S. (2012). The research university, entrepreneurship and regional development: Research propositions and current evidence. *Entrepreneurship and Regional Development, 24*(5–6), 383–404.

Smith, H. L., Glasson, J., & Chadwick, A. (2005). The geography of talent: Entrepreneurship and local economic development in Oxfordshire. *Entrepreneurship and Regional Development, 17*(6), 449–478.

Stanley, R., Lillis, K. A., Zuspan, S. J., Lichenstein, R., Ruddy, R. M., Gerardi, M. J., & Dean, J. M. (2010). Development and implementation of a performance measure tool in an academic pediatric research network. *Contemporary Clinical Trials, 31*(5), 429–437.

Sternberg, R. (2012). Do EU regional policies favour regional entrepreneurship? Empirical evidence from Spain and Germany. *European Planning Studies, 20*(4), 583–608.

Sundin, H., Granlund, M., & Brown, D. A. (2009). Balancing multiple competing objectives with a balanced scorecard. *European Accounting Review, 19*.

Taylor, J., & Baines, C. (2012). Performance management in UK universities: Implementing the balanced scorecard. *Journal of Higher Education Policy and Management, 34*(2), 111–124.

Theriou, N. G., Demitriades, E., & Chatzoglou, P. (2004). A proposed framework for integrating the balanced scorecard into the strategic management process. *Operations Research, 4*(2), 147–165.

Tseng, M.-L. (2010). Implementation and performance evaluation using the fuzzy network balanced scorecard. *Computers in Education, 55*(1), 188–201.

Turok, I. (2004). Cities, regions and competitiveness. *Regional Studies, 38*(9), 1069–1083.

Van Looy, B., Landoni, P., Callaert, J., van Pottelsberghe, B., Sapsalis, E., & Debackere, K. (2011). Entrepreneurial effectiveness of European universities: An empirical assessment of antecedents and trade-offs. *Research Policy, 40*(4), 553–564.

Verdecho, M.-J., Alfaro-Saiz, J.-J., & Rodriguez-Rodriguez, R. (2012a). Prioritization and management of inter-enterprise collaborative performance. *Decision Support Systems, 53*(1), 142–153.

Verdecho, M.-L., Alfaro-Saiz, J.-J., Rodriguez-rodriguez, R., & Ortiz-bas, A. (2012b). A multi-criteria approach for managing inter-enterprise collaborative relationships. *Omega, 40*, 249–263.

Verdecho, M.-J., Alfaro-Saiz, J.-J., Rodríguez-Rodríguez, R., & Ortiz-Bas, A. (2012c). The analytic network process for managing inter-enterprise collaboration: A case study in a collaborative enterprise network. *Expert Systems with Applications, 39*(1), 626–637.

Viale, R., & Etzkowitz, H. (2010). *The capitalization of knowledge: A tripla helix of University-Industry-Government* (1st ed.). Cheltenham: Edward Elgar.

Wong, P. K., Ho, Y. P., & Autio, E. (2005). Entrepreneurship, innovation and economic growth: Evidence from GEM data. *Small Business Economics, 24*(3), 335–350.

Wu, I.-L., & Chang, C.-H. (2012). Using the balanced scorecard in assessing the performance of e-SCM diffusion: A multi-stage perspective. *Decision Support Systems, 52*(2), 474–485.

Determinants and Interdependence of Firm Entry, Exit and Mobility

Rui Baptista and Murat Karaöz

Abstract Structural evolution of the industries involves the mechanism of market selection, which comprises entry, exit as well as market share changes, mobility, of incumbent firms. However, entry and exit studies so far have paid attention to the interdependencies only between the entry and exit. In this paper, we include the third dimension of mobility for analyses. Industrial and Regional diversification, unemployment, and unit cost deters entry, exit and mobility of incumbents. However, regional diversification deters entry only for larger firms. Unit cost does not have clear effect on incumbent mobility. Industries with smaller sized firms and with higher growth volatility experience higher entry, exit and mobility. Growth cause more entry and mobility of incumbents and fewer exits. We also test the phenomenon that entry and exit activities symmetrically, simultaneously and with delay response to each other, considering mobility. We use a longitudinal data set from Portugal covering 1986–1993 time period and 320 six-digit industries. We classify industries into growing, mature and declining parts. According to the estimation results, growing industries experience significantly more entry, exit and changes in market shares among incumbents. Although the argument that declining industries experience less entry has support, its positive significance on exit equation has some weak support. Overall, the results support that entry, exit and incumbent mobility response to unobserved common determinants. Moreover, additional estimation results suggest that also there is interdependence among entry, exit and incumbent mobility.

Keywords Firm Entry • Exit and Mobility • Interdependence • Symmetry • Simultaneity • Displacement • Replacement • Industry Life Cycle

R. Baptista (✉)
CEG-IST, Instituto Superior Técnico, Universidade de Lisboa, Av. Rovisco Pais, 1049-001 Lisbon, Portugal
e-mail: rui.baptista@tecnico.ulisboa.pt

M. Karaöz
Department of Econometrics, IIBF, Akdeniz Üniversitesi, Dumlupınar Bulvarı, Kampüs, 07000 Antalya, Turkey

T. Devezas et al. (eds.), *Industry 4.0*, Studies on Entrepreneurship, Structural Change and Industrial Dynamics, DOI 10.1007/978-3-319-49604-7_19

1 Introduction

Structural evolution of the industries involves the mechanisms of entry, exit as well as market share changes (mobility) of incumbent firms (Caves 1998; Baldwin and Gorecki 1994). From the outset to the maturity and later decline, entry, exit and mobility of firms are constantly shapes the evolution of market structure toward its long-term equilibrium level. This relationship is *interdependent*, they response to each other either *simultaneously* or with delay. Moreover, they act upon same set of exogenous factors. New entry of firm brings additional amount of capacity and, perhaps, innovative product and processes to the industry, which distracts the incumbents. It may cause more competition and some of the inefficient incumbents are forced to leave the industry. As flows of entry and exit maintain, the incumbents continuously response to this process while adjusting their market shares with expansion or contraction, if not remain same. The process of entry, exit and mobility contour the structure of the industry as it proceeds along the life cycle. Whether the industry become concentrated with a few firms or become a competitive atomistic industry or fall in somewhere between the two all depends on how various causes shape the entry, exit and mobility of firms. The causes and consequences of entry, exit and mobility have been subject of numerous empirical and theoretical studies (Orr 1974; Khemani and Shapiro 1986; Shapiro and Khemani 1987; Audretsch and Acs 1990; Baldwin and Gorecki 1991; Geroski and Schwalbach 1991; Rosenbaum and Lamort 1992; Baldwin 1998; Fotopoulos and Spence 1998; Lay 2003 among others).

While there is a significant positive correlation between entry and exit, a range of studies in the industry life cycle suggest that different stages of the cycle yield different regularities in entry and exit rates. Entry, exit and mobility continuously enlarge the winners and shrink the losers and contribute productivity growth as industry evolves over time (Caves 1998). Thus, not only entries and exits but also the incumbent firms are part of the industries evolutionary process across the product life cycle. Higher levels of entry and exit rates tend to occur in emerging and growing industries, or in industries under rapid structural change (Gort and Klepper 1982; Klepper and Graddy 1990; Klepper and Miller 1995; Agarwal and Gort 1996, 2002; Agarwal 1997; Klepper and Simons 2005; Agarwal and Audretsch 2001, among others). Large waves of entries—either bringing innovative and more competitive products to the markets or just trying out their luck—lead to large waves of exits of those competitors whose abilities are at the fringe.

In this paper, we analyze the determinants of entry, exit and changes in market shares of incumbent firms, drawing from literature on entry, exit and market share mobility. As a departure from previous studies, first, we not only include entry and exit but changes in market shares of incumbent firms as separate equations. We argue that entry and exit behaviors across the industries are also contemporaneously affected by the changes in market shares of incumbent firms. It is also possible that they response to same set of unobserved factors. So far, the literature has not considered the mobility equation into the system of market selection. Exclusion of an equation from the system may cause misspecification bias in estimated coefficients if they are part of the same system. We also search for the effects of industry life cycle on entry, exit and mobility activities. As various studies suggest, life cycle of an industry brings differing regularities on entry, exit and mobility of firms.

2 Interdependence Among Entry, Exit and Mobility of Incumbents

Modeling the evolution of industries has always been attracted attention of researchers. Market structure of an industry evolves with continuous changes in identities (entry and exit) and adjustment of size classes of firms (mobility) (Caves 1998; Baldwin and Gorecki 1994). Unfolding the gross entry, exit and mobility and their determinants and *interdependence* reveal significant information on industries. The importance and *interdependence* of entry, exit and mobility on evolution of industries can be visualized with the metaphor explained by Marshall (1920). Marshall explains that industry evolves as similar to what can be seen in forests where the old trees are replaced by young ones; some old firms exit and young ones enter and grow. In a similar way, Audretsch (1995) put forward the "revolving door" argument which posits that entry and exit are the phenomenon to explain industry evolution; however the level of entry and exit activity in industries mostly involves young firms; canonical door revolves faster at the bottom. Both metaphors shows that changes in market shares of incumbents are also prominent part of industry evolution as "young trees" grows or some young entrants survive and even grow and reach to the upper part of the canonical door, as some incumbents at the upper part of the canonical door loose their grip and shrink and even exit the industry.

Since its emergence, expansions happening in earlier periods and eventual shrinkage in decline period shape the structure of an industry through the mechanisms of entry as well as expansion, shrinkage and divestiture (Duetsch 1975; Klepper and Graddy 1990). Entry may facilitate introduction of new technologies into industry, bringing additional capacity and increase competition through selection, learning, innovation and price (Jovanovic 1982; Grossman and Helpman 1991; Hopenhayn 1992; Ericson and Pakes 1995; Baldwin and Rafiquzzaman 1996). Firms make the decision to enter. Upon entry, they continuously face updating actions about the decision to expand, contract or exit given the available characteristics of its own, as well as industry and economy. On their influential study on evolution of new industries and determinants of market structure Klepper and Graddy (1990) have indicated that the rate of growth in output attributable to change in number of producers through entry and exit as well as the rate of capacity adjustment of incumbents. Incumbent firms adjust their sizes to new conditions and the least efficient firms will exit. Theories as well as empirical evidence suggest that these three mechanisms response to each other either *simultaneously* or with delay and significantly dependent on each other as flows of entry and exit maintain, the incumbents continuously response to this process while adjusting their market shares either by expansion or contraction. Entry, exit and mobility of firms overtime may cause industry become concentrated with a few firms or cause to become a competitive atomistic industry or fall in somewhere between.

The relationship and *interdependence* among entry, exit and mobility could been explained by Structure-Conduct-Performance (SCP) paradigm. At any point in time, market structure of an industry reflects the size distribution as well as number of firms within the industry. SCP states that Structure (number and size distribution of firms) of an industry determines the conduct (strategy such as investment, advertising, research and development activity, limit pricing) of firms. The performance (outcomes, profits) patterns result from the behavioral patterns. However, it has also been signified that this relationship from structure to conduct and from conduct to performance is not mere one-way relationship and also there are feedbacks toward structure of industry both from conduct and performance. The feedback mechanism causes adjustment in structure of industry. As pointed by Weiss (1965) the adjustment taking place in market structure between two time points naturally happens through entry, exit and changes in market shares of incumbents. Entry reduces market concentration, exit increases. Expansions made by existing firms raise concentration and contractions cause the opposite. On their study with over 200 leading U.K. firms, Davies and Geroski (1997) found that changes in market shares of surviving firms are dominant influence on concentration. Demand growth, technological change, entry, exit and other endogenous and exogenous disturbances all lead to changes in market structure (Curry and George 1983; Geroski et al. 1987). In particular, Rosenbaum (1993) modelled that changes in concentration *simultaneously* determined by net entry rates (Entry minus exit) as well as profits.

As the underlying mechanisms of market selection process, entry, exit and changes in market shares of incumbents could also be explained by the notion of turbulence. The turbulence rate is defined as the sum of entry rate and exit rate (Caves and Porter 1976; Beesley and Hamilton 1984; Audretsch and Acs 1990; Dunne and Roberts 1991; Fotopoulos and Spence 1998) and it contains information on identity of entrants and exiters as well as changes in market shares of incumbents (Baptista and Karaöz 2006). Thus turbulence rate reflects the entire extent of market selection intensity as an evolutionary process. As the broadest variable, the sum actually states the level of turbulence in an industry which reflects the level of selection intensity. On the other hand turbulence produces certain amount of net residual in the form of net addition or reduction to an industry. Then the net entry rate is defined as the difference between entry and exit rates. Net entry rate represents changes in number of firms within an industry, and it is accommodated by industry growth and changes in market shares of incumbent firms depending on the stage of product life cycle. For instance positive values of net entry rate with zero industry growth suggest that net entry is accommodated by incumbents' market share reductions. In growing industries both growth rates and changes in market shares of incumbents contribute on net entry and turbulence. In declining industries, however, shrinking industry size produce negative net entry, while some incumbents exit, other reduce their market share. The difference between the turbulence and the absolute value of net entry shows the level of volatility in entry and exit rates. It contains information on changes in identity of incumbents (Fotopoulos and Spence 1998; Dunne and Roberts 1991). Fotopoulos and Spence (1998) notes that the most volatile industries are those where large numbers of new firms *displace* large number of recent entrants without affecting the number of firms in long existence (Geroski 1995), or where the revolving door throws out, in its turning, the most recent entrants (Audretsch 1995).

Agarwal and Gort (1996) indicate that high correlation between entry and exit raise, first, because entry increase competition and competition cause exits of least efficient firms. The mechanism of *displacement effect* causes the exit of least efficient firms. Second, given the high infant mortality among entering firms higher entry rates trigger higher rates of exit. The mechanism of *replacement effect* creates pressure of exits among incumbents simply because overcrowding caused by large amount of entry. Pressure particularly cause exits among young, inexperienced, small-sized entrants. Other than that, another relationship also tested arguing that there could be some common both unobservable factors make effect on entry, exit and incumbent mobility. This is called the *symmetry hypothesis* and rooted back to work by Caves and Porter (1976), Eaton and Lipsey (1980, 1981) and Harrigan (1981). High levels of sunk cost created by cost of entry, advertising, research and development activities, fast reductions in capital expenditures firm assets due to asset durability and specificity and high capital commitment may create barriers to exit. Fewer exits create less vacuum and less pressure for *replacement effect*. Entry barriers also act to create impediments for exit. Reduced entry lowers the pressure of *displacement* to exit among incumbents. Hopenhayn (1992) suggests that when entry cost is large and not recoverable, entrants will be deterred. Then exit will be less likely since there is less pressure of competition from entrants. On the other hand, ex ante knowing the fact that entrant face risk of failure, since if it occurs entry cost and liquidation of the assets may create large losses. This suggests that exit barriers also impede entry.

The *interdependence* between entry and exit has been investigated in number of studies (Baldwin and Gorecki 1991; Shapiro and Khemani 1987; Austin and Rosenbaum 1991; Rosenbaum and Lamort 1992; Evans and Siegfried 1992; Fotopoulos and Spence 1998; Lay 2003).[1] These studies implicitly involve the changes in market shares of incumbent as the third vital mechanism next to entry and exit for evolution of industries. However, Fotopoulos and Spence (1998) was the first authors explicitly incorporated the *changes of market shares of incumbents* using the notions intra-industry dynamics, turbulence and trial and error process. They pointed "understanding of the factors underlying industry turbulence . . . really needs . . .. an understanding of the determinants of intra-industry mobility",

[1]See Siegfried and Evans (1994) and Fotopoulos and Spence (1998) for extensive reviews.

lending to that the actual *displacement* of incumbents involves mostly recent entrants and older and larger firms are displaced in slower phase and longer time period. Analysis of variance performed for the entry and exit into Greek manufacturing industry also including decompositions for intra-industry variations.

Additional evidence on *interdependence* of *market share mobility* with entry and exit has been observed in various empirical studies. Baldwin and Gorecki (1991) concluded that importance of entrants depends on the probability of entry, on the size of entrants, and on their growth rate after birth and all three have to be examined to appreciate fully the role that entry and exit play. McGuckin (1972) investigated the effects of net entry on concentration change and market share stability. He partitioned the changes in concentration of *r* largest firms in the industry into net entry (difference of entry and exit rates) and *changes in market shares* of *r* largest firms that comprises of changes in value of their total shipments and changes in average firm size in industry. Duetsch (1975) pointed that although measures of entry and exit represent a changing pool of potentially strong competitors they ignore the size of entering and exiting firms. The entry of a firm which achieves all known scale economies may have considerably greater competitive impact than entry by much smaller firms. He further argued that the intensity of competitive struggle found in an industry in which zero net entry takes place may differ significantly from that found in an industry with zero net entry and zero turnover. Hypothetically assuming zero entry and exit to an industry, any changes in the rank conditions of some incumbents or mobility, changes in concentration will still be zero and will fail to reflect these changes. On the other hand, if same number of firms enters the market, while the same number of firms exit with the same size distribution, changes in concentration will still be zero. The obstacle created with zero change in market structure regarding invisibly of identities of firms could be solved by having underlying entry and exit rates. In reality entry and exit as well as mobility of firms are part of the changes in market structure. Full understanding of the underlying changes in market concentration requires all three to be included in an analysis. Many studies analyzing cohorts of entrants and products shown that, over time, some of the entrants leave the industry while the others take the exiting firms shares and expand (Muller 1976; Dunne and Hughes 1994; Baldwin and Rafiquazzaman 1996; Baldwin and Gorecki 1991; Agarwal and Gort 1996, 2002). Traditional approach to market selection assumes that exit (entry) have significant role on entry (exit). However, mere inclusion of entry and exit rates implicitly put all entrants and exiters in equal size. However both the number of entrants and their sizes differ significantly. The inequality of sizes and numbers need the mechanisms of market share changes of incumbents assuming the unfilled (excess) aggregate industry demand always will be cleared either with entry (exits) or with incumbents' size adjustments.

All in all, empirical and theoretical literature on evolution of industry suggests decisions of firms to enter and to exit as well as making adjustments in the firm size not independent from each other. Therefore, in addition to entry and exit equations, we also employ "market share changes of incumbent firms" equation as part of the mechanism in order to investigate the determinants of entry, exit and mobility of incumbents, *symmetry* and the causal *interdependence hypotheses*.

3 The Hypotheses and Models

The apparent *interdependency* and high positive correlation has also been under scrutiny in some of the recent studies on entry and exit (Austin and Rosenbaum 1991; Dunne and Roberts 1991; Shapiro and Khemani 1987; Fotopoulos and Spence 1998; Lay 2003). Initial approach involves *displacement* and *replacement effects*. We test the *displacement* and *replacement effects* of entry and exit both considering the *simultaneous* and lagged relationship setting. Both the entry and lagged entry causing exit of incumbents are called '*displacement effect*' in the literature. New firm entry creates competitive pressure and cause overpopulation, forcing the exit of those incumbents at the fringe, regardless of their age. However, it is also plausible to assume that the entry of firm does not instantly cause exit decisions among at least in some of the fringe firms and it may happen with delay. Pressure for exit among some incumbents may take time to realize, and may require period for pre-exit assessment and liquidation or transfer of assets which delays actual exits. They may even hope that the negative effect could be temporary. The literature (Shapiro and Khemani 1987; Rosenbaum and Lamort 1992; Fotopoulos and Spence 1998; Lay 2003) also suggests that similar *simultaneous* and lagged relationship occur due to exits. Exits cause vacuum and attract entry of firms and it has been dubbed as '*replacement effect'*. The *simultaneity hypothesis* suggests that exit of incumbents act to create room in the market therefore attracting new firms to enter again instantaneously. However, it is also plausible to assume that the same relationship process may also take time, since entry decision and pre-entry preparations delay actual entry. In the alternative approach, studies also investigated if entry and exit rates act in a *symmetrical* way and barriers to entry (exit) become barriers to exit (entry). This hypothesis suggests that entry and exit response to unobserved common determinants. Testing the hypotheses, we adapted similar three step approach setting used by Lay (2003). However, we additionally introduce the third equation of the changes in market shares of living firms the system and tested if *symmetry* and *simultaneity hypothesis* also embraces this third market selection phenomenon.

Below statistical equations (1) state that entry (*EN*), exit (*EX*) and as well as incumbent mobility (*MSC*) rates are functions of vector of incentives (*X*), barriers (*Z*) of industry and as well as other variables (*W*) of economy. In equations, *i* represents the industry and *t* represents time.

$$EN_{it} = f(X, Z, W) + \varepsilon_{it} \tag{1.1}$$

$$EX_{it} = f(X, Z, W) + \gamma_{it} \tag{1.2}$$

$$MSC_{it} = f(X, Z, W) + \varnothing_{it} \tag{1.3}$$

In equations 1.1, 1.2 and 1.3, it is assumed that the error terms ε_{it}, γ_{it} and $\varnothing_{it}$ are independent. We test the validity of *replacement* (*displacement*) *effects* just in lagged setting by adding the lag of exit rate (lag of entry rate) into entry equations. We further assume that, among others, expansion and contraction decisions of incumbent firms response to entry and exit of firms within the industry, and the actual expansion and contraction occurs with lag. Thus the Model (2) become,

$$EN_{it} = f(X, Z, W) + EX_{i,t-1} + \varepsilon_{it} \tag{2.1}$$

$$EX_{it} = f(X, Z, W) + EN_{i,t-1} + \gamma_{it} \tag{2.2}$$

$$MSC_{it} = f(X, Z, W) + EN_{i,t-1} + EX_{i,t-1} + \varnothing_{it} \tag{2.3}$$

As a final approach, in addition to the lagged *interdependence*, *simultaneity* of *replacement* and *displacement effects* has been tested (Rosenbaum and Lamort 1992). We include entry rate in exit equation and exit rate in entry equation. We further assume that expansion and contraction of some incumbent firms also act same as entry and exit decisions of new firms and make similar effects on new firm entry and incumbent firm exit decisions. *Simultaneous interdependence hypothesis* has been extended by also adding the market share changes of incumbents variable to both entry and exit equations. Entry and exit variables are added to changes in market shares equation. Model (3) has become

$$EN_{it} = f(X, Z, W) + MSC_{it} + EX_{it} + EX_{i,t-1} + \varepsilon_{it} \tag{3.1}$$

$$EX_{it} = f(X, Z, W) + MSC_{it} + EN_{it} + EN_{i,t-1} + \gamma_{it} \tag{3.2}$$

$$MSC_{it} = f(X, Z, W) + EX_{it} + EN_{it} + EN_{i,t-1} + EX_{i,t-1} + \varnothing_{it} \tag{3.3}$$

4 Data and Variables

For the present analysis, a longitudinal data set on industry entry and exit was built from the Portuguese SISED ("Quadros de Pessoal") database. This database is built by the Portuguese Ministry of Employment and Social Security. This information covers virtually the whole universe of Portuguese firms (sole traders are excluded). The panel covers a time span of 8 years, from 1986 to 1993, which includes 4 years of macroeconomic expansion followed by 4 years of low, sluggish growth, thus allowing for the examination of the effect of the business cycle on turbulence. The period of analysis starts in 1986, the year of Portugal's entry into the European Community, which brought about significant market de-regulation, as well as greater financing opportunities for new and existing businesses.

4.1 Dependent Variables

Entry rates (*EN*) are calculated as the ratio between the numbers of entrants in industry i at time t and the stock of firms in industry i at time $t-1$. In the same manner, exit rates (*EX*) are calculated as the ratio between the number of exits in industry i at time t and the stock of firms in industry i at time $t-1$. Changes in market shares of incumbents[2] (Instability Index, *MSC*) is calculated as $\sum_{i=1}^{n}(|m_{it} - m_{i,t-1}|)$ where m_{it} and $m_{i,t-1}$ are market shares of firms relative to industry at time t and $t-1$ (Hymer and Pashigian 1962; Mazzucato 2002).

4.2 Industry Barriers, Incentives and Other Independent Variables

Contestable markets theory (Baumol et al. 1982) emphasizes that emergence of better profit opportunities may cause some firms to exit the current industry if the exit barriers are low. However, most industries incur sunk cost which has critical role to play in effectiveness and strength of industry barriers in terms of deterrence of entry and exit as well as mobility. Some expenditures for establishing, promoting and liquidating business are sunk (Kessides 1990). Caves and Porter (1976) and Harrigan (1981) indicate that technological or structural factors related with sunk cost not only hinder entry but also exit. Entry deterrent barriers also prolong the departure from the industry. Barriers reduce the chance that the market shares of the older producers will decline as the industry grows (Gort 1963). Large value of sunk cost intensifies the strength of barriers to entry, exit and investment for growth. Barriers could be specific to an industry which arises from a distinct government regulation such as patents or license agreements or arises from natural supply conditions. Structural industrial factors such as scale economies in R&D, learning, production, marketing and distribution would act as significant sources of barriers to entry, exit and mobility. Recovery of these expenditures require voluminous and successful amount of sales activity which usually require some amount of time. Overall, entry (exit) and expansion will act positively (negatively) to market opportunities when industry barriers are modest. Also some strategic barriers to entry may arise if few large firms in a highly concentrated market act together collusively to limit entry and penetration (Strickland and Weiss 1976; Duetsch 1975; Bunch and Smiley 1992). Various intangible assets such as company image, pool of accumulated customers, common distribution channels, vertical integration, plants that are used against rivals could act as exit barriers by making firms unwilling to abandon some businesses although they cause some loss (Harrigan

[2] See Cable (1997) for a summary of market share measures.

1981). Higher industry barriers tend to lead to higher levels of concentration and stability in market shares of incumbents.

Durability of capital involves sunk costs (Kessides 1990). Fast rates of depreciation of firm's capital increase sunkness and may naturally facilitate involuntarily firm exit unless updating its out of date assets. Degree of specificity and mobility of capital brings sunk cost if the equipment and material to be sold are highly specialized. It would be hard to find an interested buyer or renter which pays a reasonable price. However, durability, specificity and mobility cause deterrence of entry and reduce entry rates at the beginning so in the end exit rates will be lower.

Four-firm concentration ratio (CR4) is used in entry and exit equations. However log of Herfindahl index (LNHERF) are used in changes in market share equation capture the entire movement of market shares. Although high concentration in industry signals economies of scale, measures of concentration are not straightforward guarantee for the monopoly power (McGuckin 1972). Seller concentration will be determined partially by the nature of entry conditions (Encaoua and Jacquemin 1980). Various studies found that concentration adjust more quickly in rapidly growing (shrinking) markets (Pashigian 1969; Scherer 1979; Caves and Porter 1980; Levy 1985; Rosenbaum 1993). Economic theory sees that the size of an industry and number of firms are negatively related to market concentration. Any high levels of market concentration, particularly available in emerging and growing industries, will be wiped out by entry and turbulence if barriers to entry are low. For the same reasons Hymer and Pashigian (1962) found that more concentrated markets are the ones also with more instability. They found that industries that are more concentrated also tend to be the fastest growing industries. Industry characteristics such as profit and growth, as well as low Minimum Efficient Size will make the entry more attractive. As industry size grows over the life cycle, market allows for more optimal number of firms and market concentration decline.

Levy (1985) suggests that the increase in rate of change in concentration reduce entry. In fact, sharp increases in market concentration at any given short time period may indicate the existence of increasing levels of economies of scale and consequent rise in industry barriers which cause less entry, exit and mobility. Gort (1963) points that barriers to entry also tend to lead to concentration; a positive relation between concentration and stability should be present. Percentage change in Herfindahl index in previous year (LCHERF) is expected to capture this effect. Increasing levels of concentration in previous year bring more stable shares among incumbents and higher levels of industry barriers to entry and exit in current year.

The share of micro-firms in the sector should influence entry and exit positively. It signals both the size heterogeneity and small average firm size in industry. Share of micro firms is measured by share of number of firms in the industry with more than 5 (SML5). Large number of small firm population cause more entry so does exit without causing much adjustment in sizes of large firms. Moreover, Caves and Porter (1977) also indicate that coexistence of various size subgroups is possible in some industries. Although entry and survival with small size may be possible, higher entry size requirements may blockade entry. Average industry per firm Entrants' Capital Labor ratio (ENTKL) has been used in entry equation. On the other hand, existence of size heterogeneity means coexistence of large and small firms together may protect large firm mobility while low levels of Minimum Efficient Scale (MES) cause mobility to accelerate. However the gap between small sized firms and average size of incumbent in industry gives the steepness of barriers to mobility that an entrant need to reach toward upward subgroups. Analogous to cost differential between small and large firms variable by Bloch (1981), the mobility variable of size gap (SZGAP) is developed as the ratio of average incumbent firm size—excluding new entrants—(INCSZ) to average size of firms with lower than MES size (LWMES) and used in entry equation. We expect positive effect of entry on SZGAP variable. Given INCSZ gets larger with the extent of economies of scale while lower values of LMES correspond to low entry threshold and survival opportunity for lower size subgroup firms without competition from firms in upper subgroups. The larger values of SZGAP cause higher entry rates. As in Fotopoulos and Spence (1998) and Pashigian (1969), MES is defined as $\sum_{i=1}^{n}(A_i/n_i)(A_i/A)$ where A represents total employment of firms. A_i denotes total employment in the ith firm size class and n_i denotes the number of firms in the ith class. Hence (A_i/n_i) ratio represents average firm size in ith industry class. It is weighted by the share of employment accounted for by firms in that class. The summation of all classes produced the MES value of the industry.

On a study for analyzing the determinants of entry, Duetsch (1975, 1984) found that extent of multiplant operations significantly limit the entry. Multiplant operations industries signal existence of economies of scale such as specialization, geographic flexibility and product differentiation which help spreading risk, reaching more customers, reducing cost and raising capital (Duetsch 1984; Mayer and Chappell 1992). Duetsch (1984) reported that there are two types of situations exits which may lead to lower average costs than running a single plant. The first type occurs when firm operate and serve in a large geographical area and incur large transportation cost. The second type occurs if multiplant operations could allow lower average cost in segmented products line. In such an industry, economies often result from specialization of plants in particular product line. Entry into industries with multiplant operations requires large amount capital commitment and experience since diversification increases with size of firm and signal economies of scale (Berry 1971; Caves 1981; Schwalbach 1987). Caves and Porter (1978) found that diversification reduce market share instability in industry. We expect entry, exit and mobility is not easy in this type of industries. For each industry, we produce Average Industrial Diversification Index (IDIVX) and Average Regional Diversification Index (RDIVX) variables (Berry 1971; Gorecki 1980, Von der Fehr 1991) in order to capture extent of the effect of multiplant operations. Those two indices are the calculated averages of related diversity indices all firm's within the industry. Industrial Diversity Index for firm has been defined as in Berry (1971), $D = 1 - \sum_{i=1}^{n} p^2$, where p is the ratio of the firm's employment in the ith industry to the firms employment in all industries. Firm's main industry has been learned from the database, which is usually taken as the industry of the largest plant. This index takes value of zero if firm is active only in single industry and approaches to unity when firm produce equally in different industries. Same formula is also used for Average Regional Diversification Index as well where this time p is the ratio of the firm's employment in the ith region (NUTS III in our case) to the firms employment in all regions. These two variables inform in more detail if industries requiring more multi-industry connections and productions (economies of scope) and regional diversification advantages on entry, exit and mobility. An additional cost measure has been developed assuming that average unit cost in the industry indicate existence of economies scale. Caves and Porter (1978) point that cost structure affect the stability of market shares in industry. Average industry wage rate (LNWAGE) has been used. We expect higher average wage signal per unit cost and deters entry and expansion decisions. However it may induce exit and downsizing as disengaging employees in routine positions may be easier in comparison liquidating tangible company assets with high sunk cost particularly in industries with small sized industries. Considering that Portuguese economy is more service oriented and specialized in more on labor intensive sectors, this proposition seem plausible. Overall, the net effect of LNWAGE on mobility of incumbents depends on negative effect of expansion activity and positive effect of employee downsizing.

High R&D intensity in industries creates barriers to entry and exit and changes in market shares of incumbents. Similarly, higher levels of advertising intensity in industry also act as a barrier to mobility and suggesting less instability in market shares of incumbents. Both R&D and advertising are not used as explanatory variables in the present study due to lack of availability of data. But international statistics show that Portugal is not an R&D intensive country. Our database records also support these findings when seeing availability of only a few the number of scientific personnel and only in some industries. Even though it is recognized that both higher R&D efforts and advertising by incumbents do generate barriers to entry and exit, other measures of such barriers are already included in the study are believed to provide an appropriate control for this particular determinant of turbulence.

International evidence supports this view that market growth and profit cause increases in entry, exit and mobility (Caves and Porter 1980; Geroski and Schwalbach 1991; Davies and Geroski 1997; Baldwin 1998; Geroski 1995). Growth of the market demand cause allowance of entry due to pressure from excess demand where it may also allow growing price levels (Martin 1979). High growth signals availability of profit opportunities and industry attracts too many small sized entrants that are unrealistically overconfident with low visibility, and their intentions about prospect of business are unknown or uncertain (Vivarelli and Audretsch 1998; Vivarelli 2004). Magnitude of industry growth is a major incentive for entry and stay. Higher rates of growth also accommodate the survival of some inefficient firms due to growing and unfilled demand. This could also be more evident in industries with more inelastic demand structure. Entry (Exit) is positively (negatively) related with industry growth rates. Mobility increases with industry growth

(Allen 1981). Industry Growth rates (EMPGRO) are calculated using the yearly employment growth rates in industry. To understand the effect of persistent growth rates on market selection process we separate industries into growing, mature and declining industries as part of industry life cycle. Dummy variables for growing and (GRDM) and declining (DCDM) industries are incorporated into analysis. The study focuses on 320 six-digit sectors which covered about 95 % of the total industry population employment in 1993. To classify industries as growing, mature and declining, we adapted the taxonomy used by Thietart and Vivas (1984). Therefore industries with an 8-year (1986–1993) average growth rate over 4.5 % is considered as growing and while the industries between 0 % and 4.5 % average growth as mature. Industries with average growth rate less than 0 % are classified as declining industries. Among 320 industries, 148 of them are classified as growing, 83 as mature and the remaining 89 as declining industries. We expect more entry, fewer exits and higher mobility in growing industries. Declining industries attract less entry and experience more exits. Mobility is expected to be higher due to mergers, shrinkage and acquisitions.

Other than growth itself, fluctuations in growth may be substantial source of turbulence. Gort (1963) and Caves and Porter (1978) found that sharp movements in levels of demand reduce the stability of market shares. Gort (1963) argued that rapid growth generates instability in two ways. First given imperfect foresight, some firms adjust their scale of production to anticipated growth faster than others and this leads to shifts in market shares. Second, assuming there are recurrent lags in the adjustment of supply to rapid changes in demand, earning will recurrently rise above a normal rate of return. This will lead to entry of firms in the industry hence to changes in market shares. Caves and Porter (1978) reasoned that instability of demand puts the oligopolistic consensus in test because fluctuation in demand affect the profit maximizing conditions and create wider plant capacity error. They also added that higher growth attracts entry which itself bring market share instability to the market shares of incumbents and creates more exits. Studies of markets in a variety of disciplines, including industrial organization and organizational ecology, have shown that demand growth volatility is a significant source of uncertainty and, therefore, of risk (see, for instance, Hannan and Freeman 1989; Geroski 1991). Geroski (1991) indicates that price and output shocks, growth and technological change are expected to increase turnover (entry and exit). In industries where the pace of technological, socioeconomic and regulatory changes is faster, or price and output shocks are evident which are reflected in high growth rates also creates higher volatility, which affects entry, exit and mobility positively (Owen 1971; Caves and Porter 1977; Sutton 1997; Caves 1998). Following Baptista and Karaöz (2006), the present study generates two average industrial growth volatility indexes for each of the 4 year periods of 1986-89 and 1990-93 from regressions of industrial employment growth (EMPGRO) rates against time. In order to calculate each of the 4-year average volatility indexes, we divide the sum of absolute values of residuals by the number of years (in the present case, four) in the regression; in order to mitigate the relatively high level of correlation between these two variables we use the natural logarithm of volatility (LNVOL) in the regression analysis. The use of the natural logarithm of volatility substantially mitigated the problem of linear correlation between growth and its volatility, so both variables can be used together in estimation.

The *displacement* and *replacement effects* are tested considering they could happen both with and without lag. In addition to the entry (EN) and exit rates (EX) we use 1-year lagged values of entry rates (ENTLAG; $EN_{i,t-1}$) and exit rates (EXTLAG; $EX_{i,t-1}$) in order to capture *displacement* and *replacement effects*, respectfully as entry also response with lag. The *displacement* (*replacement*) *effect* brought about by recent entrants is expected to be negative (positive) for entry and positive (negative) for exit equations. The level of turbulence in relation to industry size is also investigated. Measures of market size acts as a proxy for technological conditions and capital intensity (Geroski and Pomproy 1990). Various studies of entry and exit suggest that larger industries are subject to more entry and exit activities (Orr 1974; Khemani and Shapiro 1986), hence the inclusion of the share of industry employment in total employment in the economy (ECOSHR) is used. Effects of macroeconomic environment on entry, exit and mobility has been measured by the macroeconomic unemployment rates (UNEMP). We expect more entry and fewer exits when unemployment in the economy is high. Effect of unemployment on market share movement will be negative considering more entry, less exits and firms cancels their expansion plans as the macroeconomic expectations on economy is low.

Table 1 summarizes the variables used in each equation. Additional summary statistics are given in Appendix Tables 4 and 5. In basic model, we use two additional variables in entry equation and different measure of concentration in mobility equation. We use lag of entry in exit, lag of exit in entry equation.

4.3 Estimation Results

We estimate three equations, of entry, exit and changes in market shares of incumbents in individual, system as well as in simultaneous settings using the models given by (1), (2) and (3). Prior to estimation, considering the variety of explanatory variables being used and the level of industry aggregation in the data, the presence of heteroskedasticity could be potentially a serious problem for the reliability of the estimates and for the hypotheses testing due to possible large standard errors, under limited number of observations. However, Least Squares (OLS) estimates become consistent with large samples. Our sample has 409.5 mean number of firms at six-digit level, with 1017.7 standard deviation and with the minimum of 1 up to 13014 maximum. Consequently we expect significant differences in variances of the variables across the sectors. Cook-Weisberg tests (Cook and Weisberg 1983) suggest that heteroskedasticity is significant at 1 % level in all three equations, with ($\chi^2 = 747.4$) in entry, ($\chi^2 = 16.8$) in exit and ($\chi^2 = 1107.6$) in mobility equation. It is also reasonable to assume that errors are correlated over time and across the sectors. We test for serial correlation using the Breusch-Pagan

Table 1 Expected signs of variables in the models

EN	Expected Coeff. Sign	EX	Expected Coeff. Sign	MSC	Expected Coeff. Sign
IDIVX	–	IDIVX	–	IDIVX	–
RDIVX	–	RDIVX	–	RDIVX	–
SZGAP	+				
ENTKL	–				
ECOSHR	+	ECOSHR	–	ECOSHR	+
LNWAGE	–	LNWAGE	–	LNWAGE	–
CR4	?	CR4	?		
				LNHERF	?
LCHERF	–	LCHERF	–	LCHERF	–
SML5	+	SML5	+	SML5	+
EMPGRO	+	EMPGRO	–	EMPGRO	+
LNVOL	+	LNVOL	+	LNVOL	+
GRDM	+	GRDM	+	GRDM	+
DCDM	–	DCDM	–	DCDM	+
UNEMP	+	UNEMP	–	UNEMP	–
EXTLAG	+			ENTLAG	+
		ENTLAG	+	EXTLAG	+
EX	+			EX	+
		EN	+	EN	+
MSC	–	MSC	+		

At time t: EN: Industry Entry Rate, EX: Industry Exit Rate, IDIVX: Industry Diversification Index, RDIVX: Regional Diversification Index of industry, SZGAP: Ratio of average incumbent firm size – excluding new entrants – to average size of firms with lower than MES size in industry, ENTKL: Capital Labor Ratio of entrants in industry, ECOSHR: Percentage share of industry in the economy, LNWAGE: Log of average industry wage, CR4: Industry Four-firm concentration ratio, LCHERF: Lag values of Changes in Concentration ratio in industry, LNHERF: Log of Herfindahl Index in Industry, SML5: Share of firms in the industry with 5 or less workers, EMPGRO: Industry employment growth, LNVOL: Volatility in employment growth of industry, GRDM: =1 if industry growing industries, DCDM: =1 if industry in declining industry, UNEMP: Unemployment rate in the economy, EXTLAG: Lag values of Exit rate, ENTLAG: Lag values of entry rate, MSC: Changes in Market Shares of Incumbents.

Lagrange multiplier (LM) test for first-order (AR1) serial correlation (Baltagi and Li 1995; Baltagi 2005). Test results confirm the existence of serial correlation at the 1 % significance level with ($\chi^2 = 47.1$) in entry, ($\chi^2 = 10.6$) in exit and ($\chi^2 = 123.8$) in mobility equation. We consider heteroskedasticity and autocorrelation problems during the estimations of all three models.

Using Models (1) and (2), first, we estimate each equation separately. The Feasible Generalized Least Squares (FGLS) estimator corrects for AR(1) serial correlation specific to each panel and is also appropriate to deal with heteroskedasticity when needed (Beck and Katz 1995; Parks 1967; Cameron and Trivedi 2005). Panel data model individual estimation results for Models (1) and (2) with FGLS estimator are presented in Table 2. Both estimations have produced

Table 2 FGLS Estimates with and without Lag values of Entry and Exits

	Panel FGLS—No Lags			Panel FGLS—With Lags		
	EN	EX	MSC	EN	EX	MSC
IDIVX	−30.45	−10.99	−18.75	−29.09	−10.24	−17.97
	−3.36***	−2.68***	−2.97***	−3.17***	−2.56***	−2.8***
RDIVX	30.21	−8.00	−5.51	30.24	−10.77	−5.50
	4.86***	−4.27***	−1.92	4.78***	−5.3***	−1.83*
SZGAP	−0.04			−0.04		
	−1.34			−1.33		
ENTKL	0.00			0.00		
	−9.38***			−9.61***		
ECOSHR	1.84	0.57	−0.18	1.82	0.42	−0.30
	6.93***	6.12***	−1.17	6.79***	4.32***	−1.8*
LNWAGE	−8.19	−1.27	0.39	−8.03	−0.74	0.75
	−18.12***	−5.45***	1.17	−17.28***	−3.12***	2.17**
CR4	0.04	0.00		0.04	0.00	
	4.5***	−0.46		4.2***	−0.23	
LNHERF			−0.82			−0.82
			−8.00***			−7.93***
LCHERF	0.00	0.00	0.00	0.00	0.00	0.00
	−1.55	0.19	0.84	−1.58	1.74*	1.01
SML5	0.09	0.06	0.04	0.09	0.05	0.03
	12.42***	14.98***	5.74***	11.66***	13.49***	4.73***
EMPGRO	0.05	−0.01	0.03	0.05	−0.01	0.03
	6.21***	−3.56***	5.93***	6.21***	−3.28***	5.86***
LNVOL	2.24	0.24	2.46	2.29	0.08	2.37
	12.58***	2.56***	15.92***	12.75***	0.84	14.92***
GRDM	3.83	0.95	1.60	3.71	0.47	1.50
	7.36***	4.4***	4.59***	7.01***	2.15***	4.25***
DCDM	−2.36	0.20	−0.50	−2.34	0.57	−0.40
	−5.9***	0.73	−1.16	−5.73***	2.26**	−0.92
UNEMP	−0.26	−0.08	−0.73	−0.25	−0.04	−0.66
	−2.06**	−1.13	−8.32***	−1.98**	−0.66	−7.38***
ENTLAG					0.11	0.03
					9.77***	2.58***
EXTLAG				0.04		0.05
				1.39		2.44**
CONS	90.43	19.76	13.18	88.36	13.43	8.52
	16.99***	7.15***	3.41***	16.03***	4.79***	2.1**
Wald χ^2	1667.24	687.52	989.04	1647.77	840.36	994.6
Log likelihood	−4826.8	−5094.1	−5638.7	−4819.4	−5080.2	−5635.0
Groups	290	319	319	290	319	319
Obs	1569	1908	1907	1569	1908	1907

z values in second row's of each variable for coefficient estimates. (***): Significant at 1 %, (**): Significant at 5 %, (*): Significant at % 10 %

similar results for variable estimations. However, *replacement effect* in Model (2) is insignificant although positive. *Displacement effect* is significant and positive in exit equation. Both effects create positive effect in mobility equation. Having advantage of similarity of results we continue evaluating both results together. Estimates from both Models suggest that diversified industries (IDIVX) attract less entry, produce less exits and incumbent mobility.

However, entry has positive significant sign for regional diversification (RDIVX) suggesting regional diversification increases entry.[3] Regional diversification deters exit and market share mobility. Industries with higher entry size (ENTSZ) also cause entry deterrence. Industries with larger share in economy (ECOSHR) have more entry and exit since market structure become more atomistic while mobility is more evident. LNWAGE indicates that larger unit cost deters industry entry and exit. However, high worker payment causes more market share mobility in Model (1) estimates but insignificant in Model (2).[4] Positive and significant coefficient of SML5 indicates that industries with more small sized firms have more entry, exit and market share changes in industry. Employment growth (EMPGRO) causes more entry and fewer exits. Industries experiencing higher levels of growth also experience higher market share instability. Fluctuations in growth (LNVOL) cause more entry, exit and market share instability. However, coefficient in exit equation in Model (2) is insignificant. Industries with higher market concentration (CR4) have more entry. However the relevant coefficient in exit is insignificant. Market share instability is negatively affected from market share concentration (LNHERF). Lag values of increase in market concentration (LCHERF) insignificant in both Models, except positive effect in exit in Model (1). Dummy variable for growing industries (GRDM) is positive and significant in all three equations. Industries in declining industries (DCDM) have less entry and more exits in Model (1) but not in Model (2). Mobility is insignificant in both Models. Sign of unemployment rate on entry is surprisingly negative. It seems that unemployment rate signals lower profit expectations and low survival rate. This may lead to lower entry. Mobility equation is negative and significant suggesting that firms are less willing to adjust their current size in austere economic environment.

---

[3]We suspected and tested if the positive sign in entry equation is due to industries having small sized establishments. We generated interaction term between the variables of share of small firms (10 or less) in industry and regional diversification. Interaction term had clear positive effect at 1 % on regional diversification in both Models and then sign of regional diversification variable turned into significant (10 %) and negative. It becomes clear that regionally diversified industries with large firm size deters entry. Only industries with more small sized firms positively affect entry.

[4]In mobility equation, we create interaction term between share of small firms in industry with ten or less workers and LNWAGE to see this effect is relevant only for small firms. Because larger firms has more restriction to hire and fire workers due unions and other formalities. Interaction term is positive and significant at 1 % level in both equation estimates. Then, LNWAGE equation turned into insignificant suggesting that larger firms are not responsible from positive sign of incumbent mobility.

Table 3 SUR, 2SLS and 3SLS estimates

	SUR			2SLS			3SLS		
	EN	EX	MSC	EN	EX	MSC	EN	EX	MSC
IDIVX	−29.24	−12.50	−25.28	−13.32	−7.42	−27.83	−33.62	10.39	−15.67
	−2.01**	−1.64*	−1.96**	−0.23	−0.33	−1.46	−0.94	0.64	−0.97
RDIVX	24.85	−8.54	−11.04	12.68	−0.60	−10.81	9.44	1.85	−3.64
	3.11***	−2.1**	−1.59	0.22	−0.03	−0.92	0.48	0.19	−0.37
SZGAP	0.00			0.00			0.00		
	7.85***			5.44***			3.11***		
ENTKL	0.00			0.00			0.00		
	−2.5***			0.29			−0.45		
ECOSHR	1.94	0.45	0.42	0.80	0.11	0.64	1.63	−0.32	0.38
	4.72***	2.07**	1.07	0.74	0.23	1.13	1.56	−0.64	0.71
LNWAGE	−7.86	−1.30	0.18	−2.76	−0.94	−0.32	−3.39	−0.45	0.26
	−9.1***	−2.9***	0.23	−0.77	−0.6	−0.2	−1.54	−0.35	0.2
CR4	0.06	0.02		−0.03	0.04		−0.01	0.02	
	5.07***	3.4***		−0.51	1.83*		−0.29	1.34	
LNHERF			−0.44			−0.27			−0.29
			−2.11**			−0.68			−1.06
LCHERF	−0.02	0.00	−0.02	−0.02	0.00	−0.02	−0.04	0.01	−0.02
	−3.3***	0.27	−2.95***	−0.88	0.28	−1.56	−2.38**	2.04**	−2.5***
SML5	0.10	0.06	0.04	−0.03	0.02	0.06	0.03	0.00	0.01
	7.64***	7.76***	2.74***	−0.45	0.63	1.17	0.56	0.23	0.23
EMPGRO	0.03	−0.02	0.01	0.10	−0.03	0.01	0.10	−0.03	0.03
	4.67***	−4***	2.03**	1.29	−0.96	0.19	4.83***	−3.2***	2.8***
LNVOL	2.04	0.45	3.74	4.22	−1.89	3.93	7.13	−2.77	3.28
	6.37***	2.66***	12.88***	1.64*	−1.41	6.32***	4.1***	−3.8***	7.81***
GRDM	3.42	0.33	1.85	3.72	−1.10	1.91	5.05	−1.46	1.74

	5.03***	0.93	3.02***	1.57	−0.86	2.07**	2.91***	−1.64*	2.1**
DCDM	−1.66	0.27	0.20	−2.08	0.02	0.52	−1.50	0.31	−0.19
	−2.03**	0.63	0.28	−0.97	0.02	0.51	−0.81	0.36	−0.21
UNEMP	0.19	−0.44	−1.22	0.42	0.44	−1.37	−0.84	0.61	−0.82
	0.64	−2.8***	−4.54***	0.39	0.91	−3***	−0.89	1.56	−2.3**
ENTLAG		0.11	−0.06		0.05	0.03		0.10	−0.08
		8.26***	−2.54*		0.86	0.24		2.03**	−1.53
EXTLAG	0.15		0.09	−0.92		0.27	−0.22		0.07
	3.41***		2.17**	−2.09**		0.79	−0.94		0.79
EN					0.05	−0.04		0.11	−0.12
					0.39	−0.47		0.86	−0.88
EX				3.66		−0.56	2.70		0.81
				2.56***		−0.55	3.71***		2.59***
MSC				−1.12	0.71		−1.93	0.88	
				−1.4	2.02**		−4.0***	4.84***	
CONS	82.94	20.42	9.24	26.35	6.12	21.04	45.94	−1.08	6.85
	7.97***	3.71***	1.54	0.63	0.35	1.07	1.64*	−0.07	0.47
χ^2	682.6	300.7	323.8	76.5***	3.3**	35.5***	198.3	116.9	537.1
R^2	0.30	0.16	0.17						
Obs	1585	1585	1585	1585	1585	1585	1585	1585	1585

t (SUR) and z values in second row's of each variable for coefficient estimates. (***): Significant at 1 %, (**): Significant at 5 %, (*): Significant at % 10%

Table 4 Variable descriptions and descriptive statistics

Variable	Description	# Obs	Mean	Std. Dev.	Min	Max
EN	Entry Rate	2560	10.77	13.16	0	200
EX	Exit Rate	2560	8.70	7.42	0	100
MSC	Changes in Market Shares (Instability Index)	2558	15.68	10.73	0	135.01
LCHERF	Change in Herfindahl index (%)	2237	0.85	42.64	−94.54	814.37
ENTLAG	Lag values of entry rate	2560	11.06	13.73	0	200
EXTLAG	Lag values of exit rate	2560	8.65	7.18	0	100
SZGAP	Ratio of average incumbent firm size—excluding new entrants—to average size of firms with lower than MES size in industry	2534	15.45	471.91	1	21133
ENTKL	Avg. Industry Capital Labor Ratio of entrants	2130	96888	2380384	2.17	1.05E+08
CDR	Cost disadvantage ratio	2534	0.20	0.13	0.00	0.88
MES	Minimum efficient scale	2560	471	1919	0	22510
LNWAGE	Log of per firm per worker average wage	2233	10.73	0.43	9.51	12.27
LNAVEMP	Log of average employment per firm	2559	3.13	1.39	0.41	10.01
LNAVNEST	Log of average per firm number of establishments	2559	0.30	0.62	−0.69	5.41
IDIVX	Average Per Firm Multi-Industry Diversification Index	0.017	0.029	0	0.349	0.017
RDIVX	Average Per Firm Regional Industrial Diversification Index	2559	0.028	0.066	0	0.857
SML5	Share of small firms with 5 or less employees in industry (%)	2559	36.24	22.83	0	100
ECOSHR	Share of industry in economy (%)	2559	0.29	0.61	0.00	6.61
HERF	Herfindahl index	2558	1470	2214	3	10000
CR4	Four-firm concentration ratio	2559	43.44	31.31	0.96	100
EMPGRO	Employment growth (%)	2558	6.38	44.23	−93.41	1079.7*
LNVOL	Log of 4 years averages of volatility	2552	2.02	1.06	−1.48	5.93

*30 cases out of 2558 observations display growth rates above 100

Table 5 Correlation coefficients of variables in estimation

	ENTRT	EXTRT	MSC	IDIVX	RDIVX	SZGAP	ENTKL	ECOSHR	LNWAGE	CR4	LNHERF	LCHERF	SML5	EMPGRO	LNVOL	GRDM	DCDM	UNEMP	ENTLAG
EXTRT	0.21	1.00																	
MSC	0.22	0.14	1.00																
IDIVX	0.00	−0.10	−0.07	1.00															
RDIVX	0.02	−0.13	−0.09	0.30	1.00														
SZGAP	0.19	−0.04	−0.04	0.21	0.27	1.00													
ENTKL	−0.03	−0.05	0.04	−0.02	0.01	0.00	1.00												
ECOSHR	0.03	0.01	−0.06	−0.08	0.05	0.03	−0.01	1.00											
LNWAGE	−0.29	−0.14	−0.04	−0.02	0.36	0.00	0.03	−0.08	1.00										
CR4	0.09	−0.01	0.06	0.23	0.37	0.06	0.06	−0.27	0.25	1.00									
LNHERF	0.07	−0.04	0.04	0.21	0.35	0.06	0.04	−0.38	0.25	0.93	1.00								
LCHERF	−0.05	0.01	−0.01	0.02	−0.01	0.00	0.02	−0.04	0.02	0.05	0.06	1.00							
SML5	0.23	0.26	0.13	−0.20	−0.24	0.01	0.01	−0.06	−0.16	−0.34	−0.42	0.04	1.00						
EMPGRO	0.20	−0.03	0.16	−0.01	0.01	−0.01	0.16	−0.04	−0.01	0.11	0.09	−0.06	0.07	1.00					
LNVOL	0.30	0.12	0.32	0.11	0.04	−0.02	0.07	−0.25	−0.19	0.42	0.44	0.12	−0.04	0.24	1.00				
GRDM	0.23	0.13	0.20	−0.09	−0.02	−0.01	−0.01	−0.15	0.05	0.11	0.13	0.05	0.21	0.16	0.30	1.00			
DCDM	−0.12	−0.08	−0.07	0.13	0.08	0.07	−0.02	−0.01	0.01	0.20	0.23	0.02	−0.32	−0.13	−0.01	−0.19	1.00		
UNEMP	0.20	0.00	−0.05	0.08	0.01	0.05	−0.02	0.00	−0.51	0.02	0.03	−0.02	−0.02	−0.04	0.23	0.00	0.00	1.00	
ENTLAG	0.43	0.31	0.11	−0.12	−0.09	−0.03	−0.03	0.04	−0.20	−0.02	−0.03	−0.10	0.27	0.07	0.20	0.22	−0.19	0.05	1.00
EXLAG	0.20	0.42	0.14	−0.13	−0.22	−0.04	−0.04	0.00	−0.18	−0.09	−0.10	0.07	0.28	0.04	0.12	0.14	−0.13	0.02	0.33

Moreover, it is possible that there are some unobserved common factors those affect processes of Entry, Exit and Market Share Mobility. This suggests the assumption that the error terms are independent is not valid. The Seemingly Unrelated regression (SUR) estimator is used considering the instantaneous relations of error terms in Model (2). We also test the lagged *replacement* and *displacement effects*. Results for SUR estimates are presented in first three columns of Table 3. Correlation coefficient between entry, exit and mobility of incumbents and Breush-Pagan test of independence of equations support that correlations among error terms are significant at 1 % level and in fact they response to common unobserved factors. Although SUR results have altered sign and significance of some of the variables, the effects of IDIVX, SML5 and EMPGRO on all three equations sign and significance have persisted. The effects of RDIVX, ECOSHR and LNWAGE turned into insignificant only in mobility equation.

Negative effect of LCHERF turned into negative in entry and mobility equations but no effect on exit equation. High levels of CR4 caused into more entry and exit while mobility is negatively affected by market structure. Industries that are more concentrated are more stable. May be the identity of entering and exiting firms usually same and small and it does not affect much the identity and size distribution of incumbents. Although entry is easy, the upward movement into higher subgroups is not as easy. The significance partially disappeared in the system for GRDM and DCDM. LNVOL variable turned into significant in all equations. The effect of UNEMP become insignificant in entry equation and it turned into significant in exit equation. Lagged replacement effect has become significant with SUR estimates that are positively related with entry and mobility. Displacement had positive effect in exit equation and negative significant effect in mobility equation. Overall, further improvement has been achieved when we allow for correlations among the equations.

In Model (3), we allow for *simultaneous* movement of entry, exit and mobility equations. Two Stage Least Square (2SLS) as well as Three Stage Least Squares (3SLS) methods implemented. 2SLS estimates each equation in the system separately. Structural equations are estimated using the estimates for right-hand-side endogenous variables of other two equations in the system. The right-hand-side endogenous variables are estimated from reduced form equations using all exogenous variables in the system as instruments. Our Entry equation is just identified while the Exit and Mobility equations are overidentified. 2SLS using Generalized Method of Moments (GMM) estimation considering autocorrelation and heteroskedasticity is implemented (Wooldridge 2002) for estimation of each equation. Moreover Three Stage Least Squares (3SLS) also implemented. As oppose to 2SLS, 3SLS use all equations in estimation of parameter estimation. Model (3) estimation results are presented in related columns of Table 3. In general, 2SLS and 3SLS estimation results weakly support the existence of *simultaneous interdependence* of entry, exit and mobility. It seems that there is usually lag among responses of entry, exit and mobility as suggested in Model (1) and (2). In 2SLS estimation, both exit and lag of exit rate have significant impact on entry. Exit rates have positive effect on Entry and Mobility equation in 3SLS equation while lag of exit does not. Mobility has significant positive effect on entry and exit in 3SLS equation but only in exit equation in 2SLS estimation. Neither entry and exit nor lag values do not produce significant effect on mobility equation according to both estimations, except significant effect of exits in 3SLS. Moreover, both 2SLS and 3SLS estimations show that exit and mobility are not *simultaneously* affected by entry. Overall, the *simultaneity hypothesis* is moderately confirmed, suggesting that partial immediate effect of entry, exit and mobility on each other. The results indicate that exits and mobility seems to have immediate impact on the other two equations.

5 Conclusions

We analyze the determinants of entry, exit and mobility of firms using single as well as three equations system approaches. We further test if the entry and exit activities response to each other also adding mobility as the third phenomenon. *Symmetry hypothesis* which states that entry, exit and market share mobility of incumbents response to unobserved factors has been tested. *Interdependence* of entry, exit and market share mobility is tested if they exist *simultaneously* and with delay. Additionally we search for the answer if the *interdependence* and regularities in these three elements of market selection process varies along the industry life cycle. To investigate the answers for the hypotheses, we use equations of entry, exit and changes in market shares or mobility of incumbents. To test our all hypotheses, we estimate entry, exit and incumbent mobility equations by separate as well as system estimators, which follow different assumptions about the *interdependence* of error terms among system equations. We use a longitudinal data set on 320 six-digit sectors from Portugal covering a time span of 8 years, from 1986 to 1993.

We have analyzed the determinants of entry, exit and mobility. Variables related with economies of scale, namely, Industrial and Regional diversification and unit cost deters entry, exit and mobility of incumbents. However, regional diversification deters entry only for larger firms. Unit cost does not have clear effect on incumbent mobility. Industries with more small sized firms and with more growth volatility experience more entry, exit and mobility. Unemployment in the economy cause less entry, exit and mobility. Effect of concentration on entry is positive, while it is negative on incumbent mobility. Its positive effect is weak on exit. Growth cause more entry and mobility of incumbent and less exits.

We have tested *symmetry* and *interdependence hypotheses*. *Interdependence* effect includes *simultaneity* or instant effect as well as lagged affect of variables to each other. The results support that entry, exit and incumbent mobility response to unobserved common determinants. The results on entry and exit are parallel to previous findings. The mobility also makes significant contribution to the system. Estimation results also suggest that there is *interdependence* among entry, exit and incumbent mobility. However the hypothesis that they response to each other *simultaneously* is partially supported. Some of the *interdependence* during the market selection occurs with delay. It seems that while there is no immediate impact of entry on other equations, exit rates and incumbent mobility partially create immediate impact on entry and incumbent mobility. According to the estimation results, growing industries experience significantly different in entry, exit and changes in market shares among incumbents. Declining industries experience less entry has found support. Weak support has been found for positive significance of difference of declining industries on exit. No significantly different effect on in mobility has been found. This introductory finding leads that further separate analyses are fruitful for different stages of the life cycle. This will allow having detailed answers if the *symmetry* and *interdependence hypotheses* change along the industry life cycle. However this is beyond the scope of this work and could be done in further studies. Taken as a whole, this study has been the first attempt on including the mobility of incumbents equations into system of entry and exit equations and accepting that market selection process covers all three together. Effects of incumbent mobility on market selection and its interaction with entry and exit could be explored in prospective studies.

References

Agarwal, R. (1997). Survival of firms over the product life cycle. *Southern Economic Journal, 63*(3), 571–584.

Agarwal, R., & Audretsch, D. B. (2001). Does entry size matter? The impact of life cycle and technology on firm survival. *The Journal of Industrial Economics, 49*(1), 21–43.

Agarwal, R., & Gort, M. (1996). The evolution of markets and entry, exit and survival of firms. *The Review of Economics and Statistics, 78*(3), 489–498.

Agarwal, R., & Gort, M. (2002). Firm and product life cycles and firm survival. *American Economic Review, 92*(2), 184–190.

Allen, B. T. (1981). Structure and stability in gasoline markets. *Journal of Economic Issues, 15*(1), 73–94.

Audretsch, D. B. (1995). *Innovation and industry evolution*. Cambridge: MIT Press.

Audretsch, D. B., & Acs, Z. J. (1990). The entrepreneurial regime, learning, and industry turbulence. *Small Business Economics, 2*(2), 119–128.

Austin, J. S., & Rosenbaum, D. I. (1991). The determinants of entry and exit rates into U.S. manufacturing industries. *Review of Industrial Organization, 5*(2), 211–221.

Baldwin, J. R. (1998). *The dynamics of industrial competition: A North American perspective*. Cambridge University Press.

Baldwin, J. R., & Gorecki, P. K. (1991). Firm entry and exit in the Canadian manufacturing sector. *Canadian Journal of Economics, 26*(2), 1970–1982.

Baldwin, J. R., & Gorecki, P. K. (1994). Concentration and mobility statistics in Canada's manufacturing sector. *Journal of Industrial Economics, 42*(1), 93–104.

Baldwin, J. R., & Rafiquzzaman, M. (1996). Selection and learning in the post-entry process. *The Canadian Journal of Economics, 29*(2), S455–S457.

Baltagi, B. H. (2005). *Econometric analysis of panel data*. 3rd: Wiley.

Baltagi, B. H., & Li, Q. (1995). Testing AR(1) and MA(1) disturbances in and error component model with autocorrelated disturbances. *Journal of Econometrics, 68*(1), 133–151.

Baptista, R., & Karaöz, M. (2006). Turbulence in high-growth and declining industries. Mimeo, Center for Innovation Technology and Policy Research, IN+, Lisbon, 2006.

Baumol, W. J., Panzar, J. C., & Willig, R. D. (1982). *Contestable markets and the theory of industry structure*. New York: Harcourt Brace.
Beck, N. L., & Katz, J. N. (1995). What to do (and not to do) with time-series cross-section data. *American Political Science Review, 89*(3), 634–647.
Beesley, M. E., & Hamilton, R. T. (1984). Small firms' seedbed role and the concept of turbulence. *Journal of Industrial Economics, 33*(2), 217–231.
Berry, C. H. (1971). Corporate growth and diversification. *Journal of Law and Economics, 14*(2), 371–383.
Bloch, H. (1981). Determinants of the variance in market shares in Canadian manufacturing. *Journal of Industrial Economics, 29*(4), 385–393.
Bunch, D., & Smiley, R. (1992). Who deters entry? Evidence on the use of strategic entry deterrents. *Review of Economics and Statistic, 74*(3), 509–521.
Cable, J. R. (1997). Market share behavior and mobility: An analysis and time-series application. *Review of Economics and Statistics, 79*(1), 136–141.
Cameron, A. C., & Trivedi, P. K. (2005). *Microeconometrics: Methods and applications*. New York: Cambridge University Press.
Caves, R. E. (1981). Diversification and seller concentration: Evidence from changes, 1963–72. *Review of Economics and Statistics, 63*(2), 289–293.
Caves, R. E. (1998). Industrial organization and new findings on the turnover and mobility of firms. *Journal of Economic Literature, 36*(4), 1947–1982.
Caves, R. E., & Porter, M. (1976). Barriers to exit. In J. S. Bain, R. T. Masson, & D. Qualls (Eds.), *Essays in industrial organization in honor of Joe Bain* (pp. 39–70). Cambridge MA: Ballinger.
Caves, R. E., & Porter, M. E. (1977). From entry barriers to mobility barriers: Conjectural decisions and contrived deterrence to new competition. *Quarterly Journal of Economics, 91* (2), 241–261.
Caves, R. E., & Porter, M. E. (1978). Market structure, oligopoly and the stability of market shares. *Journal of Industrial Economics, 26*(4), 289–313.
Caves, R. E., & Porter, M. E. (1980). The dynamics of changing in seller concentration. *Journal of Industrial Economics, 29*(1), 1–15.
Cook, R., & Weisberg, S. (1983). Diagnostics for heteroscedasticity in regression. *Biometrika, 70* (1), 159–178.
Curry, B., & George, K. D. (1983). Industrial concentration: A survey. *Journal of Industrial Economics, 31*(3), 203–255.
Davies, S. W., & Geroski, P. A. (1997). Changes in concentration, turbulence, and the dynamics of market shares. *Review of Economics and Statistics, 79*(3), 383–391.
Duetsch, L. L. (1975). Structure, performance and the net rate of entry into manufacturing industries. *Southern Economic Journal, 41*(3), 450–456.
Duetsch, L. L. (1984). Entry and the extent of multiplant operations. *Journal of Industrial Economics, 32*(4), 477–487.
Dunne, P., & Hughes, A. (1994). Age, size, growth and survival. *Journal of Industrial Economics, 42*(2), 115–140.
Dunne, T., & Roberts, M.J. (1991). Variation in producer turnover across U.S manufacturing industries in entry and market contestability. In *Entry and Market contestability, an International Comparison* (pp. 187–203). Blackwell.
Eaton, B. C., & Lipsey, R. G. (1980). Exit barriers are entry barriers: The durability of capital as a barrier to entry. *Bell Journal of Economics, 11*(2), 721–729.
Eaton, B. C., & Lipsey, R. G. (1981). Capital, commitment, and entry equilibrium. *Bell Journal of Economics, 12*(2), 593–604.
Encaoua, D., & Jacquemin, A. (1980). Degree of monopoly, indices of concentration and threat of entry. *International Economic Review, 21*(1), 87–105.
Ericson, R., & Pakes, A. (1995). Markov-Perfect industry dynamics: A framework for empirical work. *Review of Economic Studies, 62*(1), 53–82.

Evans, L. B., & Siegfried, J. J. (1992). *Entry and exit in United States manufacturing industries from 1977 to 1982*. In David B. Audretsch, & John J. Siegfried (Eds.), Empirical studies in industrial organization: Essays in honor of Leonard W. Weiss (pp. 253–273). Kluwer Publishers.

Fotopoulos, G., & Spence, N. (1998). Entry and exit from manufacturing industries: Symmetry, turbulence and simultaneity—Some empirical evidence from Greek manufacturing industries, 1982–1988. *Applied Economics, 30*(2), 245–262.

Geroski, P. A. (1991). *Market dynamics and entry*. Oxford: Blackwell.

Geroski, P. A. (1995). What do we know about entry? *International Journal of Industrial Organization, 13*(4), 421–440.

Geroski, P. A., & Pomroy, R. (1990). Innovation and the evolution of market structure. *Journal of Industrial Economics, 38*(3), 219–314.

Geroski, P. A., & Schwalbach, J. (Eds.). (1991). *Entry and market contestability: An international comparison*. Oxford: Blackwell.

Geroski, P. A., Masson, R., & Shaanan, J. (1987). Dynamic market models in industrial organization. *International Journal of Industrial Organization, 5*(1), 93–100.

Gorecki, P. K. (1980). The Determinants of foreign and domestic enterprise diversification in Canada: A Note. *Canadian Journal of Economics, 13*(2), 329–339.

Gort, M. (1963). Analysis of stability and change in market shares. *Journal of Political Economy, 71*(1), 51–63.

Gort, M., & Klepper, S. (1982). Time paths in the diffusion of product innovations. *Economic Journal, 92*(367), 630–653.

Grossman, G. M., & Helpman, E. (1991). *Innovation and growth in the global economy*. Cambridge: MIT Press.

Hannan, J. T., & Freeman, J. (1989). *Organizational ecology*. Cambridge: Harvard University Press.

Harrigan, K. R. (1981). Deterrents to divestiture. *Academy of Management Journal, 24*(2), 306–323.

Hopenhayn, H. (1992). Entry, exit, and firm dynamics in long run equilibrium. *Econometrica, 60* (5), 1127–1150.

Hymer, S., & Pashigian, P. (1962). Turnover of firms as a measure of market behavior. *Review of Economics and Statistics, 44*(1), 82–87.

Jovanovic, B. (1982). Selection and the evolution of industry. *Econometrica, 50*(3), 649–670.

Kessides, I. N. (1990). Market concentration, contestability, and sunk costs. *Review of Economics and Statistics, 72*(4), 614–622.

Khemani, R. S., & Shapiro, D. M. (1986). The determinants of new plant entry in Canada. *Applied Economics, 18*(11), 1243–1257.

Klepper, S., & Graddy, E. (1990). The Evolution of the new industries and the determinants of market structure. *Rand Journal of Economics, 21*(1), 27–44.

Klepper, S., & Miller, J. H. (1995). Entry, exit and shakeouts in the United States in new manufacturing products. *International Journal of Industrial Organization, 13*(4), 567–591.

Klepper, S., & Simons, K. L. (2005). Industry shakeouts and technological change. *International Journal of Industrial Organization, 23*(1-2), 23–43.

Lay, T. J. (2003). The determinants of and interaction between entry and exit in Taiwan's manufacturing. *Small Business Economics, 20*(4), 319–334.

Levy, D. (1985). Specifying the dynamics of industry concentration. *Journal of Industrial Economics, 34*(1), 55–68.

Marshall, A. (1920). *Principles of economics*. New York: Macmillan.

Martin, S. (1979). Advertising, concentration, and profitability: The simultaneity problem. *Bell Journal of Economics, 10*(2), 639–647.

Mayer, W. J., & Chappell, W. F. (1992). Determinants of entry and exit: An application of the compounded bivariate Poisson distribution to U.S. industries: 1972–1977. *Southern Economic Journal, 58*(3), 770–778.

Mazzucato, M. (2002). The PC industry: New economy or early life-cycle? *Review of Economic Dynamics, 5*(2), 318–345.
McGuckin, R. (1972). Entry, concentration change and the stability of market shares. *Southern Economic Journal, 38*(3), 363–370.
Muller, J. (1976). The impact of mergers on concentration: A study of eleven West German industries. *Journal of Industrial Economics, 25*(2), 113–132.
Orr, D. (1974). The determinants of entry: A study of the Canadian manufacturing industries. *Review of Economics and Statistics, 56*(1), 58–66.
Owen, N. (1971). Competition and structural change in unconcentrated industries. *Journal of Industrial Economics, 19*(2), 133–147.
Parks, R. (1967). Efficient estimation of a system of regression equations when disturbances are both serially and contemporaneously correlated. *Journal of American Statistical Association, 62*(318), 500–509.
Pashigian, P. (1969). The effect of market size on concentration. *International Economic Review, 10*(3), 291–314.
Rosenbaum, D. I. (1993). Profit entry and changes in concentration. *International Journal of Industrial Organization, 11*(2), 185–203.
Rosenbaum, D., & Lamort, F. (1992). Entry, barriers, exit and sunk costs: An analysis. *Applied Economics, 24*(3), 297–304.
Scherer, F. M. (1979). The Causes and consequences of rising industrial concentration. *Journal of Law and Economics, 22*(1), 191–208.
Schwalbach, J. (1987). Entry by diversified firms into German industries. *International Journal Industrial Organization, 5*(1), 43–49.
Shapiro, D., & Khemani, R. S. (1987). The determinants of entry and exit reconsidered. *International Journal of Industrial Organization, 5*(1), 15–26.
Siegfried, J. J., & Evans, L. B. (1994). Empirical studies of entry and exit: A survey of the evidence. *Review of Industrial Organization, 9*(2), 121–155.
Strickland, A. D., & Weiss, L. W. (1976). Advertising concentration cost margins. *Journal of Political Economy, 84*(5), 1109–1121.
Sutton, J. (1997). Gibrat's legacy. *Journal of Economic Literature, 35*(1), 40–59.
Thietart, R. A., & Vivas, R. (1984). An empirical investigation of success strategies for businesses along the product life cycle. *Management Science, 30*(12), 1405–1423.
Vivarelli, M. (2004). Are all the potential entrepreneurs so good? *Small Business Economics, 23*(1), 41–49.
Vivarelli, M., & Audretsch, D. (1998). The link between the entry decision and post-entry performance: Evidence from Italy. *Industrial and Corporate Change, 7*(3), 485–500.
Von der Fehr, N. H. M. (1991). Domestic entry in Norwegian manufacturing industries. In P. Geroski & J. Schwalbach (Eds.), *Entry and market contestability: An international comparison*. Oxford: Basil Blackwell.
Wooldridge, J. M. (2002). *Econometric analysis of cross section and panel Data*. Cambridge: The MIT Press.

Coopetition and Co-innovation: Do Manufacturing and Service Providers Behave Differently?

Dina Pereira, João Leitão, and Tessaleno Devezas

Abstract Previous studies on coopetition considered the concept as a mix of cooperation and competition among firms, oriented towards producing innovation and generating net value added or economic benefit. The importance of studying the determinants of firms' innovative behaviour in order to produce an insightful analysis of their patent intensity behaviour based on those coopetition relationships has also warranted increasing attention by several entrepreneurship scholars. This paper tackles the issue in an innovative way, by making use of firms' behaviour in generating co-innovative products and services to reveal their innovative performance and the dynamics of coopetition targeted at open innovation. Thus, we analyse the determinant factors of firms; capacity to generate co-innovations, which is influenced by the role played by policies oriented to driving innovations among firms, cooperation with scientific stakeholders and development of the capacity to generate and transfer new products. For this purpose, we use a dataset of 3682 manufacturing firms and 1221 service firms that participated in the European Community Innovation Survey (CIS), 2008. A probit analysis is conducted separately for manufacturing and service firms and, within each sector, according to firms' category of technological intensity. The results reveal the significant influence of manufacturing and service firms' capacity to generate co-innovative products and services, such as coopetition arrangements between competing firms and other R&D stakeholders, and also firms' capacity to introduce innovations to the market. Furthermore, this study also reveals that for service firms the effects of introducing process innovations inside the firm and the existence of internal R&D activities are of major significance for creating a capacity to generate co-innovations.

Keywords Co-innovation • Coopetition • External R&D • Inside R&D

D. Pereira (✉) • J. Leitão • T. Devezas
Dept. of Electromechanics, University of Beira Interior, Covilhã, Portugal
e-mail: dina@ubi.pt

T. Devezas et al. (eds.), *Industry 4.0*, Studies on Entrepreneurship, Structural Change and Industrial Dynamics, DOI 10.1007/978-3-319-49604-7_20

1 Introduction

Innovation results from an interactive process between firms and the environment, adjusted by the absorptive capacity of the economic system and the stimulating forces of the institutions that foster and promote innovation. In this vein, governments can stimulate innovation in two ways, that is, in a technology-push orientation (in order to decrease the private costs of the innovation process) and in a demand-pull orientation (to increase the private pay-off from successful innovation through adoption of measures targeted at improving Intellectual Property-IP protection) (Nemet 2009).

As a means of fostering innovation, firms and other institutions make use of so-called coopetition, this being a compound of strategic cooperation and competition among rivals (Rusko 2011). When dealing with emerging technologies, characterised by uncertainty regarding market opportunities, firms opt for strategic coopetition (Garraffo 2002). In this sense, both incremental and radical innovations can be developed through coopetition alliances.

The strategic use of IP protection mechanisms (such as patents) has become an important tool for establishing successful coopetition arrangements between private and/or public competitors, due to the risks posed by the flow of knowledge, namely the appropriation risks. Another advantage of using IP is associated with the possibility of generating co-innovation through strategic coopetition, by using patents as an information source regarding technological position and strength, in order to define a successful direction and market orientation for co-innovative products or services.

This paper aims to measure the impact of the innovative intensity of partners involved in coopetition arrangements regarding the type of strategic coopetition; and to determine the impact of a set of factors on manufacturers and service firms' capacity to generate co-innovative products/services influenced by policies targeted at driving co-innovative behaviour among firms, scientific stakeholders and competitors.

It contributes to the empirical literature on coopetition strategy by adopting a different perspective from prior work and complementing earlier studies, deepening the level of understanding about the process of generating co-innovation based on coopetition relationships, especially by contrasting the co-innovative behaviour of manufacturing and service firms.

Several authors analysed the strategic use of coopetition by firms dealing with emerging technologies (Brandenburger and Nalebuff 1996; Gomes-Casseres 1996; Harbison and Pekar 1998). Others focused on the benefits of coopetition (Bagshaw and Bagshaw 2001; Garraffo 2002; Chien and Peng 2005; Rusko 2011).

Previous studies were also devoted to the reasons for cooperating, and proposed four types of coopetition (Garraffo 2002). Other authors studied coopetition in its different nuances, such as the dyadic features of coopetition (Bengtsson and Kock 2003), these being defined as bilateral relationships characterised by the commitment of two firms when they cooperate in upstream activities, such as research and development (R&D), buying and processing of raw materials and on multifaceted coopetition, competing also in downstream activities, namely distribution, services, product development and marketing (Luo 2004), these being defined as multilateral relationships characterized by the commitment of more than two competing firms to cooperate with each other due to public policy. The author introduces the role of the policy maker as a possible initiator of coopetition relationships between firms.

The risks of opportunistic behaviour emerging from coopetition were object of analysis (Nieto and Santamaria 2007), as well as the importance of coopetition, especially when it comes to developing incremental innovations in high-tech industries (Abernathy and Clark 1985; Fjelstad et al. 2004; Ritala and Hurmelinna-Laukkanen 2009). Some scholars concluded on the need for firms to develop absorptive capacity in order to obtain critical outcomes from coopetition (Escribano et al. 2009; Bergek and Bruzelius 2010; Cohen et al. 2000). Some authors crossed collaborative partnerships with international cooperation, based on patent data at the inventor level (Bergek and Bruzelius 2010). The technological patterns of collaborative development were crossed with international trends, also based on patent information (Archambault 2002). The risks of appropriability regarding IP and knowledge ownership in coopetition alliances were studied by a set of scholars (Seung and Russo 1996; Rammer 2002; Blomqvist et al. 2005; Dagnino and Rocco 2009; Escribano et al. 2009).

The present article intends to analyse the determinant factors of firms' capacity to generate co-innovative products/services and to produce an insightful analysis of firms' innovative behaviour, by making use of the data available in the European CIS Survey, 2008.

The remainder of this article is structured as follows. Section 2 develops the theoretical underpinnings, drawing on the literature on innovation, coopetition and patents. Section 3 presents the empirical approach. Section 4 refers to the analysis, main results and discussion. Finally, the article concludes and presents limitations, implications for policy makers and guidelines for practitioners engaged in strategic and cooperation relationships oriented to creating innovation.

2 Conceptual Framework

2.1 Innovation: From Concept to Sources

Innovation capacity is originated through an interactive process between firms and the external environment. This capacity is influenced by the dynamics of the economic system for learning and also for stimulating the strengths of institutions that support innovation (Lundvall 1985, 1988, 1992, 2007; Nelson 1993; Cooke et al. 1997; Braczyk and Cooke 1998; Cooke et al. 2000; Kaufmann and Tödtling 2001; Silva and Leitão 2009a, b).

Also pointed out by Luo (2004) is the importance of the initiator role of policy makers in fostering coopetition relations among firms in order to produce innovation. Accordingly, other authors argue for the need to stimulate public policies in modern societies that are able to promote competence building, based on the rise of the concept of “institutional specialization”, fostering the role of private and public incentives and policies to support science and technology (S&T) and innovation (Conceição and Heitor 2007).

Laranja (2008) studied the concept of Technology Infrastructure as the set of different kinds of public, semi-public and private centres and research institutes, acting as a basis for the development of technology policies and support structures for technology transfer and innovation. The emergence of a set of public policies aimed at promoting entrepreneurial activities and knowledge-based start-ups, in order to spur knowledge-based activities in general and innovation, was analysed as being the key to economic growth and employment (Leitão and Baptista 2009). The set of public policies must be reconsidered and oriented towards an early stage in the entrepreneurial and innovation process, the phase of generating business ideas, before the business is founded.

This dynamic must be a joint effort, supported by an effective mix of public support mechanisms and private incentives in order to promote knowledge networks and flows of skilled people in an uncertain environment (Heitor and Bravo 2010).

Following this line of thought, Flanagan et al. (2011) point out the emergence, take-up and use of the concept of ‘policy mix’ by innovation policy makers, policy analysts and scholars, referring to the interactions and interdependencies between different policies in the sense that they affect policy outcomes in terms of the future scope and focus of innovation.

Another mechanism enabling innovation is patent protection, which in a static model is able to foster innovation but in a sequential model tends to inhibit complementary innovation (Bessen and Maskin 2009). In this sense, innovation can be sequential when inventions are created successively based on previous inventions, or complementary, each innovator following a distinct research path.

In order to classify and label innovative firms, organizational taxonomies help to understand the diversity of innovative patterns in firms and sectors (Pavitt 1984; Archibugi 2001).

Schumpeter (1934, 1942) proposed two alternative patterns of innovation, namely the entrepreneurial and the routinized. The first was mainly concerned with the entrepreneurial activity and creativity of small and new firms. The second embraces the generation of innovation in the formal R&D activity of large and established firms.

Utterback and Abernathy (1975) presented some categorization around the generation of new technologies over the stages of the product lifecycle, as they understand that by evolving firms change the basis of their competition from product innovation to process innovation (this is also supported by Klepper 1997). In the birth stage, firms invest in product differentiation in order to compete with others. As the market matures, firms shift the focus to greater investment in manufacturing and innovative processes. The authors identified and then separated process and product innovations and related the industrial innovation pattern according to three different stages of the innovation process: the uncoordinated (where competition is based on product performance), the segmental (where the rate of product innovation decreases and radical changes are required in the production process) and the systemic (where product and process innovations diminish, being highly interdependent).

According to Nelson and Winter (1977) and Dosi (1988), the taxonomies of innovation are based on the concept of technological regime, the firm's behaviour being influenced and determined by the nature of the technologies they use.

Pavitt (1984) proposed a taxonomy of the structural characteristics and organization of innovative firms that can help to differentiate the pattern of firm innovation across sectors. Thus, firms are categorized as: science-based; specialized suppliers; supplier-dominated; and scale-intensive firms. This taxonomy is useful as a predictive mechanism regarding the determinants of firm performance, such as international competitiveness and innovative performance (Jong and Marsili 2006).

Abernathy and Clark (1985) classified innovation in four categories, namely: incremental; component; architectural; and revolutionary.

Other taxonomies of innovation also based on cognitive mechanisms were studied by Tushman and Anderson (1986) who distinguished between competence-enhancing and competence-destroying innovation.

Dosi (1988) proposed four dimensions regarding technological regime which define boundaries for what firms can achieve in the process of innovation: the level and sources of technological opportunity; the conditions for appropriating economic profits from innovation; the creation of new solutions building on prior ones; and the nature of the knowledge basis relevant for innovation.

Pavitt (1998) focused on the effects of major improvements in technology on the competencies of established firms.

McGahan (2004) also identified four phases regarding change which can make industries' activities obsolete, namely: radical; progressive; creative; and intermediating.

Furthermore, the OECD also classified industry based on the intensity of its technology production, distinguishing between high-tech and low-tech industries, measured through indicators, such as R&D intensity and technology use across sectors (Hatzichronoglou 1997). Additionally, this classification came to include non-technological dimensions as factors of production, such as intangible investments and human capital (Peneder 2002).

According to Ritala and Hurmelinna-Laukkanen (2009), incremental and radical innovations can be created through specific forms of cooperation with competitors (i.e. coopetition), especially in high-tech industries.

2.2 Demand Pull vs. Technology Push of Innovation

Nemet (2009) states that governments must address several policies to stimulate innovation. This can be accomplished by using demand pull policies, which can stimulate investment and subsequent technological improvements.

The 1960s and 1970s witnessed a debate on whether the direction and rate of technological change had been strongly influenced by changes in market demand or by developments in S&T (Nemet 2009). According to the same author, the S&T push orientation defends that advances in science determine the rate and direction of innovation, presupposing the transfer of fundamental science to applied research and product development and thereafter to commercialization. The relation between S&T and the innovation process has a long-term basis that increases complexity and uncertainty. Another constraint lies in the fact that technology-push orientation minimizes the effect of prices and changes in the economy on the outcomes of innovation. Different strands in the literature related to the technology-push paradigm have a less determining orientation.

Although these approaches consider that advances in science determine the rate and direction of innovation, they defend the importance of the interrelatedness of the technological system in promoting and guiding innovation (Frankel 1955). Others defend the existence of a clear relationship between the available exploitable technological opportunities and the rate and direction of innovation (Rosenberg 1974; Nelson and Winter 1977; Klevorick 1995). In a complementary way, it can also be stated that firms must invest in R&D to develop their capacity to absorb knowledge and exploit opportunities (Mowery 1983; Rosenberg and Birdzell 1990; Cohen and Levinthal 1990). Alternatively, firms can use the knowledge flows between sectors to overcome limitations detected in the technological system (Rosenberg 1976, 1994) or even adopt a sequential behaviour characterised by science and technology-push (Rothwell 2002).

Within the demand-pull approach, the rate and direction of innovation are driven by demand. Changes in market conditions, such as production costs, the geographical scope of demand, latent demand, or potential new markets are the main drivers of opportunities to invest in innovation (Hicks 1932; Griliches 1957; Vernon 1966; Schmookler 1966, 1979; Rosenberg 1976; Rothwell 2002). In some situations, the demand-pull approach cannot answer hidden needs in demand (Simon 1973) helping to explain incremental innovations rather than radical ones which are responsible for the most important innovations (Mowery and Rosenberg 1979; Cohen et al. 2000); and in some situations it is not able to answer hidden needs in demand (Simon 1973). Moreover, this approach is too broad to be useful (Mowery and Rosenberg 1979; Scherer 1982; Kleinknecht 1990; Chidamber and Kon 1994).

The technology-push orientation fails to recognize market conditions and in contrast, the demand-pull paradigm does not take into account technological capabilities. Although they interact simultaneously, both are needed to explain innovation (Arthur 2007).

Nemet (2009) argues that in a technology-push approach, policy-makers can foster innovation by implementing measures to decrease the private costs of the innovation process. The same author, following a demand-pull orientation, defends that policy-makers can also increase the private payoff of successful innovation. In this connection, and of particular interest in the present study, are measures to increase the protection of IP.

2.3 From Coopetition to Co-Innovation

Rusko (2011) defines coopetition as a compound of collaboration and competition among firms. According to Luo et al. (2007), this concept was introduced in the 1980s by Raymond Noorda and became the subject of several studies during the 1990s, namely the issue of dyadic coopetition (Bengtsson and Kock 2000, 2003) or multifaceted coopetition (Amburgey and Rao 1996; Tsai 2002; Luo and Slotegraaf 2006).

Brandenburger and Nalebuff (1996) consider coopetition as an alternative way to perform in business, as distinct from competition, strategically used by firms that deal with emerging technologies in innovation networks (e.g. biotechnology, information and communication technologies, electronics and semiconductor industry). Since the area of emerging technologies has a high level of uncertainty regarding market opportunities and technology developments, firms in these industries tend to manage uncertainty by establishing strategic cooperation arrangements with competitors, to share common resources and thus reduce risk (Garraffo 2002).

In the view of Bagshaw and Bagshaw (2001) coopetition allows better performance for the firms involved than competitive arrangements, as by strategically managing cooperation and competition, the relationship can evolve through controlled behavior by partners and rivals.

Chien and Peng (2005) state that inter-organizational relationships evolve into a social structure of coopetition, becoming a tool for cooperation and also for competition, acting at multiple levels, such as firms, strategic business units, departments and task groups. It can also be used for developing a corporate market strategy, reducing costs, improving firms' competitiveness and acquiring a leading market position.

Rusko (2011) defends that one of the main motivations for competitors to engage in strategic cooperation arrangements is based on the creation of greater value or benefit, in order to improve economic performance. Walley (2007) states that coopetition provides additional benefits not only to competitors but also to customers.

In the view of Garrafo (2002), the decision to cooperate with competitors usually has the following motivations: (a) to access and/or exchange new technologies and complementary knowledge; (b) to enter into new markets; and (c) to influence and/or even control technological standards. Understanding the motivations for coopetition among competitors is crucial for better evaluation of their commitment to technological developments and market creation. Thus, we can state that the option to pursue coopetition projects focused on technological development or on fostering collaborative efforts in market development depends on the partners' purpose in the arrangement.

In this connection, Jong and Marsili (2006) proposed a typology of coopetition arrangements, namely: (i) exchanges of patents and knowledge, characterized by low commitment to cooperative technology developments and low collaborative efforts in market generation; (ii) collaborative R&D activities with high commitment to cooperative technology developments and limited efforts to improve the market on a joint basis; (iii) strategic alliances for setting new standards, determined by high commitment in collaborative efforts focused on market generation and low commitment to cooperative technology development; and (iv) collaborative agreements to integrate established firms, characterized by high commitment to cooperative technology development and great collaborative efforts to access the market. These types of coopetition arrangements determine the firm´s ability to compete in the marketplace and to implement the portfolio of a firm's coopetition activities that evolves over time. When dealing with firms that work on radical innovations, definition of new standards or new converging technologies, coopetition is carried out for sizing market opportunities related to radical innovations, setting new standards, and/or integrating established firms through converging technologies.

Padula and Dagnino (2007) synthesised the two dominant paradigms, on one hand, the competitive paradigm, which underestimates the power of the positive interdependences of cooperation, and on the other hand, the cooperative paradigm which underestimates the benefits of the negative interdependences of cooperation. It embraces the sharing of mutual interests while increasing the positive sum game, in a win-win strategic scenario.

Bengtsson and Kock (2003) define coopetition as a dyadic relationship, since competition is related to output activities such as distribution, services, product development and marketing. In turn, cooperation deals with input activities, like R&D, buying, logistics and processing raw materials. In between the two, there are midstream activities, like production.

Luo (2004) introduced four strategic domains, which are also dyadic, even if they involve other agents, such as the government or the public sector. These domains are: coopetition with global rivals; coopetition with foreign governments; coopetition with strategic partners; and coopetition within a multinational company. Coopetition with the government remains coopetition, since two or more competing firms can strategically collaborate in the context of government procurement or in response to different actions taken by the government to promote this type of strategic relationship.

Different approaches to coopetition using the resource-based view of the firm and the game theory demonstrate that coopetitive arrangements can generate more innovative activities than simple collaborations between non-competitors. For Brandenburger and Nalebuff (1996), Dussauge et al. (2000) and Tether (2002), competitors engage in a collaborative scheme, through the exchange of resources that generate value for all participants. Several authors point out that the main benefit derived from collaboration between competitors is the creation of completely new products (Tether 2002; Quintana-Garcia and Benavides-Velasco 2004).

Additionally, Belderbos et al. (2004) defend that R&D cooperation between competitors generates incremental efficiency gains. On the contrary, Nieto and Santa-Maria (2007) argue that coopetition does not favour innovation, since it can promote opportunistic behaviour and minimize trust among rivals.

Ritala and Hurmelinna-Laukkanen (2009) state that coopetition helps to develop incremental innovation in current products and services, being an effective mode of generating new innovations especially in high-tech industries.

Establishing strategic partnerships between different firms in innovation projects to share risks, costs and expertise has also become an important pattern in innovation management, of interest to both scholars and practitioners (Chesbrough 2003; Huston and Sakkab 2006). This pattern results in coopetition, funded on strategic cooperation with competitors in innovation initiatives. Achieving higher absorptive capacity and forming collaboration schemes with competitive partners increases the pace of engaging in coopetition and imitation especially when dealing with incremental innovations, the emphasis on protection being fundamental (Ritala and Hurmelinna-Laukkanen 2009). Following this line of thought, radical innovations come up against less competitive pressure, since markets are more emergent and differentiation is easier due to the novelty.

Cohen and Walsh (2001) studied this process using the framework based on the concept of the firm's absorptive capacity. This concept refers to the identification of valuable knowledge in the environment, the capacity to assimilate it and align it with existing knowledge stocks and finally exploit it in internal R&D activities to achieve successful innovation.

As Cohen and Levinthal (1989) defend, the firm's knowledge base plays the role of both innovation and absorption, since its tendency to assimilate external knowledge creates an incentive to invest in R&D. Gambardella (1992) also states that firms with better in-house R&D programs are more able and prepared to absorb external scientific information. Other authors analysed the determinant role of the firm's absorptive capacity in exploiting the alliances it establishes (Arora and Gambardella 1994; Zahra and George 2002).

Zahra and George (2002) analysed the concept of absorptive capacity as a dynamic capability, creating a model of the components, antecedents, contingencies and outcomes of absorptive capacity. Their model was innovative because they substituted the component of "recognizing the value" with "acquisition" and relocated the influence of appropriability regimes. Additionally, these scholars enlarged the model with the transformation concept that follows the assimilation component, activation triggers and social integration mechanisms, and divided absorptive capacity into "potential" absorptive capacity and "realized" absorptive capacity. The process of transformation gives firms the capacity to develop changes in existing processes to be able to absorb new knowledge, assimilating it by means of interpretation and comprehension within existing cognitive structures.

Regarding that statement, Todorova and Durisin (2007) proposed that firms cannot transform their knowledge assets when they are not able to assimilate them. Furthermore, Zahra and George (2002) distinguish between potential absorptive capacity and realized absorptive capacity. The first has to do with acquisition and assimilation of new external knowledge by reconfiguring the resource base and deploying capacities, while the second deals with transformation and exploitation of new external knowledge by developing new products and processes. Potential absorptive capacity without realized capacity does not produce an effect on the firm's competitive advantage.

In addition, the authors identified the activation triggers, social integration mechanisms and appropriability regimes acting as key contingencies. Social integration mechanisms help to lower the barriers between assimilation and transformation, increasing absorptive capacity, which is understood by the proposed model as being a dynamic capacity involving a set of organizational routines (e.g. social interactions) and processes. The ability to learn and absorb depends on the capacity to value external knowledge (Zahra and George 2002).

Appropriability regimes allow moderation between absorptive capacity and its outcomes, resulting in competitive advantage. Thus, firms with low efficacy of intellectual property rights and easy replication are more prone to fail in the appropriation of innovation returns, giving open space to competitors.

Cockburn and Henderson (1998) state that the firm's ability to recognize the knowledge flow from the scientific community is determined by the close ties established with this community. They also stress the significant role of absorptive capacity in the firm's competitive advantage, since that capacity depends on its knowledge stock and resources.

George and Pradhu (2003) analysed the importance of a link between firms' absorptive capacity and the country's absorptive capacity, enabling innovation. Moreover, Cassiman and Veugelers (2006) analysed the positive impact of reliance on more basic R&D, which might proxy a firm's absorptive capacity, on the complementarity between internal and external innovation activities.

According to Rothaermel and Alexandre (2009), the greater the firm's absorptive capacity the greater its ability to fully capture the benefits resulting from flexibility in technology sourcing. Furthermore, the ability to recognize and exploit knowledge flows varies from one firm to another, resulting in unequal benefits acting as a competitive advantage. This absorptive capacity varies according to the firm's existing knowledge stock embedded in its products, processes and people. The authors also suggest that absorptive capacity plays a more determinant role in turbulent knowledge sectors and sectors with tighter and stronger IPR, and therefore governments should develop policies to foster firms' absorptive capacity in high-tech industries in conjunction with initiatives to increase IPR protection.

Silva and Leitão (2009a, b) also explore the benefits and roles of different types of relationships with external partners (including universities and research centers) to stimulate entrepreneurial innovation, resulting in product innovation.

Li (2011) examined sources of external technology, absorptive capacity and innovation capacity in Chinese state-owned high-tech firms, analysing three types of investment to acquire technological knowledge in determining firms' innovation capacity, namely: in-house R&D; importing foreign technology; and purchasing domestic technology. He concluded that importing foreign technology only promotes innovation if in-house R&D is also conducted. Nevertheless, domestic technology purchases, such as patent licensing, have a favourable direct impact on innovation. The study also finds that absorptive capacity is determined by the source or nature of the external knowledge.

Kostopoulos et al. (2011) explore the role of absorptive capacity as a mechanism to identify and translate external knowledge inflows into tangible benefits, and also as a vehicle to achieve greater innovation and time-lagged financial performance. The authors suggest that external knowledge inflows are directly related to absorptive capacity and indirectly related to innovation.

Vasudeva and Anand (2011) studied firms facing technological discontinuities and their use of alliance portfolios to gather knowledge flows. They subdivide absorptive capacity into "latitudinal" and "longitudinal" components. The first corresponds to the use of diverse knowledge and the second to distant knowledge. Their findings suggest that a firm with a moderate latitudinal absorptive capacity, which is equivalent to medium diversity in its portfolio, has a high propensity for optimal use of knowledge.

2.4 Coopetition and the Strategic Use of Product/Service Co-Innovation

According to Smith (2005), a product or service innovation can be protected via a patent which is a contract established between an inventor and a government that entitles the former to a limited time (20 years) in which a monopoly for the use and exploitation of a technical invention is ensured. The invention to be patented must be demonstrated to be a non-obvious advance in the state of the art. This confers a limited protection against competitors, to avoid copy or imitation. The patent system is considered to be a tool for promoting the creation of new economically valuable knowledge and also for dissemination of the state of the art in innovation technologies.

Macdonald (2004) argues that the patent is considered to be a means to an end, innovation being the end. In this vein, the public records that gather information on existing patents and applications can be used as an important tool for carrying out technological surveillance and monitoring, including data on granted patents with commercial viability on similar or cited patents. Chen and Chen (2011) state that patents protecting those product/service innovations are one of the firm's important intangible assets, in the sense that they can provide additional revenue to be generated towards product commercialization. Patent databases are of extreme importance for inventors to map new technologies and new products.

Griliches et al. (1991) and Chen and Chang (2010a, b) argue that patent information can provide more information than R&D information gathered in financial reports, which usually reveal very limited information (Chen and Chang 2009, 2010a, b). Also, Macdonald (2004) stated that patent data provide information on patent behaviour, by measuring performance in specific technological fields. An increasing number of firms use patent information not only to monitor competitors but also to prevent infringement.

According to Lai et al. (2007), patent analysis is able to provide information on technological innovation and its development. This is a critical issue, especially for firms that aim to make major investments and reach a competitive positioning in specific R&D sectors. In this sense, opting for strategic coopetition can be important to pursue sustainable competitive advantage. Additionally, in recent years patent rights have become a crucial tool in the coopetition strategy, since it is possible to obtain specific information regarding the technological position and to strengthen a certain sector, by detecting, forecasting and anticipating changes in terms of competitive positioning and business scope.

Wang (2010) refers to patents as important sources of technical information, since 90 % of knowledge is in the form of patent database and 70 % is disseminated through literature on patents. Acting as a source of ideas and technology resource, patents are considered to be a lever for technological innovation.

Seymour (2008) and Lee et al. (2009) argue that patent information can allow the identification of trends in technology and commercialization, in terms of patent distribution, competition status and development. In the same line of thought, Ernst (2003) defended that patent information is able to support technology management in five areas, complementing financial data when evaluating a firm's performance, namely: (i) support for R&D investment decisions; (ii) human resources and knowledge management in R&D activities; (iii) IP protection; (iv) screening and evaluation of external technological sources; and (v) maximising the value of the patent portfolio. The area of patent protection is extremely important in achieving competitive advantage, since it protects patent assignees from imitation and supports the internal use of technologies. Thus, strategic management of the patent portfolio is also important to achieve benefits and obtain competitive advantage (Grindley and Teece 1997).

According to Ernst (1998, 1999), patent information is valuable since it provides information on R&D even for firms that are not required to disclose R&D data, this information being available in several areas, such as business units, products, technological fields and inventors. This enables firms to carry out more accurate competitor analysis. Moreover, patent data can help firms to develop and guide the technological trajectory and monitor companies' R&D strategies. Patent statistics can be used to measure the outcomes of the codified knowledge from R&D activities and industrial development (Grupp and Schmooch 1999; Somaya 2003; Aoki and Schiff 2008).

Another important aspect of patent information is the data that can be collected to analyse the degree of rivalry, technology tracking and forecasting, identification of important developments, international strategic analysis and infringement monitoring. It is also critical for assessing viability of mergers and acquisitions and technological collaboration (Mogee 1991; Breitzman and Mogee 2002; Ma and Lee 2008).

Bergek and Bruzelius (2010) also point out the interest of patent data as an indicator of collaborative technological activity. The association of several international inventors suggests the existence of international cooperation (Carayol and Roux 2007; Ma and Lee 2008). In addition, patents can indicate the emergence of an international trend in a certain technological field, which in turn can contribute to revealing the evolutionary pathway in terms of collaborative development oriented to technological innovation (Archambault 2002).

In coopetition, controlling knowledge flows during joint R&D activities involves some risk, this being a critical issue in reaching success in strategic alliances oriented towards innovation activities embracing competitors. The risks of appropriability in a strategic alliance can be higher when partners are direct competitors (Park and Russo 1996). Appropriability methods can be of two types, formal and informal (Rammer 2002). Formal methods are the legal forms of protection such as patents, copyrights and trademarks, to prevent others from using the firm's patents and knowledge embedded in them, despite allowing the competing firm to access patent knowledge and learn from it. Informal methods include secrecy, complex design and lead time.

Blomqvist et al. (2005) defend the need to use formal tools to regulate collaboration between asymmetric partnerships and to protect intellectual capital. In this vein, IP rights are considered to be critical assets in knowledge-based competition, with discussions regarding ownership emerging during the collaboration process. The strategic management of intangibles is extremely important when regulating collaboration schemes, in order to prevent the incorrect appropriation of knowledge.

As mentioned by Dagnino and Rocco (2009), when coopetition occurs between public and private competitors, for instance between universities and industrial partners, in the challenging task of knowledge production two critical situations can arise: coopetition for publications and coopetition for IPRs. To overcome these problematic issues, the previous authors suggest three strategies to mitigate the competitive pressure between university and industry, namely the sequencing and sanitizing of data and joint patents. The first implies the strategic management and sequential processes of first patenting and then publishing. The second concerns the removal of data that shall not be published, in order to avoid risks when patenting. The third corresponds to the collaborative patenting of knowledge, sharing rights and duties in the patent process. Firms usually regard this type of coopetition strategy as disadvantageous, preferring exclusive rights in order to commercialize technology freely.

2.5 Research Hypotheses

Recent studies have taken coopetition as being a compound of collaboration and competition among firms, in order to produce innovation. Furthermore, patents are used, as stated by Carayol and Roux (2007) and Ma and Lee (2008), to establish collaborative technological relationships between firms and their stakeholders. Additionally, one of the main determinants for competing firms to engage in strategic cooperation arrangements and coopetition activities is the generation of value added or benefit, in order to improve their economic performance.

Studying the determinants of firms' capacity to generate innovative products/services and producing an insightful analysis of firms' innovative intensity behaviour based on coopetition relationships has been the target of several analyses. Some researchers, for instance Brandenburger and Nalebuff (1996), Dussauge et al. (2000) and Tether (2002), devoted their studies to the association between firms' innovative capacity and the coopetition arrangements they enter to generate value added and increase productivity.

Other scholars (Zahara and George 2002; Todorova and Durisin 2007; Rothaermel and Alexandre 2009; Kostopoulos et al. 2011) conducted their studies to analysing the impact of introducing process innovations inside the firm, which can be either in the production process or in organisational structure, embracing R&D positioning, such as fostering open innovation channels and absorptive capacity on the firm's capacity to generate innovations. Thus:

H1 The introduction of process innovations inside the firm has a positive and significant impact on the firm's capacity to generate co-innovative products/services.

The positive and significant impact of firms' investment in R&D activities performed inside the firm was also the subject of multiple studies, such as those by Cohen and Levinthal (1989), Gambardella (1992), Cassiman and Veugelers (2006) and Li (2011). These authors point to the major importance of the firm's possession of in-house R&D programs, of investing in the firm's basic R&D intensity, and of increasing the firm's in-house R&D performance. In this sequence, we present the following hypothesis:

H2 The performance of R&D activities inside the firm has a positive and significant impact on the firm's capacity to generate co-innovative products/services.

The introduction of innovations to the market was also the subject of several studies (Tether 2002; Quintana-Garcia and Benavides-Velasco 2004; Belderbos et al. 2004; Ritala and Hurmelinna-Laukkanen 2009) which focused on the significant effect of those on the innovation capacity of the firm.

In this vein, we formulate Hypothesis 3 as follows:

H3 The introduction of innovations into the market has a positive and significant impact on the firm's capacity to generate co-innovative products/services.

The determinant factor of establishing coopetition arrangements between competing firms for the firm's capacity to create innovations, either in products or in services, was analyzed by multiple scholars (Bradenburger and Nalebuff 1996; Bengtsson and Kock 2000, 2003; Bagshaw and Bagshaw 2001; Garraffo 2002; Belderbos et al. 2004; Chien and Peng 2005; Jong and Marsili 2006; Ritala and Hurmelinna-Laukkanen 2009; Rusko 2011; Vasudeva and Anand 2011). Thus we hypothesize:

H4 The set of coopetition relationships established between the firm and competing firms has a positive and significant impact on the firm's capacity to generate co-innovative products/services.

The impact of relationships with the scientific community as being of major importance in generating firms' innovative performance has warranted the attention of several researchers, for example, Cockburn and Henderson (1998), Li (2011), Kostopoulos et al. (2011) and Vasudeva and Anand (2011). Thus, we formulate hypothesis 5:

H5 The set of coopetition relationships established between the firm and other R&D stakeholders has a positive and significant impact on the firm's capacity to generate co-innovative products/services.

3 Methodology

3.1 Dataset, Method and Variables

This paper analyses the determinant factors of firms' capacity to generate co-innovative products and services, by making use of the data available in the European CIS Survey, 2008.

The data available is used to produce two samples related to manufacturing and service firms. The first is divided in two categories, according to the NACE classification, namely high-tech firms and low-tech firms. The second is divided into knowledge-intensive service firms and less knowledge-intensive service firms. A probit model is used to assess the probability of the independent variables explaining the factors derived from factor analysis regarding the determinants of firms' capacity to generate co-innovative products and services.

The manufacturing firm sample has 3682 respondent firms, considering all firms in the analysis since they are all statistically valid. The service sample has 1221 respondent firms, also considering all firms in the analysis since they are all statistically valid. The samples of manufacturing and service firms are submitted to a probit regression to estimate the probability associated with the different determinant factors of firms' innovative intensity performed with the independent variables presented in annex 1. For the analysis performed on the manufacturing sample we consider two more independent variables, namely the low-tech firm and the high-tech firm. Concerning the service sample, two other independent variables are also considered, knowledge-intensive firms and less knowledge-intensive firms.

The dependent variable used is product/service innovation (1 for a firm that has carried out co-innovative product/service innovation and 0 otherwise), which refers to the firm having generated and introduced into the market a new or improved product or service, by adopting a co-innovative behaviour with respect to its capacities or potential, ease of use, parts or subsystems, based on coopetition arrangements. In accordance with previous studies, the creation of new products was also used to analyse firms' innovative capacity, either through formal knowledge protection mechanisms or not (Tether 2002; Belderbos et al. 2004; Quintana-Garcia and Benavides-Velasco 2004; Ritala and Hurmelinna-Laukkanen 2009).

The binary dependent variable suggests the use of a probit model for estimation purposes. The dependent variable was used as a proxy to assess the co-innovative behavior of firms, which is revealed through empirical evidences on pro-innovation behavior of different types of firms, according to the data available on the CIS survey.

3.2 Descriptive Statistics

In the next figures (Figs. 1 and 2) we present a set of descriptive statistics for the manufacturing firm dataset consisting of 3682 firms. The major conclusions from the statistical analysis are that approximately 88 % of firms are low tech and 12 % are high tech as shown below. Additionally, 93 % are large firms.

Figure 2 shows that almost 37 % have developed product/service innovation, the authorship percentages for process innovations being distributed as follows: 30 % by the firm in isolation; 13 % by the firm in cooperation and the remaining by other forms.

As presented in Fig. 3, almost 29 % carry out inside R&D activities and approximately 14 % acquire outside R&D activities. About 11 % acquire other external knowledge (in the form of patents, copyrights and other unprotected knowledge), and 18 % show some introduction of new products to the market.

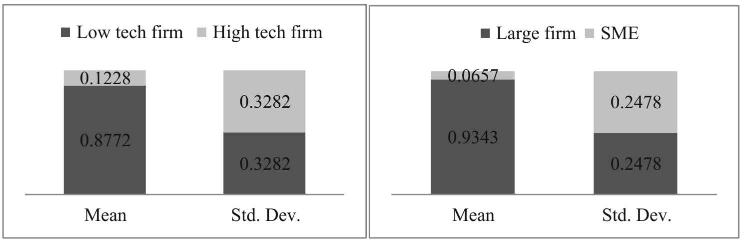


Fig. 1 Composition of manufacturing sample by technological intensity and size

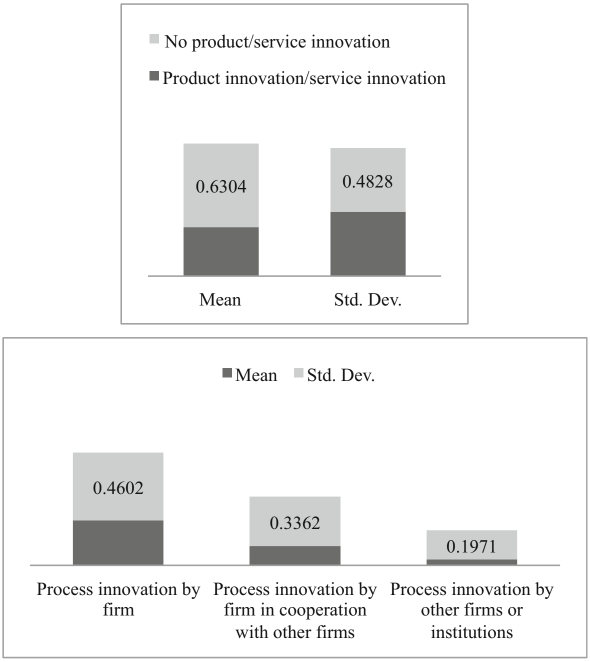


Fig. 2 Composition of manufacturing sample by product innovation performance and process innovation authorship

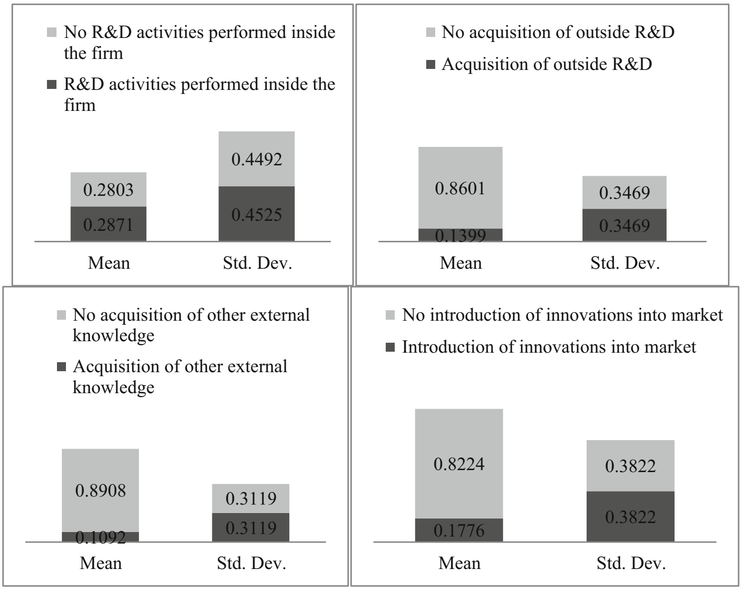


Fig. 3 Composition of manufacturing sample by R&D activities

As illustrated in Fig. 4, almost 19 % state that they cooperate in R&D activities, the preferred type of partner being public partners (83 %). In addition, only 4 % cooperated with Portuguese competitors, 2 % with European, and 1 % with American. Almost 7 % cooperated with Portuguese laboratories, 2 % with European ones, and 0.2 % with American ones. Finally, approximately 7 % cooperated with Portuguese universities, 10 % with European universities and 0.1 % with American universities.

The next figures present the descriptive statistics for the 1221 service firms. Approximately 60 % of firms are knowledge-intensive firms, and almost 91 % are large, as seen in Fig. 5.

In addition, Fig. 6 reveals that 26 % have developed product/service innovations, authorship percentages for process innovations being distributed as follows: 30 % by the firm itself; 16 % by the firm in cooperation with other firms and the remaining by other forms.

Considering Fig. 7, almost 35 % perform inside R&D activities and approximately 20 % acquire outside R&D activities. About 17 % acquire other external knowledge (such as patents, copyrights and other unprotected knowledge), and 17 % introduce new products/services to the market.

As presented in Fig. 8, almost 24 % state that they cooperate in R&D activities, showing no special preference for private or public partners. Moreover, almost 8 % cooperated with Portuguese competitors, 3 % with European, and almost 1 % with American. Approximately 4 % cooperated with Portuguese laboratories, 1 % with European ones, and 0.08 % with American ones. Almost 9 % cooperated with Portuguese consultants, 2 % with European ones and 0.4 % with American. Finally, 10 % of firms cooperate with Portuguese universities, 1 % with European universities and only about 0.7 % deal with American universities in cooperative relationships.

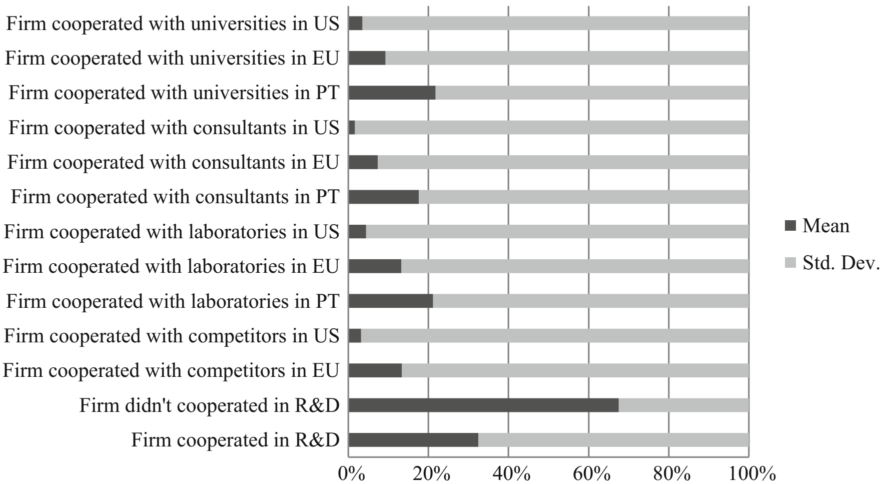


Fig. 4 Composition of manufacturing sample by cooperation activities

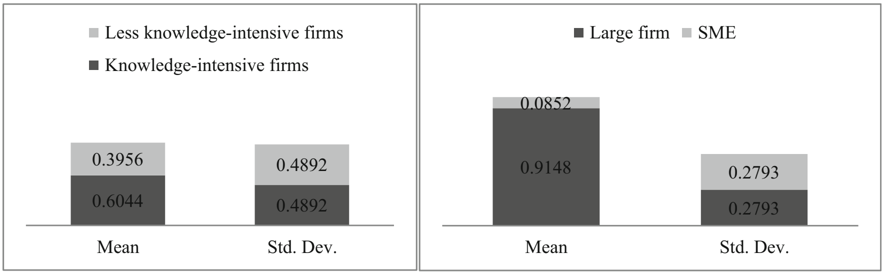


Fig. 5 Composition of service sample by technological intensity and size

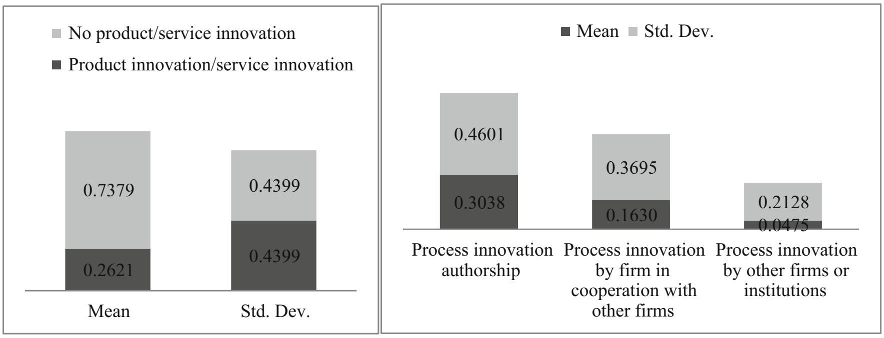


Fig. 6 Composition of service sample by product innovation performance and process innovation authorship

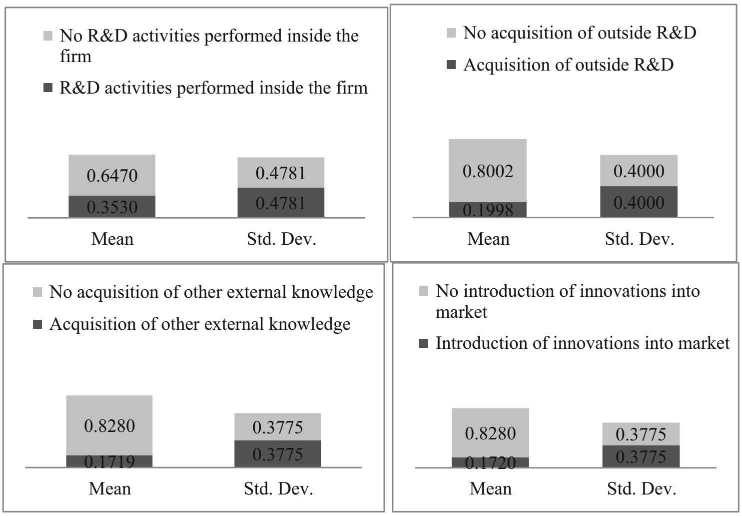


Fig. 7 Composition of service sample by R&D activities

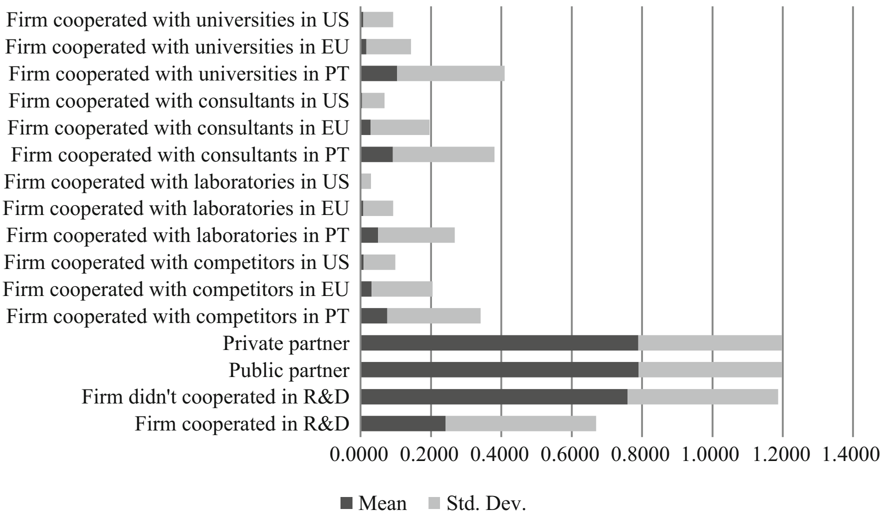


Fig. 8 Composition of service sample by cooperation activities

4 Empirical Findings

4.1 Probit Estimation Results

Probit regressions were run on manufacturing firms and service firms separately. In addition, within each sector two additional separate regressions were run based on the intensity of firms' technology. These groups are based on the NACE classification for low-tech and high-tech manufacturing firms, and knowledge-intensive firms and less knowledge-intensive firms, for the service dataset. The results of these regressions are presented in Tables 1 and 2.

Regarding the results of the probit regression for the sample of manufacturing firms, from the column of 'all firms', we can conclude that for the 3682 firms under analysis, the likelihood ratio chi-square of 356.21 with a p-value of 0.0000 tells us that our model as a whole is statistically significant, that is, it fits significantly better than a model with no predictors.

According to the values presented in Table 1, two determinant factors with a negative and significant influence on firms' capacity to generate co-innovative products or services are innovation processes implemented by the firm and non-acquisition of outside R&D services either from firms or from scientific partners.

Furthermore, the firm's cooperation both with Portuguese and European competitors also has a significant effect, although positive. Also having a significant and positive impact is firm cooperation with Portuguese and European laboratories, and Portuguese universities.

Table 1 Results of probit regressions for manufacturing firms

Co-innovative product/services	All firms	Low-tech firms	High-tech firms
Low tech firm	−0.0576045	–	–
High tech firm	–	–	–
Large firm	−0.10673	–	–
SME	–	0.0867604	0.6572347*
Product innovation/service innovation	–	0.071071	–
No product/service innovation	−0.046014	–	0.151123
Process innovation by firm	−0.1642524***	−0.1398221	−0.2718769
Process innovation by firm in cooperation with other firms	−0.0801076	−0.0461338	−0.2116772
Process innovation by other firms or institutions	−0.0909556	−0.1386163	0.2278046
R&D activities performed inside the firm	−0.0895601	−0.1275257	0.2301739
No R&D activities performed inside the firm	−0.0106024	−0.0263119	−0.0241219
No acquisition of outside R&D	−0.165559***	0.182017*	0.0541903
Acquisition of other external knowledge	0.0669793	0.0361933	0.2378885
Introduction of innovations into market	0.016341		0.5726169***
No introduction of innovations into market	–	0.0411851	–
Firm cooperated in R&D	–	–	−0.1123371
Firm did not cooperate in R&D	−0.1790235	−0.2542488	–
Public partner	–	–	−0.095345
Private partner	−0.2591099	−0.3274853**	–
Firm cooperated with competitors in PT	0.5715775***	0.6294335***	0.333516
Firm cooperated with competitors in EU	0.7269029***	0.5445594**	–
Firm cooperated with competitors in US	−1.001.694	−2.222.612**	–
Firm cooperated with laboratories in PT	0.490596***	0.4872625***	0.6268807*
Firm cooperated with laboratories in EU	0.5615345***	0.5449852***	0.8529929
Firm cooperated with laboratories in US	0.4544849	0.7596476	–
Firm cooperated with universities in PT	0.7174335***	0.6918532***	0.7266951*
Firm cooperated with universities in EU	0.510753	0.4389709	–
Firm cooperated with universities in US	−0.6025823	−0.8155625	–
Observations	3682	3267	415
Log Likelihood	−2244.5439	−1954.1143	−264.63102
Pseudo R^2	0.0735	0.0676	0.0741

*significant at 10 %|**significant at 5 %|***significant at 1 %

The last two columns in Table 1 show regressions for the sub-samples of different technological intensity.

As for the results of the probit regression for the sample of low-tech manufacturing firms, we can conclude that for the 3267 firms under analysis, and considering the likelihood ratio chi-square of 283,49 with a p-value of 0.0000 our model as a

Table 2 Results of probit regressions for service firms

Co-innovative products/services	All firms	KIS firms	LKIS firms
Less knowledge-intensive firms	−0.0351267	–	–
Large firm	0.2917284*	–	–
SME	–	−0.024813	−0.71954***
Process innovation by firm	0.6788217***	0.6425258***	0.8003994***
Process innovation by firm in cooperation with other firms	0.4931047***	0.579551***	0.5354501***
Process innovation by other firms or institutions	0.4324939***	0.314317	0.4787559
R&D activities performed inside the firm	0.5340988***	0.4726756***	0.6925766***
No R&D activities performed inside the firm	−0.0772205	−0.1096384	–
Acquisition of outside R&D	–	0.2268566	–
No acquisition of outside R&D	−0.2870978***	–	−0.0354656
Acquisition of other external knowledge	–	–	0.3181008
Introduction of innovations into market	0.5200406***	–	–
No introduction of innovations into market	–	−0.8073311***	0.0673119
Firm did not cooperate in R&D	−0.8041166***	−1.037.318***	−0.5045445
Public partner	−3.605.851	0.7028044***	−4.005.418
Private partner	4.071.048***	–	4.335.834***
Firm cooperated with competitors in PT	−0.0816495	−0.326807	0.2739777
Firm cooperated with competitors in EU	0.5535745*	1.375.734***	0.7578617
Firm cooperated with competitors in US	−1.003.039**	−1.929.241***	−1.308.725
Firm cooperated with laboratories in PT	0.3690016	0.318485	0.9656868*
Firm cooperated with laboratories in EU	−1.708.198**	−2.208.943***	–
Firm cooperated with consultants in PT	−0.0796663	−0.0186786	0.1562392
Firm cooperated with consultants in EU	−0.2792368	−0.2461181	−0.4516517
Firm cooperated with consultants in US	1.132.891	0.8281421	–
Firm cooperated with universities in PT	−0.1740137	−0.2204494	−0.2928758
Firm cooperated with universities in EU	0.7373061*	1.217.358**	0.2346324
Firm cooperated with universities in US	−0.5046289	−0.12174	–
Observations	1221	746	475
Log Likelihood	−526.22295	−318.34736	−190.09896
Pseudo R^2	0.2453	0.2957	0.1907

*significant at 10 %|**significant at 5 %|***significant at 1 %

whole is also statistically significant, that is, it fits significantly better than a model with no predictors.

Considering the sample of ‘high-tech manufacturing firms’, the likelihood ratio chi-square being 42.38 with a p-value of 0.0003 our model is also statistically significant for the 415 firms under analysis.

In Table 1, we verify that for ‘low-tech manufacturing firms’ the fact that firms do not acquire outside R&D activities impacts negatively and significantly (at 10 %

significance) on the co-innovative products or services performed by the firm. Additionally, cooperation in R&D activities with private partners also has a positive and significant effect on the dependent variable (at 5 % significance).

Other determinant factors explaining the innovation capacity of firms to generate co-innovative products or services are the firm's attitude towards cooperation with Portuguese competitors, with Portuguese and European laboratories and with Portuguese universities, which impacts positively and significantly (at 1 % significance) and with European and American competing firms (at 5 % significance).

Considering the sub-group of 'high-tech manufacturing firms', their size is revealed to be important when explaining co-innovative capacity, i.e., small and medium high-tech firms are more likely to impact positively and significantly (at 10 % significance) on product/service innovation. Also, the introduction of innovations to the market reveals a positive and significant impact on the dependent variable (at 1 % significance). Firms' cooperation with Portuguese laboratories and universities has a positive and significant effect on their capacity to generate co-innovative products or services (at 10 % significance).

Of special interest here are the major differences between the results obtained for high tech and low tech manufacturing firms.

The process innovation carried out by the firm itself or a group of firms is positively and significantly associated with co-innovative products or services for all firms, but it does not reveal any importance in the other sub-samples. In addition, non-acquisition of external R&D services in the 'all firms' sample has a negative and significant effect on the co-innovative products or services, but in the 'low-tech firms' sub-sample it has a positive and significant effect on the dependent variable.

Cooperation activities with Portuguese and European competitors and laboratories, and Portuguese universities, always show a positive association with the firm's co-innovative capacity for the 'all firms' sample and for the 'low-tech' sub-sample. In addition, for the latter sub-sample, cooperation alliances with US competitors also have a significant, though negative, impact on the firm's co-innovative capacity.

For 'low-tech firms', the effect of the type of partner is significantly, but negatively associated with co-innovative capacity.

Finally, for "high-tech firms', the variables showing a significant association with firms' product/service innovation capacity are slightly different from the other samples, namely the positive significance of firm size, revealing that SMEs are more associated with innovativeness, the capacity to introduce innovations to the market and the set of cooperation activities with Portuguese laboratories and universities.

Regarding the set of results of the probit regression for service firms in Table 2, and particularly the 'all firms' column, we can conclude that for the 1221 firms under analysis, the likelihood ratio chi-square of 356.21 with a p-value of 0.0000 confirms that our model as a whole is statistically significant, that is, it fits significantly better than a model with no predictors.

For the sample of 'all service firms', we can conclude that being a large firm impacts positively and significantly on firms' capacity to generate co-innovative products or services.

Moreover, the introduction of process innovations, either alone or in cooperation with other firms, by both 'knowledge intensive service firms' and 'less knowledge intensive service firms' has a positive and significant effect on the firm's capacity to generate co-innovative products or services (at 1 % significance). Additionally, for the sample of 'all service firms' there is a positive and significant impact of the introduction of process innovations by other institutions, on the firm's innovative capacity to create co-innovative products or services (at 10 % significance).

The fact that the firm carries out R&D activities inside the firm and introduces innovations into the market also shows a positive and significant effect on the dependent variable (at 1 % significance).

Service firms that neither acquire external R&D activities nor cooperate in R&D activities present a negative and significant (at 1 % significance) association with the firm's capacity to generate co-innovative products or services.

Private partner profile also has a positive and significant impact on the dependent variable (at 1 % significance), with European competitors and European universities being the partners with greatest positive and significant impact on co-innovative products or services (at 10 % significance).

Finally, cooperation relationships with American competitors and European laboratories have a significant, but negative, effect on the firm's capacity to generate co-innovations (at 5 % significance).

The last 2 columns show the probit regressions disaggregated into service sub--groups—'KIS' and 'LKIS'. Considering the sub-sample of 'KIS firms', in a total of 746 firms, the likelihood ratio chi-square presents a value of 267.31 with a p-value of 0.0000, suggesting a statistically significant model.

For this sub-sample, introduction of process innovations to the firm, either by the firm itself or the firm in cooperation with others, presents a positive and significant association with the capacity to generate co-innovation (at 1 % significance). Besides, the set of R&D activities performed inside the firm also has a positive and significant impact on the dependent variable (at 1 % significance).

The fact that this type of firm does not introduce innovations to the market has a negative and significant effect on the capacity to generate co-innovative products or services (at 1 % significance), giving an association between the generation of innovation and subsequent market introduction.

Also negative is the impact of the inexistence of cooperative relationships in terms of R&D on the dependent variable (at 1 % significance), a public partner being the preferred type of partner in cooperative relationships, this dummy variable having a positive and significant impact (at 1 % significance).

Cooperative relationships between the firm and European competitors and universities present a positive and significant association with the firm's capacity to generate co-innovation (the first at 1 % significance and the second at 5 % significance).

The set of cooperation agreements with a significant, though negative, impact on the firm's capacity to generate co-innovations, either product type or service type, are with American competing firms and European laboratories.

When analysing the sub-sample of 'LKIS service firms', a total of 475 firms, the estimations present a likelihood ratio chi-square value of 89.59 with a p-value of 0.0000, also suggesting a statistically significant model.

For this sub-sample, the dummy variable of SME has a negative and significant impact on the firm's capacity to generate innovations. Furthermore, introduction of process innovations by the firm itself and/or in cooperation with other firms has a positive and significant impact on the firm's capacity to generate product and/or service innovation (at 1 % significance).

R&D activities carried out inside the firm also show a positive and significant association with the firm's generation of co-innovations (at 1 % significance). For 'less knowledge intensive service firms', private partners show a positive and significant association with the co-innovative products or services (at 1 % significance), and among all partners, Portuguese laboratories are the ones showing a positive and also significant impact on those co-innovations (at 10 % significance).

The major considerations to be pointed out when comparing results for the sub-samples of 'all firms' and 'KIS' and 'LKIS firms' are the fact that introduction of process innovations in the firm presents a positive and significant association with the firm's capacity to generate co-innovations in all sub-samples.

Furthermore, size is only important for the sample of 'all firms', showing the positive impact of the large firm variable and for the 'LKIS' sub-sample showing the negative effect of the SME variable.

Carrying out R&D activities inside the firm reveals a positive and significant effect on the firm's capacity to generate co-innovations for all cases.

Considering the introduction of innovations into the market, this has a positive and significant effect on the dependent variable for the 'all firms' sample and in the opposite direction, non-introduction of innovations has a negative and significant impact on the dependent variable, for 'KIS firms'.

For 'KIS firms', the most important type of partner is the public one, this dummy variable having a positive and significant impact on the firm's generation of new products/services. In turn, for 'LKIS firms' and for the 'all firms' sample, it is the private type of partner that shows a positive and significant association with that capacity.

Additionally, EU competitors and EU universities have a positive and significant impact on the firm's capacity to produce co-innovations for the 'all firms' sample and 'KIS firms'. Conversely, US competitors and EU laboratories show a negative and significant impact on the firm's innovation generation for the 'all firms' and 'KIS firms' samples. On the contrary, for the 'LKIS' sub-sample the only important cooperation is with Portuguese laboratories, where joint actions impact positively and significantly on the dependent variable.

4.2 Research Hypotheses and Discussion

Considering the sample of manufacturing firms and the results produced by the probit regressions, we can summarise that regarding the first hypothesis of a positive and significant effect of introducing process innovations in the firm on the firm's capacity to generate co-innovations, it is possible to confirm a significant, but negative, association, when considering the 'all firms' sample. Thus, we partially fail to reject H1. This is in line with previous studies mentioned (Zahara and George 2002; Todorova and Durisin 2007; Rothaermel and Alexandre 2009; Kostopoulos et al. 2011).

Additionally for this sample, and considering H2, suggesting a significant and positive effect of performing R&D activities inside the firm on its capacity to generate co-innovative products or services, we conclude that for 'manufacturing firms' this is not of particular importance, and so reject H2. Thus, this is not in line with previous scholars' analyses, such as those of Cohen and Levinthal (1989), Gambardella (1992), Cassiman and Veugelers (2006) and Li (2011).

Also, for H3 proposing a positive and significant impact of the introduction of innovations into the market on the firm's subsequent capacity to generate co-innovations, we can point out that only for the sub-sample of 'high-tech manufacturing firms' is this effect revealed to be positive and significant, and so we fail to reject H3. Here, we follow previous scholars (Tether 2002; Belderbos et al. 2004; Quintana-Garcia and Benavides-Velasco 2004; Ritala and Hurmelinna-Laukkanen 2009).

The H4 proposes a positive and significant association between the set of coopetition relationships with a firm's competitors and its capacity to generate co-innovative products or services. For the samples of 'all firms' and 'low-tech firms' this relationship is positive and significant, especially for Portuguese and European competitors, and thus we fail to reject H4. For 'low-tech firms' we also found a significant, but negative, effect when considering American competitors, which means we partially fail to reject H4 in this case, which is in line with previous studies (Bradenburger and Nalebuff 1996; Bengtsson and Kock 2000, 2003; Bagshaw and Bagshaw 2001; Garraffo 2002; Belderbos et al. 2004; Chien and Peng 2005; Jong and Marsili 2006; Ritala and Hurmelinna-Laukkanen 2009; Rusko 2011; Vasudeva and Anand 2011).

Regarding H5 suggesting a positive and significant effect of coopetition relationships among firms and other R&D stakeholders on the firm's capacity to generate co-innovative products or services, we can confirm a positive and significant impact of Portuguese and European laboratories and Portuguese universities in the 'all firms' sample and the 'low-tech firms' sub-sample, leading us to fail to reject H5. Furthermore, when analysing the 'high-tech' sub-sample, we can corroborate such results, as Portuguese laboratories and Portuguese universities have a positive and significant impact on the dependent variable. Therefore, we also fail to reject H5 for 'high-tech manufacturing firms'. Here, we are in agreement with several studies already mentioned, for instance Cockburn and Henderson (1998), Li (2011), Kostopoulos et al. (2011) and Vasudeva and Anand (2011).

Considering the service firm dataset and taking into consideration H1, proposing a positive and significant effect of the introduction of process innovations in the firm on its capacity to generate co-innovations, we find a significant and positive association for all samples under analysis. Thus, we fail to reject H1. These results are contradictory to those obtained for manufacturing firms which tended towards a negative association.

Taking into account H2 proposing a significant and positive impact of performing R&D activities inside the firm on its capacity to generate product/service innovation, we confirm a positive and significant effect, failing to reject H2. This is also different from the manufacturing dataset, which did not reveal any association between these variables.

For the H3, which defends a positive and significant impact of the introduction of innovations to the market on the firm's capacity to generate co-innovative products or services, we verified a positive and significant effect, when considering the 'all firms' sample, and so we fail to reject H3. For the 'KIS' and 'LKIS' sub-samples such an effect is not observed. This result is in line with the one obtained for the manufacturing high-tech firms sub-sample.

Considering H4 arguing for a positive and significant association between the set of coopetition relationships with firm's competitors and its capacity to generate co-innovative products or services, we obtained a positive and significant effect for European competitor relationships, for the 'all firms' sample and the 'KIS firms', leading to failure to reject H4. In addition, we can point out a significant, though negative, impact, of US and Portuguese coopetition relations on the firm's capacity to generate co-innovations, and so we partially fail to reject H4. These results are in line with previous results achieved for the manufacturing dataset.

Finally, for H5, proposing a positive and significant effect of coopetition relationships among firms and other R&D stakeholders on the firm's capacity to generate co-innovative products or services, we confirm a positive and significant impact of European universities for the 'all firms' sample and the 'KIS firms' sub-sample, and so we fail to reject H5. Furthermore, we also detect a significant but negative effect of coopetition relationships, particularly analysing the impact of European laboratories in the 'all firms' sample and the 'KIS' sub-sample, on the dependent variable. Therefore, we also partially fail to reject H5 for the 'all firms' sample and the 'KIS firms' sub-sample.

5 Concluding Remarks

In summing up, it is important to stress some differences detected between regressions with the three samples. Considering manufacturing firm factors, such as the introduction of process innovations in firms' internal organization and procedures and the practice of internal R&D activities, they do not impact on the firm's capacity to generate co-innovation.

On the contrary, for the service firms' dataset, both these factors are of major importance for the firm's innovative capacity to create co-innovative products or services, for the 'all firms' sample and for 'KIS' and 'LKIS firms'.

Regarding the dummy variable of introduction of innovations to the market, this only reveals a significant and positive effect in 'high-tech manufacturing firms' and in the service firm dataset as a whole.

Moreover, the set of coopetition relationships between the firm and competitors is seen to have an impact on the firm's capacity to generate co-innovation in both datasets, but for manufacturing firms this importance is due to the joint actions of Portuguese plus European competitors and for service firms only European competitors show a positive and significant impact on the dependent variable. However, for 'high-tech manufacturing firms' this effect is not observed, the same being true for 'LKIS' service firms.

Taking into consideration the impact of the set of coopetition relationships between firms and other R&D stakeholders, the major difference detected between the two datasets is the fact that coopetition agreements with European laboratories and manufacturing firms, especially when dealing with the 'all firms' sample and the 'low-tech' sub-sample, have a positive and significant effect on the firm's capacity to generate co-innovative products or services. On the contrary, the significant effect of coopetition agreements with European laboratories on that capacity is revealed to be negative for service firms, specifically for the 'all firms' sample and the 'KIS firms'. Furthermore, in the manufacturing firm dataset we verified a positive and significant effect of Portuguese universities on the firm's capacity to generate co-innovation. For service firms, the positive and significant effect is also detected but with European universities.

5.1 Policy and Managerial Implications

Since public policies play a crucial role in fostering co-innovative capacities, it is important that policy-makers understand the determinants of firms' capacity to generate co-innovative products and services, and their effects on innovative performance, the generation of net value added and economic benefits.

In terms of policy implications arising from the present study, it is suggested that public policies should be guided towards the creation and consolidation of open innovation flows, and towards fostering patenting strategies in firms and in consortiums between firms, and between firms and the scientific community, securing formal channels and mechanisms directed at minimising appropriability risks.

By making use of firms' capacity to generate innovation in order to reveal their innovative performance and the dynamics of coopetition public policies oriented to open innovation, the present study can give insights to those who manage innovation policy orientations, since knowledge of the set of determinant factors of firms' innovative behaviour can be helpful in drawing up guidelines to foster and properly manage the open innovation workflows between firms and their stakeholders, and then developing the capacity to generate and transfer new co-innovative products or services to market.

Overall, the results of this analysis may provide helpful starting points for practitioners (either in firms or coopetition stakeholders) who wish to estimate the directions of their organization's R&D projects and patents. Hence, the study may increase the effectiveness of innovative behaviour among coopetion partners, namely their patenting performance, in fostering synergetic relationships. Furthermore, by facilitating the externalization and codification of technological knowledge, patenting behaviour among coopetition partners can be endangered if there are no protection measures regarding appropriability risks.

Anticipation of such risks can enhance the efficiency of technology transfer flows, and consequently stimulate the creation, diffusion and regulation of defensive mechanisms to be used as routines by the partners involved.

5.2 Limitations and Future Research

The main limitation of the present study is the lack of data on firms' innovative capacity when trying to access data on patenting behaviour and other IP rights, such as copyrights and trademarks. This is also the main limitation of the database used in this study, the European CIS Survey, 2008, with the quasi-inexistence of data regarding firms' IP performance, considering additional data on patents, copyrights and other IP rights, since the only reference to innovative products or services generated inside and by the firm that can or cannot be protected via IP formal mechanisms is the variable of product/service innovation.

Other important information about firms' patenting capacity is not included in the survey. Furthermore, this study only relates to Portuguese innovative firms, a sample that should be expanded in future research to consider cross-country differences. In this connection, future research should be focused on the factors that motivate firms to engage in patenting projects, whether coopetition patenting initiatives, technological surveillance or forecasting projects. Firms' patenting strategies and characteristics, which influence their cooperation arrangements, should also be analysed.

Acknowledgment João Leitão acknowledges Fundação para a Ciência e a Tecnologia (FCT) for support under the Sabbatical Scholarship Program (Ref.: SFRH/BSAB/127908/2016).

References

Abernathy, W., & Clark, K. (1985). Innovation: mapping the winds of creative destruction. *Research Policy, 14*, 3–22.

Agresti, A. (1996). *An introduction to categorical data analysis*. New York: Wiley.

Amburgey, T., & Rao, H. (1996). Organizational ecology: Past, present and future directions. *Academy of Management Journal, 39*(5), 1265–1286.

Aoki, R., & Schiff, A. (2008). Promoting access to intellectual property: Patent pools, copyright collectives and clearinghouses. *R&D Management, 38*, 189–204.

Archambault, E. (2002). Methods for using patents in cross-country comparisons. *Scientometrics, 54*(1), 15–30.

Archibugi, D. (2001). Pavitt's taxonomy sixteen years on: A review article. *Economic Innovation and New Technology, 10*, 415–425.

Arora, A., & Gambardella, A. (1994). Evaluating technological information and utilizing it: Scientific knowledge, technological capability, and external linkages in biotechnology. *Journal of Economic Behavior and Organization, 24*, 91–114.

Arthur, W. (2007). The structure of invention. *Research Policy, 36*(2), 274–287.

Bagshaw, M., & Bagshaw, C. (2001). Co-opetition applied to training—A case study. *Industrial and Commercial Training, 33*(4/5), 175–177.

Belderbos, R., M., C., & Lokshin, B. (2004). Cooperative R&D and firm performance. *Research Policy, 33*, 1477–1492.

Bengtsson, M., & Kock, S. (2000). Coopetition in business networks—To cooperate and compete simultaneously. *Industrial Marketing Management, 29*, 411–426.

Bengtsson, M., & Kock, S. (2003). Tension in co-opetition. *Academy of Marketing Science Annual Conference*, May 28–31, Washington, DC.

Bergek, A., & Bruzelius, M. (2010). Are patents with multiple inventors from different countries a good indicator of international R&D collaboration? The case of ABB. *Research Policy, 39*, 1321–1334.

Bessen, J., & Maskin, E. (2009). Sequential innovation, patents, and imitation. *The RAND Journal of Economics, 40*(4 (Winter)), 611–635.

Blomqvist, K., Hurmelinna, P., & Seppanen, R. (2005). Playing the collaboration game right—Balancing trust and contracting. *Technovation, 25*, 497–504.

Braczyk, H., & Cooke, P. (1998). *Heidenreich, Regional Innovation Systems*. London: UCL Press.

Brandenburger, A. M. & Nalebuff, B. J. (1996). Co-opetition. Currency, Doubleday, New York.

Breitzman, A., & Mogee, M. (2002). The many applications of patent analysis. *Journal of Information Science, 28*(3), 187–205.

Carayol, N., & Roux, P. (2007). The strategic formation of intra-individual collaboration networks. Evidence form co-invention patterns. *Annales d'Économie et de Statistique, 87/88*, 275–301.

Cassiman, B., & Veugelers, R. (2006). In search of complementarity in the innovation strategy: Internal R&D and external knowledge acquisition. *Management Science, 52*, 68–82.

Chen, Y., & Chang, K. (2009). Using neural network to analyse the influence of the patent performance upon the market value of the US pharmaceutical companies. *Scientometrics, 80* (3), 637–655.

Chen, Y., & Chang, K. (2010a). The nonlinear nature of the relationships between the patent traits and corporate performance. *Scientometrics, 82*(1), 201–210.

Chen, Y., & Chang, K. (2010b). The relationship between a firm's patent quality and its market value—The case of US pharmaceutical industry. *Technological Forecasting & Social Change, 77*(1), 20–33.

Chen, Y., & Chang, K. (2010c). Analyzing the nonlinear effects of firm size, profitability and employee productivity on patent citations of the US pharmaceutical companies by using artificial neural network. *Scientometrics, 82*(1), 75–82.

Chen, Y., & Chen, B. (2011). Utilizing patent analysis to explore the cooperative competition relationship of the two LED companies: Nichia and Osram. *Technological Forecasting & Social Change, 78*, 294–302.

Chesbrough, H. (2003). Open innovation: The new imperative for creating and profiting from technology, Boston.

Chidamber, S., & Kon, H. (1994). A research retrospective of innovation inception and success: The technology-push, demand-pull question. *International Journal of Technology Management, 9*(1), 94–112.

Chien, T., & Peng, T. (2005). Competition and cooperation intensity in a network—A case study in Tawain simulator industry. *Journal of American Academy Business, 7*, 150–155.

Cockburn, I., & Henderson, R. (1998). Absorptive capacity, coauthoring behavior, and the organization of research in drug discovery. *The Journal of Industrial Economics, 46*(2), 157–183.

Cohen, W., & Levinthal, D. (1989). Innovation and learning: The two faces of R&D. *Economic Journal, 99*, 569–596.

Cohen, W., & Levinthal, D. (1990) Absorptive capacity: A new perspective on learning and innovation. *Administrative Science Quarterly*, 35, 1, Special Issue: Technology, Organizations, and Innovation, March, 128–152.

Cohen, W., & Walsh, J. (2001). R&D spillovers, appropriability and R&D intensity: A survey based approach, the economical evaluation of technological change, June 15–16 1998, Georgetown University Conference Center Washington DC, Editor R. Spivack.

Cohen, W., Nelson, R., & Walsh, J. (2000). Protecting their intellectual assets: Appropriability conditions and why U.S. manufacturing firms patent (or not). *NBER Working Paper*, No. 7552.

Conceição, P., & Heitor, M. (2007, January). Diversity and integration of science and technology policies. *Technological Forecasting and Social Change, 74*(1), 1–17.

Cooke, P., Uranga, M., & Etxebarria, G. (1997). Regional innovation systems: Institutional and organizational dimensions. *Research Policy, 26*(4–5), 475–491.

Cooke, P., Boekholt, P., & Tödtling, F. (2000). *The governance of innovation in Europe: Regional perspectives on global competitiveness*. London: Pinter.

Dagnino, G., & Rocco, E. (2009). *Coopetition strategy: Theory, experiments and cases*. New York: Routledge.

Dosi, G. (1988). Sources, procedures and microeconomic effects of innovation. *Journal of Economic Literature, 26*(3), 1120–1171.

Dussauge, P., Garrette, B., & Mitchell, W. (2000). Learning from competing partners: Outcomes and durations of scale and link alliances in Europe, North America and Asia. *Strategic Management Journal, 21*, 99–126.

Ernst, H. (1998). Industrial research as a source of important patents. *Research Policy, 27*, 1–15.

Ernst, H. (1999). Patent portfolios for strategic R&D planning. *Journal of Engineering Technological Management, 15*, 279–308.

Ernst, H. (2003). Patent information for strategic technology management. *World Patent Information, 25*, 233–242.

Escribano, A., Fosfuri, A., & Tribó, J. (2009). Managing external knowledge flows: The moderating role of absorptive capacity. *Research Policy, 38*, 96–105.

Ferrão, M. (2003). *Introdução aos modelos de regressão multinível em educação*. Campinas: Komedi.

Fjelstad, A. D., Becerra, M., & Narayanan, S. (2004). Strategic action in network industries: An empirical analysis of the European mobile phone industry. *Scandinavian Journal of Management, 20*, 173–196.

Flanagan, K., Kieron, E., & Laranja, M. (2011, June). Reconceptualising the 'policy mix' for innovation. *Research Policy, 40*(5), 702–713.

Frankel, M. (1955, June). Obsolescence and technological change in a maturing economy. *The American Economic Review*, *45*(3), 296–319.

Gambardella, A. (1992). Competitive advantages from in-house basic research. *Research Policy, 21*, 391–407.

Garraffo, F. (2002). Types of coopetition to manage emerging technologies. *2nd Annual Conference, Innovative Research in Management*, Stockholm, Sweden.

George, G., & Prabhu, G. (2003). Developmental financial institutions as technology policy instruments: Implications for innovation and entrepreneurship in emerging economies. *Research Policy, 32*, 89–108.

Godinho, M. (2003) Inovação: Conceitos e Perspectivas Fundamentais. In M. Rodrigues, A. Neves, & M. Godinho (Orgs.), Para uma Política de Inovação em Portugal, Dom Quixote, Lisboa: Biblioteca de Economia & Empresa, 29–51.

Gomes-Casseres, B. (1996). *The alliance revolution*. Cambridge, MA: Harvard University Press.

Griliches, Z. (1957). Hybrid corn: An exploration in the economics of technological change. *Econometrica, 25*(4), 501–522.

Griliches, Z., Hall, B., & Pakes, A. (1991). R&D, patents and market value revisited: Is there a second (technological opportunity) factor? *Economic Innovation New Technology, 1*(3), 183–201.

Grindley, P., & Teece, D. (1997). Managing intellectual capital: Licensing and cross-licensing in semiconductors and electronics. *California Management Review, 39*, 8–41.

Grupp, H., & Schmooch, U. (1999). Patent statistics in the age of globalization: New legal procedures, new analytical methods, new economic interpretation. *Research Policy, 28*, 377–396.

Harbison, K., & Pekar, P. (1998). *Smart alliances*. San Francisco: Jossey-Bass.

Hatzichronoglou, T. (1997). Revision of the high-technology sector and product classification. *OECD STI Working Paper Series* No. 1997/2, Paris.

Heitor, M., & Bravo, M. (2010, February). Portugal at the crossroads of change, facing the shock of the new: People, knowledge and ideas fostering the social fabric to facilitate the concentration of knowledge integrated communities. *Technological Forecasting and Social Change, 77* (2), 218–247.

Hicks, J. (1932). *The theory of wages*. London: MacMillan.

Huston, L., & Sakkab, N. (2006). Connect and develop. *Harvard Business Review, 84*(3), 58–66.

Jong, J., & Marsili, O. (2006). The fruit flies of innovations: A taxonomy of innovative small firms. *Research Policy, 35*, 213–229.

Kaufmann, A., & Tödtling, F. (2001). Science-industry interaction in the process of innovation: The importance of boundary-crossing between systems. *Research Policy, 30*, 791–804.

Kleinknecht, A. (1990, August). Demand and innovation: Schmookler re-examined. *Research Policy, 19*(4), 387–394.

Klepper, S. (1997). Industry life cycle. *Industrial and Corporate Change, 6*(1), 145–181.

Klevorick, A. K. (1995). On the sources and significance of inter-industry differences in technological opportunities. *Research Policy, 24*(2), 185–205.

Kostopoulos, K., Papalexandris, A., Papachroni, M., & Ioannou, G. (2011). *Absorptive capacity, innovation, and financial performance*. In Press, Corrected Proof.

Lai, K., Su, F., Weng, C., & Chen, C. (2007). Co-opetition strategy from the patent analysis perspective: The case of the stent market. *International Journal of Innovation and Technology Management, 4*(02), 137–153.

Laranja, M. (2008, January). The development of technology infrastructure in Portugal and the need to pull innovation using proactive intermediation policies. *Technovation, 29*(1), 23–34.

Lee, S., Yoon, B., & Park, Y. (2009). An approach to discovering new technology opportunities: Keyword-based patent map approach. *Technovation, 29*, 481–497.

Leitão, J., & Baptista, R. (2009). *Public policies for fostering entrepreneurship: A European perspective*. London, New York: Springer.

Li, X. (2011). Sources of external technology, absorptive capacity, and innovation capability in Chinese state-owned high-tech enterprises. *World Development, 39*(7), 1240–1248.

Lundvall, B.-A. (1985). Product innovation and user-producer interaction. *Industrial Research*, No. 31, Aalborg: Aalborg University Press.

Lundvall, B.-A. (1988). Innovation as an interactive process: From user-producer interaction to the national system of innovation. In G. Dosi, C. Freeman, R. Nelson, G. Silverberg, & L. Soete (Eds.), *Technical Change and Economic Theory, Chap. 17* (pp. 349–369). London: Pinter.

Lundvall, B.-A. (1992). *National systems of innovation: Towards a theory of innovation and interactive learning*. London: Pinter.
Lundvall, B.-A. (2007). National innovation systems—Analytical concept and development tool. *Industry and Innovation, 14*(1), 95–119.
Luo, Y. (2004). *Coopetition in international business*. Denmark: Copenhagen Business School Press.
Luo, X., & Slotegraaf, R. (2006). Cross-functional coopetition: The simultaneous role of cooperation and competition within firms. *Journal of Marketing, 70*(2), 67–80.
Luo, X., Rindfleisch, A., & Tse, D. K. (2007). Working with rivals: The impact of competitor alliances in financial performance. *Journal of Marketing Research, 44*, 73–83.
Ma, Z., & Lee, Y. (2008). Patent application and technological collaboration in inventive activities: 1980–2005. *Technovation, 28*(6), 379–390.
Macdonald, S. (2004). When means become ends: Considering the impact of patent strategy on innovation. *Information Economics and Policy, 16*(1), 135–158.
McCullagh, P., & Nelder, J. (1989). *Generalized linear models* (2nd ed.). London: Chapman & Hall.
McGahan, A. (2004). *How industries evolve: Principles for achieving and sustaining superior performance*. Boston. MA: Harvard Business School Press.
Mogee, M. (1991). Using patent data for technology analysis and planning. *Research Technology Management, 34*(4), 43–49.
Mowery, D. (1983). The relationship between intrafirm and contractual forms of industrial research in American manufacturing 1900–1940. *Explorations in Economic History, 20*(4), 351–374.
Mowery, D., & Rosenberg, N. (1979). The influence of market demand upon innovation: A critical review of some empirical studies. *Research Policy, 8*(2), 102–153.
Nelson, R. (1993). *National systems of innovation: A comparative analysis*. Oxford: Oxford University Press.
Nelson, R., & Winter, S. (1977). In search of a useful theory of innovation. *Research Policy, 6*(1), 36–76.
Nemet, G. (2009). Demand pull, technology push, and government-led incentives for non-incremental technical change. *Research Policy, 38*(5), 700–709.
Nieto, M., & Santamaria, L. (2007). The importance of diverse collaborative networks for the novelty of product innovation. *Technovation, 27*(6–7), 367–377.
Padula, G., & Dagnino, G. (2007). Untangling the rise of coopetition. *International Studies and Management & Organization, 37*(2), 32–52.
Park, S., & Russo, M. (1996). When competition eclipses cooperation: An event history analysis of joint venture failure. *Management Science, 42*(6), 875–890.
Pavitt, K. (1984). Sectoral patterns of technological change: Towards a taxonomy and theory. *Research Policy, 13*(6), 343–373.
Pavitt, K. (1998). Technologies, products and organization in the innovating firm: What Adam Smith tells us and Joseph Schumpeter doesn't. *Industrial and Corporate Change, 7*(3), 433–452.
Peneder, M. (2002). Intangible investment and human resources. *Journal of Evolutionary Economics, 12*(1), 107–134.
Quintana-Garcia, C., & Benavides-Velasco, C. A. (2004). Cooperation, competition, and innovative capability: A panel data of European dedicated biotechnology firms. *Technovation, 24*(12), 927–938.
Rammer, C. (2002). Patente Und Marken Als Schutzmechanismen Für Innovationen. *Studien Zum Deutschen Innovationssystem*, 11-2003, Mannheim.
Ritala, P., & Hurmelinna-Laukkanen, P. (2009). What's in it for me? Creating and appropriating value in innovation-related coopetition. *Technovation, 29*(12), 819–828.
Rosenberg, N. (1974). Science, invention and economic growth. *The Economic Journal, 84*(333), 90–108.

Rosenberg, N. (1976). The direction of technological change: Inducement mechanisms and focusing devices. *Economic Development and Cultural Change, 18*, 1–24.

Rosenberg, N. (1994). *Exploring the black box: Technology, economics and history*. Cambridge: Cambridge University Press.

Rosenberg, N., & Birdzell Jr., L. E. (1990). Science, technology and the Western miracle. *Scientific American, 263*(5), 42–54.

Rothaermel, F., & Alexandre, M. (2009). Ambidexterity in technology sourcing: The moderating role of absorptive capacity. *Organization Science Archive, 20*(4), 759–780.

Rothwell, W. (2002). Putting success into your succession planning. *Journal of Business Strategy, 23*(3), 32–37.

Rusko, R. (2011). Exploring the concept of coopetition: A typology for the strategic moves of the Finnish forest industry. *Industrial Marketing Management, 40*(2), 311–320.

Scherer, F. (1982). Inter-industry technology flows and productivity growth. *The Review of Economics and Statistics, 64*(4), 627–634.

Schmookler, J. (1966). *Invention and economic growth*. Cambridge, MA: Harvard University Press.

Schmookler, J. (1979). Fuentes económicas de la actividad inventiva. In N. Rosenberg (org.). Economía del cambio tecnológico. México: *Fondo de Cultura Económica*, 107–125.

Schumpeter, J. (1934). *The theory of economic development*. Cambridge, MA: Harvard Economic Studies.

Schumpeter, J. (1942). *Capitalism, socialism and democracy*. New York: Harper & Row.

Seung, P., & Russo, M. (1996). When competition eclipses cooperation: An event history analysis of joint venture failure. *Management Science, 42*(6), 875–890.

Seymour, R. (2008). Platinum group metals patent analysis and mapping. *Platinum Metals Review, 52*(4), 231–240.

Silva, M., & Leitão, J. (2009a). Reinforcing the entrepreneurial innovation capacity of industrial SMEs with networks. In J. Leitão & R. Baptista (Eds.), *Public policies for fostering entrepreneurship: A European perspective*. London, New York: Springer.

Silva, M., & Leitão, J. (2009b). Cooperation in innovation practices among firms in Portugal: Do external partners stimulate innovative advances? *International Journal of Entrepreneurship and Small Business, 7*(4), 391–404.

Simon, H. (1973). Does scientific discovery have a logic? *Philosophy of Science, 40*, 471–480.

Skyrms, B. (2004). *The stag hunt and evolution of social structure*. Cambridge: Cambridge University Press.

Smith, K. (2005). Measuring innovation. In J. Fagerberg, D. Mowery, & R. Nelson (Eds.), *The Oxford Handbook of Innovation* (pp. 148–179). Oxford: Oxford University Press.

Somaya, D. (2003). Strategic determinants of decisions not to settle patent litigation. *Strategic Management Journal, 24*(1), 17–38.

Tether, B. S. (2002). Who co-operates for innovation, and why: An empirical analysis. *Research Policy, 31*(6), 947–967.

Todorova, G., & Durisin, B. (2007). Absorptive capacity: Valuing a reconceptualization. *Academy of Management Review, 32*(3), 774–786.

Tsai, W. (2002). Social structure of coopetition within a multiunit organization, competition and intraorganizational knowledge sharing. *Organization Science, 13*(2), 179–190.

Tushman, M., & Anderson, P. (1986). Technological discontinuities and organizational environments. *Administrative Science Quarterly, 31*(3), 439–465.

Utterback, J., & Abernathy, W. (1975). A dynamic model of process and product innovation. *Omega, 3*(6), 639–656.

Vasudeva, G., & Anand, J. (2011). Unpacking absorptive capacity: A study of knowledge utilization from alliance portfolio. *The Academy of Management Journal, 54*(3), 611–623.

Vernon, R. (1966). International investment and international trade in the product cycle. *The Quarterly Journal of Economics, 80*(2), 190–207.

Walley, K. (2007). Coopetition—An introduction to the subject and an agenda for research. *International Studies of Management & Organization, 37*(2), 11–31.

Wang, X. (2010). Worldwide patent analysis and mapping of combine harvester innovation. *African Journal of Agricultural Research, 5*(24), 3493–3499.

Zahra, S., & George, G. (2002). Absorptive capacity: A review, reconceptualization, and extension. *Academy of Management Review, 27*(2), 185–203.