Infrastructure Investments

Politics, Barriers and Economic Consequences

Economic Issues, Problems and Perspectives

Gisele Ferreira Tiryaki
André Luís Mota dos Santos
Editors

NOVA

ECONOMIC ISSUES, PROBLEMS AND PERSPECTIVES

INFRASTRUCTURE INVESTMENTS

POLITICS, BARRIERS AND ECONOMIC CONSEQUENCES

Economic Issues, Problems and Perspectives

Additional books in this series can be found on Nova's website under the Series tab.

Additional e-books in this series can be found on Nova's website under the e-book tab.

ECONOMIC ISSUES, PROBLEMS AND PERSPECTIVES

INFRASTRUCTURE INVESTMENTS

POLITICS, BARRIERS AND ECONOMIC CONSEQUENCES

GISELE FERREIRA TIRYAKI
AND
ANDRÉ LUÍS MOTA DOS SANTOS
EDITORS

nova science publishers
New York

Library of Congress Cataloging-in-Publication Data

ISBN: 978-1-53610-808-8 (eBook)

Published by Nova Science Publishers, Inc. † New York

CONTENTS

PREFACE

The quality of infrastructure services is key to the production and trade of goods and services. Efficiency in energy provision, in transport, in telecommunication and in water and sewerage serivces contributes to increasing the productivity and competitiveness of an economy. In addition, efficient infrastructure services have positive spillover effects on the well being of the population, which also boosts labor productivity. Altogether, infrastructure investment tends to promote greater economic growth.

Infrastructure provision involves investment in capital intensive assets with long-term maturity and high levels of sunk costs. Thus, identifying the risk factors and designing policies conducive to such initiatives are essential. Macroeconomic stability, improved governance institutions, appropriate regulatory measures are some of the aspects which need to be addressed in order to avoid that infrastructure bottleneks compromise long-term economic growth. Political and regulatory matters are particularly relevant to the performance of infrastructure initiatives, specially to those that involve private sponsors, due to the large economies of scale and scope that are often present in most infrastructure ventures.

This book aims at providing a throughful review of the fundamental issues being lately discussed with regards to initiatives in energy, transport and water and sewerage sectors. Relying on statistical and econometric analysis, as well as on case studies and on indepth literature surveys, this text hopes to provide a diagnostic of the relevant aspects to be addressed when planning and executing long-term infrastructure investments. The contributions provided by each author to this book aimed, at last instance, to provide guidance on the best strategies to eliminate barriers to efficient infrastructure provision.

The first two chapters of the book focus on the importance of politics, institutions and regulatory measures to promote public-private partnerships. Both chapters rely mainly on data from developing countries to establish the conection between governance, regulation and private investment in infrastructure. While the first chapter provides a general review on the influence of the enabling environment for public-private partnerships in infrastructure, the second chapter uses econometric analysis to establish the link between governance institutions and business conditions for private investment in renewable energy projects.

The third chapter addresses the importance of regulatory policies to the more specific case of electricity provision in Brazil. The authors use seemingly unrelated regressions and frontier analysis to verify the adequacy of the regulatory measures adopted in Brazil to the prevailing technology of the companies operating in that country's electricity generation segment.

The fourth and fifth chapter also focus on case studies related to the energy sector in Brazil. The fourth chapter provides an indepth analysis of the recent events of the country´s oil sector, including the discovery of large off-shore oil reserves, the changes in the regulatory framework to deal with this new scenario and the political and financial problemas faced by the state-owned company responsible for tapping this bounty. The fifth chapter, in turn, considers whether government revenues raised from taxing oil exploration actitivies promote economic growth and the well-being of the population at the local level. This chapter undertakes an econometric analysis using data from several municipalities located at the state of Bahia, one of Brazil's main oil producing regions.

The next two chapters turn to the analysis of the transport sector. These chapters focus on planning and policy design that foster better transport infrastructure provision. The sixth chapter provides a throughful review of the literature on urban transport development, analyzing the empirical findings, the future trends in urban transportation and policy implications. The seventh chapter, in turn, uses optimization and machine learning to evaluate the value of infrastructure investment in railways, using freight transport from and to the Port Botany located at Sydney, Australia, as a case study.

Lastly, the eighth chapter focuses on the provision of wastewater services. The authors consider the case of sewerage provising in South Africa to base their conclusions related to the best strategies in pricing to finance wastewater investments.

In: Infrastructure Investments
Editors: Gisele Ferreira Tiryaki et al.
ISBN: 978-1-53610-792-0

Chapter 1

THE INSTITUTIONAL AND REGULATORY ENVIRONMENT FOR PUBLIC-PRIVATE PARTNERSHIPS IN INFRASTRUCTURE IN EMERGING MARKETS AND DEVELOPING ECONOMIES

Fernanda Ruiz-Nuñez and Zichao Wei*
The World Bank

ABSTRACT

The literature on Public-Private partnerships (PPP) suggests that a strong and well established regulatory and institutional framework corresponds with a successful infrastructure PPP investment environment, although the empirical evidence is still limited to case studies. Using the Economist Intelligence Unit (EIU)'s Infrascope Index that assess countries' readiness to do PPPs, and the World Bank's Private Participation in Infrastructure (PPI) database that tracks the volume of investments; this chapter shows that there is a positive and significant correlation between the quality of PPP regulations and institutions in Latin America and Caribbean (LAC) region and the average volume of PPP investments four years later. While results do not imply causality

* Corresponding author: Fernanda Ruiz-Nuñez. Email: fruiznunez@worldbank.org; zwei @worldbank.org.

and they remain limited as the sample size is relatively small, this chapter underlines the importance of using data to inform decision making.

Keywords: infrastructure, public-private parterships, regulatory and institutional framework

INTRODUCTION

Governments have long acknowledged the key role infrastructure plays in economic growth and poverty reduction. As countries face growing demand for infrastructure, Public-Private Partnerships (PPPs)[1] continue to play a crucial role in improving the efficiencies in delivering public services, one of the key elements to tackling the infrastructure gap. This becomes even more important as history shows that shifting the development, maintenance and operational risk onto the private sector bears potentially higher quality and overall better results than government provisioning.

The benefits of PPPs have been well recognized among practitioners. Under the right circumstances, PPPs can mobilize additional sources of financing for infrastructure. By subjecting assumptions to the market test of attracting private finance, PPPs can go some way to improving project selection. Countries with relatively long PPP histories have found that PPPs manage construction better than traditional government procurement, with projects coming in on time and on budget more often—typically attributed to the incentives created by the PPP structure. Finally, the longer-term investment perspective under PPP contracts can also help to ensure adequate maintenance and keep assets in a serviceable condition (World Bank, 2014).

Using the Private Participation in Infrastructure (PPI) Database, this chapter analyzes the historical trend of PPPs in infrastructure. The PPI database gathers data on investment commitments in infrastructure projects with private sector participation, including different contractual agreements. It does not distinguish pubic from private investments. Therefore, the analysis that follows includes PPP arrangements only, and excludes divestitures and merchant projects. This chapter covers the energy, transport, and water and

[1] PPP is defined as, "A long-term contract between a private party and a government entity, for providing a public asset or service, in which the private party bears significant risk and management responsibility, and remuneration is linked to performance." It includes brownfield and greenfield projects as well as performance-based management contracts. Source: World Bank. 2014. "Public-private partnerships: reference guide version 2.0" Washington, DC: World Bank Group.

sanitation sectors, and excludes the telecom sector. All monetary values are expressed in US dollars at 2015 prices (adjusted by the US Consumer Price Index).

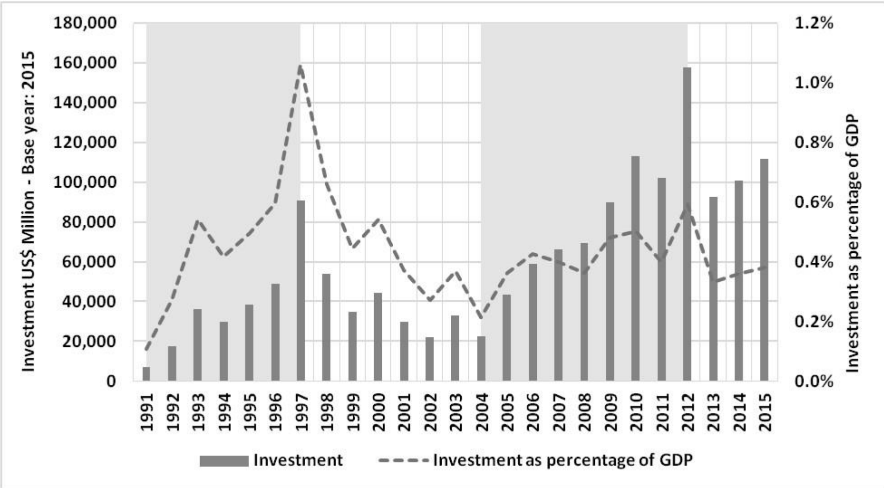
While investment in infrastructure remain predominantly dominated by the public sector, the role of the private sector has been growing over time. The use of the PPP models in the 1990s was quickly adopted and rather notable as investment in infrastructure in developing countries surged.

From 1991 to 2015, investment in PPP infrastructure projects has been ambitious as developing countries pushed foreword with the construction of roads, bridges, light and heavy rail, airports, power plants and energy and water distribution networks. Over these 25 years, investment commitments totaled US$1.5 trillion in over 5,000 infrastructure projects in 121 low- and middle-income countries[2]. During that period, there were two notable periods of expansion and one period of contraction. Strong growth took place leading up to the 1997 Asian Financial Crisis as the private sector's role in financing and delivering infrastructure services grew steadily from 16 projects totaling US$7 billion in 1991 to 230 projects totaling US$91 billion in 1997. This period of brisk activity reflected a growing global economy and structural reforms in many developing countries.

But the Asian financial crisis in 1997/1998 and the economic crisis in Argentina in 2001/2002 proved to sour investor appetite as commitments steadily declined from 1997 to 2004. At the onset of the Asian financial crisis, year-over-year investments decreased from US$37 billion in 1997 to just US$11 billion in 1998. The number of new projects also dropped from 91 to 40 year-over-year. And because East Asia and Pacific (EAP) captured a disproportionately large share of total investment (56%) among the six World Bank regions in the years previous to the crisis (1994-1997), this proved to be especially impactful on global totals.

In the case of Argentina, a brewing economic crisis in 2001/2002 further exacerbated the already vulnerable investment landscape. As a global leader up to that point, Argentina rapidly went from a massive success story—consistently commanding no less than US$3 billion annually from 1992-2000—to not reaching the US$1 billion mark in any given year since 2008.

[2] "Infrastructure" refers to energy, transport, and water projects, excluding oil and gas extraction but including natural gas transmission and distribution. "Projects," in turn, refers to brownfield concessions, greenfield projects, and management and leases, but excludes merchants, as defined by the Private Participation in Infrastructure Database (available at http://ppi.worldbank.org/resources/ppi_methodology.aspx). Finally, countries were identified as low- or middle-income as classified by the World Bank.



Source: PPI database World Bank, as of Nov 2015.

Note: Data covers projects in energy, transport, and water and sanitation that reached financial closure from 1991-2015, excluding telecom, divestitures and merchants. All investment is 2015 US$ price level.

Figure 1. Investments in PPP Infrastructure Projects 1991-2015.

The global decline from 1997 to 2004 was observed across all sectors: energy fell from US$51.8 billion to US$14.9 billion; transport declined from US$25.1 billion to US$5.4 billion; and water commitments dropped from US$14.0 billion to US$1.6 billion. Among income groups, low income countries suffered the greatest declines (-90%), followed by lower middle income countries (-85%) and upper middle income countries (-71%).

When adjusting the data by the size of the economy, the Asian financial crisis proves to be an inflection point (see Figure 1). The trend in infrastructure PPP investments as a share of GDP grew solidly all through the 1990s—from 0.1% in 1991 to 1.1% in 1997. Post-crisis, however, this figure steadily declined over the next seven years (from 1.1% to 0.2%), reflecting investment falling at a faster rate than GDP.

Nonetheless, a wave of structural reforms, favorable growth policies and a global economy that picked up steam in the mid-2000s resulted in a second growth phase that witnessed a seven-fold increase in total commitments. Indeed, this eight-year expansion from 2005 to 2012 culminated in record investment of US$158 billion. Commitments in energy grew by 414%; transport by 166%; and water by 96%.

The global financial crisis in 2008 had a much smaller impact on investment than the Asian Financial Crisis in 1997/1998. Many countries increased the share of public financing in infrastructure projects to help boost investment after the global financial crisis, particularly in infrastructure with strong MDB support. In addition, lower interest rates and higher oil prices continued to propel investment in energy, which increased over six fold from US$10 billion in 2004 to US$92 billion in 2012. Power generation projects in South Asia and, more notably in India, flourished from 2006 to 2010, growing each year before culminating in record investment of US$35 billion in 2010.

While total global investments in PPPs in absolute terms experienced a second expansion (2005-2012) with levels of investments much higher than the previous growth phase, infrastructure investments as a percent of GDP remained relatively low between 0.2% and 0.6% and it did not surpass the previous record of 1.1% in 1997.

A significant declined occurred in 2013 when total investments decreased from US$158 billion in 2012 to US$ 93 billion in 2013. Investments as percentage of GDP also dropped from 0.6% to 0.3% year-over-year, a level not seen since 2004. After 2013, investments has been growing slowly in absolute terms (7% annual growth rate compared to 27% during the 2004-2012 period) and with almost no increase as percentage of GDP.

Using the Economist Intelligence Unit (EIU)'s Infrascope Index that assess countries' readiness to do PPPs, and the World Bank's Private Participation in Infrastructure (PPI) database, this chapter aims to provide empirical evidence on the links between the quality of PPP regulation and institutional framework and investment in PPP infrastructure projects in LAC region.

The rest of the chapter is organized as follows. Section 2 presents the related literature and the motivation of the research. Section 3 discusses an empirical analysis on the correlation between PPP enabling environment performance and average volume of PPP investments. Session 4 concludes the chapter and discusses the relevant sources and research areas going foreword.

The Enabling Environment: Regulatory and Institutional Framework for PPPs

The literature on the determinants of investment confirms that macroeconomic and institutional and regulatory conditions of a country are critical for PPP markets to grow. In fact, given the high costs and risks investors face, numerous criteria must be met, particularly in Emerging Markets and Developing Economies (EMDEs) where economic and financial conditions are often more tenuous. For example, there must be peace and stability, a rule of law, good governance with accountability and transparency, clear property rights and enforceable contracts, just to name a few. Hammami et al. (2006), for example, use the World Bank PPI Database to analyze the determinants of PPI and conclude that lower levels of corruption and more effective rule of law are associated with more Public-Private Partnership projects. This study focuses on capturing the effect on the number of projects committed rather than investment levels per se. It breaks down the number of projects by sector, but not the levels, leaving room for further study, especially if we consider that bigger projects (committing more resources) may be more sensitive to the risk of the country.

Yet another key element to attracting private investment is instilling confidence in investors. One way to do this is by maintaining a stable environment in which both domestic and foreign investors can operate with limited risk and unpredictable circumstances. Moszoro et al. (2014) shows that PPI investment in infrastructure is highly sensitive to conditions such as freedom from corruption, rule of law, quality of regulations, and the number of disputes in a sector. Moreover, decreasing corruption by 10 points can increase PPI by 15%; improving rule of law by one standard deviation can increase PPI by three percent; improving quality of regulation by one standard deviation can increase PPI by four percent; and having one more project going to court decreases investment by four percent. Clearly, then, EMDEs can greatly influence private sector involvement by offering stable and reliable political conditions. When reliable political conditions suffer, investment will too.

These factors are especially important as empirical evidence confirms that country sovereign risk is negatively correlated with PPI investment in infrastructure: the riskier the country, the lower the investment by the private sector. One example of this is the length of time it takes to attract notable investment in a conflict-affected country. While it can often take six to seven years to reestablish private investment once a conflict has ceased, commitments in sectors where assets are more difficult to secure—such as water, power distribution, or roads—are often slower to come back or, in some cases, never materialize. In addition, PPI seems to be more sensitive to country risk than foreign direct investment (FDI) as shown in Figure 2 (Araya et al. 2013).

EMPIRICAL ANALYSIS

The previous discussion shows that when analyzing the drivers of PPP investments in infrastructure, the literature has concentrated on the importance of macroeconomic and institutional and regulatory conditions at the country level, but it has specifically not analyzed the importance of PPP regulations and institutions. A country could perform strongly in terms of rule of law at the national level but still have not adequate regulations to prepare and procure PPPs.

Many countries, particularly in emerging economies, have made significant efforts to improve and enable their investment environment for PPPs. Anecdotal evidence suggests that a favorable regulatory and institutional framework corresponds with a successful PPP investment environment, despite limited data on the subject. One such gauge is the Economist Intelligence Unit (EIU)'s Infrascope Index, which attempts to measure a countries' readiness and capacity for sustainable, long-term PPP projects (transport, energy and water and sewerage). By scoring various aspects of the enabling environment for doing PPPs, the index provides an assessment of a country's attractiveness to private investors.

EIU's Infrascope Index is composed by six sub-indexes, each described as follows:

1. investment climate index, which that assesses if a country has a strong business environment, characterized by stable, effective and transparent government, policy commitment to enabling PPPs, and a healthy overall market;
2. operational maturity index, which assesses if a country has a track record of implementing a strong pipeline of projects to a high standard, adhering to sound principles at all stages;
3. financial facilities index, which assesses if a country faces low government payment risk, ample sources of long-term finance for infrastructure from mature local financial markets, and effectively-targeted subsidies for low-income users of infrastructure;
4. regulatory index, which assesses if a country has a sound legal/policy basis for enabling and regulating the selection and oversight of PPPs;
5. institutional index, which assesses if a country has a clear allocation of roles and responsibilities with regard to PPP selection and monitoring, with provision for independent checks and backed up by a reliable judiciary; and,

6. subnational domestic PPP activity index, which evaluates whether infrastructure concessions can be carried out at a regional, state or municipal level, and the relative success and consistency of these frameworks.

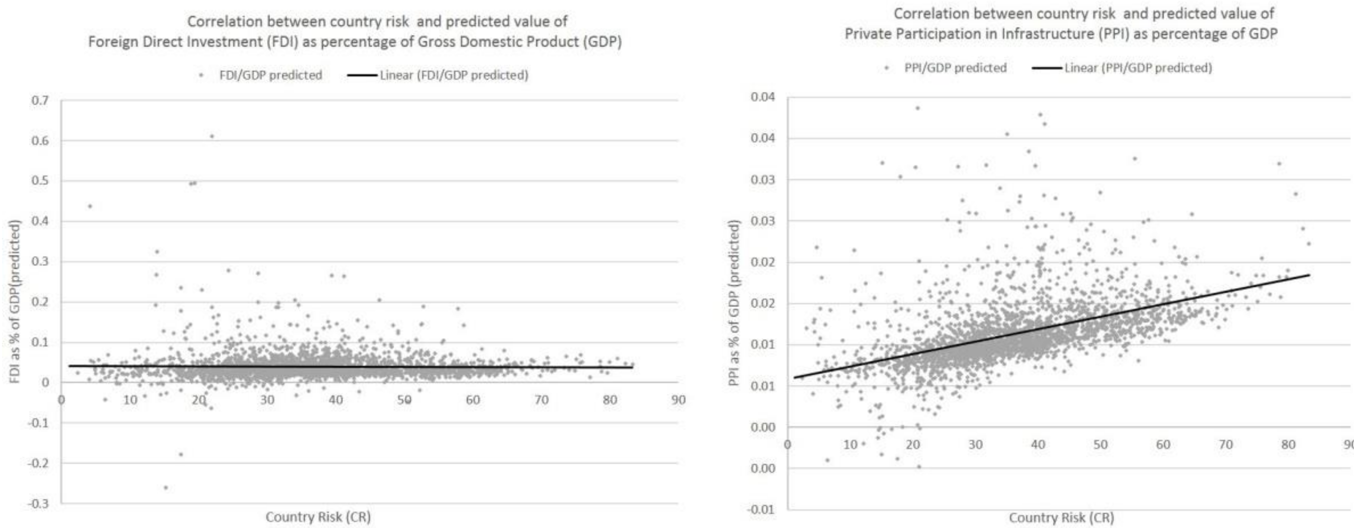
So far, eight Infrascope reports have been produced on four regions: Latin America and the Caribbean (2009, 2010, 2012 and 2014), Asia-Pacific (2012 and 2014), Eastern Europe (2012) and Africa (2015). These reports include data for 73 countries.

One of the most prominent measures of success for investment in PPP projects is a country's history of achieving financial closure. The World Bank Private Participation in Infrastructure (PPI) database presents data on commitments to over 6,400 infrastructure projects with private sector participation for the 1990-2015 period in low- and middle-income countries, covering projects in the energy, telecommunications, transport, and water and sewerage sectors. The database also allows distinguishing PPP projects from divestitures.

Using the latest Infrascope score available (LAC 2014, Asia 2014, ECA 2012 and Africa 2015) and PPI data between 2010 and 2014, we are able to correlate average PPP investment as a percent of GDP with a country's readiness to carry out PPPs, as measured by the Infrascope index. Figure 3 shows a positive relation between overall Infrascope score[3] and investments in PPP as percentage of GDP. Nascent markets (Infrascope score: 0-30) and developed markets (Infrascope score: 60-80) are more sensitive to having a robust enabling environment than emerging markets (Infrascope score: 30-60). The same pattern is observed when looking at the sub-index for investment climate.

The operational maturity and the financial facilities sub-indexes show a linear positive relation but mainly for mature and developed markets. When looking at the regulatory and institutional framework sub-indicators of the Infrascope index, there is not a clear relationship pattern with PPP investments as percentage of GDP. This could indicate a lag effect; for example, it may take several years for reforms to spur investment (preparing and financially closing a project make take a few years). This was confirmed by data from the Latin America and the Caribbean region (LAC), as shown in Figure 4.

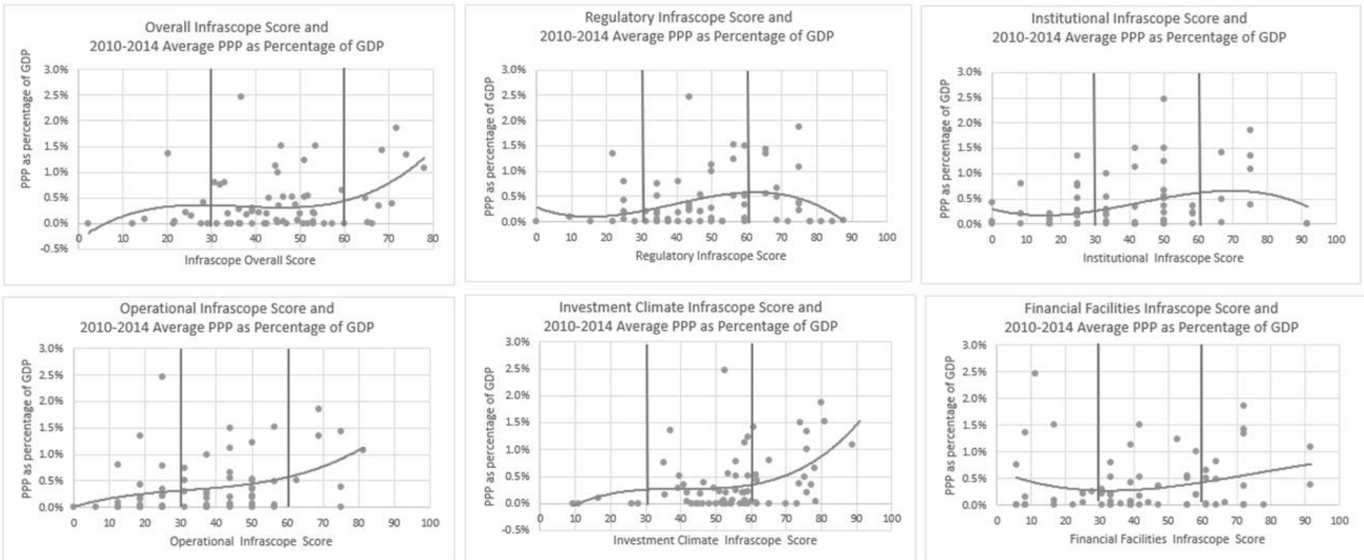
[3] The overall indicator and the operational maturity indicator of Infrascope exclude number of projects from PPI database to avoid spurious correlation.



Source: Araya, Andres and Schwatz (2013).

Note: The panel data covers 124 developing countries within 1990 and 2010. The country risk ratings were obtained from Euromoney. For the figure on the left, the predicted Foreign Direct Investment (FDI) as percentage of GDP was the result of a random effect model. The graph fits the following equation: (FDI/GDP) = α + β_1 CR + β_2 Growth + β_3 (FDI/GDP)[-1]. The graph on the right, in turn, fits the same equation but excluded eleven outliers for which PPI/GDP was greater than 0.04.

Figure 2. Correlation between Country Risk and Economic Variables.

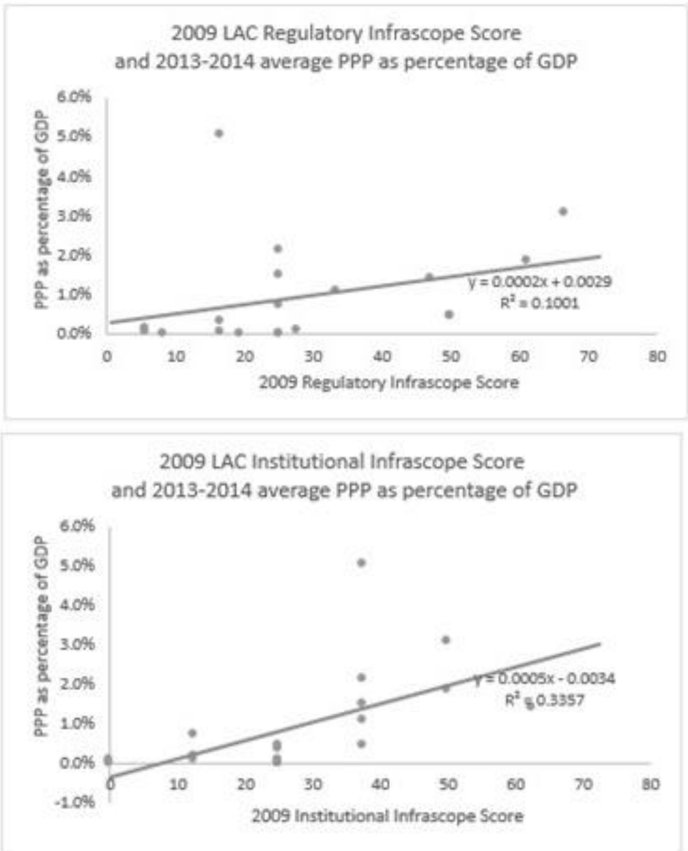


Source: PPI database World Bank PPI Database and Economist Intelligence Unit (EIU) Infrascope.

Note: Data covers projects in energy, transport, and water and sanitation that reached financial closure from 1990-2014, excluding telecom, divestitures and merchants. The operational maturity indicator of Infrascope excludes number of projects from PPI database to avoid spurious correlation.

Figure 3. Infrascope Score and 2010-2014 Average PPP Investments as Percentage of GDP.

Outside of LAC, there is not historical Infrascope data to test the lag effect of the regulatory and institutional framework on infrastructure PPP investment. As a result, assumptions must be made based on a limited sample of 19 countries in LAC. Figure 4 shows a positive and significant correlation between the quality of regulations (32%) and institutions (58%) in LAC, when using the Infrascope data for the regulatory and institutional framework sub-indicators in 2009 and the average PPP investments as percentage of GDP five years later (2013-2014).



Source: PPI database World Bank, Infrascope Economist Intelligence Unit (EIU).
Note: Data covers projects in energy, transport, and water and sanitation that reached financial closure from 1990-2014, excluding telecom, divestitures and merchants.

Figure 4. 2009 LAC Infrascope Score and PPP as Percentage of GDP.

This association does not imply causality but the graphs above provide insight on the potential importance of having a robust enabling environment. However, results remain biased as the sample size is relatively small. As such, efforts are underway to collect more frequent and reliable data as in the World Bank Benchmarking PPP procurement (World Bank, 2015).

The ongoing World Bank Benchmarking PPP procurement (BPPP) aims at assessing key aspects of a country's PPP institutional and regulatory framework and practices focusing on the different stages of the PPP procurement process. It collects comparable data on the institutional and regulatory framework as well as on the procurement process of PPPs. The BPPP is a diagnostic tool that aims to identify and capture areas for improvement in the PPP procurement process. It is also intended to help guide policy makers in the regulatory cycle by highlighting key aspects of a country's PPP legal and regulatory framework.

CONCLUSION

Developing infrastructure through PPPs presents new challenges. Without the appropriate regulatory provisions in place, PPPs could be used to bypass public financial management controls, creating undetected fiscal risks. PPPs can also be poorly selected and planned if not adequately integrated within the broader context of public investment.

In 2015, a BPPP pilot exercise (World Bank, 2015) revealed sizable disparities in how countries implement their legal framework for PPPs. After carefully examining 10 EMDEs—Cameroon, Colombia, the Arab Republic of Egypt, Ghana, Kenya, Nigeria, Peru, South Africa, Tanzania and Tunisia—, results show that while some countries espouse specific PPP laws and regulations, others have merely amended existing regulatory frameworks that allow for PPPs. Among the surveyed pilot countries, specific laws for PPPs have been adopted in Cameroon, Colombia, the Arab Republic of Egypt, Kenya, Peru, and Tanzania. Excluding the Arab Republic of Egypt and Kenya, every country's PPP laws are complemented by the broader public procurement regulatory framework.

In the Arab Republic of Egypt, the law expressly excludes the application of the public procurement law. In Kenya, too, the law abrogates provisions that include "concessioning" within the scope of the public procurement law. South Africa, on the other hand, does not have a specific PPP law. PPPs and overall public procurement are regulated by the Treasury Regulations and numerous "Practice Notes" governing both the approval and procurement process and the suggested content of the PPP contract.

Clearly, the variation in legal and regulatory frameworks shows just how difficult it is to execute a complex PPP transaction. It requires extensive legal, technical and financial knowledge, making the need for specialized skills imperative for successful implementation.

While data from the pilot program is limited, ongoing efforts to carry out the benchmarking exercise for 80 countries. This is a big step to better understanding links between institutional and regulatory frameworks and PPP infrastructure investments.

The current global landscape is not very promising. Macroeconomic factors, such as depreciating emerging market currencies and a commodities deflation have recently affected the investment environment and country-specific issues, such as the liquidity crunch coupled with land acquisition issues, have affected investment in India for the fifth straight year. Political distress has also inhibited investor sentiment in developing countries, particularly in Brazil.

The post-2015 development agenda offers an opportunity to think more broadly about development finance and the role of PPPs. One of the targets of the Sustainable Development Goals (SDGs) is to "Encourage and promote effective public, public-private and civil society partnerships, building on the experience and resourcing strategies of partnerships" (United Nations, 2012). If public private partnerships are to play a prominent role in the post-2015 development agenda, PPP investment as a percent of PPPs must increase significantly. This will require all EMDE countries to develop their PPP markets.

Significant efforts have been made by EMDE governments and MDBs to facilitate PPP investment in infrastructure, but there are still areas that could be further enhanced. As indicated by Infrascope index, there is room to improve the countries' readiness and capacity for sustainable, long-term PPP projects, particularly the institutional and regulatory framework.

References

Araya, Gonzalo, Jordan Schwartz, and Luis Andrés. 2013. "The Effects of Country Risk and Conflict on Infrastructure PPIs." *World Bank Policy Research Working Paper No 6569*. Washington, DC: World Bank.

Economist Intelligence Unit. 2014. "Evaluating the Environment for Public-Private Partnerships in Africa: 2015 Infrascope." Extracted from http://www.eiu.com/Handlers/WhitepaperHandler.ashx?fi=Africa-Infrascope-2015.pdf&mode=wp&campaignid=AfricaInfrascope2015.

Hammami, Mona, Jean-Francois Ruhashyankiko, and Etienne B. Yehoue. 2006. "Determinants of Public-Private Partnerships in Infrastructure." *IMF working paper*. WP/06/99. International Monetary Fund.

Harris, Clive, and Pratap V. Kumar. 2008. "What Drives Private Sector Exit from Infrastructure? An Analysis of the Cancellation of Private Infrastructure Projects." Washington, DC: World Bank.

Lamech, Ranjit and Saeed Kazim Saeed. 2002. "Private Power Investors in Developing Countries: Survey 2002 — Preliminary Findings." Washington, DC: Energy Forum.

MIGA (Multilateral Investment Guarantee Agency). 2013. "World Investment and Political Risk 2013." Washington, DC: MIGA, World Bank Group.

Moszoro, Marian, Gonzalo Araya, Fernanda Ruiz-Nuñez, and Jordan Schwartz. 2014. "Institutional and Political Determinants of Private Participation in Infrastructure." In Public Private Partnerships for Transport Infrastructure: Renegotiations, How to Approach Them and Economic Outcomes. Washington, DC: ITF Roundtable.

United Nations. 2012. Sustainable Development Goal 17.17. Sustainable Development Solutions Network (SDSN) http://indicators.report/targets/17-17/.

World Bank. March 2009. "What Drives Private Sector Exit from Infrastructure? Economic Crisis and Other Factors in the Cancellation of Private Infrastructure Projects in Developing Countries." *Gridlines, Note No. 46*. Washington, DC: World Bank.

World Bank. 2014. "Public-private partnerships: reference guide version 2.0" Washington, DC: World Bank Group. http://documents.worldbank.org/curated/en/2014/01/20182310/public-private-partnerships-reference-guide-version-20.

World Bank. 2015. "Benchmarking PPP Procurement: A Pilot in 10 Economies." Extracted from http://bpp.worldbank.org/~/media/WBG/BPP/Documents/Reports/BenchmarkingPPPProcurement-2015.pdf?la=enS.

In: Infrastructure Investments
Editors: Gisele Ferreira Tiryaki et al.
ISBN: 978-1-53610-792-0

Chapter 2

GOVERNANCE INSTITUTIONS AND THE INVESTMENT COSTS OF NON-CONVENTIONAL RENEWABLE ENERGY INITIATIVES IN DEVELOPING COUNTRIES

Daiana Dalla Vecchia
and Gisele Ferreira Tiryaki, PhD*
Department of Economics
Universidade Federal da Bahia
Salvador, BA, Brazil

ABSTRACT

Growing concerns about climate change, energy security, self-sufficiency and competition have fostered the use of renewable sources of energy, particularly in the last decade. Nevertheless, the investment costs of wind, solar, biomass and other non-conventional renewable sources of energy remain high, particularly in comparison to power projects which rely on conventional sources of energy. Therefore, identifying the factors that encourage such renewable energy initiatives is therefore imperative. This chapter aims at verifying the importance of governance institutions to private investment in non-conventional renewable energy projects in developing countries. Using cross-section data from private commitments

* Corresponding Author: Gisele Ferreira Tiryaki. E-mail: gtiryaki@ufba.br

to more than 1,000 initiatives in 51 developing economies during the 2002-2014 period, the results from a robust linear model indicate that strong governance institutions encourage private investment in non-conventional renewable energy projects.

Keywords: renewable energy, governance

INTRODUCTION

The availability of energy at a low cost is essential to the promotion of economic development. Yet, the growing global population and the consequent increase in energy use, particularly after the Industrial Revolution in the late eighteenth century, have brought along significant environmental costs associated with pollution, acid rain and climate changes. The world still relies heavily on non-renewable conventional energy sources but these negative environmental impacts have fostered the use of renewable resources, particularly in the past decade. The use of energy sources which are abundant, affordable and economically less harmful to the environment has contributed to overcome the ecological contradictions underlying the current economic framework.

The growing reliance on renewable energy brings about several benefits, such as higher resources' productivity, greater competitiveness due to the efficient use of energy and lower burden on land, air and water resources (see OCDE, 2011). Investment in renewable energy generates both public returns arising from the decrease in the emissions of polluting substances and private returns stemming from the market sale of energy.

In recent years, electricity generated from non-conventional renewable energy sources - wind, solar, biomass and small hydroelectric-power plants - has increasingly contributed to global energy consumption[1]. This trend has been driven by several factors, such as (i) the need for technological diversification; (ii) the search for self-sufficiency; (iii) the need to provide access to energy to people that still do not benefit from modern energy services[2]; and (iv) environmental concerns arising from the growing emission of greenhouse gases.

[1] Non-conventional renewable energy generation projects exclude large hydro power plants because of the environmental impacts of such projects.

[2] Decentralized production of electricity from biomass, biogas, solar and wind generators have grown substantially over the past years.

The transition to a sustainable system still faces challenges as renewable energy projects demand large irreversible investment commitments with long-term maturity, which contrains the availability of external financing to these initiatives. In this context, governments play a crucial role in providing support for renewable energy projects through the provision of subsidies, direct financing and other incentive schemes. The barriers faced by renewable energy initiatives are particularly significant in developing coutries, whose governments commonly face budget restrictions. According to Percoco (2014), the lack of infrastructure investment in these countries has significantly hurt their economic growth potential.

Several incentive mechanisms have been designed to encourage private investment in renewable energy generation. According to OCDE (2015), the most common incentives are feed-in tariffs[3], tax credits, tax deduction and the use of renewable energy certificates. Overall, these policies have paid off: data from the World Bank (2016a) indicates that, between 1990 and 2014, approximately 20% of the total investment in new privately-sponsored electricity generation projects in developing countries were directed to non-conventional renewable energy ventures.

Although private investment in renewable energy projects has grown significantly in recent decades, it is important to identify the barriers to such initiatives. A growing literature has emphasized the importance of institutional governance as key to promoting private investment in infrastructure and, in particular, in the energy sector. According to Tiryaki (2008), political and economic instability, lack of respect to the legal system, corruption, poor property rights' protection and inefficiency in the provision of public services likely reduce private sector participation in infrastructure projects.

Jacqmin (2015) claims, in turn, that investment costs tend to rise when institutions are weak, as procedures become more time-consuming due to administrative inefficiencies and the complexities in the regulations. Additionaly, the availability of credit reduces as the transactions between private parties are undertaken in an unstable environment with rules that lack clarity and enforcement. Renewable energy projects are especially affected by the lack of robust governance institutions, since their costs per unit of capacity are usually higher than the cost of initiatives based on conventional sources of energy. The author also points out that corruption and poorly defined property

[3] The establishment of different electricity tariffs depending on the source of energy being traded. In Brazil, for example, the public bidding for power purchase agreements with electricity companies are segmented by source of energy being promoted.

rights increase the uncertainty associated with a project's payback both in the short and in the long run.

In recent years, several studies have sought to identify the importance of the institutional framework for energy and environmental policy decisions. There has been a growing consensus that institutional governance has a significant role in reducing poverty and stimulating a sustainable development. Economies with more robust institutions are more likely to regulate and enforce environmental friendly policies. Fredriksson and Svensson (2003) and Cole (2007), for example, established the theoretical and empirical relationship between political instability and corruption with the implementation of environmental policies. The authors concluded that corruption reduces the stringency of environmental standards.

Culas (2007) and Di Vita (2009), in turn, stressed the importance of institutions that enforce property rights and a country’s legal origin for its environmental policy enforcement. Similarly, Castiglione et al. (2013) showed that countries with greater protection of property rights emit less pollutants, while Gallego et al. (2014) presented empirical evidence for the relationship between environmental performance and institutional factors, represented in particular by government initiatives aimed at controlling corruption. Therefore, the quality of institutions seems to influence the participation of society in environmental preservation and recovery processes.

Other studies have sought to establish the link between institutional aspects and the use of renewable energy sources. Radulovic (2003), for example, examined the role of the state in promoting photovoltaic solar energy use in India, while Bernard et al. (2011) analyzed the role of institutions in encouraging wind power ventures in the United States, Germany and Spain. These authors concluded that the greater the robustness of the institutional framework of a country, the more representantive the share of renewable sources in the energy matrix is.

Cadoret and Padovano (2015), in turn, analyzed the role played by political factors in the promotion of renewable energy in the European Union. Among other results, the authors showed that the quality of institutional governance contributes to a more widespread use of renewable energy sources. Jacqmin (2015) conducted a similar work analyzing the role of institutions in promoting renewable energy initiatives in the European Union and concluded that monetary stability and effective management processes are crucial for increasing the participation of renewable sources of energy.

Finally, Mehrara et al. (2015) focused on whether consumers are also more concerned with the environment in countries with more robust governance institutions. Using data from ten member countries of the Organization for Economic Cooperation and Development (OECD), the authors found that the consumption of renewable energy is higher in countries with better governance institutions.

The present chapter aims at analyzing the importance of governance institutions for private investment in non-conventional renewable energy projects. Using a large dataset of projects undertaken in 49 developing countries, the present chapter intends to contribute to the literature that focuses on the relationship between institutions and environmental quality. In particular, the study used a larger, more heterogenous sample of countries to provide evidence on the link between intitutions and environmental initiatives.

A robust linear model was estimated using cross-section data from more than 1000 privately-sponsored non-conventional renewable energy generation projects implemented between 2002 and 2014. Controlling for other possible determinants of private sector participation in such projects, the results indicate that the quality of governance institutions and the business environment positively affects the magnitude of the investment per unit of capacity in projects of non-conventional renewable energy.

Besides this introduction, this chapter includes four additional sections. The second section briefly presents an overview of the private investment in non-conventional renewable energy projects in developing countries between 2002 and 2014. The third section, in turn, describes the variables and the econometric methodology employed in the empirical analysis. The fourth section discusses the results obtained in the estimations, while the fifth section presents the final remarks.

PRIVATELY SPONSORED NON-CONVENTIONAL RENEWABLE ENERGY INITIATIVES IN DEVELPING COUNTRIES

The composition of a country's energy matrix depends largely on economic aspects, such as cost concerns and the availability of natural resources. More recently, energy planning in many countries has also focused on environmental issues. Although several developed countries have embraced renewable energy initiatives for more than a decade, investment in renewable sources of energy has also experienced a significant growth in developing countries since the second half of the 2000s.

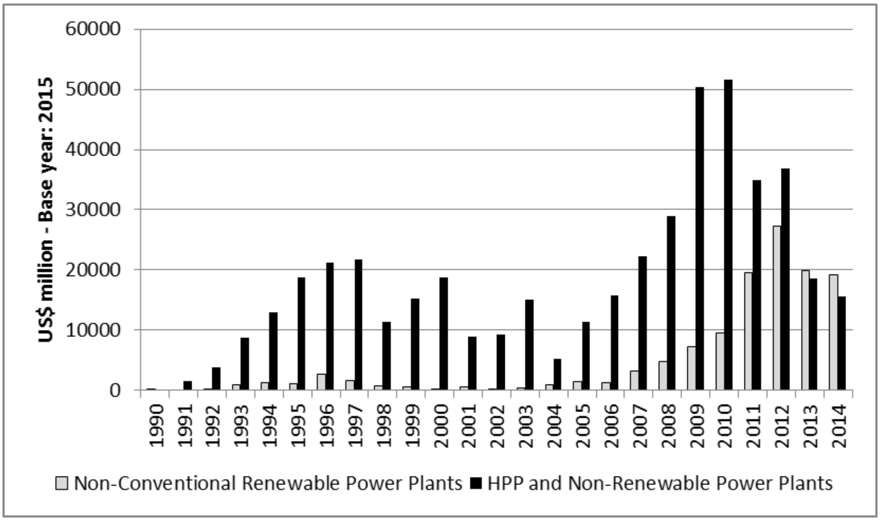
The role of the private sector in renewable energy generation projects in developing countries is particularly relevant, and several incentive mechanisms have been designed to attract private interest to such projects. In addition to bringing gains in terms of production and operation efficiency, private investment also guarantees a constant flow of investment in a context of significant government budget restrictions. Between 1990 and 2014, arround US$ 123.6 billion was invested in new non-conventional renewable energy power plants in emerging countries, or about 20% of the total invested in new generation units (see World Bank, 2016a)[4].

The share of non-conventional renewable energy generation projects has grown rapidly particularly after 2007, although investment in thermoelectric power plants based on fossil fuels still dominate (coal, diesel and natural gas fueld power plants account for nearly two-thirds of the total investment that took place in developing countries between 1990 and 2014). Since 2013, investment in greenfield projects involving solar, wind, biomass, geothermal and small hydropower plants outgrew the investment in large hydropower plants and fossil fueled power plants (see Figure 1). Nonetheless, these latter type of power plants are still more representative in terms of generating capacity.

According to Sawin (2013), this growing interest in non-conventional renewable energy projects may be the result of several factors, such as the growing interest of investors in emerging markets, the increasing demand for energy, the greater access to renewable resources and the reduction in the cost of some of these technologies.

According to the data from the World Bank (2016a), the Latin America and Caribbean region (LAC) has attracted the greatest volume of investment directed to non-conventional renewable energy generation projects: around 46% of the total investment and 45% of the new generating capacity allocated to such projects in developing countries took place in LAC (see Figure 2). Between 2007 and 2014, the flow of investments in the region grew exponentially, particularly in Brazil, which drew nearly two-thirds of the investment in new generating capacity from non-conventional renewable in LAC.

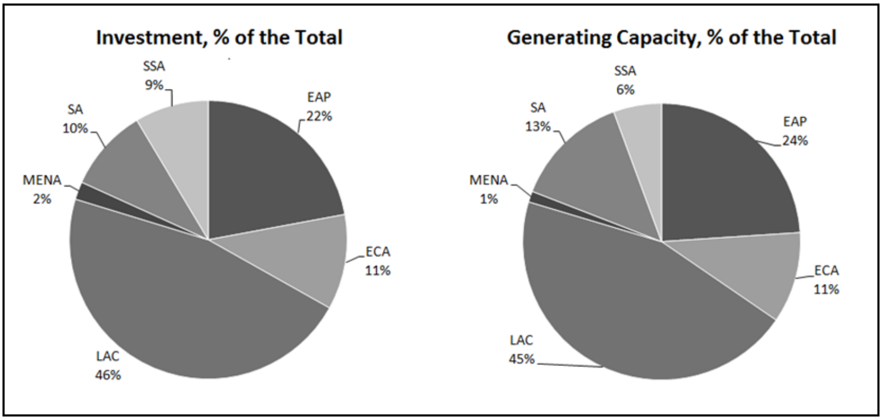
[4] The data described in this section and used in the econometric analysis was obtained from the PPI Database (2016a). The investment values refer to greenfield energy generation projects with private participation. The values were adjusted to inflation using 2015 as the base year, as described in the next section. Details of the methodology adopted by the PPI Database are available at http://ppi.worldbank.org.



Source: World Bank (2016a).

Note: HPP stands for large hydro power plants.

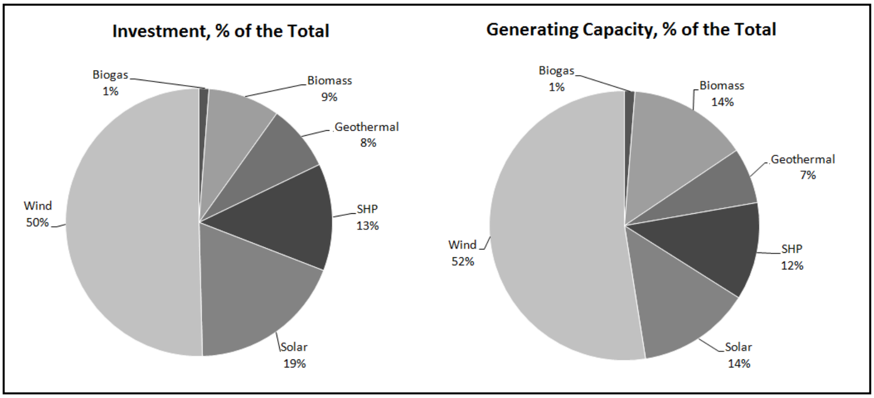
Figure 1. Private Investment in Power Generation Greenfield Projects in Developing Countries: 1990-2014.



Source: World Bank (2016a).

Note: EAP - East Asia and the Pacific; ECA - Eastern Europe and Central Asia; LAC - Latin America and the Caribbean; MENA - Middle East and North Africa; SA - South Asia; and SSA - Sub-Saharan Africa.

Figure 2. Non-Conventional Renewable Energy Greenfield Projects in Developing Countries between 1990 and 2014, by Region.



Source: World Bank (2016a).
Note: SHP stands for small hydro power plants.

Figure 3. Non-Conventional Renewable Energy Greenfield Projects in Developing Countries between 1990 and 2014, by Source.

The countries in LAC are rich in natural resources, which attracts private investment to the energy sector in the region. Additionally, the high prices of electricity in much of the region, the growing demand for energy, the problems related to energy security and the export potential of some countries have created favorable investment opportunities in the renewable energy sector of this region.

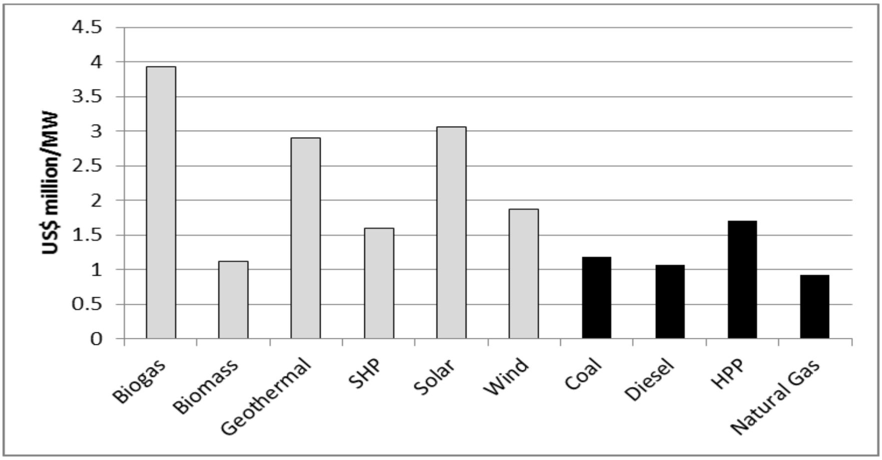
Besides LAC, the East Asia and Pacific (EAP) region has also attracted a sizeable amount of investment to non-conventional renewable energy generation projects: 22% of the total investment and 24% of the generating capacity in greenfield projects. China is the leading country as it accounts for nearly half of the total investiment in such projects in the region.

Private participation in non-conventional renewable energy projects is concentrated in the wind power segment. As shown in Figure 3, more than half of the investment and of the generating capacity of non-conventional energy initiatives in developing countries was associated with wind power plants. In recent years, solar photovoltaic power generation has also experienced a significant growth, representing 20% of the investment and 14% of the new capacity.

In the past, wind and solar technology use was considered inaccessible to developing countries due to its high costs and the lagging stage of development of these technologies compared to conventional fossil fueled and hydro power projects. Economic and regulatory incentives were required to easy the technical and economic barriers. Lately, these incentive mechanisms have become less essential, particularly in the wind segment. Despite its high initial cost, non-conventional renewable energy initiatives have low operating costs and reduced social and environmental impacts. As societies grew more concerned about the negative externalities associated with the use of fossil fuels and the development of large scale hydro power plants, these social costs have been internalized to the costs of these conventional energy ventures, increasing the competitiveness of non-conventional renewable energy projects.

Yet, initiatives aimed at reducing the high costs of renewable energy technology still need to be encouraged. The data from the World Bank (2016a) shows that biogas, solar and geothermal energy projects have the highest investment value per unit of capacity, compared to other non-conventional renewable sources (see Figure 4). Biomass fueled power plants, on the other hand, are more competitive than other non-conventional sources, with the average investment per unit of capacity comparable to the investment costs of the conventional energy generation projects.

The data also indicates that despite the leadership of LAC in terms of investment and installed capacity in greenfield projects, investment per unit of capacity in the region exceeds that of the East Asian and Pacific (EAP) region, except for biomass and biogas fueled power projects. Solar, wind and small hydro power projects have much lower costs in EAP, which reflects the competitiveness and leadership of China.



Source: World Bank (2016a).
Note: SHP - small hydro power plants; HPP - Large hydro power plants.

Figure 4. Median Investment per Unit of Capacity in Greenfield Power Generation Projects, by Technology Type.

Despite the increasing share of renewable energy generation in developing countries in recent years, there are still economic, financial, technical and institutional barriers that prevent countries from improving their performance. Better governance institutions have the potential for promoting gains in efficiency and reducing the costs associated to renewable energy projects and should therefore be encouraged.

THE ECONOMETRIC ANALYSIS

As previously stated, this chapter aims at verifing the importance of governance institutions to the private investment in non-conventional renewable energy generation in developing countries. To this end, an econometric analysis was conducted using a cross-section dataset of more than 1000 projects carried out in 49 developing countries between 2002 and 2014. The following sections describe the variables and the methodology used in the econometric analysis.

Description of the Variables

Data on annual investment and installed capacity of electricity generation projects sponsored by the private sector in developing countries were obtained from the PPI Database (World Bank, 2016a). The investment values, which were originally denominated in nominal United States dollars, were converted to real values using the United States consumer price index (CPI) with 2015 as the base year. CPI data was obtained from St. Louis FED (2016).

The natural logarithm of the real investment per unit of capacity (US$ million/MW) was used as the dependent variable, *INVMW*. For projects implemented in phases, the median investment per MW in all phases was used, while referring to the most recent year as the reference for the project implementation date.

The choice of the countries included in the sample was based on the existence of non-conventional renewable energy projects implemented between 2002 and 2014. The period of analysis in turn was determined by the availability of information[5].

[5] Although privately sponsored non-conventional renewable energy projects in developing countries date back to the early 1990s, they were implemented in very few countries. The

Institutional Governance and Business Environment Variables

There is a consensus that the existence of a robust institutional framework is crucial for the proper functioning of the market, contributing to both the economic and social progress of a country. Clear rules, respect for the law and for property rights, control of corruption, political stability and regulatory robustness help reducing the cost of the investment in infrastructure projects by decreasing the risk premium required by investors. In addition, the citizens of countries with consolidated governance institutions are likely concerned with environmental issues, which should encourage more investment in renewable energy initiatives.

Several indicators were used to measure the quality of governance institutions and the easiness of doing business in the countries under study. These variables were obtained from the Worldwide Governance Indicators (WGI) database (World Bank, 2016b) and from the Heritage Foudation Index of Economic Freedom (IEF, 2016.)

The governance indicators were estimated based on the opinions of a large number of companies, citizens and research experts interviewed in both industrialized countries and developing countries (see Kaufmann et al., 2001.) The indicators relied on more than 30 sources of individual data produced by a variety of research institutes, think tanks, non-governmental organizations, international organizations and private sector companies. Kaufmann et al. (2001) argue that the use of subjective data, which are based on the perceptions which the economic agents hold of the quality of institutions, should be favored because of the following reasons:

- The perceptions and expectations of the individuals on how a country's institutions operate guide their actions;
- Some aspects of governance lack concrete data to allow for the establishment of a quantitative metric. In these cases, an alternative to the researcher is the conversion of the individuals' perceptions of such aspects into measurable scales; and
- The use of information gathered directly from a country's legislation code to estimate the level of governance may not be appropriate, because law enforcement is often weak.

choice of 2002 as the starting year of the analysis was made because the activity started to pick up in a greater number of countries by that year.

The data used to build the governance indicators were scaled and combined using the methodology known as unobservable components modelling. The derived indicators range from -2.5 to 2.5, with higher values indicating more robust institutional governance. The WGI database includes six dimensions of governance, which are:

- Corruption Control (CONTCOR): this variable is an estimate of the extent to which the public officials use their power for private gain;
- Government Efficiency (GOVEF): this variable measures the public's perception of the effectiveness in the government's provision of public services and in the formulation and implementation of public policies, as well as the credibility on the government's commitment to such policies;
- Political Stability (POLSTAB): this indicator measures the perception of economic agents regarding the likelihood that a government will be destabilized or removed, possibly unconstitutionally or violently;
- Quality of the Regulation (QREG): this indicator captures the perception of economic agents regarding the government's ability to formulate and implement sound policies and regulations which promote private sector initiatives;
- Voice and Accountability (VOICE): this variable reflects the extent to which economic agents are able to participate in selecting their government, as well as the freedom of expression, of association and of the press; and
- Rule of Law (LAW): this variable measures the degree to which individuals trust and respect the rule of law, in particular with regards to the respect to contracts, the protection of property rights, the effectiveness of the judiciary system and the control of crime and violence.

Regarding the business environment, the present study used data from the Index of Economic Freedom (IEF). This index ranges from 0 to 100, with higher values indicating greater economic freedom. It includes information on the following aspects related to the business environment:

- Property Rights: the degree to which the laws of a country protect private property rights, and the degree to which these laws are actually enforced;

- Freedom from Corruption: the level of corruption in a country;
- Fiscal Freedom: the degree of the tax burden imposed by the government, including all forms of direct and indirect taxation of all levels of government;
- Government Spending: the level of government spending as a percentage of gross domestic product (GDP);
- Investment Freedom: the easiness of starting, operating, and closing a business;
- Labor Freedom: a measure of various aspects related to the labor market legislation restrictions;
- Monetary Freedom: a combined measure of price stability and an assessment of price controls (price stability without microeconomic intervention is ideal for a free market);
- Trade Freedom: the burden of tariff and nontariff barriers which affect both the imports and exports of goods and services;
- Investment Freedom: the degree of restrictions on capital investment flow; and
- Financial Freedom: a measure of the efficiency of the banking system and the degree of government interference with the financial sector (ideally, the government interferes minimally, there is supervision of central bank independence, and the regulation of financial institutions is limited to the enforcement of contractual obligations and the prevention of fraud).

Table 1 illustrates the cross-correlation between the various governance indicators and the index of business environment using a unbalanced panel data set of the countries under study with time series observations between 2002 and 2014[6]. The majority of the variables are significantly correlated, indicating the possibility of multicolinearity if all are included simultaneously in the econometric estimations. Factor analysis was used to avoid this problem and to reduce the number of variables used in the model. The underlying assumption of factor analysis is the existence of latent factors that influence the behavior of the variables under study.

[6] The data are available for most of the 49 countries between 2002 and 2014. Only the Index of Economic Freedom is missing for very few countries in the early years.

Table 1. Cross-Correlation

	CCORRUP	GOVEF	POLSTAB	QREG	LAW	VOICE	IEF
CCORRUP	1.000						
GOVEF	0.780	1.000					
POLSTAB	0.532	0.364	1.000				
QREG	0.715	0.832	0.336	1.000			
LAW	0.872	0.851	0.535	0.760	1.000		
VOICE	0.614	0.486	0.486	0.567	0.586	1.000	
IEF	0.681	0.674	0.305	0.847	0.666	0.555	1.000

The first step in Factor Analysis is to verify the adequacy of the sample data set. Three key aspects are considered: sample size, the degree of cross-correlation among the variables and the measure of sample adequacy. The ratio of the number of observations to the number of variables is 90 to 1, which is considered high (some researchers find that a ratio of 15:1 is suitable). Regarding the cross-correlations, the requirement is that they are mostly above 0,30, which is also a criteria met by the data used in the present study. Finally, the Kaiser Measure of Sample Adequacy (MSA), which stood at 0.848, also indicates that the data is well suited for conducting Factor Analysis (see Favero et al., 2009).

The next step in conducting Factor Analysis is factor extraction. The most commonly used method is principal factors, which was used along with parallel analysis as the method for indicating the number of factors to be retained[7]. The Factor Analysis results and the goodness of fit test results are presented on Table 2.

Since only one factor was retained, there was no need to proceed with the rotation of factors, which is the final step in a Factor Analysis. The goodness of fit tests results indicate the model was adequately specified: the RMSR remained below the 0.08 limit and the incremental adjustment indexes were above the 0.95 lower bound (see Hu and Bentler, 1999). The retained factor was named *GOVBUSENV*, and it was interpreted as a measure of both the institutional strength and the easiness of doing business of a country. The information obtained from this indicator was assigned to each project implemented in a country for a given year.

[7] See Costello and Osborn (2005) for details on the general procedures when conducting Factor Analysis.

Control Variables

Besides governance institutions and the business enviroment, other variables which likely influence the investment costs in non-conventional renewable energy projects were included in the cross-section analysis. These control variables were obtained from the World Development Indicators database (World Bank, 2016c).

Table 2. Factors Analysis:
Governance and Business Environment Indicator

Method: Principal Factor			
Number of factors: Parallel Analysis (nreps = 100. Rng = kn. Seed = 1058418958)			
Number of observations: 632 (unbalanced panel data set with 49 countries, 2002 - 2014 period)			
	Factor 1	Communality	Uniqueness
CCORRUP	0.892	0.795	0.205
GOVEF	0.876	0.768	0.232
POLSTAB	0.515	0.266	0.734
QREG	0.889	0.791	0.209
LAW	0.917	0.842	0.158
VOICE	0.665	0.442	0.558
IEF	0.810	0.656	0.344
Fator	Variance	Cummulative	
F1	4.559	4.559	
Total	4.559	4.559	
Model Adequacy Indicators			
	Estimated Model	Model of Independence	Saturated
Parameters	14	7	28
Degrees of freedom	14	21	---
Reason Parsimony	0.667	1	---
Absolute Adjustment Indexes			
Discrepancy	0.119	8.683	0.000
RMSR	0.075	0.643	0.000
Incremental Adjustment index			
Bollen Relative (RFI)	0.980		
Bentler-Bonnet Normed (NFI)	0.986		

Note: according to Hu and Bentler (1999), the optimal values for the adequacy ratios are RMSR lower than 0.08 and incremental fit indices greater than 0.95. The parsimony ratio indicates that the estimated model is 33% more efficient than the independence model (no common factors).

There is consensus among many authors that the environmental performance, which is also associated with more investment in renewable energy generation, depends on the economic performance of a country. People's concerns with sustainability is more evident in countries with higher income levels, and the consequent social pressure likely leads to more favorable local legislations towards investment in renewable energy iniciatives (e.g., lower taxes, concession of subsidies). Esty and Porter (2005), Scruggs (1999 and 2001) and Cracolici et al. (2010) argue that per capita income is positively related to environmental performance. The income level indicator used in the present analysis was the natural logarithm of the GDP per capita at constant 2005 international prices in the year the project was carried out, GDPPC.

The importance of industrial activities to a country GDP was also added as a control variable. According to Sadorsky (2013), most authors believe that industrialization leads to higher energy use. The introduction of new machinery, equipment and techniques, which increase the industrial activity, normally uses more energy than traditional agriculture or service activities. In countries with a more thriving industrial sector, demand for energy is higher, which allows for the development of larger power generation projects with the potential to benefit from economies of scale. Therefore, it is possible that power generation projects in industrialized countries involve lower costs per unit of capacity. For this reason, a control variable for the degree of industrialization in a given year was included in the econometric analysis (the value added of industrial activities as a percentage of GDP in the year of project implementation, *INDUSTRY*).

Economic stability is also an important factor for investment projects. In more stable economies, the risk premium required by investors is lower, reducing investment costs. Two variables were included as proxies for economic stability:

- Exchange rate volatility (EXCFLEX): exchange rate stability is important for non-conventional renewable energy projects, which often depend on the import of essential components to the generating unit. This variable is represented by the standard deviation of the exchange rate in the year the project was implemented.
- Annual inflation rate (INFLATION): price stability is also essential for promoting investment. This variable was measured by the percentage change in the consumer price index in the year the project was carried out.

Finally, in order to account for qualitative characteristics that may affect a project investment cost, dummy variables for the country where the project took place, for the year of implementation and for the technology type were also included. Dummy variables indexing countries were added to control for unobservable characteristics of the countries where the projects were undertaken; dummy variables indexing time were used to control for cyclical variations that simultaneously affect projects in a given period; and technology dummy variables capture the cost differentials associated with different types of non-conventional renewable projects (SHP, biomass, biogas, solar, wind and geothermal).

The Econometric Methodology

The econometric model was estimated using a sample of 1024 non-conventional renewable energy generation projects, which were implemented in 49 developing countries between 2002 and 2014. The following equation was estimated:

$$\boldsymbol{INVMW = \beta_0 + \beta_1 GOVBUSENV + \beta_2 X + \gamma D_T + \delta D_C + \varphi D_Y + \varepsilon}$$

where X is the matrix of control variables (GDPPC, INDUSTRY, EXCFLEX and INFLATION), D_T is the matrix of dummy variable representing the technologies (SHP, biomass, wind, solar and biogas), D_C is the matrix of dummy variables representing the countries, and D_Y is the matrix of dummy variables representing the year of the projects' implementation[8].

The model was initially estimated using ordinary least squares (OLS) controlling for heteroskedasticity, which is common in cross-section analysis. However, the presence of outliers detected by influence statistics indicated the need to use robust least squares, RLS (see Figure 5). Rousseeuw and Leroy (2005) and Johnson (2006) provide a detailed discussion on the advantages of conducting robust estimations instead of merely excluding the outliers. The exclusion of these extreme values would be arbitrary and could exclude relevant information of the population characteristics.

[8] To avoid singularity, the dummy variable representing geothermal projects was excluded from D_T, the dummy representing Indonesia was excluded from D_P, and the dummy representing the year of 2014 was excluded from D_A.

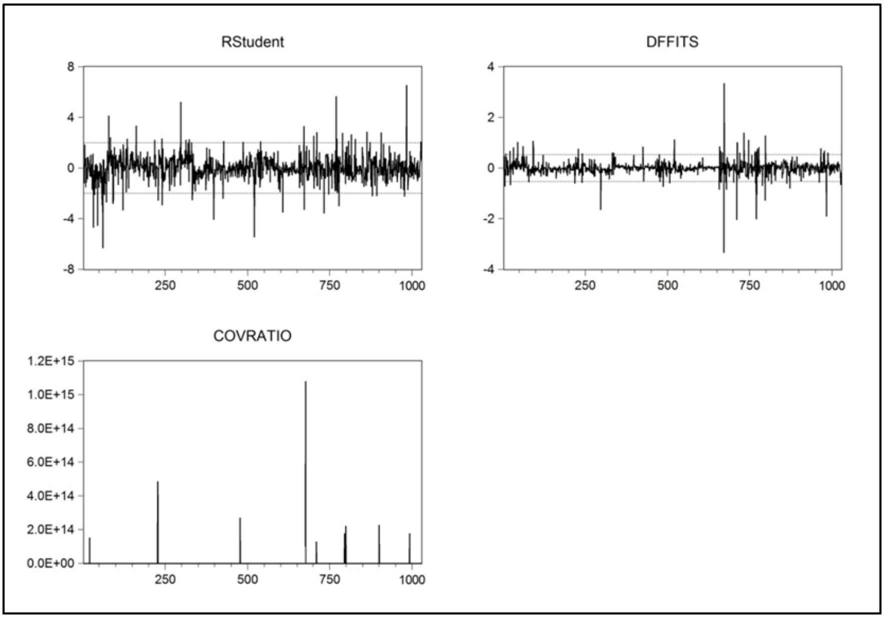


Figure 5. Influence of Statistics.

According to Rousseeuw and Leroy (2005), the robust model can be estimated by correcting for the presence of outliers in the dependent variable (type M estimation), for the presence of outliers in the independent variables (type S estimation) or the presence of outliers in the dependent and independent variables (type MM estimation). The presence of independent binary dummy variables renders the estimation of type S and type MM models impossible. Therefore, the M-type estimation proposed by Huber (1973) was used in the present study.

RESULTS

The results of the econometric analysis are presented on Table 3. For reasons of space, the coefficients of D_P and D_A (the vectors δ and φ, respectively) were not displayed, although the relevant results are discussed below[9]. The coefficient of GOVBUSENV is statistically significant at a level of 10% when using OLS, but its statistical relevance rises to 1% once the

[9] Full results can be requested from authors.

results are obtained from the RLS estimation. The sign of the coefficient, however, is opposite to the expected: more robust governance institutions and business environment is associated with greater the investment value per unit capacity.

The inclusion of the technology and country dummy variables may be the reason for this unexpected result. It is possible that, when controlling for the impact of different technologies and the specific characteristics of a country, the actual relationship being estimated between INVMW and GOVBUSENV is indicating the willingness of private sponsors to invest per MW. One possible way to verify this hypothesis would be to use a panel data set of countries, instead of cross-section data of projects[10].

Table 3. Econometric Analysis: Results

OLS			RLS		
Variable	Coeficient	Estat. t	Variable	Coeficient	Estat. z
C	0.974	0.476	C	-0.199	-0.164
GOVBUSENV	0.262*	1.768	GOVBUSENV	0.303***	2.847
GDPPC	-0.042	-0.160	GDPPC	0.144	0.928
EXCFLEX	0.0002	1.378	EXCFLEX	0.0002	1.039
INDUSTRY	0.011	0.854	INDUSTRY	0.006	0.645
INFLATION	-0.015**	-2.005	INFLATION	-0.007	-1.457
D_SOLAR	-0.003	-0.019	D_SOLAR	-0.026	-0.203
D_WIND	-0.444***	-3.291	D_WIND	-0.489***	-3.811
D_BIOGAS	0.234	1.472	D_BIOGAS	0.330**	2.287
D_BIOMASS	-0.709***	-4.944	D_BIOMASS	-0.677***	-5.188
D_HYDRO	-0.386***	-2.826	D_HYDRO	-0.429***	-3.375
R^2	0.443		R^2	0.433	
Adj. R^2	0.402		Rw^2	0.650	
F-Stat	10.838		Rn^2 Stat.	1278.347	
Prob. F-Stat.	0.000		Prob. Rn^2 Stat.	0.000	
#Obs.	1024		#Obs.	1024	

Note: *, ** and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively; settings used for the estimation of type M: Weight = Bisquare, line = 4.685, range = MAD (centralized median (standard errors and covariance of Huber Type I).

[10] Tiryaki (2008) uses this approach when evaluating the impact of institutions on infrastructure investment involving private sponsors.

In addition to being statistically significant, the coefficient of GOVBUSENV is also economically relevant. Using data from the robust estimation, consider a biomass project implemented in Brazil in 2013, with an investment of US$ 2.1 million/MW. If GOVBUSENV were 10% higher that year, the volume of investment per MW would be 12.8% greater.

Except for the dummy variables, the other control variables included were not statistically significant. The coefficients of the technology dummy variables were statistically relevant, particularly those representing biomass, wind and small hydropower projects. The results show that projects with these technologies involve less investment per MW.

Regarding the country dummy variables, only the coefficients of the variables indexing the projects undertaken in Albania, Bangladesh, Cambodia, Chile, China, Malaysia, Nicaragua, Sri Lanka, Thailand and Vietnam showed statistical significance. The signs of the coefficients were all negative, except for Bangladesh and Cambodia. It is also noteworthy that the magnitude of the coefficient for Chile was significant, indicating that projects in this country attracted relatively less investment per MW.

Finally, the coefficients of time dummy variables were all positive and only showed statistical significance after 2007. This result may be capturing the increase in investment in non-conventional renewable generation projects observed after the second half of the 2000s, as discussed previously.

CONCLUSION

The importance of governance institutions and the business environment for growth and economic development as well as for more environmental friendly energy policy decisions has been studied by a growing literature. The empirical analysis conducted in this chapter showed that increased institutional development also boosts investment in privately sponsored non-conventional renewable energy projects in developing countries.

Renewable energy projects require a significant volume of resources and involve irreversible investments with long-term maturity. A stable institutional environment affect investment costs by reducing the risk premium required by investors and encouraging them to raise the level of investments committed to renewable energy projects.

The econometric analysis indicated that the coefficient of the governance institutions and business environment variable was statistically significant and economically relevant. Yet, its positive sign seems puzzling, as it was expected that more robust institutions and business environment would imply lower investment costs per unit of capacity. This result may have arisen because the variable INVMW could be proxing for the investor willingness to invest per MW instead of investment cost per unit of capacity. An empirical evaluation using country panel data may be more effective in identifying the direct relationship between institutional governance and business environment and the cost of non-conventional renewable energy generation projects.

In the context of government budget constraints, which is usually seen in developing countries, encouraging the involvement of the private sector in the provision of infrastructure may bring greater operational efficiency. This is particularly important for non-conventional renewable energy projects, which normally involves advanced technologies and high costs. The promotion of renewable energy has become increasingly important due to the need for technological diversification, energy security and for environmental concerns. Additionally, a more widespread use of non-conventional renewable energy may impact on productivity, costs and competitiveness of the economies and should be considered crucial to promoting sustainable economic development.

References

Bernard, Michelle; Craig, Michael.; Sened, Itai. 2011. "The role of institutions in the implementation of wind energy." In *Crosssector leadership for the green economy*, edited by Marcus, Alfred., Shrivastava, Paul., Sharma, Sanjav. and Pogutz, Stefano, 3-18. New York: Palgrave-MacMillan.

Cadoret, Isabelle.; Padovano, Fabio. 2015. "The Political Economy of Renewable Energies." Working Paper 2015-01, 1-13, Centre Condorcet for Political Economy, Université de Rennes1, France.

Castiglione Concetta.; Infante Davide; Smirnova Janna. 2013. "Institutional enforcement, environmental quality and economic development." A panel VAR approach. In: *Proceedings of the IV CICSE Conference on Structural Change, Dynamics and Economic Growth*, 12-14. Livorno, Italy.

Cole, Matthew A. 2007. "Corruption, Income and the Environment: an Empirical Analysis," *Ecological Economics*, 62(4): 637–647.

Costello, Anna B.; Osborn, Jason W. 2005. Best Practices in Exploratory Factor Analysis: Four Recommendations for Getting the Most From Your Analysis. *Practical Assessment Research and Evaluation*, 10(7): 1 - 9.

Cracolici, Maria. F.; Cuffaro, Miranda.; Nijkamp, Peter. 2010. "The measurement of economic, social and environmental performance of countries: A novel approach." *Soc. Ind. Res*. 95, 339–356.

Culas, Richard. J. 2007. "Deforestation and the Environmental Kuznets Curve: an Institutional Perspective," *Ecological Economics*, 61(2-3): 429–437.

Di Vita, Giuseppe. 2009. "Legal Families and Environmental Protection: Is There a Causal Relationship?" *Journal of Policy Modeling*, 31(5): 694–707.

Esty, Daniel. C.; Porter, Michael. E. 2005. "National environmental performance: An empirical analysis of policy results and determinants." *Environ. Dev. Econ.* 10, 391–434.

Fávero, Paulo. L.; Belfiore, Patrícia. P.; Silva, Fabiana. L.; Chan, Betty. L. 2009. *Análise de dados: modelagem multivariada para tomada de decisões*. Rio de Janeiro: Elsevier.

Fredriksson, Per.; Svensson, Jakob. 2003. "Political instability, corruption and policy formation: the case of environmental policy." *Journal of Public Economics*, 87, 1383–1405.

Gallego, A. Isabel.; Galindo V. Maria. P.; Rodríguez R. Miguel. 2014. "Environmental Performance in Countries Worldwide: Determinant Factors and Multivariate Analysis." *Sustainability*, 6, 7807-7832.

Hu, Li-Tze; Bentler, Peter. M. 1999. "Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives." Structural Equation Modeling: A Multidisciplinary Journal, vol. 6, no. 1, p. 1 - 55.

Huber, Peter. J. 1973. *Robust regression: Asymptotics, conjectures and Monte Carlo*. *Annals of Statistics*, vol. 1, p. 799-821.

IEF. 2016. "2016 Index of Economic Freedom." The Wall Street Journal and The Heritage Foundation. Accessed May 20, 2016. http://www.heritage.org/index/.

Jacqmin, Julien. 2015. "The role of market-oriented institutions in the deployment of renewable energies: evidences from Europe." Available at SSRN 2690169.

Johnson, Paul. 2006. "The Hat Matrix and Regression Diagnostics." University of Kansas, Manuscript, p. 1-14.

Kaufmann, Daniel; Kraay, Aart; Mastruzzi, Massimo. 2011. "The worldwide governance indicators: methodology and analytical issues." *Hague Journal on the Rule of Law*, v. 3, n. 02, p. 220-246.

Mehrara, M. S.; Rezaei, D. Razi, H. 2015. "Determinants of Renewable Energy Consumption among ECO Countries; Based on Bayesian Model

Averaging and Weighted-Average Least Square," *International Letters of Social and Humanistic Sciences*, Vol. 54, pp. 96-109.

OECD. 2015. "Policy Guidance for investment in clean energy infrastructure: expanding access to clean energy for green growth and development." Accessed March 28, 2016. http://www.oecd.org/environment/policy-guidance-for-investment-in-clean-energy-infrastructure-9789264212664-en.html.

OECD. 2011. "OECD green growth studies: Energy." OECD publishing. Accessed May 20, 2016. http://www.oecd.org/greengrowth/greening-energy/49157219.pdf.

Percoco, Marco. 2014. "Quality of institutions and private participation in transport infrastructure investment: Evidence from developing countries." *Transportation Research Part A: Policy and Practice*, Volume 70, Pages 50-58.

Radulovic, Verena. 2003. "Institutions matter, but are they sufficient for enhancing all markets? Examining the role of the state in promoting off-grid PV technology for productive uses in Punjab, India." Master's thesis, London School of Economics.

Rousseeuw, Peter J.; LEROY, Annick M. 2005. "Robust regression and outlier detection." New York: John Wiley and Sons.

Sadorsky, P. 2013. "Do urbanization and industrialization affect energy intensity in developing countries?." *Energy Economics*, 37, 52-59.

Sawin, Janet. L. 2013."Renewables 2013, Global Status Report." Renewable Energy Policy Network for the 21st Century, Paris, Tech. Rep.

Scruggs, Lyle. 1999. "Institutions and environmental performance in seventeen western democracies." *Br. J. Polit. Sci*. 29, 1–13.

Scruggs, Lyle. 2001. "Is there really a link between neo-corporative and environmental performance? Updated evidence and new data for the 1980s and 1990s." *Br. J. Polit. Sci*. 31, 686–692.

St. Louis FED. FRED Economic Data. St. Louis: Federal Reserve Bank of St. Louis, 2016. Accessed March 28, 2016. https://research.stlouisfed.org/fred2/.

Tiryaki, Gisele F. 2008. "Desenvolvimento institucional e o envolvimento do setor privado na provisão de infra-estrutura." *Economia Aplicada* 12(3), 499-525.

Tiryaki, Gisele F. 2009. "Aspectos de governança, ambiente para negócios e o investimento privado no setor de energia de países em desenvolvimento." *Revista Brasileira de Energia* 14(2), 27–45.

World Bank. 2016a. "PPI Database." Washington. Accessed March 8, 2016. http://ppi.worldbank.org.

World Bank. 2016b. "Governance indicators database." Washington. Accessed March 8, 2016.http://info.worldbank.org/governance/wgi/index.asp.

World Bank. 2016c. "World development indicators." Washington. Accessed March 8, 2016. http://databank.worldbank.org/ddp/home.do?Step=12&id=4&CNO=2.

In: Infrastructure Investments
Editors: Gisele Ferreira Tiryaki et al.
ISBN: 978-1-53610-792-0

Chapter 3

ELECTRIC POWER GENERATION IN BRAZIL: EVALUATING THE REGULATION MODEL'S BASIC ASSUMPTION

***Mauricio Marins Machado, PhD*[1,*]**
***and Maria da Conceição S. de Sousa, PhD*[1,2]**

[1]Department of Economics, University of Brasília, Brasília, DF, BR
[2]Department of Economics, Federal University of Paraíba,
João Pessoa, PB, BR

ABSTRACT

Like other infrastructure sectors in Brazil, the electricity industry was reshaped during the 1990s in order to increase competition and private investment. Further regulation reforms were carried out after the electricity shortage in 2001. The new incentives brought about by these reforms represented a retrogress from the previous trend towards a more competitive market. Their implicit assumption was the existence of large scale economies in the electric power generation. This assumption and its policy implications are evaluated in this study. To accomplish such task, the cost structure of electricity production is analyzed based on a panel data of 21 companies during the 2000–2010 period. Two econometric methods were estimated with the data: a simultaneous equation model with fixed effects and a stochastic cost frontier. Both methodologies do

[*] Corresponding Author Maurício Marins Machado. Email: mmmaaur@gmail.com.

not reject the hypothesis that economies of scale are a typical feature of the generation market in Brazil. Regarding the companies' institutional and technical characteristics, it is not clear whether thermoelectric power plants and vertically integrated companies have cost advantages. On the other hand, the results consistently suggests the existence of cost disadvantages for state-owned companies, which still hold the largest market share in the power generation segment, despite the efforts to encourage private investment with the reforms.

Keywords: electricity generation, economies of scale, efficiency measurement, SUR, random parameter model

INTRODUCTION

The tradeoff between free-market and regulation in the electric power generation has been analyzed since the classical works of Nerlove (1963) and Christensen and Greene (1976). They were the first to point out the gains driven by competition, and further studies confirmed such benefits (Huettner and London, 1977; Goto and Tsutsui, 2008). Based on these findings, several countries have reshaped the sector by promoting policies and iniciatives that encourage greater competition among firms in order to boost efficiency (Wolak, 1997; Joskow, 1997).

The aforementioned researches and policies have not ceased the debate over the tradeoff between regulated and competitive models in electric power production. There is also evidence of indivisibilities over the relevant output range, which likely restricts the efficiency gains brought about by free-market competition. In such context, the sector technology may grant advantages to natural monopolies, cutting down potential benefits from a more competitive market structure (Hisnanick and Kymn, 1999; Berry and Mixon, 1999).

In light of these evidences, it is impossible to rule out substantial constraints on the benefits arising from competition in the electric power generation field. Moreover, there are experiences of competitive policies that were restructured after imbalances and other unexpected results occurred. One of these experiences was the California electricity crisis which took place in the early 2000s. The crisis followed deregulation policies in the sector, which were adjusted after a stunning price upsurge (Borenstein et al., 2002). Similar circumstances happened in Brazil, where the electricity shortage in 2001 led to regulatory reforms that curtailed the liberalization policies carried out during the 1990s (Pinto et al., 2007).

A relevant issue in this debate is the presence of economies of scale in the electric power production. If scale economies are significant, lower gains may be achieved by free-market incentives. Therefore, the tradeoff in question can be addressed by appraising the production levels that exhaust economies of scale in the electricity generation activity. It is mainly an empirical analysis, given the regional and temporal specificities that may affect the scale economies' magnitude.

This chapter contributes to the aforementioned discussion by investigating the degree of economies of scale in Brazil’s electricity generation segment. Examining this issue is particularly critical when evaluating the current regulatory model in the industry. The cost structure of electricity generation is analyzed based on a panel data of 21 companies during the 2000–2010 period. Two econometric methods were estimated with this data, both representing translog cost functions that impose no restrictions on the production technology. First, a simultaneous equation model with fixed effects was estimated based on the work of Machado et al. (2016); the main findings are reproduced here along with few new results. The second estimated method was a stochastic frontier that controls for potential bias due to aggregating inefficiency and other unobserved phenomena. Random coefficients were also estimated to control for possible heterogeneity bias related to observable firms’ institutional and technical characteristics. Hence, the stochastic frontier model adds two piece of information: i) time-varying efficiency measures; and ii) cost advantages related to observable firms’ characteristics.

The chapter has seven sections, besides this introduction. The second section reviews the electricity regulatory framework in Brazil, while the third section describes the econometric models used in the analysis of the cost structures of electricity generating utilities. The fourth section details the data and the variables used in the empirical exercises. The fifth section, in turn, explains the results of the simultaneous equation model, whereas the sixth section discusses the results of the stochastic frontier model. At last, the seventh section summarizes the main conclusions.

ELECTRIC POWER GENERATION MARKET IN BRAZIL

The Brazilian electricity producers engage in nationwide competition, made possible by the extension of the electricity transmission system (Trindade, 2012). Although dominated by hydroelectric power plants, which accounted for 77% of the installed capacity in 2010, an thermoelectric companies have also shown an increasing role, being responsible for 19% of the aforementioned capacity. The remaining 4% of the capacity is provided by other technologies, such as wind power, whose market share has also increased in recent years (Empresa de Pesquisa Energética, 2011). Due to such technology diversity, utilities with different sizes coexist in this industry.

Following the international trends, the electricity industry in Brazil has substantially changed over the last two decades. The industry restructuring, launched in the mid-1990s, included an extensive privatization program. Besides attracting private investment, the reform intended to foster competition and to break-up vertically integrated utilities (Santana and Oliveira, 1999; Ramos-Real et al., 2009). The sector reform underlying hypothesis was that the gains from competition would outweigh the losses from deverticalization and from relinquised economies of scale.

The incentives to competition and “unbundling” demanded regulation and coordination mechanisms that were provided by two institutions: the Electricity National Agency (ANEEL) and the National Electricity System Operator (ONS). The former is the regulatory agency responsible for mediating conflicts between the agents in the industry, preventing and charging fines to opportunistic behaviors and overseeing the several contracts between the federal govnernment and the distribution, transmission and generation utilities. The ONS coordinates supply and demand through the interconnected national electricity transmission network. The power suppliers cannot a priori deliver their production to a given customer in this network; they control only the electricity they add to the transmission system; likewise, a consumer cannot choose a specific supplier, demanding only the energy from the network. When matching demand and supply, the ONS identifies lower cost producers, directing demand instantaneously towards those suppliers. Whenever there is a mismatch between previously established contracts between producers and electricity distribution companies and large industrial consumers, financial compensations are settled in the spot market managed by the Electric Energy Trading Chamber, CCEE (Pinto et al., 2007; Carpio and Pereira, 2007).

The electricity rationing that occurred in 2001 demonstrated that previous reforms were not sufficient to adjust the power supply to a rapidly growing demand. In 2005, new regulatory measures were introduced to cope with the power shortage. In particular, two distinct electricity trading environments were established: the Regulated Contracting Chamber (ACR) and the Free Contracting Chamber (ACL).

The creation of the ACR regulated market was the main distinctive aspect of the regulatory reform and represented a step back in the move to a more competitive model pursued by the reforms which took place during the 1990s. This functioning of the ACR was inspired by the “single-buyer” model where an entity buys electricity from producers and sells it to distributors. Such coordination is carried out by the ANEEL. It registers the distribution utilities’ electricity demand and acts as a Walrasian Auctioneer, so that: i) the aggregated demand of the distribution companies – instead of the individual ones – is brought from the power suppliers; and ii) the matching between supply and demand is implemented trough public auctions. The lower-cost suppliers sign bilateral contracts with distributors from up to a five-year horizon for the physical delivery of electricity.

The segmentation into free and regulated markets brought about by the 2005 reform was based on the assumption that indivisibilities in electricity generation are significant enough to justify the aggregation of the electricity demand. This intertemporal aggregation through the regulated market may lead to welfare gains because it reduces the firm’s transaction costs. This aggregation allows that electric power producers, especially the largest ones, sign a single contract to sell their whole output instead of writing separated (and costlier) sale contracts with many buyers. Furthermore, the ACR market reduces the uncertainties faced by the producers, by assuring that they have minimum contracted out sales for up to a five-year period.

Regarding the ACL market, it deals with final electricity consumers (companies, electro-intensive manufacturers, etc.) that individually demand substantial amounts of electricity (in comparison to households and small business). The contracts established in this market are not mediated by the regulatory authority (ANEEL). The ACL works as a back up to the regulated market, filling the gaps between predicted demand and supply in the ACR market. Moreover, the ACL has two important roles: i) it provides electricity to consumers unable to access the power transmission and distribution network, and to large final consumers which are not willing to pay for the costs of these infrastructures and be subject to regulations; and ii) it is an useful signalling mechanism of electricity scarcity (for private and government agents), as the long term contracts established in the ACR may be unable to reflect sharp changes in the supply and demand balance.

METHODOLOGY

To evaluate the presence of scale economies in the Brazilian electric power generation, this chapter examines the cost structure of electricity generation companies. Such analysis is based on the estimation of a translog cost function. Because of its flexibility and convenient properties – it imposes no restrictions on the production technology – the translog functional form has been widely used to estimate cost functions. This function, which constitutes a local, second-order approximation to an arbitrary cost function, may be written as:

$$\ln C_{it} = \beta_0 + \beta_q \ln q_{it} + \sum_k \beta_k \ln p_{it}^k + \beta_T T + \frac{1}{2}\sum_k \sum_l \beta_{kl} p_{it}^k p_{it}^l + \frac{1}{2}\beta_{qq}(\ln q_{it})^2$$
$$+ \frac{1}{2}\beta_{TT}T^2 + \sum_k \beta_{kq} \ln p_{it}^k \ln q_{it} + \sum_k \beta_{kT} \ln p_{it}^k T + \beta_{qT} \ln q_{it} T + \varepsilon_{it} \quad (1)$$
$$with\ k,l = r,w,f;$$

where C_{it} is the cost of firm i in period t to produce output q_{it}; ε_{it} is the composed error term and T is a time trend. The input prices are given, respectively, by p^r_{it}, p^w_{it} and p^f_{it}; the first two represent, respectively, the prices of the production factors capital and labor; and p^f_{it} stands for fuel and water input prices, including eventually the unitary cost of electricity purchased from other utilities.

Economic theory requires that a cost function should be concave, linearly homogeneous in input prices and non-decreasing in input prices and output. These assumptions, together with the symmetry hypothesis, imply that the following restrictions should be imposed on the parameters of equation (1):

$$\sum_k \beta_k = 1; \quad \beta_{kl} = \beta_{lk}; \quad for\ all\ k,l;$$
$$\sum_k \beta_{kl} = \sum_l \beta_{lk} = \sum_k \beta_{kq} = \sum_k \beta_{kT} = 0. \quad (2)$$

The translog cost function is estimated using two econometric models: Zellner's iterative method for seemingly unrelated regressions (SUR) and a stochastic frontier model. The first model was originally estimated by Machado et al. (2016), and their main findings are reproduced here along with few new results. Its estimation requires equation (1) and the following factor share equations derived from Sheppard's lemma:

$$s_k = \frac{\partial \ln C_{it}}{\partial \ln p^k_{it}} = \beta_k + \sum_l \beta_{lk} \ln p_l + \beta_{kq} \ln q + \beta_{kT} T \tag{3}$$

The second econometric model separates the error term ε_{it} into two components: one of them, v_{it}, reflects statistical noise and the other, u_{it}, represents the time-varying firm inefficiency which will be estimated. A particular version of this approach, the random parameter stochastic frontier model, also allows the firm's technology idiosyncrasy to explain some cost function parameters besides the constant term (Greene, 2005). This version of the model is estimated, and the following equations clarify how its generalization is carried out:

$$\begin{aligned}
&\ln C_{it} = \ln c(\beta_i x_{it}) + v_{it} + u_{it}, \quad \varepsilon_{it} = v_{it} + u_{it} \\
&\beta_i = \beta + \Delta z_i + \Gamma w_i \\
&v_{it} \sim N(0, \sigma^2{}_v), \quad v_{it} \perp u_{it} \\
&u_{it} = |U_{it}|, \quad U_{it} \sim N(\mu, \sigma^2{}_u), \quad \mu > 0
\end{aligned} \tag{4}$$

where v_i has a standard normal distribution, centered at zero, with constant variance and independent identical distribution (iid). The technical inefficiency u_i is also iid with constant variance ($\sigma^2{}_u$), but it is a non-negative variable and follows a one-sided distribution (a half-normal distribution): $|N(0, \sigma^2{}_u)|$. In addition, x_{it} is the vector of ones, input prices and output; β_i stands for the vector of random parameters, which is a function of the observable and unobservable technology heterogeneity represented by z_i and w_i, respectively (w_i is usually assumed to be normally distributed, and this hypothesis is also assumed here). Besides the constant term β and the matrix Δ, the unrestricted lower triangular matrix Γ is also to be estimated. The other parameters and variables were already explained.

The parameters are estimated based on the following likelihood function (Greene, 2005):

$$L(\Lambda | y_{it}, x_{it}, z_{it}) = \int_{w_i} \prod_{t=1}^{T_i} g(\Lambda | y_{it}, x_{it}, z_{it}) h(w_i) dw_i \tag{5}$$

where Λ is the full set of parameters to be estimated; $g(\Lambda | y_{it}, x_{it}, z_{it})$ is a general density function of the observation (i,t); and $h(w_i)$ is the density function of the unobservable idiosyncrasy w_i. The integrals in equation (5) will not exist in closed form, which prevents estimating the parameters by maximizing this equation. However, the integrals may be approximated by simulation under conditions usually existing on the densities applied in the stochastic frontier models. In this case, the parameters are estimated by maximizing a simulated likelihood function. To compute the function, primitive draws are simulated from the distribution assumed for the unobservable w_i. The use of these draws and the general form of the simulated log-likelihood function are illustrated in the next equation:

$$\log L_s = \sum_i \log \frac{1}{R} \sum_{r=1}^{R} \prod_{t=1}^{T_i} g(\Lambda | y_{it}, x_{it}, z_{it}) \tag{6}$$

where R is the number of simulated sample draws. The simulation is implemented over R draws on w_i through β_i as defined by (4). Because this estimator is computed by maximizing (6) over the full set of parameters Λ, it is known as the maximum simulated likelihood estimator.

DATA AND VARIABLES

To apply the econometric models described in the previous section, the present study used data from the utilities' financial accounting reports published by the Bolsa de Valores, Mercadorias e Futuros S.A. (the Brazilian stock market), as well as the information released by the state-owned company Centrais Elétricas Brasileiras (2011). The sample consists of annual observations from 21 firms between 2000 and 2010. The data set is an unbalanced panel of 198 observations because only after 2005 regulatory rules compelled the integrated utilities to separately identify the accounts for the generation activities from those of the distribution and transmission segments in their financial statements (Law 10.848, 03/15/2004). Hence, for some of these firms, there is no comparable data before 2005. Moreover, there are some newcomer firms to the generation segment, which lack data for the early years. Finally, the financial statements are unavailable for some of the incumbents for the first years of the period under evaluation.

Regarding the variables used in the examined cost structure, labor prices (p^w) were obtained by dividing annual labor (l) costs by the number of employees; fuel and water prices (p^f) were computed by dividing annual fuel and water (f) costs by their respective output. The price of capital (p^r) was based on the weighted average cost of capital (WACC), which corresponds to the cost associated with a firm's capital structure. The two main components of the WACC are the following: i) debt (d), which is includes all the loans informed as liabilities in the “Liabilities and Shareholders' Equity” section of the balance sheets; and ii) equity (eq), which is informed as the amount of capital in the same section of the balance sheets. The WACC is computed by using the following equation:

$$wacc = \delta_d k_d + \delta_{eq} k_{eq} \tag{7}$$

where δ_i (i = d, eq) represents the relative weights of each component of the capital structure in total capitalization, and k_i (i = d, eq) is the component’s cost (amount of interest and dividends).

Table 1 shows the main descriptive statistics of the variables used in the empirical exercises, and the data that supported their computation.

Table 1. Data Summary Statistics: 2000/2010

Variables (annual data)[a]	Median	Mean	Standard error	Maximum	Minimum
Electricity Production (GWh)	11,659.82	21,288.38	23,348.12	94,344.52	339.00
Cost of Capital	10.07%	11.24%	7.37%	48.17%	0.50%
Average Wage (US$ 1,000)	71.88	77.00	37.33	397.92	14.60
Capital costs (US$ 1,000)	219,225.99	385,137.62	705,250.42	4,991,048.17	1,094.09
Labor Costs (US$ 1,000)	40,592.33	73,997.06	99,787.99	563,489.93	524.07
Fuel and Water Costs (US$ 1,000)	96,765.25	310,052.59	768,634.89	6,654,485.23	646.70
Total Cost (US$ 1,000)	383,852.33	769,187.26	1,184,047.89	7,078,474.39	6,822.57
Capital share (%)	53.77%	49.76%	27.03%	91.61%	1.32%
Labor share (%)	11.47%	16.67%	16.96%	83.05%	0.26%
Fuel and Water share (%)	24.54%	33.57%	24.68%	95.02%	3.63%

Source: Bolsa de Valores, Mercadorias e Futuros S.A. (2011), Centrais Elétricas Brasileiras (2011).

[a] Price Index: IPCA/IBGE (December 31, 2012 = 100). Exchange rate (R$/US$): 2.0435 (12/31/2012).

THE SUR RESULTS

Most of the results shown in this section were originally presented in Machado et al. (2016). Similarly to these authors, the SUR was employed to estimate the system of equations described by equations (1) and (3), taking into account the theoretical hypothesis of linear homogeneity by dividing total cost, capital and labor input prices by the price of fuel and water inputs (pf), whose share equation was excluded. Symmetry hypotheses (equation 2) were also imposed. Since total cost and most of the explanatory variables are in logarithms and have been normalized around the industry's sample median, the first order coefficients represent cost elasticities evaluated at the sample median.

Table 2 provides the estimated parameters of the translog cost function. Two models were estimated: Model 1 presents the unrestricted translog model; Model 2 differs from Model 1 by adding dummies related the 2001 electricity shortage and the regulatory reform. Both models take into account the firms' fixed effects. The core variables' coefficients for the cost model – output and input prices– are statistically significant and have the expected signs, pointing out to a well-behaved cost function which is increasing in output and input prices.

Consistent with a capital-intensive industry, the coefficient of the capital cost is larger than the one attached to wages. Values estimated for the first order coefficients, β_r and β_w, indicate that capital and labor represent, respectively, 51% and 18% of total production costs at median production. The capital share slightly decreases over the period analyzed; indeed, the coefficient $\beta_{rt} = -0.007$ is negative and significant. Moreover, the estimated value for β_{wt}, although negative, is not significant, indicating that the labor saving hypothesis is rejected in the period studied.

Fuel and water costs represent approximately one third of total generation costs as shown by the positive and statistically significant coefficient $\beta_f = 0.31$. This share is rising in the period analyzed as attested by the positive and significant coefficient $\beta_{ft} = 0.009$. The increasing use of thermoelectricity, which is relatively more fuel intensive (Marreco and Carpio, 2006), may explain this finding. For hydroelectric power companies, environmental costs (financial and bureaucratic ones), by augmenting the water costs, may constitute the reason for the increase in the share of fuel and water on total costs (Sternberg, 2010).

Table 2. SUR Cost Function Parameter Estimates

Parameters	Model 1		Model 2	
	Coefficient	D. Default	Coefficient	D. Default
βr	0.509721*	0.0112835	0.50948*	0.011284
βw	0.1828783*	0.0093898	0.182911*	0.009388
βf	0.3074007*	0.0083294	0.3076089*	0.0083261
βq	0.5344760*	0.0929164	0.538854*	0.092987
βt	-0.0463376*	0.0050993	-0.04522*	0.006082
βrr	0.0672424*	0.0040257	0.067371*	0.004025
βww	0.03349280*	0.0046616	0.033683*	0.004658
βff	0.0795013*	0.0036542	0.0798016*	0.0036606
βwr	-0.0106169*	0.0038982	-0.01063*	0.003892
βwf	-0.0228759*	0.0026465	-0.0230569*	0.0026539
βrf	-0.0566254*	0.0025197	-0.0567447*	0.0025213
βqq	0.1029704*	0.0301362	0.102679*	0.03013
βtt	-0.0006176	0.0008565	-0.00045	0.000969
βwt	-0.0020677	0.0024724	-0.00208	0.002475
βrt	-0.0074966*	0.0022703	-0.00737*	0.002272
βft	0.0095643*	0.0015961	0.0094539*	0.0016074
βrq	0.0015432	0.0066501	0.002032	0.006674
βwq	-0.0470939*	0.0059276	-0.04736*	0.005937
βfq	0.0455507*	0.0051371	0.04533*	0.0051385
βtq	-0.0020214	0.0018303	-0.00211	0.001831
DQ	0.5935506*	0.1646027	0.585755*	0.164782
D2001			-0.0274152	0.0319943
D2005/2010			-0.010988	0.0285104
Constant	0.0716965**	0.0355094	0.080791**	0.041947

Source: the authors' own computations (Model 2) and Machado et al. 2016 (Model 1).
Note: levels of significance: * 1% and ** 5%.

The dummy variable DQ separates the firms into two groups: small firms producing at most 5000 GWh per year – below the output of the median firm – and those that produce at least 10,000 GWh. DQ takes a value of 1 if the firms are in the first group and zero otherwise. Because this variable interacts with the variable *lnq*, one can test whether economies of scale vary between the groups. The positive and significant coefficient attached to the variable DQ indicates that small firms exhaust their economies of scale at a lower levels of production. Further, the significant and positive coefficient *βqq* shows that scale economies are decreasing with the level of production for both groups.

Model 2 provides additional information concerning the two intercept dummies that intend to evaluate potential effects related to the electricity shortage in 2001 and the regulatory reform. The first dummy is equal to one in 2001 and the second is equal to one between 2005 and 2010, when the new regulation rules became valid. Both dummies are not significant, and the other parameters are consistent and similar to those estimated by the Model 1. Therefore, the estimated cost structure was not affected by these two events.

Economies of Scale

In Brazil, large and medium-sized hydroelectric power firms coexist with small utilities known as PCH (Pequena Central Hidroelétrica – Small Hydro Power Plants). The latter ones benefit from legal incentives as PCHs are thought of having smaller environment impacts. On the other hand, indivisibilities may arise in the power supply market due to the presence of economies of scale. Possible sources of these economies are related to environment regulations that substantially increase companies' sunk costs (such as licensing fees and project costs) as well as capital fixed costs.

Because the potential to increase the hydroelectric power production in the South and Southern regions of the country is largely exhausted, many of the new utilities have been installed in the remote Amazon region, where the environmental related expenses are higher, which contributes to raise the sunk costs (Ferreira, 2000). Besides the environmental costs, other potential sources of scale economies are the building and maintenance costs of the dam, the steel lined pressure shaft, the power house and the turbines (Filippini and Luchsinger, 2007). As these expenses are fixed capital costs, smaller utilities are thought to be in a disadvantaged position relatively to the larger ones, which operate several power plants. Moreover, the uncertainty inherent to hydroelectric power technology (related to meteorological factors for instance) may lead to higher costs for the smaller firms unable to diversify those risks among plants and locations.

Therefore, it is unclear if it is better to generate electricity in small and medium-sized firms or in a few, large utilities with many power plants, as there is no a priori indication of the magnitude of scale economies prevailing in the electric power production in Brazil. To account for these issues, the translog functions allows testing for long-run scale economies (ES), holding output and input prices constant. These economies are given by equation (8):

$$ES = 1 - \frac{\partial \ln C_{it}}{\partial \ln q_{it}} \tag{8}$$

where $\frac{\partial \ln C_{it}}{\partial \ln q_{it}}$, the cost elasticity with respect to output, is obtained by differentiating equation 1 with respect to *lnq*:

$$\frac{\partial \ln C_{it}}{\partial \ln q_{it}} = \frac{\partial C_{it}}{\partial q_{it}} \frac{q_{it}}{C_{it}} = \beta_q + \beta_{DQ} DQ + \sum_k \beta_{kq} \ln p^k{}_{it} + \beta_{qq} \ln q_{it} + \beta_{qT} T \tag{9}$$

Marginal costs for the jth firm may be computed by using equation (9):

$$\frac{\partial C_{it}}{\partial q_{jt}} = \frac{\partial \ln C_{it}}{\partial \ln q_{jt}} \frac{C_{it}}{q_{jt}} = [\beta_q + \beta_{DQ} DQ + \sum_k \beta_{kq} \ln p^k{}_{it} + \beta_{qq} \ln q_{jt} + \beta_{qT} T] \frac{C_{jt}}{q_{jt}} \tag{10}$$

From equations (8) and (10), ES may be written as:

$$ES = 1 - \beta_q - \beta_{DQ} DQ - \sum_i \beta_{iq} \ln p_i - \beta_{qq} \ln q - \beta_{tq} t \tag{11}$$

Thus, a positive value for ES denotes the presence of economies of scale; a zero value stands for constant returns to scale, whereas a negative value indicates diseconomies of scale. Notice that ES is simply ($1 - \beta_q - \beta_{DQ}$DQ) at the median value for the variables. Evaluating ES for the Brazilian power supply utilities at the median level, the cost elasticity with respect to output (EL) is given by β_q, whose estimated value is 0.5345 (Table 2). The value for ES = ($1 - \beta_q$) is 0.47 for the utilities producing above the sample median output level, indicating the presence of economies of scale for this group. Hence, increasing production would lead to a less than proportional rise in costs, thus benefitting the largest utilities.

In the sample under stury, ITAIPU’s production is above 20% of total electricity generated, which corresponds to eight times the median output. To investigate whether the economies of scale of the largest companies were driven by this outlier, we excluded ITAIPU, and used the new sample to re-estimate Model 1 (Table 2). The results of this estimation in Table 3 show that the estimated cost elasticities and scale economies are very similar to the ones estimated for the whole sample. Therefore, the significant economies of scale of the largest companies are not influenced by this outlier.

Table 3 also shows that the estimated elasticity and scale economies are quite different between the two groups being examined. The elasticity for the smallest companies in the complete sample is given by EL = ($\beta_q + \beta_{DQ}$), whose estimated value is 1.1280. The respective value for ES = 1 − ($\beta_q + \beta_{DQ}$) is −0.1280 suggesting that, at the median output and prices, this group faces diseconomies of scale.

Table 3. Cost Elasticity[a] (EL) with Respect to Output by Group and Sample, at the Median Output of Each Sample

Group	EL computation	Sample			
		Complete		Without ITAIPU	
		Coefficient	St-error	Coefficient	St-error
Firms producing below sample median output	EL = $\beta q + \beta_{DQ} DQ$, where $DQ = 1$.	1.128*	.101	1.099*	.098
Firms producing above sample median output	EL = βq, since $DQ = 0$.	.534*	.093	.507*	.089

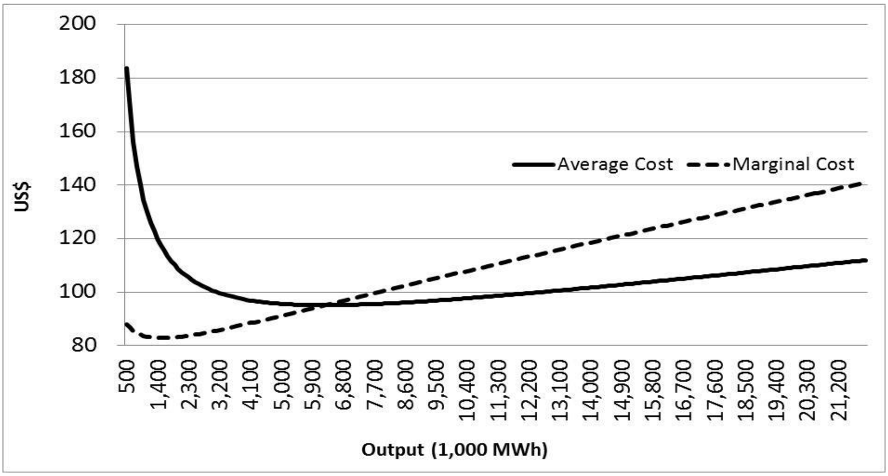
Source: Machado et al. (2016).

Note: Level of significance: * 99%.

[a] – Elasticity computed at each sample median output, price and time levels.

In addition, Figure 1 plots the marginal and average cost curves for the simulated levels of production of firms producing less than the median output. As required by the translog cost function, those curves as well as the ones shown in Figure 2 were fitted by normalizing the explanatory variables around the median of each mentioned group. The two figures are also based on the parameters estimated by Model 1 (Table 2). Up to 6000 MWh, marginal costs are lower than the average costs in Figure 1, so that unit costs are decreasing. For higher output levels, the average costs are slightly increasing, and marginal costs continue to rise, driving unit costs upwards. For output levels between 6000 and 12,000 MWh, close to the median output of the complete sample (11,700 MWh), the average cost curve is slightly increasing. For this production range, average costs lie between US$ 95 and US$ 100 per MWh, so that the hypothesis of a relatively flat average cost curve in this output range cannot be rejected.

For companies producing above the median output, the picture is rather different, as Figure 2 highlights. It shows that the average cost curve declines over the relevant output range, but remains always above the marginal cost curve. This result confirms the other results detailed in this section, which point out to the existence of indivisibilities in the Brazilian electricity generation system. Such indivisibilities and their respective economies of scale indicate the changes brought about by the 2005 reforms went in the right direction, particularly those related to the creation of the ACR market.

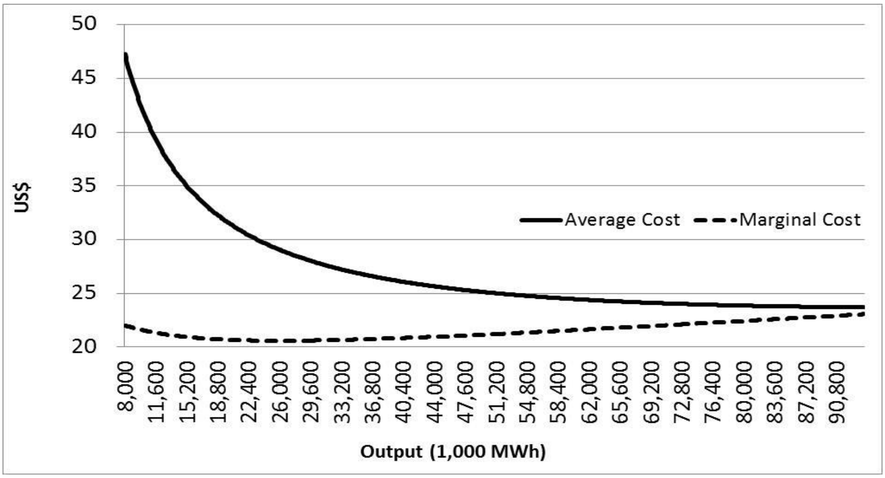


Source: Machado et al. (2016).

Note: Price Index: IPCA/IBGE (December 31, 2012 = 100). Exchange rate (R$/US$): 2.0435 (12/31/2012).

Figure 1. Average and Marginal Costs for the Group of Companies with Average Production Below Median Output.

The aggregation of electricity sales through this market tend to provide welfare gains because it reduces transaction costs for the larger utilities and, in doing so, allows companies to benefit from substantial scale economies. The reduction on transaction costs arises because sales take place throught a single contract negotiation in the ACR auctions, instead of through several searches, bargains and other business processes in the market required to sell the same output without the aforementioned intermediation mechanism. Because this single negotiation in the ACR is supervised by regulatory institutions, it has greater enforcement and lower risks.



Source: Machado et al. (2016).

Price Index: IPCA/IBGE (December 31, 2012 = 100). Exchange rate (R$/US$): 2.0435 (12/31/2012).

Figure 2. Average and Marginal Costs for the Group of Companies with Average Production Above Median Output.

It is worth mentioning that the economies (diseconomies) of scale vary with respect to the pattern of input utilization. A significant and negative (positive) value for the coefficient β_{iq} (i = w, r, f) indicates the presence of economies (diseconomies) of scale with respect to the use of that particular factor. Hence, when the power supply grows, a reduction (rise) in the input–output ratio with respect to the input is observed. Table 2 presents the estimated values for these coefficients. The coefficient β_{rq} is not statistically significant, but the coefficients β_{wq} and β_{fq} are significant (− 0.047 and 0.046, respectively), indicating that labor creates economies of scale whereas fuel generates diseconomies of scale. This result shows that the natural resources used in the electric power production, primarily water, represent a constraint to the scale economies.

STOCHASTIC FRONTIER MODEL RESULTS

The maximum simulated estimator is applied to estimate the stochastic cost frontier model given by equations (1) and (4)[1]. The random parameters are

[1] The estimation was carried out with Limdep 8.0 (Greene, 2003).

the first order terms of the translog function, whereas the others are common to all companies in the panel. It means that all the cost elasticities in the median sample are estimated by random parameters, since the variables in the model are centralized at this point.

To reduce the difference between the maximum likelihood function and its simulated version given by equation (6), the literature suggests the use of several simulated sample draws. The simulation was carried out by Halton sequences. This method provides draws that correspond to almost ten draws produced by a random number generator. The maximum likelihood simulation was implemented with 1000 Halton draws, which corresponds roughly to several thousand draws obtained by the pseudorandom numbers method (see Bhat, 1999; and Train, 2003).

Table 6 reveals the cost function parameters estimated by the stochastic frontier model. Like the SUR model, the core variables of the stochastic cost frontier have coefficients that are statistically significant and show the expected sign. Therefore, they represent once again a well-behaved cost function, which is increasing in output and input prices. Other similarities are the large capital intensity and the increasing share of expenditures related to the use of natural resources (water and fuel). The estimation also reports a significant difference between the inefficiency measurement and the statistical noise[2], which is the reason for using stochastic frontier models.

The estimated stochastic cost frontier model, whose results are shown in Table 6, additionally tests if some observable characteristics of the companies support eventual technology differences among them. Such characteristics are represented by the following dummy variables:

i. ST is equal to one if the company is state-owned, and zero otherwise;
ii. VER is equal to one whenever the firm is vertically integrated, and zero otherwise;
iii. TR is equal to one if there is a widespread use of thermal power plants, and zero otherwise.

[2] Given by the t-test for σ and λ, where: i) σ is the square root of a joint variance composed by the variance of the inefficiency (σ^2_u) and the variance of the random error (σ^2_v): $\sigma = \sqrt{\sigma_u^2 + \sigma_v^2}$; and ii) λ is the rate between the standard deviation of the inefficiency (σ_u) and the standard deviation of the random error (σ_v): $\lambda = \frac{\sigma_u}{\sigma_v}$.

Table 6. Stochastic Cost Frontier: Estimates from a Random Parameter Model with Observable Heterogeneity

Parameters	Mean		Observable heterogeneity					
			TR		VER		ST	
	Coeff	S-Err	Coeff	S-Err	Coeff	S-Err	Coeff	S-Err
β_0	-.299*	.012	-.031	.0289	-.394*	.015	.031***	.018
$\sigma_{\beta 0}$	.244*	.007						
βr	.525*	.018	.067*	.025	-.015	.019	-.151*	.016
$\sigma_{\beta r}$	.072*	.004						
βw	.243*	.018	-.238*	.0339	-.066*	.026	.083*	.024
$\sigma_{\beta w}$	.093*	.005						
βf	.232*	.025	.171*	.055	.081***	.048	.068	.046
$\sigma_{\beta f}$	0.118*	0.006						
βq	.591*	.011	.211*	.020	-.027	.017	.134*	.013
$\sigma_{\beta q}$	.116*	.003						
βt	-.072***	.004	-.029*	.006	.014*	.004	.023*	.005
$\sigma_{\beta t}$	.043*	.002						
βrr	.063*	.006						
βww	.0005	.011						
βff	.118*	.025						
βwr	.027***	.015						
βwf	-.028	.019						
βrf	-.090*	.016						
βqq	.015*	.006						
βtt	-.0004	.0004						
βwt	.003	.003						
βrt	-.013*	.002						
βft	.010*	.004						
βrq	-.013*	.006						
βwq	-.053*	.010						
βfq	.066*	.012						
βtq	.015*	.001						
σ	.116*	.003						
λ	3.15*	.361						
Parameters	Mean		Observable heterogeneity					
			TR		VER		ST	
	Coeff	S-Err	Coeff	S-Err	Coeff	S-Err	Coeff	S-Err
Log LF	118.16							

Source: the authors' own computations.

Note: levels of significance - * 99%, ** 95%, *** 90%.

ST and VER stand for institutional characteristics of the companies, whereas TR is a technology choice. Table 6 shows that all of these dummies are mostly statistically significant explanatory variables for the technology differences among the firms. For instance, a state-owned company is more labor intensive, since its cost elasticity for this input increases from 24%, the overall mean of the model, to 33% at the sample median output, price and time levels. The same happens with its output cost elasticity, which grows from 59% to 72%. Therefore, state-owned companies benefit less from economies of scale. Both results may arise due to their longer existence in the electricity generation market and their consequent reliance on older power plants (Pinto, 2007). Indeed, the test points out that, different from the private companies in the sample, less of their cost reduction is driven by technical progress.

The observable heterogeneity test also shows that the state-owned companies have substantial cost advantages in the use of capital. The cost elasticity of capital is 32%, much lower than the 52% found for their private counterparts. This advantage may indicate institutional failures of the Brazilian capital market.

Regarding the vertically integrated companies, they have cost advantages in the use of labor, given that the cost elasticity for this input decreases from 24% to 18% at the sample median output, price and time levels. This result suggests that increases in labor productivity driven by gains from specialization might not be relevant in the electricity industry. The heterogeneity test also indicates cost advantages for the vertically integrated companies, given by the constant term coefficient. Notwithstanding, such advantages are undermined by the inefficiency measurements explained in the next section.

The third test of observable idiosyncrasy is the use of thermal power technology. The test shows that thermal power plants are far more intensive in fuel usage and have much smaller economies of scale. The share of fuel in their total cost is 40% at the sample median output, price and time levels, far larger than the 23% of their non-thermal counterparts. For the economies of scale, they achieve 20% at the mentioned median point, whereas their counterparts reach 41%. These results indicate cost advantages of smaller-sized thermal power plants, confirming the fact that thermal power technology is known for having lower fixed costs but higher operational costs, primarily driven by fuel expenses, when compared to hydroelectric power technology. Therefore, the thermal power plants are more suited to smaller electric power production projects (Santana, 2005; Barroso et al., 2006).

Economies of Scale

In this section, new estimations are employed to investigate the scale economies in the Brazilian electric power generation. The analysis is based on cost elasticity computations, which are just another way to appraise scale economies, as equation (8) indicated. Table 7 presents the output cost elasticities and the average cost for each company in the sample. They are computed at the sample median input prices and time levels, applying the parameters shown on Table 6. The output cost elasticities are close to 1.00 for the firms which use thermal power technology. This result implies that they exhaust their economies of scale, or come close to it, at the median output. For the firms using primarily hydroelectric power technology, the output cost elasticity mainly declines by 12,500 GWh, and from then on it rises, pointing out to the decrease of scale economies, which are completely exhausted at 89,000 GWh. Figures 3 and 4 sketch the output cost elasticities for both groups of companies.

Regarding the average costs, Table 7 indicates that, among the firms whose production is lower than the sample median output (11,660 GWh), the ones using thermal power technology have lower average costs. This reflects the better scale efficiency of these firms in the lower range of electric power production, as argued previously. Note also that the average cost tends to decline by the output level of 30,000 GWh: between this level and 50,000 GWh, there is no noticeable change in the average cost. Nonetheless, at 89,000 GWh, there is a substantial increase in the average cost due to the exhaustion of scale economies.

The aforementioned results confirm the results obtained from the SUR model. Both indicate that, in the Brazilian electric power generation, natural monopoly characteristics prevail over efficiency gains driven by the competitive market. These results provide support to those who believe that the largest firms have significant cost advantages over the smallest ones in the submarkets where a large demand for electricity allows the exploitation of scale economies. Lastly, our findings point out to the limits to the gains of scale in the production of electricity at the upper output range.

Table 7. Median Output, Output Cost Elasticity and Average Cost by Firm

Firm (Main Technology)	Median Output (GWh)	Output Cost Elasticity: β_q (S-Error)	Average Cost (R$/MWh)
CELESC (Hydroelectric P.)	495.238	0.882* (0.024)	124.416
CJORDÃO (Hydroelectric P.)	1,186.100	1.007* (0.057)	78.642
CERON (Thermal Power)	1,413.000	0.972* (0.053)	83.812
EMAE (Hydroelectric P.)	1,695.000	0.517* (0.033)	198.160
ITAPEBI (Hydroelectric P.)	2,075.000	0.727* (0.001)	123.684
CGTE (Thermal Power)	2,182.000	1.131* (0.002)	61.869
TERMOPE (Thermal Power)	3,986.000	0.941* (0.030)	80.014
CEEE (Hydroelectric P.)	4,060.000	0.724* (0.069)	100.440
CPFL (Hydroelectric P.)	4,574.000	0.508* (0.004)	118.572
LIGHT (Hydroelectric P.)	4,967.000	0.554* (0.081)	109.225
DUKE (Hydroelectric P.)	9,908.000	0.544* (0.065)	79.545
TERMONC (Thermal Power)	12,415.000	0.992* (0.002)	73.796
AES-TIETE (Hydroelectric P.)	12,475.000	0.441* (0.035)	71.097
COPEL (Hydroelectric P.)	19,111.000	0.656* (0.062)	62.528
TRACTABEL (Hydroelectric P.)	29,822.000	0.577* (0.052)	50.292
CEMIG (Hydroelectric P.)	30,412.000	0.567* (0.104)	49.446
ENORTE (Hydroelectric P.)	38,871.000	0.694* (0.077)	52.205
CESP (Hydroelectric P.)	40,725.000	0.701* (0.094)	51.980
CHESF (Hydroelectric P.)	49,911.000	0.734* (0.038)	51.728
FURNAS (Hydroelectric P.)	55,743.000	0.761* (0.085)	52.708
ITAIPU (Hydroelectric P.)	89,082.000	1.078* (0.010)	92.100

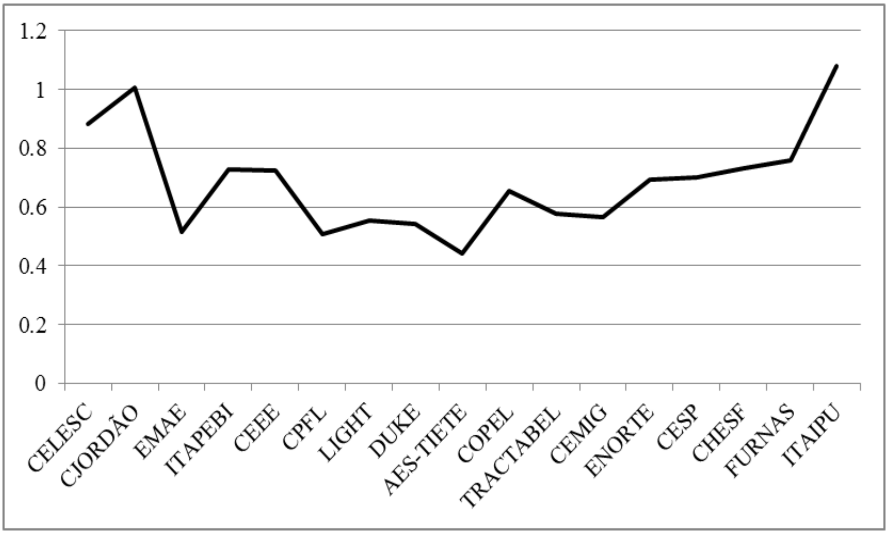
Source: the authors' own computations.
Note: levels of significance - * 99%.

Efficiency Measurement

This section aims at measuring the electric power companies' efficiency. This measure is the distance between the firm's expenditures and the industry's cost function, which is the efficiency frontier. The next equation details this measure:

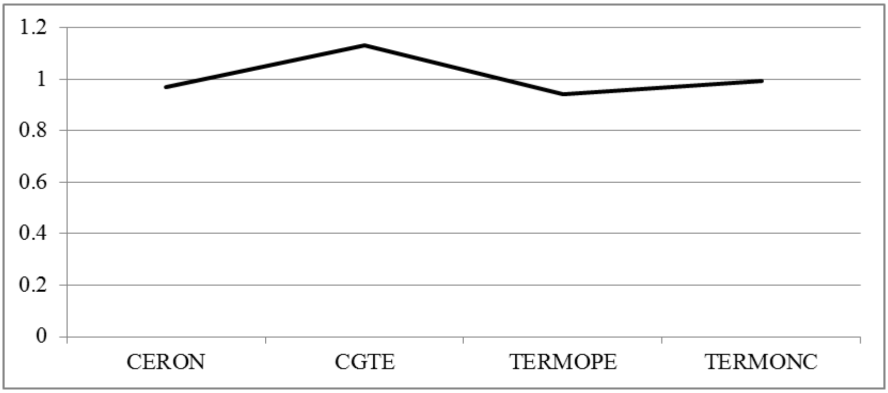
$$EF_{it} = \frac{E(c_t|q_{it}, p_{it})}{E(C_{it}|u_{it} > 0, q_{it}, p_{it})} = \frac{E(C_t|u_{it} = 0, q_{it}, p_{it})}{E(C_{it}|u_{it} > 0, q_{it}, p_{it})} = \exp(-u_{it}) \tag{12}$$

EF_{it} varies between 0 and 1, which are the minimum and maximum efficiency levels, respectively. It is computed from the parameters of the stochastic cost frontier presented in Table 6. The descriptive statistics for the overall sample and the observable groups, represented by the dummies ST, TR and VER, are displayed in Table 8. In addition, Table 9 presents the t-test for the difference in the groups' mean.



Source: the authors' own computations.

Figure 3. Output Cost Elasticity of Firms Using Hydroelectric Power Technology.



Source: the authors' own computations.

Figure 4. Output Cost Elasticity of Firms Using Thermal Power Technology.

According to the results shown in Table 8, the thermal power technology, the state property/management and vertical integration seem to lower the efficiency of the electric utilities. The t-test shown in Table 9 confirms these outcomes for at least two observables idiosyncrasies. It does not reject the hypotheses that vertically integrated and state-owned companies are associated with lower cost efficiency at a 1%-significant level. For this level of significance, the test rejects the hypothesis that the thermal power technology lowers the efficiency. Rejection of this hypothesis would require levels of significance higher than 8.14%. Thus, it is not clear if the thermal power technology affects the efficiency, as it is for the aforementioned institutional specificities.

Table 8. Cost Efficiency by Observable Characteristics: Descriptive Statistics for the Random Parameter Model

Statistics	Overall sample	Thermal power companies	State-owned companies	Vertically integrated companies
Mean	.827	.771	.807	.697
Standard error	.012	.036	.016	.016
95% conf. interval	[.803, .851]	[.698 .844]	[.776 .839]	[.665 .729]
Minimum	.511	.511	.511	.511
Maximum	.995	.991	.995	.995
# of Observations		35	135	74

Source: the authors' own computations.

Table 9. T-test for the Difference in the Mean Efficiency (ME)

Statistics	Thermal *versus* hydroelectric power companies	State-owned *versus* private power companies	Vertically integrated *versus* non-integrated power companies
Difference in the Mean Efficiency	-.068	-.063	-.209
t-statistic	1.785	2.804	10.386

Source: the authors' own computations.

The lower efficiency of state-owned utilities in the Brazilian electricity industry was already measured by Tannuri et al. (2009), but for the distribution segment. Notice that the state-owned companies still have the largest market share in the generation segment, despite the previous efforts to enlarge the private investment in the cited reforms (ANEEL, 2013). Therefore, the gains from improving efficiency through greater private sector involvement may have not been exhausted yet. In addition, the aforementioned regulation strength is better suited to a market dominated by private and not state-owned companies – considering the conflicts of interest that arise when the state has to enforce the regulation to its controlled companies.

Regarding the vertically integrated firms, the previous section suggested that they have cost advantages. However, these advantages are offset by the described efficiency losses. Taking into account such lack of consistency, it is not clear whether or not the vertically integrated companies exploit cost advantages.

Productivity Evolution

This section examines the evolution of the cost efficiency and the cost reduction driven by institutional and/or technical progress over the 2000 - 2010 period. The cost efficiency is defined by equation (12) and the institutional and/or technical progress is derived from the stochastic cost frontier defined in equations (1) and (4), as follows:

$$\frac{\partial \ln C_{it}}{\partial T} = (\beta_T + \beta_{TT} T + \sum_k \beta_{kT} \ln p^k{}_{it} + \beta_{qT} \ln q_{it}) \tag{13}$$

The described institutional and/or technical progress is composed by three elements:

i. the neutral institutional and/or technological change: $T_1 = \beta_T + \beta_{TT} T$;
ii. the non-neutral institutional and/or technological change: $T_2 = \sum_k \beta_{kT} \ln p^k{}_{it}$; and,
iii. the scale-augmenting effect: $T_3 = \beta_{qT} \ln q$.

The neutral institutional and/or technical progress accounts for factors such as managerial improvements and learning by doing. The non-neutral technological change, in turn, reflects the decrease or upsurge in the use of a particular input. Factor–using (factor-saving) technical changes are indicated by positive (negative) values for β_{kT} defined in equation (1). Finally, the scale-augmenting effect accounts for cost reductions (increment) due to more (less) efficient scale of production. If $T_3 > 0$, the companies' scale efficiency decreased on average over time; $T_3 < 0$ is a sign of an increase in the scale efficiency. Moreover, if $\beta_{qT} > 0$ ($\beta_{qT} < 0$), the economies of scale have decreased (increased) during the period studied.

Table 10 shows that the cost efficiency increased by an average of 0.97% over 2000-2010 period, while the technical progress led to cost reductions of 2.97%. Most of the mentioned cost reduction resulted from neutral technological changes. The scale effect also contributed to cost reductions, in a context of decreasing economies of scale given by the positive coefficient β_{qT} = 1.5%. On the other hand, non-neutral technological progress prevented further cost reductions, which reflected the dominance of the significant fuel-using effect ($\beta_{fT} \ln p_{it}^{f}$) over the significant capital-saving effect ($\beta_{rT} \ln p_{it}^{r}$). This dominance is consistent with the increasing environmental constraints on the supply of natural resources for the electricity production in Brazil.

Table 10. Evolution of Cost Efficiency, Institutional and/or Technological Progress: 2000/2010

Year	Cost efficiency		Institutional and/or technical progress			
	Index	Variation	TC	T_1	T_2	T_3
2000	78.71%		-2.87%	-3.08%	0.49%	-0.28%
2001	81.71%	3.81%	-1.66%	-2.73%	1.26%	-0.19%
2002	80.31%	-1.71%	-2.94%	-3.31%	0.45%	-0.08%
2003	83.46%	3.93%	-2.84%	-3.16%	0.61%	-0.29%
2004	84.24%	0.93%	-2.60%	-2.99%	0.78%	-0.39%
2005	82.55%	-2.00%	-2.79%	-2.99%	0.53%	-0.33%
2006	79.92%	-3.19%	-3.37%	-3.15%	0.30%	-0.51%
2007	81.50%	1.98%	-3.39%	-3.15%	0.24%	-0.47%
2008	84.31%	3.45%	-3.42%	-3.31%	0.17%	-0.27%
2009	84.01%	-0.36%	-3.42%	-3.31%	0.14%	-0.25%
2010	86.38%	2.82%	-3.37%	-3.31%	-0.09%	0.02%
2000-2010	**82.46%**	**0.97%**	**-2.97%**	**-3.14%**	**0.44%**	**-0.28%**

Source: the authors' own computations.

CONCLUSION

This study has examined the cost structure of the electricity generation companies in Brazil during the 2000–2010 period by using two econometric methods: SUR and stochastic frontier analysis. Both models converge in the main results regarding economies of scale and other features of these firms' cost structure. There are also tests for potential group effects related to observable characteristics of the firms. Once again, the tests confirm previous results concerning scale economies, cost elasticities and productivity change.

The SUR, in particular, points out to the existence of substantial economies of scale only for the group of the largest companies. Firms producing below the median output exhaust their economies of scale at relatively low production levels. For utilities producing above the median output, economies of scale prevail over the relevant production range. Hence, the hypothesis that economies of scale are a characteristic of the generation market in Brazil cannot be rejected, and such economies are not exhausted at lower levels of production. This result, likely related to the fact that the structure of the electric power industry is still very concentrated, suggests that the economies of scale predominate over the efficiency gains driven by competition from the new entrants. Here we have the main policy implication of the study: the presence of indivisibilities restricts welfare gains derived from market competition in the Brazilian electricity generation market. Our results endorse the measures undertaken under the last restructuring in the industry, which were based on the indivisibility assumption.

On the other hand, the diseconomies of scale related to fuel and water use suggest that the endowment of natural resources – mainly the water used by the hydroelectric power plants – constitutes a restriction to further increases in the use of these inputs to produce electricity, implying that the cited resources are increasingly limiting the exploitation of scale economies.

The existence of substantial scale economies in the Brazilian electric power generation is confirmed by the stochastic cost frontier estimations. The model additionally identifies that these economies are related to the use of hydroelectric power technology. The effect of the companies' other observed characteristics on the cost structure were also tested. They could not be implemented in the SUR model given the individual fixed effects used to control potential unobserved heterogeneity bias.

Regarding the aforementioned characteristics, the stochastic cost frontier indicates that the state-owned companies are more intensive in labor, have lower economies of scale and show less technical progress. These specificities are likely related to the older technology used by the cited companies. Regarding the vertically integrated firms, their production technology exhibits labor cost advantages, suggesting that increases in labor productivity driven by gains from specialization might not be relevant in the electricity industry.

Lastly, the thermal power technology exhibits reduced economies of scale, greater technical progress, and is more intensive in fuel. These results confirm the expectations that this technology choice is determined by the market share and the relative price of fuel predicted by the firm.

The stochastic frontier model estimations provide no clear conclusion regarding the eventual cost advantages of the thermal power technology or of the vertically integrated firms. On the other hand, the results are consistent in indicating the lower cost efficiency of the state-owned companies. Considering that state-owned enterprises still have the largest market share in the generation segment, regulatory efforts to attract private investment may still be an option to improve efficiency. In addition, the adopted regulation measures adopted in Brazil is better suited to a market dominated by private, rather than state-owned companies.

Another implication of the estimated stochastic cost frontier is that the productivity gains over the 2000-2010 period were mainly driven by institutional and/or technical progress, instead of efficiency gains. Finally, the productivity gains were restricted by increasing fuel and water expenditures, which is consistent with the new environmental constraints and the consequent scarcity of natural resources and their related inputs to the electricity production in Brazil.

REFERENCES

ANEEL, 2013. Available at: < http://www.aneel.gov.br/aplicacoes/AgenteGeracao/SelecaoDaEmpresaPasso1.asp? Access in 09/03/13.

Arocena P., Cost and quality gains from diversification and vertical integration in the electricity industry: a DEA approach. *Energy Econ.* 30 (1), 2008, 39–58.

Arocena P., Saal D. and Coelli T., Vertical and horizontal scope economies in the regulated U.S. electric power industry. *Ind. Econ.* 60 (3), 2012, 434–467.

Bhat, C., (1999) "Quasi-Random Maximum Simulated Likelihood Estimation of the Mixed Logit Model," *Manuscript, Department of Civil Engineering*, University of Texas, Austin.

Berry D.M. and Mixon J.F., Multiproduct outputs and scale economies in electric power production: some new estimates. *Rev. Ind. Organ.* 15, 1999, 65–76.

Blackorby C. and Russell R.R., Will the real elasticity of substitution please stand up? A comparison of the Allen/Uzawa and Morishima elasticities, *Am. Econ. Rev.* 79 (4), 1989, 882–888.

Bolsa de Valores and Mercadorias e Futuros S.A., 2011, Available at: < http://www.bovespa.com.br/Principal.asp, Access in 10/25/11.

Carpio L.G.T. and Pereira A.O., Economical efficiency of coordinating the generation by subsystems with the capacity of transmission in the Brazilian market of electricity, *Energy Econ.* 29 (3), 2007,454–466.

Centrais Elétricas Brasileiras, 2011, Available at: http://www.eletrobras.com/elb/ri/main.asp?Team=7B7152FB662D45B72D44CE2D9FAD2DEDC8E45986567D, Access in 09/23/11.

Christensen, L.R. and Greene, W.H., Economies of scale in electric power generation, *J. Polit. Econ.* 84 (4),1976, 655–676.

Empresa de Pesquisa Energética, Balanço Energético Nacional 2011: Ano Base 2010, 2011, Available at: https://ben.epe.gov.br/downloads/Relatorio_Final_BEN_2011.pdf, Access in 10/01/12.

Ferreira C.K.L., A Privatização no Brasil: Privatização Do Setor Elétrico no Brasil, In: BNDES Working Paper, 2000, (Available at: <http://www.bndes.gov.br/SiteBNDES/export/sites/default/bndes_pt/Galerias/Arquivos/conhecimento/ocde/ocde06.pdf>. Access in 10/15/13).

Filippini M. and Luchsinger C., Economies of scale in the Swiss hydropower sector, *Appl. Econ. Lett.* 14 (15), 2007,1109–1113.

Gilsdorf K., Vertical integration efficiencies and electric utilities: a cost complementarily perspective, *Q. Rev. Econ. Financ.* 34 (3), 1994, 261–282.

Goto M. and Tsutsui M., Technical efficiency and the impacts of deregulation, *Energy Econ.* 30 (1), 2008, 15–38.

Greene, W., 2003. Limdep Version 8.0. New York: Econometric Software.

Greene, W., 2005. Reconsidering heterogeneity in panel data estimators of the stochastic frontier model. *Journal of Econometrics* 126, 269–303.

Hisnanick J.J. and Kymn K.O., Modeling economies of scale: the case of US electric power companies, *Energy Econ.* 21, 1999,225–237.

Huettner D.A. and Landon J.H., Electric utilities: scale economies and diseconomies, *South. Econ. J.* 44, 1977,883–912.

Jorgenson D.W. and Griliches Z., The explanation of productivity change, *Rev. Econ. Stud*. 34 (3), 1967, 249–283.

Joskow P.L., Restructuring, competition and regulatory reform in the US electricity sector, *J. Econ. Perspect.* 11 (3), 1997, 119–138.

Joskow P.L. and Schmalensee R., Markets for Power: An Analysis of Electric Utilities Deregulation, 1983, MIT Press; Cambridge, US.

Kamerschen D., Klein P.G. and Porter D.V., Market structure in the US electricity industry: a long-term perspective, *Energy Econ*. 27 (5),2005, 731–751.

Kaserman D.L. and Mayo J.W., The measurement of vertical economies and the efficient structure of the electric utility business, *J. Ind. Econ*. 39 (5), 1991, 483–503.

Kwoka J.A., Vertical economies in electric power: evidence on integration and its alternatives, *Int. J. Ind. Organ*. 20 (5), 2002, 653–671.

Lee B., Separability test for the electricity supply industry, *J. Appl. Econ*. 10 (1), 1995, 49–60.

Marreco J.M. and Carpio L.G.T., Flexibility valuation in the Brazilian power system: Aa real options approach, *Energ Policy* 34 (18), 2006, 3749–3756.

Nemoto J. and Goto M., Technological externalities and economies of vertical integration in the electric utility industry, *Int. J. Ind. Organ*. 22 (1), 2004, 67–81.

Nerlove M., Returns to scale in electricity supply, In: Christ C.F., (Ed), *Measurement in Economics*, 1963, Stanford University Press; California, US.

Pinto H.Q., Almeida E.F., Bomtempo J.V., Iotty M. and Bicalho R.G.,Economia Da Energia: Fundamentos econômicos, evolução histórica e organização *Industrial*, 2007, Elsevier; Rio de Janeiro, BR.

Ramos-Real F.J., Tovar B., Iotty M., Almeida E.F. and Pinto H.Q., The evolution and main determinants of productivity in Brazilian electricity distribution 1998–2005: an empirical analysis, *Energy Econ*. 31 (2), 2009, 298–305.

Santana E.A. and Oliveira C.A., A Economia dos Custos de Transação e a Reforma na Indústria de Energia Elétrica do Brasil, *Estud. Econ*. 29 (3), 1999,367–393.

Sternberg R., Hydropower's future, the environment, and global electricity systems, *Renew. Sust. Energ. Rev.* 14 (2), 2010, 713–723.

Tannuri-Pianto, M.E., Sousa, M.C.S., Arcoverde, F.D., 2009. Fronteiras de Eficiência Estocástica para as Empresas de Distribuição de Energia Elétrica no Brasil: uma análise de dados em painel. *Estudos Econômicos* 39(1), 222–247.

Train, K. (2003). Discrete Choice Methods with Simulation. Cambridge University Press, Cambridge.

Trindade P.D., Definição de Mercado Relevante no Setor elétrico e articulação Entre a Agência Nacional de Energia Elétrica (Aneel) e Os órgãos Do Sistema Brasileiro de Defesa Da Concorrência, In: *Thesis of Master in Public Economics*, 2012, University of Brasília.

Wolak F.A., Electricity Market Structure and Pricing: an International Comparison, In: *Energy Modeling Forum* 15, 1997, University, California, US; Stanford.

In: Infrastructure Investments
Editors: Gisele Ferreira Tiryaki et al.
ISBN: 978-1-53610-792-0

Chapter 4

TAPPING THE BOUNTY: CHALLENGES TO RESTORE THE INVESTMENTS IN THE BRAZILIAN OIL INDUSTRY

Edmar de Almeida[1,*], Luciano Losekann[2] and Breno Medeiros[1]

[1]Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil
[2]Universidade Federal Fluminense, Niterói, RJ, Brazil

ABSTRACT

This chapter aims at analyzing how the changes on the domestic and international oil and gas landscape are hindering investments in the Brazilian oil and gas (O&G) industry. The chapter also discusses possible measures to overtake these challenges. The discovery of several giant oil reservoirs in the Pre-salt area has put the country in a new perspective in the world oil industry, and, in a few years, Brazil may become a relevant oil exporter. Investments in the Brazilian O&G industry have increased rapidly in the last years, partially because of the development of the Pre-salt area. However, the O&G sector in Brazil has been suffering an important setback since 2014 due to international and domestic changes in the oil market. In the second half of 2014, in particular, the oil prices decreased sharply, forcing an adjustment in the oil industry worldwide. The adjustment to the Brazilian oil industry has been much more severe,

* Corresponding Author: Edmar de Almeida. E-mail: edmar@ie.ufrj.br.

as a result of domestic factors, such as Petrobras' financial and political problems, restrictions for private companies to operate in the Pre-salt area, regulatory costs associated with the diversity of contracts and fiscal regimes in same area and fields, the rigidity of the local content regulatory framework, and the lack of transparency in the domestic pricing mechanism for oil products.

Keywords: Pre-salt, Brazil, oil industry

INTRODUCTION

The discovery of the Pre-salt oil reservoirs in 2006 has completely repositioned Brazil in the international oil and gas market, allowing the country to change its role from an oil and gas importer to an important potential oil exporter, with significant impacts for its economy and for the world oil balance. This new environment attracted the attention of the largest international oil companies and had completely changed the business strategy and ambition of Petrobras, the Brazilian state-owned oil company[1].

The federal government granted Petrobras with the right to be the sole operator of the Pre-salt fields and created a specific regime for the oil exploration and production (E&P) in those areas. Petrobras along with international partners have more than doubled their upstream investments between 2008 and 2013, triggering a euphoric period for the oil industry in Brazil.

In 2014, the scenario started to change when a large corruption scheme involving Petrobras, several shipyards and EPC companies was denounced. This event has seriously affected the Brazilian oil supply chain. Additionally, Petrobras suffered from heavy losses on impairments tests, which resulted in negative adjustments, totaling US$16.8 billion[2]. These facts, along with the consequences of the political misuse of the company in the previous years[3], contributed to place the company in a delicate financial situation. Since then, Petrobras has reduced significantly its investment plan and implemented a

[1] Petrobras is controlled by the Federal Government but has floated its shares both in the Brazilian and the American stock markets.

[2] Three types of losses in different areas of Petrobras were indicated by the impairments test, which is an accounting analysis of the company's assets/future cash flow: (i) refining: the postponement of projects and assets; (ii) E&P: the decline of oil prices; and (iii) petrochemical: the reduction in oil demand and the company's margins.

[3] The Government tried to hold inflation by controlling fuels' prices via Petrobras.

large divestment plan in an attempt to generate cash flow and better manage its huge debt.

The international oil market also radically changed in 2014. The unconventional revolution[4] in North America affected the oil market balance while the growth in demand for oil in emerging economies stalled. Since 2008, the US has increased its oil production four million barrels per day or 80 percent (Bordoff and Losz, 2015). When Saudi Arabia refused to cut further its oil production to accommodate the new supply, oil prices collapsed at the end of 2014. Even though the high E&P cost of the Pre-salt reserves has decreased over the years, the oil price collapse still represents a challenge for the competitiveness of Pre-salt layers oil exploration.

This chapter sheds light on how the changes in the domestic and international oil and gas landscape are hindering investments in the Brazilian oil and gas industry. The chapter also discusses possible measures to overtake these challenges. This chapter is divided into three sections, including this introduction and in addition to the conclusion. The second section is dedicated to the analysis of the recent evolution of the Brazilian oil industry, from the euphoria of the Pre-salt layer oil discoveries to the current crisis. The third section presents a brief analysis of the changing international landscape in the oil and gas market and discusses the challenges and measures to restore the investments in the Brazilian oil industry.

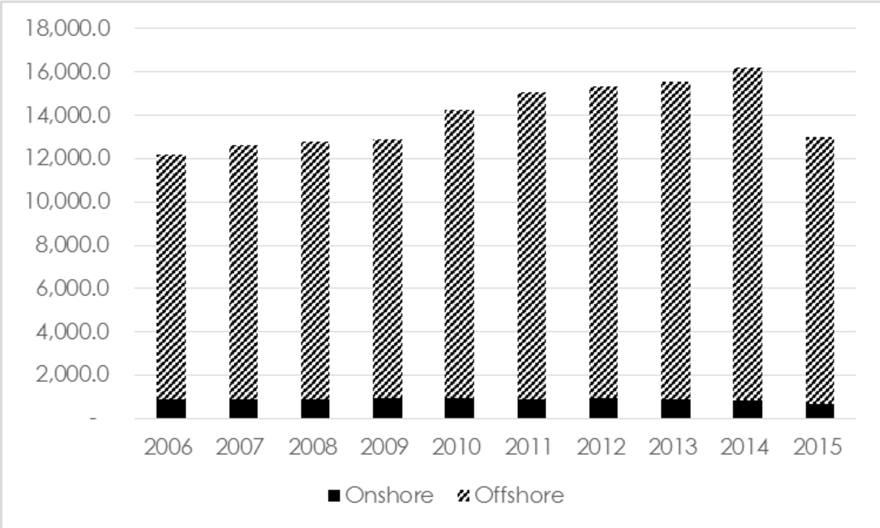
THE OIL PRODUCTION IN BRAZIL: CHALLENGES AND PROSPECTS

Brazil has substantial oil reserves in its 7.5 million km^2 sedimentary area. Indeed, in December of 2015, the proven reserves were 13 billion barrels of oil and 429.5 billion cubic meters (bcm) of natural gas. From these totals, 95 percent of oil and 84 percent of gas are offshore. In 2014, Brazil was placed 15th in the ranking of the countries with the highest proven oil reserves in the world and 31st in the ranking of proven natural gas reserves (ANP, 2015).

Brazil produced about 890 million barrels of oil and 35.1 bcm of natural gas in 2015. In 2014, the country was ranked 13th among the highest producers of oil (ANP, 2015). The Brazilian O&G production has a strong focus on

[4] Bordoff and Losz (2015, p. 195) technically explain this American oil revolution as follows: “The technological innovation of combining horizontal drilling with hydraulic fracturing first gave rise to a dramatic increase in American natural gas production, and the technology was then applied to extract oil from shale and other tight rock formations.”

offshore, mainly deep-water activities. The E&P in Brazil has been historically concentrated in the Campos Basin, but the Santos Basin has more recently been increasing it share, with the development of Pre-salt fields (Figure 1).



Source: ANP (2016).

Figure 1. Evolution of Proven Oil Reserves in Brazil per location: 2006 – 2015 (Millions of barrels).

Brazil is also active in the international O&G market, with significant trade, especially exporting oil (heavy oil) and importing natural gas. In 2015, Brazil exported 269 million barrels of oil and imported 19 bcm of gas, particularly from Bolivia, through pipelines (ANP, 2015).

Petrobras is the major player in the Brazilian O&G industry, producing 86 percent of the Brazilian oil and 82 percent of the country's natural gas in 2014 (ANP, 2015). However, foreign companies are increasing their presence in the country, and they tend to work either by themselves or in partnerships with Petrobras. There are circa 100 companies with E&P concessions in Brazil, 55 of them are foreign companies and 45 of them are Brazilian companies. The majority of Brazilian oil companies are operating onshore or marginal or mature offshore fields.

In the late 2000s, the discovery of large amount of oil in the Pre-salt reservoirs enhanced the potential growth of oil production in the country. In fact, field development plans approved by the National Petroleum Agency – ANP, the Brazilian oil and gas regulator, indicate that oil production could reach 4.3 mb/d by 2021. If this potential growth materializes, the Brazilian oil exports may significantly affect the global supply. Accounting for 2.7 percent of global production today, Brazil may double its shares to more than 6 percent by 2035. The oil production increase depends mainly on the Pre-salt exploration competitiveness in the new international oil market context and Petrobras's investment capability.

Considering that a significant share of the country's investment is directed to the oil sector, this rise in oil production should have significant positive spillover effects for the private sector (supply chain) and for the government (taxes/royalties). Despite the 25.5 percent decrease in Petrobras' planned investment[5], the oil sector represents 26 percent of the total investment in Brazil and 53 percent of the investment in the industrial sector (BNDES, 2016). IBP and CNI (2012) estimated that the oil sector represents about 12 percent of the Brazilian Gross Domestic Product (GDP).

The Pre-Salt Revolution

The discoveries of oil in the Pre-salt layers were the result of a technological development process carried out by Petrobras, which has its roots in the 1980s, especially with the E&P increasingly moving from shallow waters to deeper waters. Petrobras has won three times the prize for the most renowned company at the Offshore Technology Conference (OTC)[6]. The liberalization process in the Brazilian oil and gas sector during the 1990s was also important for this favorable performance, as it facilitaded the entry of international and experienced oil companies, which partnered with Petrobras and shared the risks of investments in oil prospects with high risk and large potential.

The Role of the Technological Learning Process

It is important to notice that Brazil has encountered great success in deep water offshore exploration. In the 1970s and 1980s, following the oil shocks, the country strove to reduce the external oil dependency that had negatively

[5] This decrease in investment predicted in the oil sector is a consequence of the actual crisis in the sector, which will be discussed further. The 2011-2014 investment plan was R$398 billion and the 2016-2019 plan is R$296 billion.

[6] Petrobras received the prize for distinct technique/technology at the OTC in 1992, 2001 and 2015. The last prize was related to the group of technologies developed to the Pre-salt. Available at: www.petrobras.com.br/en/news/petrobras-receives-highest-award-in-global-oil-industry.htm

affected its balance of payments. In order to explore the Campos Basin's shallow waters, Petrobras, backed by the national government, attempted to gain the technological expertise required to successfully carry out this task.

Since the 1980s, Brazil has made tremendous economic and technological efforts in the offshore sector in order to obtain oil self-sufficiency. As a result of this effort, Petrobras became a leading company in offshore technology and it is currently the largest operator of deep water oil fields in the world. The Brazilian O&G sector reached its prime in the late 2000s with the discovery of oil in the Pre-salt layer.

Technological advances and the focus on deeper water exploration fostered to some extent a knowledge revolution related to geological formations, such as the Pre-salt layer, and to the operation expertise in these environments. It has enabled continuous advances in the development of new materials, supplies, equipment and procedures and has helped overcome the challenges of deep and ultra-deep water exploration, which has opened a new world of opportunities to the oil industry.

Brazil plays a key role in the international technological and operational expertise of deep offshore field exploration. This expertise has led the country to great discoveries of hydrocarbons in the last decade: seven out of the top ten discoveries in deep waters occurred in Brazil (Table 1). Eleven out of the thirty-five largest discoveries in 2001-2011, all containing more than a billion barrels, took place in Brazil. Considering the total volume of oil and gas in these fields, Brazil would have discovered the equivalent to one-third of the total resources in this ranking, accounting for 35 billion barrels of oil equivalent (VIEGAS, 2011).

Petrobras employs over 1,600 people at its research center (CENPES). The company is also well ranked in Research and Development (R&D) investments among the international oil companies. In 2014, it invested US$1.1 billion in R&D and made partnerships with about 100 universities and research centers in Brazil and 35 abroad. Most of these investments is associated with the new challenges in deep waters. There is a concession contract clause requiring that 1 percent of the gross revenues from the high productive fields should be spent in R&D activities in the country. Given that the most important recent discoveries were in the Pre-salt area, other companies in the sector are following Petrobras' lead and establishing global research centers in Brazil, such as Statoil, Schlumberger and GE[7].

[7] See: www.statoil.com/brazil/pt/technologyinnovation/researchcenterrio/pages/default.aspx, www.slb.com/about/rd/research/sbr.aspx and www.geglobalresearch.com/locations/rio-de-janeiro-brazil.

Table 1. Major Discoveries of Hydrocarbons in the 2000s

Country	Field	Discovery	Operator	Estimated reserves
Brazil	Libra	2010	Petrobras	8 - 12 billion boe
China	Nanpu	2007	PetroChina	8.6 billion boe
Brazil	Tupi/Lula	2006	Petrobras	5-8 billion boe
Brazil	Jupter	2008	Petrobras	up to 8 billion boe
Brazil	Franco	2010	Petrobras	4.5 billion boe
Brazil	Iara	2008	Petrobras	3.4 billion boe
Brazil	Jubarte	2001	Petrobras	1.77 billion
Iran	Ferdowsi	2010	POGC	1.7 billion boe
Brazil	Mexilhão	2001	Petrobras	227 Bm3 gas/ 200 million boe
Iraq	Shaikan	2009	G. Keystone	1.5 billion boe
Ghana	Jubilee	2007	Tullow	1.5 billion boe
Brazil	Guará	2008	Petrobras	1.1-2 Billion boe
Nigeria	B. Sothwest	2001	Shell	1 billion

Source: Rigzone, ANP and Petrobras. Apud: Viegas, 2011, adapted.

The Role of the Liberalization Process

The liberalization process in the Brazilian O&G sector has played an important role for the discovery of the Pre-salt oil resevoirs. The constitutional reform that ended with Petrobras' monopoly took place in 1995. In 1997, a new oil and gas law (Act 9,478) consolidated the liberalization process and established an independent regulatory agency for the petroleum sector (ANP). Since 1997, the ANP has organized 13 bidding rounds for exploratory blocks, attracting a large amount of private companies to the country.

Several reasons prompted the government undertake this liberalization process. Firstly, the country was striving to lower its external dependency, and oil had a heavy weight on the country's imports and foreign currency spending. Secondly, the goal was to increase investments in order to make the most out of the country's potential natural resources, so as to meet growing demand.

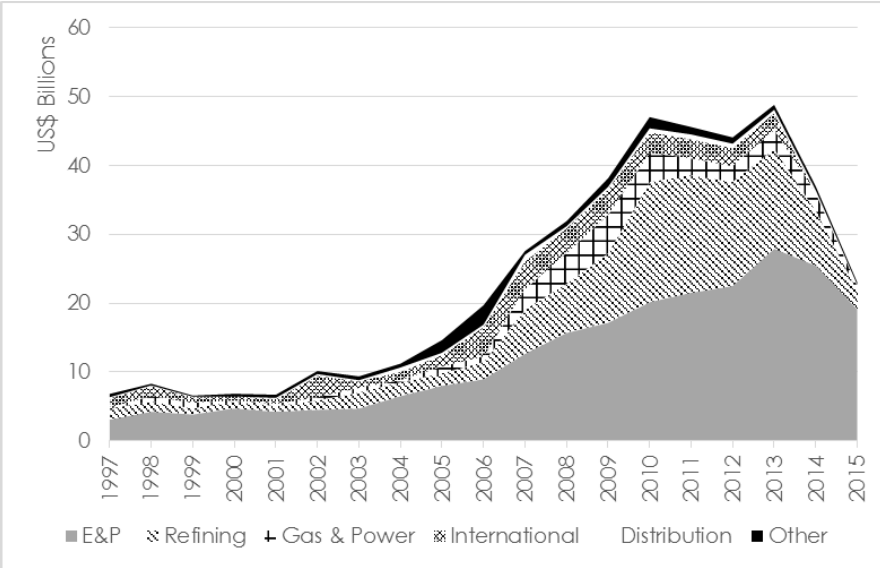
This decision has proved successful, as it produced a radical change in the dynamics of investments in the E&P of O&G in Brazil. Annual bidding rounds enabled new international and experienced players to join the local industry, mostly via consortia with Petrobras. The state-owned company could now count on its new partners' capital and technology to carry out a multitude of projects, in addition to modernizing itself by streamlining its industrial processes and control mechanisms.

It seems that the floating of its shares and the opening of the O&G market was very beneficial for Petrobras. Its market value, as well as its investment capacity, increased substantially. Figure 2 highlights the evolution of Petrobras E&P investments after the 1997 oil market opening. Additionally, foreign companies based in Brazil were able to participate in the bidding rounds, which would grant them the right to explore the oil fields in a designated area and during period of time. In these concession agreements, companies had to pay to the government, in addition to taxes, royalties (5 - 10 percent) and special participation[8] (0 - 40 percent), depending on the field's productivity.

The liberalization of the oil and gas industry in Brazil was crucial to allow an increasing exploratory effort in deep offshore areas. Partnerships with other large international oil companies allowed Petrobras to share costs and risks of investing in new geological frontiers. These associations fostered the exploration investment in Pre-salt reservoirs, with the first exploratory wells costing US$240 million (Reuters, 2008).

Moreover, new companies brought with them new strategies, expertise and methods for evaluating economic and geological potential, thus benefiting the sector as a whole. A striking example is the Peregrino field, which was found by Petrobras in 1994, but it was not in the interest of Petrobras to develop it. After the market was liberalized, the block containing the field (BM-C-7) was offered again in the second concession round in 2000. It was acquired by Encana (Canadian) and Anadarko (American), which ended up finding substantial oil reserves (300 to 600 million barrels). Statoil currently operates this field.

[8] Special participation is a kind of windfall profit tax chorged only to large and highly productvive oil fields.



Sources: Elaborated by the authors with data from IBP and Petrobras.

Figure 2. Evolution of Petrobras Investment since the market openness (Real - US$ Billion).

The Discovery of the Bounty - The "Winning Ticket"

The so-called Pre-salt area is located approximately 300 km off the coast of the state of Rio de Janeiro in the Santos Basin. It is 200 km wide and holds substantial oil reserves, trapped under a thick layer of salt. When discovered in 2006, it was seen as a blessing, as it increased Brazil's already abundant natural resources.

Three major technical barriers thwart the oil exploration in the Pre-salt area: the water depth (around 2,000 m), the distance from the coast (300 Km) and a thick salt layer (2,000 m). With the country making significant progress in both geological knowledge and deep-water technologies in the past decades, it was only a matter of time before exploration was made feasible in the Santos Basin. The first discovery of the Pre-salt oil occurred in 2006 - at the Tupi block located in the Santos Basin, which had an estimated recoverable reserve of 5 - 8 billion barrels. The first well produced good quality oil with 28 API density. This was the first of several discoveries, which today sums up to 40 billion barrels of recoverable oil.

Discoveries of the Pre-salt reservoirs came at an impressive pace, stimulated by the fact that the exploration well drilling in the Santos Basin had a success rate higher than 80 percent during the first five years. This leads many to believe that various reservoirs are yet to be discovered in this high-potential area. In addition, the productivity of most fields is higher than predicted. The first evaluation of oil discoveries assumed an average well productivity of 10 thousand barrels per day. The average well productivity in the Pre-salt Santos Basin is 25 thousand barrels. Nine out of the ten wells with highest production in Brazil are found in this area. The most productive well is in the Lula field, producing 36 thousand barrels per day. In May 2016, the Pre-salt O&G production reached 1.15 million boed[9].

In order to seize this unique opportunity, Brazil has act strategically. The country's fast rise in the oil production, projected until 2035, relies heavily on the Pre-salt oil reserves. In the World Energy Outlook of the International Energy Agency (WEO 2015), the main scenario estimated the oil production in Brazil to reach 3.1 mb/d by 2020, 4.7 mb/d by 2030 and a peak of 5.3 mb/d by 2035. The Brazilian government is more optimistic, estimating that the oil production should reach 5.1 mb/d by 2024, with a surplus of 2.1 mb/d (EPE, 2015).

With the expansion of the global deep-water E&P and Brazil's leading experience and state-of-the-art in offshore technology, the country has also the possibility to establish itself as the largest market for the deep-water industry, as well as the technology pioneer in the area. This could foster the development of competitive local suppliers, fostering the country's capability to export to international markets.

This high level of optimism by the Brazilian Government is due to the fact that it did not consider the serious problems with key local suppliers' and the consequences of the critical financial situation faced by Petrobras. Various Brazilian suppliers are involved in the corruption scandal known as "Operation Carwash" and were forbidden to participate in new Petrobras's bids, while others went bankrupt. These events should postpone the deliveries of important equipment and system, such as offshore vessels and platforms.

The Crisis in the Brazilian Oil Industry: The Road from Euphoria to Depression

Along with the liberalization of the oil market in Brazil, a Local Content (LC) policy was approved. A minimum share of LC – which refers to a certain

[9] See: http://www.petrobras.com.br/en/news/our-oil-output-in-the-pre-salt-cluster-surpasses-1-million-barrels-per-day.htm.

amount of goods and services to be provided by suppliers located in Brazil to a given project – is compulsory to every oil E&P project in the country. Moreover, the minimum share of LC offered by a bidder is one of the key elements evaluated at the bidding rounds. The government main goal was to stimulate the national industry and to promote technological development, as well as to create new jobs.

The promotion of LC in the oil sector began with the Petrobras's program named "PROCAP 1000" in 1986. This program aimed to cope with the technological challenges to produce oil in deep waters (1,000 m). The Petroleum Act (1997) and the first bidding round (1999) introduced a new phase in the Brazilian Local Content Policy (LCP). The government decided to adopt LC commitments as one of the criteria for the bidding rounds.

Until the fourth round, LC accounted for 15 percent of the criteria to win the bid and the other 85 percent was the signing bonus. From the fifth round, the Minimum Exploratory Program (MEP)[10] was introduced as another criteria to win a bid. On the fifth and sixth rounds, LC was the criteria that weighted most in the bids, weighting 40 percent, and since then its weight has been reduced to 20 percent (Medeiros, 2016).

In the fifth round in 2005, new regulations were introduced with the adoption of minimum percentages of LC commitments. In the same year, the government launched a program to promote local suppliers' technological capability (PROMINP program). In the seventh round in 2007, new regulations introduced a detailed certification process, along with a detailed manual for LC in oil and gas projects. In 2010, the Brazilian Development Bank (BNDES) created the program "INOVA PETRO" for the financing of R&D projects and technological developments by local supplier companies.

From the start of LCP, ANP has been regulating and overseeing local content in Brazil, imposing the penalties provided by law when operators do not comply with the regulations. In the case of non-compliance with LC, the criterion for fines were set since the first round which took place in 1999, even though the fines charged so far are mainly related to fields granted from the fifth round onwards (Medeiros, 2016). Risks associated with noncompliance of LC and costs overrun are fully assumed by the operators, according to the contracts signed with ANP.

[10] MEP refers to the set of activities aimed at the fulfillment of the contractual obligations at the exploration phase, which are carried out in a concession area. Each activity is computed quantitatively according to their nature and scope, which has equivalence in work units and corresponds to the winning bid parameter for a bidding area.

This possibility of fines due to LC noncompliance is a concern for the operators, as the commitments are determined at a time when they do not know precisely its needs. If an operator fails to meet the level of LC that was agreed upon, it can be subject to heavy fines, which are calculated proportionally to the level of noncompliance.

There is a waiver clause that allows the operators to have authorization to import equipment/systems in certain circumstances: whenever the delivery time or the price of the local suppliers is deemed too high in comparison with international standards or in the absence of the technology in the country. However, it seems that ANP has not authorized any waiver request so far (Medeiros, 2016).

The waiver clause could be an important instrument to balance possible excesses of the LCP and to stimulate the local suppliers to be more competitive. It should be mentioned that the money gathered from the fines charged in case of noncompliance with LCP is directed to the Federal Government with no pre-established destination. New legislation should require that these monetary sources remain in sector to assist the development of the weaker and the strategic links in the supply chain (e.g., where operators are having more difficulties to find competitive suppliers).

The great potential for wealth represented by the newly discovered Pre-salt oil reserves led to changes in the institutional framework related to oil exploration in the Pre-salt zone. The core of this change focused precisely on the expansion of state control over the exploitation of the Pre-salt resources. The main stated goal is to maximize the benefits from such exploitation under a long-term strategic perspective that transcends the boundaries of the oil industry.

The changes in legislation were not just related to increasing the state's share in oil revenues, but also in providing the government with greater control over the process for generating this income. The aim was to increase the state's oversight not only on all the oil production stages, but also on the destination of this oil and all connections made by the oil industry with other industries and services.

In 2008, two years after the discovery of the first Pre-salt field, the Federal Government initiated a political negotiation process to approve a new oil law. In fact, in 2010, three new laws were enacted. These laws created a new regulatory framework for the Pre-salt oil exploration. The main changes were the following:

(i) adoption of a Production Sharing Contract (PSC) for all new exploratory deals in the Pre-salt area;
(ii) creation of a new 100% state-owned company to represent the State in the new PSCs, named Presal Petróleo S.A. (PPSA)[11];
(iii) granting of exclusivity rights to Petrobras to operate new projects at Pre-salt areas;
(iv) creation of a sovereign fund, named Social Fund, with resources from the federal government royalties and other participations on the Pre-salt deals; and,
(v) creation of a special concession contract to allow Petrobras to find and develop up to 5 billion barrels of oil in areas surrounding the existing Petrobras exploratory blocks.

PPSA plays an important role as a mechanism for state's control over the Pre-salt production and represents the interests of the Federal Government in the PSC and in the Operating Committees of the exploration projects, with veto power in the decision making process. PPSA's main responsibilities include monitoring and auditing of costs and investments in PSC. PPSA will also trade the oil owned by the government in PSC.

In the Pre-salt exploration areas, PSC is the only type of contract that may be established apart from the existing concession contracts. In all PSC, costs and risks of E&P are split between Petrobras, which is the preferential operator with at least 30 percent of the equity, and its partners[12]. It is important to mention that under the current oil law, the Brazilian government is allowed to allocate a Pre-salt block exclusively to Petrobras without any type of auction. However, if it does decide to offer the block to the market in an open competition, a bidding round should be organized. The winning consortium will be the one offering the largest share of the oil profit to the State. The signing bonus is set by the government in the bidding rules.

The first block of the Pre-salt Santos Basin auctioned under the production sharing system was Libra. This field, from which there is good geological information available, has an estimated 8 to 12 billion recoverable oil barrels, and it is the second biggest discovery of the 21st century (the Kashagan field, in Kazakhstan, is the biggest one). The government charged US$7 billion as

[11] PPSA has some similarities to Petoro from Norway.
[12] From 2010 to October 2016, Petrobras was the sole operator in Pre-salt PSC contracts. The sole operator rule has been changed by a "preference rule," in which all new exploratory blocks in Pre-salt area will be offered to Petrobras to operate. Those blocks where Petrobras is not interested in operating will be auctioned in a competitive process.

signing bonus. A single consortium, formed by Petrobras (40 percent), Royal Dutch Shell (20 percent), Total (20 percent), CNOOC (10 percent) and CNPC (10 percent), bid and won the round.

In this new regulatory framework, Petrobras has a very key role to determine the pace of development of Pre-salt area. Petrobras being the preferential operator could hold new concession rounds, due to its financial situation, which will be discussed further.

The Rise and Crisis of the Brazilian Independent Oil Companies'[13]: The OGX Case

OGX is one of the companies created by Eike Batista, a well-known Brazilian business tycoon, who decided to launch several large new businesses in the aftermath of a very successful experience in the mining sector[14]. Mr. Batista decided to create an oil company in 2007, aiming at repeating the experience with MMX, his mining company. OGX's ambitious plan was to become a smaller version of Petrobras, as Eike Batista said himself. In order to achieve this objective, he counted on his "dream team," composed of an experienced group of executives and technical staff, many of them former employees of Petrobras (Infomoney, 2013).

OGX was created swiftly in 2007 to participate in the 9th bidding round organized by ANP. OGX won 21 exploratory blocks, becoming the operator of 14 of them. It was an enormous challenge for an E&P startup company. Beyond this fact, the market expectation for OGX's performance was considerably optimistic: it was the biggest IPO that Bovespa[15] had ever seen, raising R$ 6.7 billion worth of domestic and foreign capital, or approximately US$ 4 billion (Infomoney, 2013).

The optimism led to an overvaluation of OGX's shares, even though the company was not producing any oil or gas yet. Reports were misinterpreted by the market and/or were presented in a positive manner. One of the most emblematic anecdotes of this situation was the confusion made by OGX and the financial market with regard to reserves and resources concepts, which made the market agents believe that the company had an immense 10 billion

[13] The term "independent" refers to nonintegrated oil companies, mainly focused either on E&P or downstream activities.

[14] Mr. Batista was very successful in acquiring mining exploration areas that had no value before the commodity boom that took place in the 2000 and selling these assets during the boom.

[15] Bovespa is the Brazilian stock exchange market.

barrel reserve[16], which OGX later called “potential resources.” The numbers were clarified, and the 10 billion barrels’ mark was an extremely optimistic evaluation of OGX, based on the geologic report which was prepared by a specialized geological consulting firm (Infomoney, 2013).

OGX’s oil production started in 2012, which was later than expected and at levels lower than those previously announced by the company: the estimated daily production by OGX was around 20 thousand barrels in the beginning of the project, but it latter produced only 5-10 thousand barrels. The market’s confidence on the company’s performance collapsed as OGX’s indebtedness grew rapidly. According to Infomoney (2013), OGX filed for judicial recovery in 2013. By 2014, Eike Batista divested part of its shareholdings in OGX, and the company changed its name to OGPar.

Despite this tragic ending, the most successful and lucrative asset of OGX was not related to oil production, but to natural gas exploration: a subsidiary of OGX, called OGX Maranhão, focused on the exploration of the Parnaíba Basin, located in the Northeast of Brazil. The basin is located onshore, and it is rich in conventional gas with advantageous characteristics[17], making this natural gas highly competitive. Since there was no pipeline infrastructure in the basin area and the investment to build one was not economically feasible, OGX created a pioneer project to develop power plants next to the production fields. MPX, another company from Eike Batista’s conglomerate, was responsible for the construction and operation of the power plants integrated to the gas production.

Meanwhile, other independent companies were exploring mainly onshore basins. HRT (now PetroRio) and Petra Energia were some of these companies, which assumed the risk of exploring the Solimões Basin (in the Amazon region) and the São Francisco Basin (in the state of Minas Gerais), respectively. As new E&P companies, the access to credit was a fundamental factor to be dealt with in order to develop their projects, because there was no cash flow available for the required investments. With the international crisis in 2009, these companies suffered from credit scarcity and, as a consequence, the projects were abandoned or slowed down[18].

[16] Just as a way of comparison: according to BP (2015), the Brazilian proved oil reserves by the end of 2014 were 19.2 billion barrel oil equivalent. Thus, the estimate claimed by OGX was considerably high.

[17] It is a large scale field located near to already built electricity transmission lines connected to the national network.

[18] See: http://exame.abril.com.br/mundo/noticias/mais-da-metade-das-desistencias-da-11a-rodada-e-do-parnaiba.

Petrobras in Trouble

Petrobras is already a relevant oil company, but the expectation was that it could become a major player in the international oil market. Since the discoveries of the Pre-salt, Petrobras itself assumed this new role by planning heavy investments over the years. Figure 2 showed the investment made by the company in different areas. It is clear that total investment increased expressively in the last ten years. The most relevant investment was made in E&P and refining activities: E&P is the flagship of Petrobras, and refineries were a way for Petrobras to stop relying on the imports of oil products. After 2010, the investment increased to levels almost 10 times greater than the levels undertaken between 1997 and 2001, from about US$ 4 billion to more than US$ 40 billion per year (Medeiros, 2016).

Keeping the level of investment illustrated in Figure 3 was not feasible. Petrobras' debt soared after 2010, as the company free cash flow was not sufficient to finance a large part of the investment. One of the mains reasons why Petrobras could not finance its investment was the government decision to halt increased in the price of gasoline in order to keep inflation down[19].

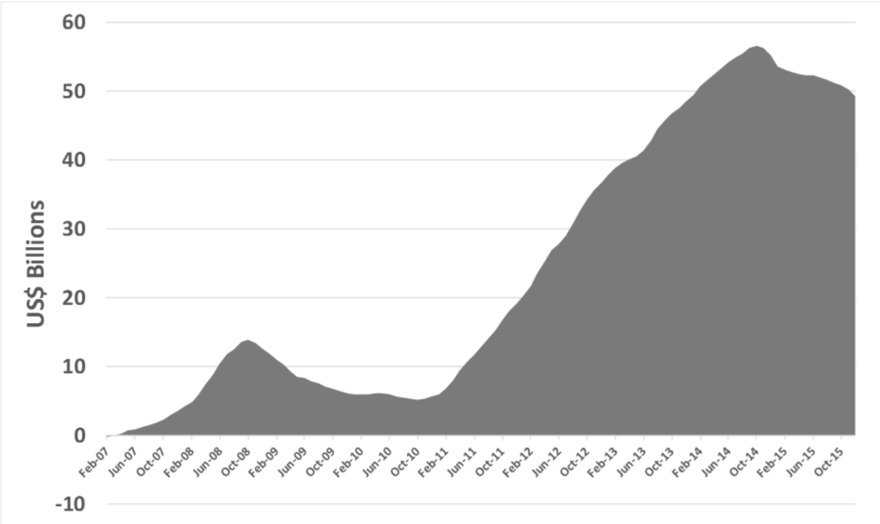
Until 2011, the price strategy for oil products was directed to a long-term convergence to international prices levels. The daily volatility of international prices would not influence the domestic prices immediately, but eventually domestic and international prices would be realigned.

After 2011, however, the Brazilian government decided to keep oil products prices under control to avoid pressuring the inflation rates, which were gaining pace at the time. Although oil products' prices are not subject to price regulation in Brazil, the government may set their levels indirectly via Petrobras, as the company controls 98.2 percent of the refining capacity of the country (ANP, 2015).

From 2011 to 2013, gasoline and diesel prices in Brazil were lower than the international levels. Additionally, the demand and imports of these fuels rose considerably in Brazil after 2010, following the growth in GDP. This had huge impacts for Petrobras and for the organization of the downstream segment in Brazil (Almeida and Oliveira, 2015). As Petrobras imported gasoline and diesel to meet the Brazilian market demand over the last years, it created a direct impact on its finances, as the importation cost was higher than the domestic sale price. Thus, a major impact on the company's finances is due to the "hand over" income, i.e., the difference between the observed

[19] In Brazil, the Minister of Finance and the chairman of Petrobras board of directors were the same person, which seemed to be a conflict of interest. This happened indeed in the last years.

income and the one it would obtain, if the international prices were set in the Brazilian market (Figure 3).



Source: Elaborated by the authors with data from Petrobras.

Note: Hand over Income = Oil product sales * international prices - Oil product sales * observed domestic prices. It includes Gasoline, Diesel and LGP.

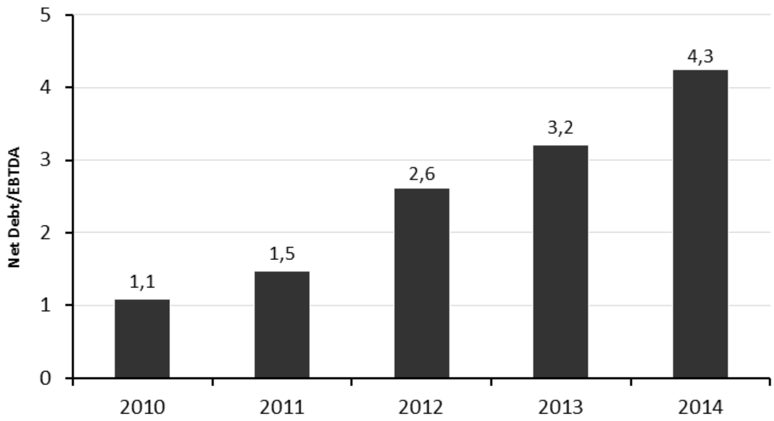
Figure 3. Hand over Income of Petrobras due to the Government Control of Oil Products' Prices.

This hand over income was determinant to the Petrobras financial deterioration. It was almost equivalent to the increase in Petrobras' net debt during the same period. In other words, if oil products' prices were aligned with the international benchmarks, there would be no need for Petrobras to increase its debts during 2011-2014. In this scenario, Petrobras would be in a much better condition to invest in the Brazilian Pre-salt oil explorating ventures.

On the other hand, after the end of 2014, when the price of oil products came down internationally following the crude oil price trends, Petrobras kept its prices levels virtually unchanged in the domestic market and started to recover part of its losses. It should be pointed out that the margins of the downstream operations are generally smaller than E&P, regardless of the presence of price controls.

Most of the impressive amount of investment undertaken by the company over the last years was based on debt. Figure 4 shows the Debt Index for

Petrobras calculated from 2010 until 2014. This index is used to evaluate the leverage of the companies, and Petrobras's index is very high for oil companies' standards[20]. This fact, in addition to the late release of the audited 2014 results and the cash flow problems faced by the company in the last years caused a reduction in the rating of Petrobras.



Source: Elaborated by the authors with data from Petrobras.

Figure 4. Debt Index for Petrobras (2010-2014).

Other example of a negative intervention made by the government in the management of Petrobras was the default that Petrobras had to accept from Eletrobras, the Brazilian state-owned electricity company. Eletrobras did not pay almost R$ 9 billion related to fuel supplied by Petrobras to its thermal power plants[21].

The economic situation faced by Petrobras deteriorated specially after the corruption scandal that commenced by the end of 2014. The Federal Police secretly launched the "Operation Carwash" on July 2013, and it was initially focused on illegal money brokers. In 2014, a senior executive of Petrobras and a money dealer entered into a plea agreement and started to cooperate with the investigation. The investigation initially was concentrated on the scheme involving several EPC companies and shipyards, local and international,

[20] For instance, in 2016, although Shell was downgraded by the rating agency Moodys, its ratio of net debt/EBITDA was estimated at about 1.8. See: https://www.moodys.com/research/Moodys-downgrades-Royal-Dutch-Shell-to-Aa2-negative-outlook--PR_346324.

[21] See (in Portuguese): http://oglobo.globo.com/economia/petroleo-e-energia/eletrobras-vai-pagar-divida-bilionaria-com-petrobras-15626224.

which had contracts with Petrobras. By 2016, the case reached other Brazilian companies and several high level political figures.

The "Operation Carwash" has already discovered R$ 23 billion (US$ 8.9 billion) in suspect payments, bribes and kickbacks paid by Petrobras to politicians. In addition, as the investigation continues, more evidence of corruption and embezzled money should surface. The consequences to Petrobras have been disastrous, and the company is in a process of recovering its management and market confidence. Due to the investigation and the need to account for the damages cause by the payment of bribes, Petrobras' auditor refused to sign-off on the company's third quarter earnings in November 2014. The late issuance of the financial results and the other consequences of the unveiled corruption scheme to the company are detailed in the next section.

Petrobras' Financial Statements for 2014: Discovering the Damage

Petrobras 2014 results were released in 2015. The net loss in 2014 was valued at US$ 7.5 billion, which was explained by different factors, including the "Operation Carwash" and impairment issues. The net loss due to write-offs associated with overpayments indicated by such operation was estimated at US$ 2.5 billion. This value was calculated considering 3 percent of the payments associated to contracts signed with 27 companies accused of taking part in the reported corruption scheme. These contracts were signed between 2004 and 2012.

Table 2. 2014 highlights of Petrobras Results

Financial
• *Sales revenues: R$ 337.3 billion* • *Net results of (-) R$ 21.6 billion* • *Adjusted EBITDA of R$ 59.1 billion* • *Investments of R$ 87.1 billion*
Operational
• *5 percent growth in total oil and natural gas production, reaching 2,669 thousand boe/day in 2004* • *Proven reserves of 16.6 billion boe (ANP/SPE criteria)* • *Refining throughput (Brazil and abroad) of 2,269 thousand bbl/day, the 6th larger oil refiner worldwide*

Source: Petrobras (2015a).
Note: the exchange rate in the end of 2014 was R$ 2.66/US$.

However, the most relevant loss was generated by the impairment tests that resulted in negative adjustments, totaling US$ 16.8 billion. Three sources of loss in different sectors of the company were indicated: i) Refining: the postponement of projects and assets; ii) E&P: the decline in oil prices; iii) Petrochemical: the reduction in demand and margins. Table 2 highlights Petrobras's main results in 2014.

The consequences were vast and complex. Petrobras late audited balance sheet led to a growing lack of market confidence, thus reducing the company credit ratings and damaging its reputation.

Suppliers and partners of Petrobras have been suffering significant constraints and financial problems. Nearly two dozen companies were banned from doing business with Petrobras, since they were involved in the corruption scandal. Some of these companies (e.g., Engevix, OAS, UTC and Mendes Júnior) were responsible for the construction of the Floating Production, Storage and Offloading (FPSO) vessels and drilling rigs for the Pre-salt projects. Most of these companies have been facing credit restrictions, which may delay the construction projects, and some of them have been dealing with judicial recovery processes.

The company created to be responsible for the leasing of deep offshore drilling rigs, Sete Brasil, was deeply affected by the "Operation Carwash" corruption scandal. The company initially had a portfolio with 29 drilling rigs that would be constructed by five different shipyards and operated by seven different operators. Due to the shortage of resources, Sete Brasil has delayed the payment of the contracted shipyards, which has slowed down the construction of the drilling rigs. Sete Brasil filed for judicial recovery (chapter 11) at the end of April 2016 (Valor Econômico, 2016).

Petrobras: New Strategies to Face the Crisis

In March 2015, Petrobras announced its new Business Plan for the period 2015-2019 with very important changes to the company's strategy in dealing with the new global landscape for the oil and gas market, as well as its own financial problems. The most important change in this new business plan was a 37 percent reduction in the expected investment commitments. The total investment was reduced from 220 billion dollars, as indicated in the 2014-2018 Plan, to 130 billion dollars stated in the new 2015-2019.

The total US$ 90 billion reduction was divided as follows: US$ 45.3 billion was cut from the upstream activities; US$ 25.9 from the downstream (mainly in the refinery sector); and US$ 19 billion from the other areas (gas, power and international). Investment in this new plan is concentrated on E&P projects, accounting for 83 percent of the total estimated investment. Petrobras will focus on oil production projects in Brazil, particularly in the Pre-salt area.

In addition to a very significant reduction in the level of investment, Petrobras launched the most important divestment plan of its history. The company plans to divest US$ 57 billion by selling assets and restructuring business areas. This aggressive divestment plan is part of a strategy to reduce leveraging, preserve cash flow and prioritize investments in the oil and gas production in areas of high productivity and return in Brazil.

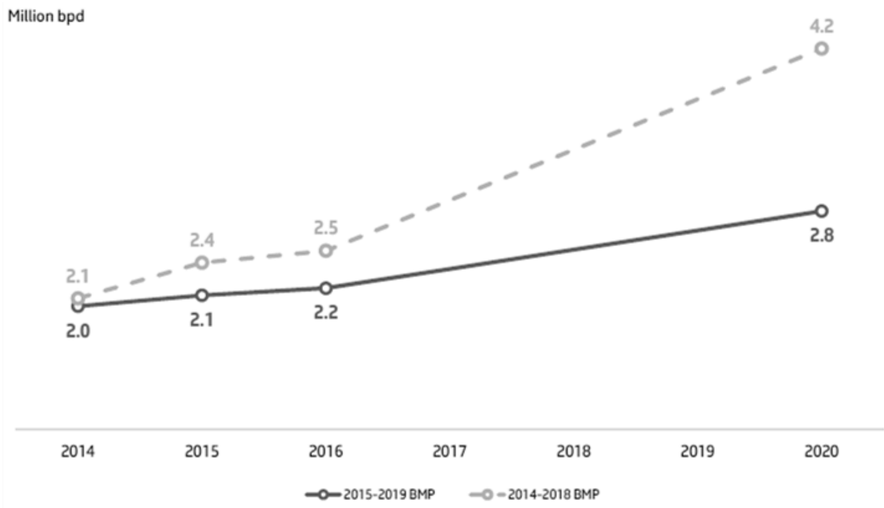
From this US$ 57 billion in total divestment, US$ 15.1 billion is planned to take place in 2015 and 2016. This short-term disinvestment plan includes E&P assets in Brazil and abroad (30 percent), as well as assets in the downstream (30 percent) and in its gas and energy segments (40 percent). The company is already focused on achieving its divestment goals and has started to sell its shareholdins in its gas distribution and transportation subsidiaries. Its operations in Argentina (PESA) has been already sold to Pampa Energia. The Presalt Carcara field was sold to Statoil for US$ 2.5 billion and the Southeast gas transport network was sold to Brookfield for US$ 5.2 billion.

If the company reaches its objective to divest US$ 57 billion, the Brazilian oil and gas landscape will change significantly, with a much more diversified industry structure and higher degree of competition. However, the main challenge is to close satisfactory deals in an unfavorable business environment due to low oil prices. Political opposition from unions and left wing parties can be a significant obstacle to accomplish this divestment plan.

Another important issue dealt with in the new plan is the adjustment of the offshore projects and its consequence to Petrobras's production curve. As discussed in the previous section, Petrobras supplier companies are facing severe financial problems and postponing the delivery of FPSOs. Furthermore, some of the not yet tendered platforms were rescheduled to be delivered after 2020, exceeding the planned horizon.

Actually, some of these projects that are not in production yet may be sold to help achieving the divestment plan related to the E&P segment. These combined factors explain the difference between the oil and natural gas liquids (NGL) production curve observed in Figure 5.

Petrobras is trying to recover from the crisis and to return to its previous growth path. However, the current Brazilian weak economy and political crisis scenario may lead to more challenges ahead for the company. The recent political instability with the impeachment of the president may affect negatively investors' decision to direct resources to Brazil, despite the opportunities brought by the announced divestment plan. Recently, Brazil lost its investment grade and, as a consequence, Petrobras and many other companies have also been downgraded. As a result, access to credit, which Petrobras still depends on to achieve its investment goals, may remain limited.



Source: Petrobras Business and Management Plan 2015-2019.

Figure 5. Comparison between the previous and the current Petrobras Business Plan– Oil and NGL production in Brazil (mb/d).

THE CHALLENGES TO RESTORE THE INVESTMENTS IN THE BRAZILIAN OIL INDUSTRY

The discovery of the Pre-salt bounty has repositioned Brazil in the world oil and gas market. From an oil and gas importer, Brazil has now the potential to become a large oil exporter. Nevertheless, to turn this potential into reality, Brazil will have to face important challenges to mobilize the necessary financial, technological and human resources.

After the liberalization of the oil and gas industry in Brazil in 1997, the level of investments increased at a very fast pace until 2013. The dynamics of investments during this period was characterized by an increase in the number of players in the industry, with maintaining Petrobras its dominance. A new phase of the Brazilian oil and gas is now about to start, as production increases and more players enter the market. The huge geological potential of the Pre-salt, on the one hand, and the important economic crisis of Petrobras, on the other hand, may give a larger role for private players in the Brazilian oil and gas sector.

The increase in oil production by other companies in Brazil will likely reduce Petrobras' market share, which is 90% nowadays[22]. This growing competition may be beneficial to the country in some relevant matters, such as: i) the reduction in the monopsony power of Petrobras, making suppliers less vulnerable to one dominant buyer; ii) the lower government vulnerability (at all levels) regarding taxes revenue generation; iii) the increased competition may bring more technological development to the industry. However, the role of private investors will depend very much on the attractiveness of Brazilian upstream in the new oil and gas global landscape, as discussed in the next section.

Brazilian Upstream Attractiveness in the New Global Landscape

Attracting private investment depends on the cost of production vis-à-vis the expected oil price. In this regard, Brazil is facing a completely new business environment since 2014. It is more and more clear that it will be very difficult for the international oil market to come back to a high price equilibrium, at least in the coming years. IEA (2016) states that the oversupply of oil in 2014 and 2015 tends to continue in 2016. Only after 2017, IEA (2016) sees a better balance between the supply and demand of oil, but the huge accumulated inventories tend to slow down the oil price recovery. This market environment will be very challenging for Brazil due to the high production cost of the Pre-salt.

Pre-salt reservoirs are mainly at ultra-deep offshore zones (about 2,000 meters in water depth), with a salt layer that can reach up to an additional 2,000 m. Drilling cost in this harsh environment is extremely expensive. The first oil wells in Pre-salt cost 200 million dollars. This cost has been reducing due to the improvement in technology and operation skills, but it is still very high (about US$ 80 million per well).

[22] In Norway, an oil market with a historically high level of State participation, the counterpart of Petrobras, Statoil, is responsible for about 70 percent of the oil and gas production of the country (Statoil, 2012). Thus, the state oil company in Norway has a market share which is almost 20 percent below Petrobras' market share.

Since drilling cost represents a large share of the total cost, the well productivity is essential to determine the economic viability of the Pre-salt oil production. As mentioned previously, the well productivity in the Pre-salt fields, developed so far, is very high. Table 3 shows the average well productivity for the platforms in operation. The well productivity in the Pre-salt area can be from 3 to 5 times higher than the post-salt fields. Nevertheless, Petrobras has stated recently that the breakeven price for the Pre-salt fields located at the Santos Basin Cluster is 54 dollars per barrel (Petrobras, 2015b).

Table 3. Well productivity in the largest oil platforms in Brazil, by oil production (April 2015)

Platform	Nº of wells	Tota Platform Production (b/d)	Production per well (b/d)
Petrobras 52	17	145,376	8,552
FPSO Cidade de Paraty	6	119,85	19,975
FPSO Cidade de São Paulo	5	119,817	23,963
FPSO Cidade de Mangaratiba	3	98,05	32,683
Petrobras 57	15	96,539	6,436
FPSO Cidade de Angra dos Reis	5	89,697	17,939
Petrobras 53	15	83,561	5,571
FPSO Cidade de Anchieta	5	82,329	$\boxed{16,466}$
Petrobras 54	13	80,592	6,199
FPSO Cidade de Ilha Bela	2	67,883	33,942

Source: Prepared by the authors, based on MME.

Note:

Presalt in Santos Basin

Presalt in Campos Basin

Considering that Petrobras is now focusing on the best fields in the Pre-salt area, costs associated with the breakeven of 54 dollars per barrel can be considered relatively high to guarantee its competitiveness in the current market environment. Therefore, it is very important that players involved with Brazilian Pre-salt exploration and the government work together to reduce production costs. The reduction in production cost in Brazil will require important changes in companies' strategies and government regulations concerning the oil and gas market in Brazil.

The Roadmap to Increase the Brazilian Upstream Attractiveness

In order to allow significant cost reductions in E&P and to foster investors' interest to the Brazilian upstream activities, it is fundamental to deal with several economic and regulatory challenges, such as: i) the elimination of the current restrictions for private companies to operate in the Pre-salt zone; ii) the reduction of the regulatory costs associated with the diversity of contracts and fiscal regimes in the same field; and iii) the improvement in local content regulatory framework.

The excess of market power by Petrobras has the following consequences: i) hinder competition in the upstream, discouraging cost reduction and technological development; ii) keep the supply chain depending on single operator (monopsony power); iii) reduce the level and the pace of investments in the Pre-salt area, and thus impacting negatively on the country's investment, on the stability of the taxes revenue generation (in all levels of government) and also on the market for suppliers.

The Ministry of Finance in Brazil highlighted the excess of the country's investment dependency to Petrobras, reinforcing our argument of vulnerability. The Ministry also estimated that a reduction of 33 percent in Petrobras investments announced recently could have a negative impact of more than 2 percent of the GDP in 2015 (Brasil, Ministério da Fazenda, 2015).

The lack of competition in the upstream may significantly reduce the pace of technological development and cost reduction. As different operators compete, alternative technological approaches and strategies may be confronted, possibly vertical supply chain competition could occur, thus accelerating the pace of innovations. In addition, diversifying the market share of operators contributes to reducing the market risk for the players in the supply chain. With a more competitive E&P of O&G, in terms of operators (market share), the suppliers can better diversify their supply risk in dealing with the end costumers.

As Petrobras currently faces huge financial restrictions, the company cannot afford to increase its upstream projects' portfolio. Therefore, it will be key to attract new investors to Pre-salt exploration. This was the main reason behind the decision to allow other players besides Petrobras to operate in Pre-salt areas.

Some of the Pre-salt oil fields that are currently under a concession regime are subject to unitization[23] of their resources. Part of the identified resources is located out of the ring fence in non-contracted areas, therefore under the ownership of the Federal Union. By law, those resources must be developed under a PSC. It is necessary to combine two fiscal regimes for the same oil field, which creates uncertainty and reduces the economic attractiveness of those projects.

As PPSA, which represents the state interests in a PSC, has no assets, it cannot invest in those fields. The concessionaries must recover their investment relative to the non-contracted area in oil when production starts. However, they cannot include the financial cost of anticipating capital. Therefore, the internal rate of return (IRR) is much lower than in projects where unitization is not necessary. IRR is also affected by projects' delays.

Regulatory uncertainties about how to make the unitization process in fields with two types of fiscal regimes is hindering negotiations for farmout agreements, which are crucial for allowing Petrobras and some partners to raise capital to the large investments required for developing these fields.

About 8.5 billion barrels of total oil reserves in the Pre-salt fields are subject to unitization on those terms (part under concession plus part without a contract)[24]. The total investment in those projects is US$ 92 billion. The average share of reserves that are not under a contract is 13 percent, but it reaches 50 percent in some fields. The estimated IRR on projects where 20 percent of resources is out of the ring fence is 7.2 percent, which suggests that if the rules are not changed some projects will not take place. On the other hand, it seems that the regulatory agency ANP and the Ministry of Mines and Energy are already implementing some initiative sto facilitate this unitization process (Brasil, Ministério da Fazenda, 2016)[25].

Local content policy also plays a fundamental role in the Brazilian O&G market. Local content policy is the way the country is seeking to integrate the

[23] Unitization is a process of considering an area, usually a reservoir, which is explored under different leases or licenses into a single one, thus defining a single operator and enhancing the performance of production of O&G (Inkpen and Moffett, 2011).

[24] Botelho et al. (2016).

[25] Other relevant actions that the authorities claim to be working on are: extension of the contracts for the concession round zero that would end in 2025 for most fields; and stimulus to operators to restart production in fields where they were interrupted. It is estimated that the two actions should generate investments of about US$ 80 billion during the next decade (Brasil, Ministério da Fazenda, 2016).

expansion of the oil sector with its industrial development. Despite the significant evolution of this policy in recent years, several improvements are needed to promote industrialization with competitiveness in the oil exploration and production activity in Brazil.

Investors in Brazilian have increasingly perceived local content regulations in upstream as an important risk factor. Local content commitments are assumed long time before the acquisition of the products and services for the projects. The risk of cost overruns, delays, and low quality of products is a key consequence of the current policy, thus affecting the attractiveness of Brazil's oil and gas industry.

Until April 2015, ANP has levied 86 fines for noncompliance of the local content requirements, totaling around US$ 90 million in fines (Medeiros, 2016). However, operators expect that the amount of fines could reach more billions of dollars if the current compliance methodologies present in the regulation are not modified.

Operators argue that it is not feasible to comply with all the local content commitments in the current market context of suppliers in Brazil. Upstream investment increased by 5 fold between 2003 and 2013. Some segments of domestic oil and gas supply industry were not able to cope with this demand growth. In addition, the current macroeconomic slowdown and the corruption scandal involving Petrobras and its suppliers have contributed to weaken even more the domestic supply chain.

Since the last quarter of 2013, the Brazilian government has been discussing alternatives for overcoming the side effects of the local content regulation on oil and gas projects. Operators and other industry players have submitted suggestions to ANP. However, the local content regulation was not changed for the 13th bidding round in 2015. Improvements in the current local content policy could include the following directions.

In order to increase the Brazilian upstream competitiveness, it is fundamental to pursue a better balance between the local content incentives and the risks of oil projects. In this sense, it is essential to design and implement mechanisms to allow flexibility in establishing local content commitments. As an alternative to the current processes, where the full commitment is established in the bidding round, a more successful model would allow instead the definition and refining of local content targets in connection with the preparation of the production development plan. Thus, companies could make feasible commitments based on the best knowledge of their goods and services needed taking into account the context of the supply chain in the domestic market.

The LC commitments could be individually addressed in the exploration and production phases. A previous negotiation between the government and operators would allow the parties to have strict control over the compliance of their commitments. Government and operators would discuss the development plans regarding the blocks, taking into account the strategies related to local content. Specifically, they could analyze the best manner of balance the efficiency related to the projects' costs and the participation of local supplier companies with high value added activities.

With the inclusion of local content as an important dimension of the field development plan, ANP would strengthen its monitoring capacity by knowing the local content strategy of the concessionaire. Moreover, ANP would have a better understanding of the local industry goods and services constrains by empowering itself to address the task of assessing applications for companies with exemptions based on the local purchase commitments, due to the lack of competitive offerings as provided by the legislation.

The purpose of the current local content policy is to punish companies that do not comply with minimum requirements, thus representing an institutional risk to the E&P projects. However, the outcome could be improved by introducing new incentive mechanisms – strengthening the local content policies via incentives to companies that exceed their commitments. Several incentives could be considered as a way to encourage operators to seek increasing levels of local content, such as: (i) reduction of specific duties; (ii) competitive advantages in the bidding rounds for exploration blocks; and (iii) reduction in fines for noncompliance with commitments made in other concession contracts (Almeida and Martinez-Prieto, 2015).

ANP could take into account the performance of the local suppliers when awarding new contracts or renewing the existing ones. This initiative would allow operators with higher percentage of local content to overcome the governmental goals of LC and therefore be favored in subsequent bidding rounds. Similarly, different incentives could be defined with local suppliers that reach the levels of competitiveness expected by the government, such as discounts in the governmental participations.

Other important point is the new discussion and simplification of the rules on the local content, including high transparency, reducing the costs and red tape. It is necessary to simplify the local content certification process, including the preparation of the spreadsheet, where the real local content of E&P activities is assessed, and discuss with operators the possibilities of negotiation about the contractual commitments, when the local suppliers are not provided with the infrastructure to compete with their foreign pairs. This important instrument of waiver clause already exists, but it seems that it has not been used so far.

Finally, it is important to consider the contracts in force. Many of them have included local content commitment levels that are not attainable and/or may result in high costs for E&P projects, given the current context of the goods and services market in Brazil. In this scenario, a pragmatic approach to identify the best path for the country is warranted. The simple application of heavy fines for noncompliance will certainly not contribute towards accelerating the development of a competitive goods and services industry in Brazil.

Currently, the fines must be collected by the Treasury, and they do not generate any benefits for the supply industry, as the money has no defined destination. It is important to discuss how to turn possible contractual penalties into mechanisms that encourage the development of the goods and services industry. Possible fines could turn into investment commitments in the supply chain, including research and development and contributing to solve the bottlenecks in the supply chain, either by the operators themselves or by independent organizations in suppliers' development programs.

Some studies suggest that for the LC to be more productive, it should focus on the supply chain segments with higher global competitive potential (PWC, 2015). Additionally, local purchases directed to projects overseas could be considered for rebates in the commitments to projects developed in Brazil. Such mechanisms would help reduce the potential fines and promote exports of domestic goods and services.

The Downstream Challenge

Although international and domestic prices are closer nowadays, the pricing strategy for oil products is still an issue for Brazil. Due to the financial deterioration and the need to focus on upstream investments, Petrobras decided to review its planning for new refineries. Petrobras gave up two refineries projects (Premium I and Premium II) and postponed the conclusion of Comperj and Abreu e Lima refineries. Thus, the Brazilian refining capacity will stall and the imports of oil products will continue to increase.

The entrance of private partners would be the solution to expand refining capacity in Brazil. However, no investor would be attracted to a segment where prices can be lower than the benchmark price for long periods. Therefore, it is essential to define a new transparent price mechanism to oil products, avoiding the use of those prices to control inflation. It is important that the new pricing mechanisms allow refineries to align their prices with the international prices of oil products, so that the refining market risk becomes bearable to private investors. A deeper step would by to change the legislation in order to improve Petrobras's governance practices.

With this type of change in regulation, Petrobras would be able to raise more capital to invest in the Pre-salt areas by selling participation on existing refineries, in addition to finding partners to help the company cope with the huge amount of capital necessary to complete the refineries under construction.

CONCLUSION

The discovery of the Pre-salt resources in the late 2000s increased tremendously the growth potential of the oil and gas industry in Brazil. From an oil importer, Brazil has a potential to become an important oil exporter, if the country manages to create an attractive business environment for private capital. Once this is materialized, it should create positive spillover effects on the sector -- on suppliers, on technological development, on tax generation and on the country's investment and economy.

The current international and domestic oil sector context is cutting down the investment required to tap the Pre-salt bounty in Brazil. The collapse of the international oil price has reduced the Pre-salt attractiveness and the availability of financial resources for private investors. In addition, Petrobras financial crisis, partially created by the "Carwash" scandal and partially by bad management, caused a dramatic reduction in the company´s ability to invest.

Given Petrobras' severe financial crisis, private players will have larger role in the Brazilian oil and gas sector, at least in the middle term. However, private investors' participation will depend very much on the attractiveness of the Brazilian upstream sector. This chapter has presented and discussed the necessary policy changes to increase the Brazilian upstream attractiveness.

We consider that the roadmap to increase the Brazilian upstream attractiveness and boost the investment in the oil sector requires the following actions: i) reduction of the restrictions to the private capital in Pre-salt exploration; ii) reduction in the regulatory costs associated with the diversity of contracts and fiscal regimes in the same fields; and iii) improvement in the local content regulatory framework. In the downstream segment, there is also the challenge to revise the pricing mechanism, thus allowing more players to enter in the segment.

Therefore, if these measures are implemented successfully, we believe that they should considerably contribute to restore oil investments in Brazil in a sustainable way, with a more diversified and competitive industry.

REFERENCES

Almeida, Edmar, and Martinez-Prieto, Diana. 2015. "The Impacts and Effectiveness of Local Content Policy in Oil Exploration and Production in Brazil." Oxford Energy Forum: 28-30. Accessed: October 12, 2016. https://ora.ox.ac.uk/objects/uuid:eb005044-d9c2-47e2-8295-684bb963 fb02.

Almeida, Edmar, and Oliveira, Patricia. 2016. "Controle de Preços da Petrobras: Chegou a Hora de Pagar a Conta." Infopetro, March 28. https://infopetro.wordpress.com/2016/03/28/controle-de-precos-da-petrobras-chegou-a-hora-de-pagar-a-conta/.

ANP. 2015. "Anuário Estatístico." Last modified July 29. http://anp.gov.br/?pg=78136&m=&t1=&t2=&t3=&t4=&ar=&ps=&1466487549703.

Bordoff, Jason, and Losz, Akos. 2015. "Oil Shock - Decoding the Causes and Consequences of the 2014 Oil Price Drop." *Horizons Journal*, Spring 3: 190-206.

Botelho, Felipe, and Losekann, Luciano and Braga, Yanna. 2016. "As Oportunidades e os Desafios Econômicos do Processo de Unitização no Pré-Sal." Paper presented at the X Congresso Brasileiro de Planejamento Energético – CBPE, Gramado, Rio Grande do Sul, September 26-28.

BP. 2015. "BP Statistical Review of World Energy." Accessed: October 12, 2016. www.bp.com/en/global/corporate/energy-economics/statistical-review-of-world-energy.html.

Brasil. Ministério da Fazenda. Secretaria de Política Econômica. 2015. "Impactos da Redução dos Investimentos do Setor de Óleo e gás no PIB." Nota à Imprensa. Accessed: October 12, 2016. http://www.spe.fazenda.gov.br/noticias/fffcimpactos-da-reducao-dos-investimentos-do-setor-de-oleo-e-gas-no-pib.

Brasil. Ministério da Fazenda. Secretaria de Política Econômica. 2016. "Medidas Regulatórias no Setor de Óleo e Gás Elevarão o Crescimento da Economia." Accessed: October 12, 2016. http://www.spe.fazenda.gov.br/noticias/nota-sobre-medidas-regulatorias-no-setor-de-oleo-e-gas.

Instituto Brasileiro de Petróleo, Gás e Biocombustíveis - IBP and Confederação Nacional da Indústria - CNI. 2012. A Contribuição do Setor Brasileiro de Petróleo, Gás e Biocombustíveis para o Desenvolvimento Sustentável no País. Brasília: CNI. http://arquivos.portaldaindustria.com.br/app/conteudo_18/2013/09/23/4970/20131003104228715851e.pdf.

EPE. 2015. Plano Decenal de Expansão de Energia. 2024. Brasilia: MME/EPE. http://www.epe.gov.br/PDEE/Relat%C3%B3rio%20Final%20do%20PDE%202024.pdf.

Infomoney. 2013. "Relembre a Trajetória da OGX, da Fundação à Recuperação Judicial." Infomoney, October 30. www.infomoney.com.br/ogxpetroleo/noticia/3027950/relembre-trajetoria-ogx-fundacao-recuperacao-judicial.

Inkpen, Andrew C., and Moffett, Michael H. 2011. The Global Oil and Gas Industry: Management, Strategy and Finance. Oklahoma: Pennwell.

International Energy Agency - IEA. 2013. World Energy Outlook - WEO. Paris: IEA.

International Energy Agency - IEA. 2015. World Energy Outlook - WEO. Paris: IEA.

International Energy Agency - IEA. 2016. Medium-Term Oil Market Report 2016. Paris: IEA.

Mcnally, Robert. 2014. "Stability and Volatility in the Global Oil Market: Past, Present, and Future." Paper presented at the workshop The Changing Global Energy Landscape: Impacts for Brazil, Rio de Janeiro, Rio de Janeiro, November 12.

Medeiros, Breno, B. 2016. " Policies to Develop the Oil and Gas Supply Chain: Essays with Emphasis on the Lessons and Contributions from the Norwegian Experience to Brazil." PhD diss., Federal University of Rio Grande do Sul. www.lume.ufrgs.br/handle/10183/139340.

PETROBRAS. 2015a. "2014 Financial Statements (USD)." Accessed: October 12, 2016. http://www.petrobras.com.br/en/news/check-out-the-publication-of-our-audited-financial-statements.htm.

PETROBRAS. 2015b. "Petrobras Update." Accessed: October 12, 2016. http://pt.slideshare.net/petrobrasri/visao-geral-da-petrobras.

PWC. 2015. "Mudar é preciso. Revistando o Conteúdo Local para a Indústria de Óleo e Gás." Accessed: October 12, 2016. www.strategyand.pwc.com/reports/mudar-e-preciso.

Reuters. 2008. "Custo de extração do pré-sal é 'econômico', diz Petrobras." Acessed October 27, 2016. http://g1.globo.com/Noticias/Economia_

Negocios/0,MUL733262-9356,00-CUSTO+DE+EXTRACAO+DO+PRESAL+E+ECONOMICO+DIZ+PETROBRAS.html.

Statoil. 2012. "Annual Report 2012." Accessed: October 12, 2016. www.statoil.com/annualreport2012/en/ouroperations/businessareas/developmentandproductionnorway/pages/developmentandproductionnorway(dpn).aspx.

Sete Brasil. 2015. "Management Presentation. Rio de Janeiro, March 2015." Accessed: October 12, 2016. http://ri.setebr.com/enu/34/SeteBrasilManagementPresentation_Mar2015_eng.pdf.

US Energy Information Administration - EIA. 2015. "Year in Review: Crude Oil Prices 2014." Accessed: October 12, 2016. www.eia.gov/finance/review/annual/.

Valor Econômico. 2016. "Sete Brasil Protocola Pedido de Recuperação Judicial no TJRJ." Valor Econômico, April, 29. www.valor.com.br/empresas/4544495/sete-brasil-protocola-pedido-de-recuperacao-judicial-no-tj-rj#.

Viegas, Thales. 2011. "Exploration and Production of Oil and Gas in Deep Waters: Evolution and Trends – Part II." Infopetro, October, 24. https://infopetro.wordpress.com/2011/10/24/exploration-and-production-of-oil-and-gas-in-deep-waters-evolution-and-trends-part-ii/.

Yergin, Daniel. 1991. The Prize: The Epic Quest for Oil, Money and Power. New York: Free Press.

In: Infrastructure Investments
Editors: Gisele Ferreira Tiryaki et al.
ISBN: 978-1-53610-792-0

Chapter 5

NATURAL RESOURCES AND ECONOMIC GROWTH: AN ANALYSIS OF THE MUNICIPALITIES LOCATED AT THE STATE OF BAHIA, BRAZIL

Luiz Gustavo Araújo da Cruz Casais e Silva, André Luís Mota dos Santos and Rodrigo Carvalho Oliveira*
Department of Economics, Universidade Federal da Bahia
Salvador, BA, Brazil

ABSTRACT

The abundance of natural resources and its relationship with economic growth has been documented by the theoretical and empirical literature. This literature maintains that natural resources may be a blessing or a curse depending on the robustness of a country's institutions. We followed the recommendation of van der Ploeg (2011) and use an intra country panel data set in order to deal with multicolinearity and simultaneity. The analysis is restricted to 417 municipalities located at the state of Bahia, Brazil. The data cover four intervals of three years: 2002-2004, 2005-2007, 2008-2010, and 2011-2013. We use a two-step generalized method of moments system

* Corresponding author: André Luiz Mota dos Santos. Email: andrelms@ufba.br.

estimators. Our focus is on the abundance of two types of natural resources: oil and natural gas; additionally, water resources are discussed. We try to contribute to the previous literature in two ways: by introducing a new variable to proxy for institutional quality, and by introducing a renewable resource. The results are compatible with the interpretation that the abundance of natural resources has a positive effect on the municipalities' growth if its exploration is associated with better institutional quality. Yet, this effect is only observed for oil and natural gas, not for water resources.

Keywords: exhaustible resources, renewable resources, economic growth

INTRODUCTION

The main characteristic of non-renewable resources is the physical limitation of its stock in nature. If the resource is indiscriminately extracted in the present, it will not be available for future generations. Two issues arise from this observation. First, it is important to find out the optimal rate of resource extraction (Gray 1914; Hotelling 1931). Second, it is also necessary to identify a sustainable level of consumption among generations, even after the natural resource is depleted (Solow 1974, Hartwick 1977; Dasgupta, Eastwood, and Heal 1978). Thus, the rule for optimal intertemporal depletion, known as the Hotelling rule, is associated with the rule for intergeneration consumption sustainability, known as the Hartwick rule, transforming the exhaustible natural resource in a reproducible asset (such as physical capital, human capital or foreign assets). Basically, the Hartwick rule states that the total stock of capital (including natural capital) should not diminish over time.

The nature of non-renewable resources is different from that of renewable resources due to a timing characteristic. The replacement rate of a non-renewable resource is too slow in nature that their stock can be considered available only once. A renewable resource, in turn, adjusts its stock level faster through a dynamic system whose complexity varies according to the type of natural resource[1] (Sweeney 1993). The use of renewable resources modifies the original dynamic system. The extent of this modification is variable, and the degree of change is so great in some situations that it becomes difficult to recognize the original system (e.g., livestock or large-scale agriculture).

[1] Systems predator-prey type, reforestation systems and natural cycles are examples of dynamic systems.

It is also possible to establish optimal exploitation rules for renewable resources, although the fact that the renewable resource is not normally depleted does not change the indication for sustainable consumption across generations. Many developing economies, which are rich in both renewable and nonrenewable resources, often move away from these optimization rules. The most trivial economic explanation for this departure from optimality is the existence of poorly defined property rights (Brown 2000). Weak institutions foster rent seeking behavior, which undermines incentives to optimally exploit the resources or to transform exhaustible resources in reproducible assets (Mehlum, Moene, and Torvik 2006; Deacon 2011).

After Sachs and Warner (1995), the expression “resource curse” started being used to characterize the relationship between natural resource abundance and lower growth rate of per capita GDP. Basically, the authors use cross-sectional data to estimate a growth convergence equation controlling for the abundance of resources (measured by the proportion of natural resources in the GDP). They found a negative relationship between resource abundance and growth. Historically, there are as many cases of natural resources being a blessing as there are of them being a curse. Furthermore, the institutional hypothesis has become theoretically and empirically more relevant for explaining differences in economic performance among different countries. Therefore, institutions should also be relevant to explaining the performance of resource-rich countries, although other transmission channels of the impact of natural resources’ abundance on economic growth may still be present.

In the theoretical and empirical literature, the appreciation of the real exchange rate and de-industrialization, for example, continue to be consequences of a resource boom (van der Ploeg and Venables 2012; Mironov and Petronevich, 2015). Theoretically, the Dutch disease is a phenomenon that occurs within a trading model similar to a Heckscher-Ohlin economy (Corden and Neary 1982). Specialization, in the presence of a resource boom, generates trade gains. The problem lies in the adverse effects associated with specialization in the production of natural resources. If the dynamic sector of the economy in terms of innovation and externalities is the one that loses ground with the expansion of the natural resource production, economic growth is hampered. The same happens if the specialization and the volatility of resource revenues lead to restrictions on liquidity and to volatility of the exchange rate (van der Ploeg and Poelhekke 2009; Sinnott, de la Torre, and Nash, 2010, 21-40; van der Ploeg 2011).

Additionally, the effects of the Dutch disease will likely be worse if institutions are weak. Institutions would be relevant to ensure, for example, economic diversification, which reduces product and exchange rate volatility (Sinnott, de la Torre, and Nash 2010). The volatility is also higher in economies with lower levels of financial development (van der Ploeg 2011). Moreover, institutional quality seems to influence the level of financial development (Acemoglu, 2009, 726-28). Thus, institutions would be relevant to pursue optimal rules as those of Hotelling and Hartwick, or to avoid transmission channels which brings about the curse of natural resources' abundance associated with the Dutch disease.

Many authors have tested the resource curse following Sachs and Warner (1995), such as Williams (2010) and James and Aadland (2011), among others. The majority of these works has been heavily criticized for the way in which institutional controls are internalized into the estimated equations, as the estimations likely suffer from reverse causality or omitted variable bias. If an institutional quality index is used interacting with a measure of the abundance of natural resources, the curse may disappear (Mehlum, Moene, and Torvik 2006). If an instrument is used for the abundance of resources, the relationship between natural resources' endowment and economic growth may disappear or even be positive (Brunnschweiler and Bulte 2008; van der Ploeg and Poelhekke 2010). To cope with the omitted variable bias, one solution is to use panel data (see Parente and Prescott 1994; Islam 1995; Caselli, Esquivel, and Lefort 1996). Studies using data within a country have been encouraged to avoid the problems associated with highly correlated explanatory variables, such as the degree of openness and financial development (van der Ploeg 2011).

In principle, the same theoretical insights can be thought of applying countries, regions or municipalities (Caselli and Michaels 2009)[2]. Therefore, the aim of this chapter is to verify whether there is empirical evidence for a relationship between resource abundance and economic growth among municipalities in Brazil, using the state of Bahia as a case study. A new proxy for institutional quality was created, but the lack of data avalability limited the analysis to this single state. We estimate an enhanced version of the convergence equation present in Mankiw, Romer, and Weil (1992) using a panel data set with 417 municipalities and four periods (every three years between 2002 and 2013). To avoid the endogeneity problem, we used the two-step generalized method of moments system estimator developed by Arellano

[2] There are theoretical models that consider regional or local characteristics. See, for example, Takatsuka, Zeng, and Zhao (2015).

and Bover (1995) and Bundell and Bond (1998). We have also introduced interaction terms with the institutional quality indicator.

The impact of the exploitation of natural resources on the income and quality of life of Brazilian cities has already been subject of research. Postali (2007) suggests the presence of a natural resources' curse in municipalities that benefited from oil and gas' royalties due to the enactment of the Law No. 9478/1997. Caselli and Michaels (2009), Postali and Nishijima (2011) and Carnicelli and Postali (2014), in turn, conclude that the production of oil and natural gas and the associated royalties have not contributed to either the provision of municipal services or for economic development. On the other hand, there are no known studies that test the contribution of renewable resources for economic growth at the municipal level.

In addition to oil and natural gas royalties, this chapter uses the royalties from the use of water resources, primarily from water reservoirs for the production of hydroelectricity. Sometimes, the classification of the water stock as a renewable or exhaustible resource is not clear cut. This is the case for the water present in aquifers, for example. Surface water storage reservoirs, however, are generally classified as renewable resources (Sweeney 1993).

The empirical literature has emphasized that different resources may have different impacts on economic growth. For example, rent seeking behavior is normally associated with point-source resources, according to Deacon (2011) and van der Ploeg (2011). Thus, using two different types of royalties allow checking the differences in impact of each resource type, one renewable and another exhaustible, on economic growth.

The results obtained from the econometric analysis are consistent with the interpretation that the abundance of natural resources has a positive effect on the economic growth at the municipal level, although this outcome holds only under certain circunstamces described in detail in the empirical section of this chapter.

Besides this introduction, this chapter has three other sections. The variables and data used in the econometric analysis are described in the next section. The third section discusses the results of the empirical strategy, while the last section provides the concluding remarks.

DATA

Following Durlauf, Johnson, and Temple (2005), we estimated the following empirical equation of growth:

$$\Delta y_{it} = \beta ln y_{i,0} + \gamma X_{i,t} + \pi Z_{i,t} + \varepsilon_{i,t} \quad (1)$$

where Δy_{it} is the growth rate of output per capita; $ln y_{i,0}$ is the natural logarithm of per capita GDP in the initial period; $X_{i,t}$ includes the other traditional determinants of growth (physical capital, according to the Solow model (1956); and human capital, according to the model of Mankiw, Romer, and Weil (1992)); $Z_{i,t}$ represents other determinants of growth. Convergence equations are based on some version of the Solow model with technical progress, so the population growth rate (n), the depreciation rate (δ) and the rate of technical progress (g) are always included. In the simplified convergence equation (1), the sum $n + g + \delta$ is included in $X_{i,t}$. This variable should have a negative effect on growth, while physical and human capital capital are expected to have positive effects. Conditional convergence means that the estimated value of β is significant and negative (Acemoglu 2009, 80-105).

Following Barro (1991) and Mankiw, Romer, and Weil (1992), Sachs and Warner (1995) estimate a convergence equation with cross-section data. They control for the abundance of natural resources by using the share of exports of primary goods in the GDP. The empirical strategy of Sachs and Warner (1995) is slightly modified by Mehlum, Moene, and Torvik (2006). The latter authors add an institutional quality measure as an interaction term: the product of the abundance of resources and institutional quality measures. With a panel of Brazilian municipalities, we included interaction terms similar to the Mehlum, Moene, and Torvik (2006) in $Z_{i,t}$. This is done with two different measures of resource abundance, one of exhaustible resources (oil and gas) and the other from a renewable resource (water in surface reservoirs). We also include institutional quality directly.

The variables are arranged in a four-period panel, each period with intervals of three years: 2002-2004, 2005-2007, 2008-2010, 2011-2013. Although there is no consensus in the literature, using a range of three years appears to balance the trade-off between reducing the influence of short-term shocks in economic activity and making increasing the the temporal dimension of panel data (Bonnefond 2014).

The data are obtained from the state of Bahia municipalities. Bahia is located in the northeast region of Brazil, has the highest GDP of this region and the 7th GDP among the Brazilian states. The first oil well in the country was discovered at the city of Salvador, capital of Bahia, in 1939, in the Reconcavo Basin. Today, the state has the fifth largest domestic production of oil and natural gas in Brazil. In terms of electricity production, Bahia is the largest electricity producer in the Northeast region and the fourth in Brazil. Much of the production comes from hydroelectric power plants located at the basin of the São Francisco River (ANEEL, 2016).

The variables considered for the model estimation are:

$$\Delta y_{it} = f\begin{pmatrix} y_{i,t-1}, n+g+\delta, Investment, Education, \\ Public\ expenditures, Institutional\ quality, \\ Royalty\ 1, Royalty\ 2, Onshore, Interaction \end{pmatrix} \quad (2)$$

The growth rate of per capita GDP, Δy_{it}, is the difference between the natural logarithm of per capita GDP of the final year and the first year of each period. The initial product, $y_{i,t-1}$, is the natural logarithm of per capita GDP in the start year of each period. $Investment$ is the rate of growth in electricity consumption in kilowatt/hour per capita. The electricity consumption is commonly used as a proxy for capital stock. Thus, we take their variation as a proxy investment. The source of these data and the population growth rate is obtained from the Superintendence of Economic and Social Studies of Bahia (in Portuguese, Superintendência de Estudos Econômicos e Sociais da Bahia; SEI, 2016). We follow Mankiw, Romer, and Weil (1992) and assume that $g + \delta = 0{,}05$.

In the Northeast region, only 8.1% of people aged 25 or older had higher education by 2014. In Bahia, the percentage of the population aged 18 to 24 who was enrolled in higher education in 2013 was only 8.7% (Observatório do PNE, 2016). There is a consensus that the numbers are worse in small towns. For this reason, we decided to use the high school education to build a measure of human capital. $Education$ is the participation of the population of formal workers with a high school degree. The data source is the Annual Social Information Report (in Portuguese, Relação Anual de Informações Sociais; RAIS, 2016).

$Public\ expenditures$ is the average share of budget speding in GDP over each three year period. The data source is the file Finance in Brazil (FINBRA), of the National Treasury Secretariat (Secretaria do Tesouro Nacional 2016).

$Royalty\ 1$ is the average share of oil and natural gas royalties to GDP over each three year period. The data source is the National Agency of Petroleum (ANP 2016). The distribution of these royalties is subject to various criteria. For example, municipalities affected by transport operations receive royalties, even if they are not producers. Additional details are available from ANP's website.

$Royalty\ 2$ is the average share of royalties for water resources usage to GDP over the three year period. These royalties are known as compensation for the use of water resources; in Portuguese, Compensação Financeira pela Utilização dos Recursos Hídricos (CFURH). They are basically reimbursements for the occupation of areas by reservoirs of hydroelectric power plants and payments for water use in power generation. The data source is the National Electric Energy Agency (ANEEL 2016).

We additionally created the following dummies: i) $Onshore$ assumes value 1 if the production of oil and/or natural gas is onshore and 0 if production is offshore; $Large$ equals 1 for municipalities whose share of royalties to GDP is above the median, and 0 otherwise. The inclusion of the $Onshore$ variable is intended to verify whether the oil and gas onshore production produces greater positive externalities than offshore production. These externalities are usually associated with increased productivity due to: i) the consolidation of the production chain; ii) increased competition in the market; and, iii) improvement in local infrastructure. The $Large$ variable, in turn, aims at observing the differences in being a major beneficiary of royalties. The data source is the ANP (2016).

$Institutional\ quality$ is a measure based on the Court of Auditors' reports of the State of Bahia municipalities; in Portuguese, Tribunal de Contas dos Municípios do Estado da Bahia (TCM-BA 2016). The TCM-BA oversees the management of the budgetary resources by the state's municipalities. It reports on the budgets of municipalities and establishes three possible views: "rejection," "approval with reservations" or "approval." The construction of the variable for each of the four periods obeyed the following criteria: i) $Institutional\ quality$ has a value of 0 if the opinion of the TCM-BA is "rejection" in at least one of the three years; ii) $Institutional\ quality$ equals 1 if the opinion is "approval with reservations" or "approved" in all three years. Although subject to a common federal legislation (tax liability law), the analysis of the budgets of municipalities is different in each state, because there are specific state regulations. So the variable is built only to municipalities of one state. The variable is a measure of the quality of the municipal budget management, since the Court's opinions take into account each municipality adherence to the following principles of sound fiscal management: i) compliance with deadlines for the establishment of estimates of revenue and expenditure; ii) establishing targets for the fiscal indebtedness; iii) mandatory application of resources to uses established by law (for example, specific tax revenues should be spent only on education); iv) compliance with rules for creating additional credits and reallocation of funds from one programming category to another; v) legality of bidding procedures, contracts and agreements. A potential problem with the variable is that only the final opinion is made available to the public, not the full report.

$Interaction$ refers to four interaction terms:

i) $Royalty\ 1 \times Institutional\ quality$;
ii) $Royalty\ 2 \times Institutional\ quality$;
iii) $Royalty\ 1 \times Institutional\ quality \times Large$;
iv) $Royalty\ 2 \times Institutional\ quality \times Large$.

All monetary values were deflated using the 1997 Consumer Price Index (IPCA). This price index is calculated by the Brazilian Institute of Geography and Statistics (IBGE 2016). All non-categorical explanatory variables are in natural logarithms, except $Royalty\ 1$ e $Royalty\ 2$.

Table 1 presents descriptive statistics of the variables in equation (2). At the top portion of the table, the information refers to the municipalities that do not receive royalties; the bottom part refers to the municipalities that receive royalties. The data set is an unbalaced panel with 417 municipalities, of which 269 receive oil royalties and natural gas and 30 water resources. Municipalities which are not beneficiaries have lower average economic growth, slightly higher level of concentration around the mean of $y_{i,t-1}$, $Investment$, $Public\ expenditures$, $Institutional\ quality$ and standard deviation greater of $Education$ e $n + g + \delta$ variables. The averages of the variables $Institutional\ quality$ and $Public\ expenditures$ are greater for beneficiaries municipalities. The high standard deviations of the $Royalty\ 1$ and $Royalty\ 2$ variables reflects the distribution criteria. For example, the city of Paulo Afonso receives on average 23,000 times more royalties for water resources than the municipality of Mucuri, because much of its territory is flooded by the water surface of a large reservoir.

EMPIRICAL RESULTS

Using panel data allows us to handle two common inconsistencies found in cross-section convergence equations: (i) the existence of omitted variables that are correlated with the explanatory variables; (ii) the existence of endogeneity, since there is reverse causality between economic growth and many of the explanatory variables (Parente and Prescott, 1994; Islam 1995; Caselli, Esquivel and Lefort 1996). Panel data allows us to consider fixed country characteristics that affect both economic growth and explanatory variables. This removes part of the simultaneity bias (Acemoglu, 2009). With panel data, the empirical literature on convergence has made use of developments Holtz-Eakin, Newey, and Rosen (1988), Arellano and Bond (1991), Arellano and Bover (1995) and Blundell and Bond (1998). Based on Roodman (2009) and following Bonnefond (2014), we use the two-step generalized method of moments system estimator with Windmeijer correction method (2005) for the covariance matrix-variance. The instrumentalized variables are $y_{i,t-1}$, $n+g+\delta$, $Investment$, $Public\,expenditures$, $Education$, and $Institutional\,quality$. For the instruments, we use two periods of lag, following the suggestion of Roodman (2009).

Table 1. Descriptive statistics

Variables	Number of Observations	Average	Standard-deviation	Minimun	Maximun
Non-beneficiary municipalities of royalty					
Δy_{it}	621	0,0944	0.1730	-0,6159	0,9351
$y_{i,t-1}$	621	8.1281	0,5220	7.1741	10.7948
Investiment	568	-2.8902	0,8184	-7.0099	-0,3086
Education	616	-2.3965	1.0573	-6.3780	3.2095
$n+g+\delta$	599	-2.9963	0,5913	-8.6529	-1.2687
$Public\ expenditures$	621	-1,4166	0,4425	-3,3306	-0,4922
$Institutional\ quality$	621	0,5668	0,4959	0	1
Beneficiary municipalities of royalty					
Δy_{it}	1047	0,1301	0,2018	-1,0231	1,7395
$y_{i,t-1}$	1047	8,2501	0,5958	7,0920	12,3310
Investiment	923	-3,0999	0,8565	-9,4361	-1,0259
Education	1043	-2,1140	0,8805	-6,2265	0,7798
$n+g+\delta$	1000	-2,9642	0,5748	-8,3248	-1,4850
$Public\ expenditures$	1047	-1,4840	0,4780	-3,5942	0,0011
$Institutional\ quality$	1047	0,4479	0,4975	0	1
$Royalty\ 1$	1003	0,2100	1,0588	0,000058	17,0738
$Royalty\ 2$	122	0,8306	1,5329	0,000047	7,0366
Total	1668				

Several alternative specifications were tested. The most relevant results are in Table 2. The generalized method of moments system estimator assumes that the twice-lagged residuals are not autocorrelated. That estimator requires that there is first-order serial correlation, but that there is no second-order serial correlation in the residuals. For all models, Arellano-Bond tests results point to the presensence of first-order autocorrelation and the absence of second-order autocorrelation in all estimated models. On the other hand, the Hansen test shows that we cannot reject the null hypothesis that the error term is uncorrelated with the instruments just for models 2 and 4, i.e., only for those models the validity of the instruments is confirmed. Therefore, considering the results of both Hansen and Arellano-Bond tests, the other models are not valid, and we can not claim anything from their results (Roodman, 2009). Since the $Onshore$ variable enters only in models 5 to 8, we can not observe any evidence that onshore production yields more positive externalities than offshore production.

The valid models, 2 and 4, present results expected to convergence, since the coefficient associated to the variable $y_{i,t-1}$ is negative and significant at the 1% level. The coefficient associated to the variable $n + g + \delta$ is significant at the 10% level and has the expected negative sign. The coefficient of the $Public\ expenditures$ variable is negative and significant at the 1% level[3]. Both models do not present, however, expected results for the $Investment$ and $Education$ variables, as their coefficients are not statistically significant. It is likely that the distortion with the $Investment$ variable arises from the fact that it includes both the electricity consumption of households and of the industry. The $Education$ variable, in turn, only captures the human capital associated with the education of formal workers. There are two potential problems with this variable. The first is that it fails to account for informal workers, which represents a significant fraction of the Brazilian labor force. The second problem is that it is similar to other measures used to imperfectly measure human capital, according to the product approach (Bils and Klenow 2000). Such measures are imperfect because they may not reflect the effectiveness of human capital for several reasons: different quality of education, different performance during school hours or during accumulation of experience in the labor market, etc. However, if these problems are not strongly enough to affect the measurement, the interpretation of the obtained result is that the level of accumulated human capital in high school is of too low quality to affect growth.

---

[3] The empirical literature shows mixed effects of public spending on growth.

Regarding the influence of the $Institutional\ quality$ variable, the results deserve special attention. Let's start with the most consistent results. In model 2, the interaction term $Royalty\ 1 \times Institutional\ quality$ has a positive and significant coefficient at the 1% level. Similarly, on the model 4, the interaction term $Royalty\ 1 \times Institutional\ quality \times Large$ has a positive

Table 2. Estimations. Dependent variable: Δy_{it}

	Model 1	Model 2	Model 3	Model 4
$y_{i,t-1}$	-0.532***	-0.557***	-0.532***	-0.557***
	(-3.55)	(-3.47)	(-3.56)	(-3.47)
$n + g + \delta$	-0.0965*	-0.0873*	-0.0965*	-0.0874*
	(-1.83)	(-1.78)	(-1.83)	(-1.78)
$Investiment$	-0.0411	-0.0366	-0.0413	-0.0367
	(-0.75)	(-0.68)	(-0.75)	(-0.68)
$Education$	0.0178	0.0206	0.0178	0.0206
	(0.67)	(0.76)	(0.67)	(0.76)
$Public\ expenditures$	-0.615***	-0.654***	-0.615***	-0.654***
	(-3.11)	(-3.08)	(-3.11)	(-3.09)
$Institutional\ quality$	-0.139**	-0.153**	-0.139**	-0.153**
	(-2.37)	(-2.41)	(-2.38)	(-2.41)
$Royalty\ 1$	0.0626**	0.0204		
	(2.44)	(0.97)		
$Royalty\ 2$	0.0254**	0.0154		
	(2.47)	(0.74)		
$Royalty\ 1 \times Large$			0.0626**	0.0207
			(2.44)	(0.99)
$Royalty\ 2 \times Large$			0.0256**	0.0158
			(2.48)	(0.77)
$Royalty\ 1 \times$ $Institutional\ quality$		0.0652**		
		(2.03)		
$Royalty\ 2 \times$ $Institutional\ quality$		0.0148		
		(0.65)		
$Royalty\ 1 \times Large \times$ $Institutional\ quality$				0.0647**
				(2.03)
$Royalty\ 2 \times Large \times$ $Institutional\ quality$				0.0144
				(0.64)
$Onshore$				
$Intercept$	3.390***	3.608***	3.392***	3.607***
	(3.47)	(3.42)	(3.48)	(3.43)
Nb of observations	1414	1414	1414	1414
Nb of Instruments	36	38	36	38
Hansen test	33.82	32.23	33.80	32.24
	(0.088)	(0.122)	(0.088)	(0.121)
Arellano-Bond test for AR(1)	-3.89	-3.80	-3.89	-3.81
	(0.00)	(0.00)	(0.00)	(0.00)
Arellano-Bond test for AR(2)	0.33	0.39	0.33	0.39
	(0.743)	(0.699)	(0.744)	(0.700)

Table 2. (Continued)

	Model 5	**Model 6**	**Model 7**	**Model 8**
$y_{i,t-1}$	-0.527***	-0.542***	-0.527***	-0.542***
	(-3.43)	(-3.32)	(-3.43)	(-3.32)
$n + g + \delta$	-0.0919*	-0.0833*	-0.0919*	-0.0834*
	(-1.83)	(-1.71)	(-1.83)	(-1.72)
$Investiment$	-0.0379	-0.0318	-0.038	-0.0319
	(-0.70)	(-0.59)	(-0.70)	(-0.59)
$Education$	0.0176	0.0198	0.0176	0.0198
	(0.66)	(0.74)	(0.66)	(0.74)
$Public\ expenditures$	-0.614***	-0.633***	-0.614***	-0.633***
	(-2.97)	(-2.94)	(-2.98)	(-2.94)
$Institutional\ quality$	-0.142**	-0.149**	-0.143**	-0.149**
	(-2.33)	(-2.34)	(-2.33)	(-2.34)
$Royalty\ 1$	0.0679**	0.0253		
	(2.35)	(1.23)		
$Royalty\ 2$				
$Royalty\ 1 \times Large$			0.0680**	0.0257
			(2.35)	(1.25)
$Royalty\ 2 \times Large$				
$Royalty\ 1 \times Institutional\ quality$		0.0593*		
		(1.89)		
$Royalty\ 2 \times Institutional\ quality$				
$Royalty\ 1 \times Large \times Institutional\ quality$				0.0589*
				(1.89)
$Royalty\ 2 \times Large \times Institutional\ quality$				
$Onshore$	-0.053	-0.0237	-0.0532	-0.0241
	(-0.87)	(-0.49)	(-0.88)	(-0.50)
$Intercept$	3.379***	3.537***	3.691***	3.536***
	(3.37)	(3.30)	(6.03)	(3.30)
Nb of observations	1414	1414	1414	1414
Nb of Instruments	36	37	36	37
Hansen test	33.62	33.48	33.60	33.47
	(0.092)	(0.094)	(0.092)	(0.094)
Arellano-Bond test for AR(1)	-3.82	-3.85	-3.82	-3.85
	(0.00)	(0.00)	(0.00)	(0.00)
Arellano-Bond test for AR(2)	0.31	0.42	0.31	0.42
	(0.756)	(0.671)	(0.757)	(0.673)

Notes: Standard errors are reported in the parenthesis. Although for the tests, p-values are in the parenthesis. Level of statistical significance: 10%***, 5%**, and 1%*. Time d*ummies* are omitted.

and significant coefficient at the 1% level. This is consistent with the perception that municipalities that have better institutions can more efficiently manage royalties in order to positively impact economic growth. But only oil and gas royalties have this effect associated with institutional quality, since the similar variables for water royalties has no significant coefficients. Since there is no evidence of a significant impact when the royalties' variables are included alone in the regressions, the results may indicate that (i) good institutions which discourage rent seeking behaviors are linked to the exploitation of natural resources (Mehlum, Moene, and Torvik 2006); and/or (ii) this effect depends on the characteristics of resources.

In both models, the stand alone $Institutional\ quality$ variable has a negative and significant effect at the 5% level. Surely this is an unexpected result. Perhaps an explanation for this inconsistency is related to the construction of the variable. The TCM-BA may reject the accounts of municipalities for several reasons. It is possible that some of these reasons is in fact associated with the quality of the municipal budget management and rent seeking behavior. As the budget is a fiscal variable, there is a fiscal component to the $Institutional\ quality$ variable. So, for example, if a rejection occurs because municipal expenditures were higher than they should be, a negative coefficient of $Institutional\ quality$ variable indicates that lower expenditures hinder growth. Higher expenditures than allowed or scheduled may be associated with contingencies unrelated to rent seeking behavior. In fact, on average, municipalities with rejected bills normally have greater budget constraints.

Another possibility is that a municipalitly is able to spend more when it better manages its budget. This is the case, for example, in Salvador, Bahia's state capital. In recent years, their spending as a proportion of GDP was lower when their accounts were rejected by the TCM-BA, and higher when their accounts were approved[4]. One explanation for this is that a municipality with approved accounts may have access to more funds.

A final interpretation is that although the $Institutional\ quality$ variable indicate when legislation is being complied with, the legislation itself may not be conducive to growth. In this case, the terms $Royalty\ 1 \times Institutional\ quality$ e $Royalty\ 1 \times Institutional\ quality \times Large$ indicate that the abundance of resources interacts with legislation creating a positive effect on growth. The abundance of natural resources may relax any restrictions imposed by law. In addition, compliance with the legislation remains an institutional quality indicator by the municipal administrations. There is a mix of components in $Institutional\ quality$ variable that has a negative effect on growth, but this effect is reduced by the interaction with natural resources. Institutions can be the key component of this interaction.

[4] On average, however, there seems to be no substantial difference between municipalities with approved accounts and municipalities with rejected accounts with regard to the $Public\ expenditures$ variable.

CONCLUSION

The aim of this chapter was to check whether there is empirical evidence for the relationship between natural resource abundance and economic growth for municipalities in Brazil. In order to address the endogeneity issue, which is often observed in growth empirics, we use recent techniques of dynamic panel data models. As we introduce a proxy for institutions, the analysis is restricted to the state of Bahia municipalities, to ensure some degree of homogeneity in the construction of this proxy. Our focus is on the abundance of two types of natural resources: i) oil and natural gas; and ii) water resources. We use royalties to develop the variables corresponding to the abundance of these resources. We try to advance the previous literature in two ways: i) introducing a new variable of institutional quality; and ii) introducing a type of renewable resource.

Elaborated based on the opinions of the Court of the State of Bahia Municipalities (TCM-BA) on municipal budgets, our proxy for institutions may capture the institutional quality of municipal management. One problem is that this variable has an associated fiscal component. Despite this difficulty, the results are consistent with the interpretation that the abundance of natural resources has no significant effect on the growth of municipalities, except if it is associated with better institutional quality. In the case under study, this positive effect remains only for oil and natural gas. But it is not possible to observe any evidence that onshore production produces a positive effect that offshore production.

There is also no evidence that water resources, even interacting with good institutions, are linked to economic growth in the municipalities of Bahia. Thus, different resources may have different effects. In the present study, the difference is between exhaustible resources (oil and gas) and a renewable resource (water in surface reservoirs). The latter is the upstream of the generation chain and distribution of electricity. In Brazil, about 65% of installed capacity for electricity generation is hydroelectric origin and there are more than 1200 plants.

The literature has pointed to differences with bases in other characteristics of natural resources; for example, if the resource is point-source or not. Then, a further step is to take into account other types of resources. Royalties for exploration of other minerals, for example, may be used. In addition, to generate more robust evidence, the variable that measures institutions should be improved. Perhaps access to the complete Court reports may allow the construction of a better institutional quality variable. Currently, access to the reports via web is only possible for a few municipalities, usually the larger ones, and for a few years only.

References

Acemoglu, Daron. 2009. Introduction to modern economic development. Princeton: Princeton University Press.

Agência Nacional de Energia Elétrica (ANEEL). 2016. "Capacidade Instalada por Estado" [Capacity Installed by State]. Accessed October 30. http://www2.aneel.gov.br/aplicacoes/ResumoEstadual/ResumoEstadual.cfm.

Arellano, Manuel, O. Bover. 1995. "Another Look at the Instrumental Variable Estimation of Error-components Models." *Journal of Econometrics* 68: 29-51.

Barro, Robert. 1991. "Economic growth in a cross section of countries." *The Quarterly Journal of Economics* 106: 407-43.

Bils, Mark, and Peter J. Klenow. 2000. "Does Schooling Cause Growth?" *American Economic Review* 90: 1160-83.

Bonnefond, Céline. 2014. "Growth dynamics and conditional convergence among chinese provinces: a panel data investigation using system GMM estimator." *Journal of Economic Development* 39.

Blundell, Richard, Stephen Bond. 1998. "Initial Conditions and Moment Restrictions in Dynamic Panel Data Models." *Journal of Econometrics* 97: 115-43.

Brown, Gardner M. 2000. "Renewable Natural Resource Management and Use without Markets." *Journal of Economic Literature* 38: 875-914.

Brunnschweiler, Christina N., and Erwin H. Bulte. 2008. "The Resource Curse Revisited and Revised: A Tale of Paradoxes and Red Herrings." *Journal of Enviromental Economics and Management* 5: 248-64.

Carnicelli, Lauro, and Fernando A. S. Postali Fernando Antonio Slaibe. 2014. "Royalties do Petróleo e Emprego Público nos Municípios Brasileiros" [Royalties of Petroleum and Public Employment in Brazilian Municipalities]. *Estudos Econômicos* 44: 469-95.

Caselli, Francesco, Gerardo Esquivel, and Fernando Leforte. 1996. "Reopening the Convergence Debate: a New Look at Cross-country Growth Empirics." *Journal of Economic Growth* 1: 363-89.

Caselli, Francesco, and Guy Michaels. 2012. "Do Oil Windfalls Improve Living Standards? Evidence from Brazil." *NBER Working Paper* 15550.

Corden, W. Max, and J. Peter Neary. 1982. "Booming Sector and De-Industrialisation in a Small Open Economy." *The Economic Journal* 92: 825-48.

Deacon, Robert T. 2011. "The Political Economy of the Natural Resource Curse: A Survey of Theory and Evidence." *Foundations and Trends in Microeconomics* 7: 111-208.

Dasgupta, Partha, Robert Eastwood, and Geoffrey M. Heal. 1978. "Resource Management in a Trading Economy." *Quarterly Journal of Economics* 92: 297-306.

Durlauf, Steven N., and Paul A. Johnson, and Jonathan R. W. Temple. 2005. "Growth Econometrics." In: *Handbook of Economic Growth,* edited by Philippe Aghion, Steven Durlauf, 555-677. Amsterdam: North-Holland.

Gray, Lewis Cecil. 1914. "Rent Under the Assumption of Exhaustibility." *The Quarterly Journal of Economics* 28: 466-89.

Hartwick, John M. 1977. "Intergenerational Equity and the Investing of Rents from Exhaustible Resources." *The American Economic Review* 67: 972-74.

Hotelling, Harold. 1931. "The economics of exhaustible resources." *Journal of Political Economy* 39: 137-75.

Holtz-Eakin, Douglas, Whitney Newey, and Harvey S. Rosen. 1988. "Estimating Vector Autorgressions with Panel Data." *Econometrica* 56: 1371-95.

Islam, Nazrul. 1995. "Growth Empirics: A Panel Data Approach." *The Quarterly Journal of Economics* 110(4): 1127-70.

James, Alex, and David Aadland. 2011. "The Curse of Natural Resources: An Empirical Investigation of US Counties." *Resource and Energy Economics* 33: 440-53.

Mankiw, N. Gregory, David Romer, and David Weil. "A Contribution to the Empirics of Economic Growth." *The Quarterly Journal of Economics* 107: 407-37.

Mehlum, Halvor, Karl Moene, and Ragnar Torvik. 2006. "Institutions and the Resource Curse." *Economic Journal* 116: 1-20.

Mironov, V.V., and A. V. Petronevich. 2015. "Discovering the Signs of Dutch Disease in Russia," *BOFIT Discussion Papers 3/2015,* Bank of Finland, Institute for Economies in Transition.

Parente, Stephen, and Edward C. Prescott. 1994. "Barriers to Technoly Adoption and Development." *The Journal of Political Economy* 102: 298-321.

Postali, Fernando A. S. 1997. "Efeitos da Distribuição de Royalties do Petróleo sobre o Crescimento dos Municípios no Brasil: Utilizando a Lei do Petróleo como um Experimento Natural" [Effects of the Distribution of Oil Royalties on the Growth of Municipalities in Brazil: Using the Oil Law as a Natural Experiment]. Paper presented at the annual meeting for the Encontro Nacional de Economia, Recife, Pernambuco, December 4-7.

Postali, Fernando A. S., and Marislei Nishijma. 2011. "Distribuição das Rendas do Petróleo e Indicadores de Desenvolvimento Municipal no Brasil nos Anos 2000" [Distribution of Oil Rents and Municipal Development Indicators in Brazil in the 2000s]. *Estudos Econômicos* 41: 463-85.

Roodman, D. 2009. "How to do xtabond2: An Introduction to "Difference" and "System" GMM in Stata." *Stata Journal* 9: 86-136.

Sachs, Jeffrey D., and Andrew M. Warner. 1995. "Natural resource abundance and economic growth." *NBER, Working Paper* 5398.

Sinnott, Emily, Augusto de la Torre, and John Nash. 2010. Recursos Naturais na América Latina: Indo Além das Altas e Baixas [Natural Resources in Latin America: Going Beyond the Highs and Lows]. Rio de Janeiro: Elsevier.

Solow, Robert M. 1974. "The Economics of Resources or the Resources of Economics." *The American Economic Review* 64: 1-14.

Sweeney, James L. 1993. "Economic Theory of Depletable Resources: An Introduction." In: *Handbook of Natural Resource and Energy Economics* Volume III, edited by A.V. Kneese, and James L. Sweeney, 759-854. Amsterdam: North-Holland.

van der Ploeg, Frederick. 2011. "Natural Resources: Curse or Blessing?" *Journal of Economic Literature* 49: 366-420.

van der Ploeg, Frederick, and Steven Poelhekke. 2010. "The Pungent smell of "Red Herrings": Subsoil Assets, Rents, Volatility and the Resource Curse." *Journal of Environmental Economics and Management* 60: 44-55.

van der Ploeg, Frederick, and Anthony J. Venables. 2012. "Natural Resource Wealth: The Challenge of Managing a Windfall." *Annual Review of Economics* 4: 315-337.

Williams, Andrew. 2011. "Shining a Light on the Resource Curse: An Empirical Analysis of the Relationship Between Natural Resources, Transparency, and Economic Growth." *World Development* 39: 490-505.

In: Infrastructure Investments
Editors: Gisele Ferreira Tiryaki et al.
ISBN: 978-1-53610-792-0

Chapter 6

DEMAND FOR PUBLIC TRANSPORTATION AND THE DESIGN OF OPTIMAL PUBLIC POLICIES

Leticia Klotz[1], André Trindade[2,*] and Edmilson Varejão[3]
[1]University of Minnesota, MN, US
[2]FGV/EPGE, Rio de Janeiro, Brazil
[3]FGV/CERI, Rio de Janeiro, Brazil

ABSTRACT

Transport and mobility are key factors affecting the quality of life in cities all around the world. In order to choose an optimal transportation system, it is vital to understand well how commuters respond to different designs of the public infrastructure. In this chapter, we attempt to address that question by reviewing the relevant empirical literature in transportation economics. Our focus is to discuss research that models and estimates demand for the use of public transportation. With that in mind, we first detail the typical approach and specifications used. Then we summarize the main results found in that literature. Finally, we discuss our findings in the context of the future of the sector, taking into account the new challenges, innovations and emerging business models.

Keywords: public transportation, demand, literature review, public policies

[*] Corresponding Author: André Trindade. Email: andre.trindade@fgv.br.

INTRODUCTION

Understanding consumer demand for public transportation is crucial to carefully evaluate how government policies in the sector affect transportation users, and hence assess the costs and benefits associated with each new project. To correctly determine the best policies, it is vital to quantify clearly how commuters will respond to different designs of the public transportation infrastructure. That is exactly the main goal of estimating the demand for transportation.

In this chapter, we review important papers on the transportation literature, focusing in particular on empirical papers that estimate demand for public transportation. Our review is not exhaustive but attempts to focus on some of the empirical findings that can be useful for policy makers. Our goal is to help governmental and private sector agents in this field of public transportation infrastructure to be able to make informed decisions about the future.

We start our review by discussing the work by McFadden. His 1974 work is one of the first papers to highlight the importance of studying travelers' behavior in demand estimation for public transport. By using a random utility maximizing framework, McFadden rationalized aggregate demand estimation as an outcome of individuals' decision-making behavior that is affected by key factors. His view has influenced academics and policy-makers to draw attention to which factors impact travelers' behavior and hence urban travel demand.

Next, we discuss the work of Paulley et al. (2006). They consider four main factors that influence transport users' decision-making: fares and fuel prices, car ownership, income, and quality of service, and discuss the mechanisms through which these attributes affect different types of travelers.

We then turn our attention to relevant parameters estimated by the transport demand literature and present some key parameters and literature findings, commenting further on policy simulations and some relevant counterfactuals. Finally, we address the state-of-art of literature on transport demand in Brazil and conclude by discussing the future of the Brazilian transportation sector.

FACTORS AFFECTING THE DEMAND FOR URBAN TRANSPORTATION

Discussing relevant factors that influence the demand for urban transportation is crucial to better understand current traveler's behavior and, at the same time, how passengers will respond to changes in the sector. Thus, it is paramount for academics and practitioners to address which aspects determine the demand for public transportation in order to properly assess whether different policies will yield desired outcomes.

Several studies and surveys have been carried out to examine which factors are relevant for estimating urban travel demand. Paulley et al. (2006) presents the findings of a comprehensive collaborative study conducted in the UK that highlights the influence of four main factors: fares and fuel prices, car ownership, income, and quality of service. In this section, we comment on the expected impacts of those factors on travelers' behavior and discuss how they influence urban travel demand[1].

When accessing local transportation systems, customers frequently face clear monetary costs, and hence it is natural to assume that such costs are taken into account when travelers decided their mode of transportation. As pointed out by Taylor and Fink (2003), fares and fuel prices are the most obvious examples of such monetary costs, and this is the reason most authors analyzing transit ridership consider these factors,. According to Paulley et al. (2006), increases in fares of one mode of transportation will reduce patronage for the same mode and generally increase demand for alternative modes due to substitution effects. Fuel prices affect private demand similarly, as stated by Graham and Glaister (2002), Hughes, Taylor and Fink (2003), Hughes, Knittel and Sperling (2008), and Donna (2015).

The effects of fares and fuel prices in turn depend heavily on local population idiosyncrasies. According to Small and Verhoef (2007), fuel prices are closely related to activities that occur on a particular location and to local demographics, and therefore local population idiosyncrasies should be considered when estimating the demand for urban transportation. McFadden (1974), for instance, proposes to estimate a discrete choice model where the utility of each individual is a function of the characteristics of each alternative,

[1] In this chapter, we do not address how macroeconomic variables such as unemployment and inflation rate, aggregate GDP growth directly influence the demand for urban transportation, focusing only on their indirect effects via microeconomics variables, such as local income and prices. For an analysis on how transport demand is affected directly by changes in macroeconomic variables, see Hughes, Knittel and Sperling (2008).

demographic variables and an independent and extreme value distributed error term. In this seminal paper, the author finds that population density, the number of minutes a traveler can arrive late[2] and car ownership are statistically significant variables.

A comprehensive analysis of overall urban travel demand should address not only demand for public transportation, but also private transportation in the form of car ownership[3], accounting for interactions and substitution effects of one segment on the other. Car owners have more alternative modes to choose from than users of public transportation systems, leading to a reduction in the demand for public transportation, as indicated by Paulley et al. (2006).

Car ridership is the most frequently used mode of transportation in several countries, considering both private and shared rides. In particular, as reported in Lucinda et al. (2015), drive and shared ride combined accounted for the greatest share (36.15%) of the total of surveyed trips in Brazil's largest city, Sao Paulo, in 2015. Rides can take the form of carpooling, which should also be considered when estimating transport demand as it may constitute a non-conventional alternative to public transportation for drivers that have a strong preference for traveling in a private vehicle. According to Bento, Hughes and Kaffine (2013), fuel and other driving costs per passenger are reduced by carpooling when compared to the costs of drivers that ride alone. Carpooling riders also benefit from lower traffic congestion in high-occupancy vehicle (HOV) lanes, which leads to faster trips and lower in-vehicle time.

Car ownership and carpooling in turn are strongly correlated with travelers' income, as summarized by Paulley et al. (2006). According to these authors, an increase in income could lead to an increase in car ownership and, as discussed previously, it would in turn result in a reduction in the demand for public transportation. Holmgren (2013), on the other hand, argues that a higher income stimulates greater mobility, and hence an increase in the number of public transportation trips is likely to occur. Using a reduced-form model, the author shows that, although relevant to explain overall urban travel demand, income has opposite effects on public transport demand, and the direction of such effects is a priori unknown.

[2] The choice of when arriving to work can be determined endogenously through a discrete choice model that depends on demographic and work characteristics, as developed in Small (1982). According to the author, distinct family status, occupation, transportation mode, and employer's policy toward work-hour flexibility yield different effects on scheduling times.

[3] Here we use the term "car ownership" indiscriminately to refer to privately owned vehicles, assuming vehicle type choice as exogenous in the short-run. In Hensher et al. (1990), travelers face alternative vehicle technologies and choose the one that maximizes the joint utility of vehicle choice and use.

Fare and fuel prices, car ownership and income are easily measurable variables that influence urban travel demand, as discussed previously, and hence should be included as covariates in any transport demand estimation. Nonetheless, as pointed out by Paulley et al. (2006), quality of service frequently involves a wide range of subjective attributes that affect travelers' choice of mode transportation, and thus defining suitable proxies is crucial for correctly estimating overall transport demand.

Moreover, as indicated by Small and Verhoef (2007), the most typical aggregate-level proxies for service quality are annual vehicle-miles or vehicle hours of service. Following Paulley et al. (2006), distance to and from the station, service intervals and in-vehicle time can be easily incorporated in transport demand forecasts. Considering individual travel behavior, time-related measures, such as in-vehicle, walking and waiting times also influence travelers' transportation mode choice, as postulated by Currie (2005), and hence impact on overall transport demand. An increase in overall time spent in a particular transportation mode tends to reduce demand for that same mode and increase the demand for alternative ones, due to substitution effects. According to Small and Verhoef (2007), estimating how travelers value time is vital for police-making, since it is a key parameter of interest in transport demand estimation, as will be discussed in the next section[4].

MAIN EMPIRICAL RESULTS IN THE LITERATURE

One of the main goals of estimating the demand for transportation is to assess the overall response of travelers to changes in relevant attributes. In order to accomplish such aim, it is vital to define adequate measures of travelers' response. In this section, we discuss the measures most frequently found in the literature and present the direction and magnitude of these findings.

Elasticities are the most frequently used metric to measure demand response in the transportation sector. According to Goodwin (1992), this might be explained by their ease of understanding, and because they are tested by experience. For this reason, we dedicate most of this section to analyze

[4] In this section, we restrict our analysis to factors that are most frequently cited in the literature to influence urban demand transportation. We admit there might be other attributes that affect demand for urban transport that were not contemplated in this chapter.

different types of elasticities calculated by the literature and to discuss how they change under distinct scenarios.

As detailed previously, changes in public fares and fuel prices both affect public and private demand transportation. Most authors seeking to understand the impact of prices and fares on urban travel demand calculate own and cross elasticity, i.e., the impact of a change in price on, respectively, the demand for the same mode and alternative mode of transportation. Moreover, the authors often distinguish between short- and long-run elasticities[5], as pointed out by Goodwin (1992) and Paulley et al. (2006).

Goodwin (1992), Graham and Glaister (2002) and Paulley et al. (2006) report different short- and long-run estimates of own price elasticities, but all authors reach the same conclusion: both measures are negative and short-run elasticities are often lower in absolute value than the long-run counterparts. On the other hand, Bento, Hughes and Kaffine (2013) find positive short-run elasticities of fuel prices in high-occupancy vehicle (HOV) lanes, suggesting that an increase in fuel prices do not necessarily lead travelers to reduce the number of car trips, but instead to choose to carpool in the short-run. According to Goodwin (1992), the ratio of long-term to short-term elasticities ranges from 1.5 to 4.0, depending, among other factors, on the level of aggregation of the data and if cross section or time series variation is used for identification. Travelers are more sensitive to fuel prices and fares in the long-run, because not only can they adapt to other mode alternatives, such as carpooling (Bento, Hughes and Kaffine, 2013) and bus rapid transit systems (BRT) (Currie, 2005), but they might also choose more efficient cars that reduce fuel consumption.

Nonetheless, long-run cross-elasticities might be in fact lower in absolute value if there are costs in substituting one mode for another. Donna (2015) assumes that travelers face costs from switching from one mode to an alternative, i.e., travelers incur in switching costs. According to the author, switching costs for travelers might associate to, for instance, obtaining new information on transit services and automobile routes, as well as purchasing ridership tickets and searching for new parking facilities. His estimates show that long-run cross-elasticities both for car and for public transportation are less than 20% of their respective estimates obtained if switching costs are not considered.

[5] The notions of short- and long-run usually exhibit some variation according to each author's choice.

Fare and fuel price elasticities estimates also vary according to local population characteristics, such as location (if close to rural or urban areas), different trips purposes (work or not), and time of day (rush hours), as exemplified in Goodwin (1992), Paulley et al. (2006) and Small and Verhoef (2007). Travelers located in urban areas have fewer alternative modes of transportation to choose from, and hence fare elasticities tend to be greater. On the other hand, trips made for work purposes often face schedule and flexibility constraints, as argued by Small (1982), and hence off-peak fare elasticities are often greater than peak values, being about twice as greater for the UK, according to Paulley et al. (2006).

Another important type of elasticity frequently estimated in the transportation literature is income elasticity of demand. As highlighted in the previous section, changes in income have opposite effects on public transport demand due to direct positive effects and competing negative effects on car ownership. Holmgren (2013) disentangles these opposite effects, yielding estimates of direct income elasticity for public transportation of 0.34 and estimates of car ownership elasticity with respect to income of 0.21, which in turn result in total income elasticity of public transport demand of 0.052, i.e., the total effect is close to zero in this case.

In spite of being straightforward and tested by experience, the different types of elasticities discussed previously fail to capture how changes in key factors affect travelers' welfare, and how they value such attributes in monetary terms. Considering the relevance of those measures to assess the impact of distinct government policies and structural changes in the sector, detailed in the next section, we present and comment here how the literature of transportation have defined and estimated such measures.

Marginal rates of substitution are especially suitable to characterize substitution patterns for the same level of consumers' welfare[6]. In the context of the transportation sector, this concept is particularly applied to time-related attributes. Small (1982) estimates marginal rates of substitution of different time schedules by each mode of transportation, finding that on average urban commuters shift their schedules by 1 to 2 minutes earlier than previously scheduled, or by 1/3 to 1 minute later, in order to save one minute of travel time.

The concept of value of time, mentioned in the previous section, can be expressed quantitatively as a particular type of marginal rate of substitution.

[6] The term welfare in this paragraph refers loosely to indirect utility levels that are assumed constant when calculating marginal rates of substitution.

Lam and Small (2001) and Wardman (2004) define value of time as the ratio of marginal utilities of time and money, which equals to the marginal rate of substitution of travel time for money, i.e., how much travelers are willing to pay for to save one hour of travelling. Small and Verhoef (2007) argue that value of time, similarly to elasticities, also depend, among other factors, on demographics and local population characteristics, trip purpose, and total duration of the trip.

In order to assess the impact on transport demand of actual or hypothetical changes in the sector, recent papers in the literature have frequently calculated welfare estimates in terms of Consumer Surplus (CS) and Equivalent Variation (EV). As defined in Mas-Colell et al. (1995), the concept of equivalent variation (EV) expresses how much, in monetary values, a significant change in key sector variables is equivalent in terms of welfare impacts. In the following subsection, we detail how welfare measures are estimated in the transport literature to evaluate relevant changes and alternative scenarios in the sector and describe their policy implications.

Counterfactual Scenarios and Policy Implications

One of the main purposes of estimating transport demand is to carefully understand how government policies and structural changes in the sector directly affect travelers' behavior and then analyze travelers' response. Here we discuss briefly the goals of some transport policies[7] and their underlying mechanisms, addressing how these different policies and structural changes in the sector affect urban travel demand.

Transportation policies should be designed to achieve economic efficiency, i.e., should create an environment that maximizes overall welfare of users and providers and reduce the impact of market failures. Externalities are the most significant market failures that affect transportation users by introducing additional costs to the travelers' decision-making process, and hence transportation polices should act to alleviate these negative effects.

According to Bento et al. (2014), congestion is one type of negative externality prevalent in the transportation sector, and, as argued by Lucinda et al. (2015), it leads to lower productivity levels, increased fuel consumption and negative environmental effects. Academics and policy makers have criticized and proposed several mechanisms designed to mitigate problems

[7] We partially reproduce here policy objectives indicated by Balcombe et al. (2004).

associated with congestion, such as increasing fuel and congestion taxes[8], and introducing rotation systems, as discussed by Parry and Small (2005) and Lucinda et al. (2015).

Parry and Small (2005) argue that fuel taxes are appealing because are administratively simple and well-established. Nonetheless, as the authors argue, fuel taxes are an imperfect instrument for controlling distance-related externalities like congestion, because travelers can respond to increases in fuel taxes by purchasing more fuel-efficient vehicles rather than reducing miles travelled. Parry and Small (2005) propose instead a second-best optimal fuel tax that reflects, among other components, external costs of congestion and air pollution, and show that actual United States and Britain fuel taxes differ dramatically from their second-best value.

Two possible alternatives are either charge road usage directly, via imposing an urban toll fare per trip, or introduce rotating systems to reduce total number of circulating vehicles, as suggested in Lucinda et al. (2015). The authors compare these two policy alternatives that yield the same traffic reduction, estimating an equivalent variation (EV) for urban toll rates of about 2.28 Brazilian Reais, while for the rotation system EV is around 3.5 Brazilian Reais. Therefore, although yielding the same results, rotation systems makes travelers worse off.

Finally, policy-making in the transportation sector should also address the effects of new transportation modes before their introduction and the impacts of new traffic and environmental initiatives on travelers' welfare. For example, McFadden (1974) used logit estimates to forecast the demand for rapid transit system in the San Francisco Bay Area (BART) before the full implementation of the system, whereas Bento et al. (2014) study the implementation of Clean Air Vehicle Stickers, an initiative that encourages the ultra-low emission vehicles by allowing solo-hybrid drivers to use high occupancy vehicle lanes (HOV). Authors show that welfare effects of this policy are negative due to increased congestion costs, illustrating interaction effects of different types of externalities in the transportation sector.

[8] Bento, Hughes and Kaffine (2013) argue that fuel prices have been used in the transportation sector as an instrument to reduce traffic congestion, but it has also been used by policymakers to reduce greenhouse gas emissions and in optimal taxation analysis (Hughes, Knittel and Sperling, 2008).

THE FUTURE OF TRANSPORTATION

The transport industry going through a huge transformation process. The industry innovation has traditionally been promoted by changes in technologies of the vehicles or infrastructures. Presently, the source of transformation are the changes on social trends, business models, types of service and technological developments, especially in the area of the Information and Communication Technologies (ICTs). This ongoing trend, which is shared with others network industry, is developing as a co-evolution between technological change and institutional change (see for example, Nelson 1994, Nelson and Sampat 2001, Von Tunzelmann 2003, Murmann 2003, Reinstaller 2005 and Fatas-Villafranca et al. 2008). They are together paving the way for the establishment of new disruptive business models. Examples of these models are the UBER platform, which provides over 1 million rides per day worldwide, or the significant investment in self driving technologies by important players such as Google, Tesla or UBER.

The empirical evidence on how consumers respond to incentives in the transportation sector is ever more relevant as the sector faces new challenges that will require new public policies to be created and old rules to be adjusted. In this section we detail those new challenges.

On the technological field, communications and information technology, systems, and applications have evolved at a rapid rate. These factors ultimately led to innovative research initiatives and an explosion of new transportation apps, often combining the use of vehicles as probes with enhanced geographic location and mapping systems in the form of user-friendly mobile and in-vehicle user interfaces. Increasingly, ITS applications are considered in two contexts—for automated purposes and/or for connected vehicle purposes. In addition to these technologies, commercial applications based on geolocation and cell phones, such as Waze and Uber, are influencing the ITS market and are part of a larger trend of shared mobility (U. S. Department of Transportation, 2016). Shared mobility—the shared use of a car, bicycle, or other low-speed mode of transportation—is one aspect of the sharing economy and enables users to obtain short-term access to transportation modes as needed rather than requiring ownership.

Regarding the institutions nowadays, the systems of established and prevalent rules that structure the market and technological systems operation are increasingly becoming a governance rule; i.e., the rules, norms and actions of the industry are produced, sustained, regulated and held no longer by government, but by other entities, private and/or international. This trend was possible because the ICT technologic innovation is developing from more centralized to distributed and decentralized.

Therefore, the result from the institutions and technology development is a playing field for the increase of new disruptive business models and market actors. Existing relationships between actors in the value chain are in the process of being replaced by new organizational forms, which allow for the creation of dynamic networks among individuals and transport services firms. The focus is on the customers rather than on the infrastructure or on the providers of mobility. In other words, historically the different transportation modes were separated from the view of the companies, users or regulation. With digitalization, a digital layer can connect information about multiples transport mode and consumers, enabling developers to build quite sophisticated transport services for the end users to compare prices, routes and qualities of different or combination of modes.

As the data that is generated is based on the clients, the user does not need to be a transport operator to generate data about transport. That is, the clients no longer need to be linked to the operator, they are linked to the platform. The innovation is the user will ultimately buy mobility and no longer transport. New technologies, as self-driving cars, buses and drones, can accelerate this process. For transport users, the focus is no longer on the transport mode, but rather on mobility. As a consequence, mobility will be increasingly seen as an information service with physical transportation products rather than a transportation product with additional services.

This growing role of the ICTs developed a new concept of transport, the 'Mobility-as-a-Service' (MaaS). It is a mobility distribution model in which a customer's major transportation needs are met thanks to a single interface with services offered by a single integrated service provider combining transportation infrastructure, transportation services, information and payment services, among other services. The model works out the best option for every journey – whether that is a taxi, public transport, a rental car or a bike share. From office commutes to weekend getaways, it manages daily travel in the smartest way possible. For extra convenience, MaaS may include value added services like deliveries services (Finger, 2015).

Few studies have been carried out to simulate the impacts of news trends in urban transportation. ITF (2015) models the impact of replacing all car and bus trips in a city with shared mobility delivered by a fleet of six-seat vehicles ("Shared Taxis") that offer on-demand, door-to-door shared rides in conjunction with a fleet of eight-person and 16-person mini-buses ("Taxi-Buses") that serve pop-up stops on demand and provide transfer-free rides.

Rail and subway services keep operating in the current pattern. The objective is to simulate effects on congestion, CO_2 emissions, use of public space and social inclusion measured by the level of accessibility of jobs, schools and health services.

That particular study is implemented by simulating the daily mobility of the city, including the movement of people and vehicles, environmental effects and changes in transport infrastructure. The alternative shared trip routing is based on an algorithm that generates the lowest time path between any pair of nodes of the network, managed by a central dispatcher system that uses the location of shared vehicles, their current occupancy level and the location of clients as its main inputs. The author uses Viegas and Martínez (2010) synthetic travel simulation model. Each trip is characterized by its time of occurrence, origin and destination, as well as other traveler's information, as age, gender, income, the possession of a driving license, car pass, motorcycle, parking spaces at home and at work (Moura et al. 2007; Martinez and Viegas, 2009; Santos et al. 2011). Having the full characterization of the trip as an input, the probability of choosing each mode is assigned to the client by simulation where modes with higher probability will be chosen more often (Eiró and Martínez, 2014).

The main findings are that, holding the number of daily trips constant, congestion disappears and traffic emissions are reduced by one third. The car fleet needed would be only 3% in size of the current fleet, reducing the required space for public parking in 95%. Without congestion and with higher efficient use of fleet capacity, journeys' prices in a city could be at least 50% below the current levels even without subsidy. The inequality in the access to jobs, schools or health services across the city, a measure of social inclusion, virtually disappeared.

Even though this process is already at a fast pace in developed countries, it is reasonable to expect that there will be a sort of catching up for developing regions. In Brazil, for instance, UBER and others *e-hailing* app's are already working on several capitals of the country, and the competition has already triggered angry reactions among taxi drivers. As most of the innovations are on business models and, considering that the regulation is held no longer by government but by other entities - private and/or international -, one of the main challenges for the dissemination of the service steems from regulatory issues. In the city of Rio de Janeiro, for example, while the Olympic Games took place, the UBER service was being provided through injunction, because the mayorship had previously banned the provision of services through UBER in the city. Almost at the same time, the mayorship of Sao Paulo laid out regulations to allow the use of ride-hailing apps in exchange for a mileage fee of R$ 0,10 per kilometer to be paid by the drivers, among other quality criterias.

Encouraged by the high cost of public transport fares and new social trends, some sharing economy iniciatives in transport are taking place in Brazil. Ride-sharing apps, that connect car drivers with room in their vehicles with members of the public in search for a ride, are attracting commuters especially in longer distance travels. Peer-to-peer carsharing apps, whereby existing car owners make their vehicles available for others to rent for short periods of time, are starting operation in Brazil's main cities.

There are other iniciatives being promoted by the city halls. Some cities around the country already counts with bike sharing systems in which users can acess the service via an app and search the nearest station with available bike. The city of Fortaleza, capital of the state of Ceara, went further and is operating an electric car-sharing system for public use on a paid subscription basis.

CONCLUSION

Transport and mobility are key factors in determining the quality of life in the cities all around the world. Historically, the evolution of urbanization and urban transportation system has a fundamental linkage. By 2050, more than 6 billion people, or around two thirds of population, will be urban residents. In many cases, cities respond to growth in the demand for mobility by expanding the transportation supply, building new transport infrastructure. However, physical limitation due to restricted space for new roads or budget constraints for expensive investments forces the public sector to better plan the strategy to accommodate the demand for transportation, current and future, in order to maximize the population's well-being.

Urbanization, new social trends and innovations in business models and technologies, especially in the area of the ICTs, are dramatically changing the transportation system and utilities. This offers a new way for city planners to deal with mobility problems, moving away from the traditional approach of increasing capacity by building costly new infrastructure. ICTs have a disruptive effect on the transport sector: new technologies enable a new integrated mobility system that focusses on the customers rather than on infrastructure or on the providers of mobility. Furthermore, new technologies such as automated vehicles are entering the transportation system at a steady pace. In a nutshell, there are several new challenges connected to transport that regulation has to take into account.

The existing transport sector's structure requires new perspectives in transport regulation and governance. The transformation of business models, especially when it involves disruptive changes, entails regulatory and legal accommodation. Regulatory interventions are needed to guide investments towards a cleaner renewable generation mix, creating conditions for innovation and transformation to generate benefits for all stakeholders. Moreover, it is necessary to mobilize political leadership to ensure the design of appropriate public policies that should include ambitious outcome-based targets.

To draw public policies intelligently, it is necessary to understand the traveler's behavior and reactions to those policies. Demand systems provide an important component of incentives for market responses to many policy and environmental changes. This chapter summarized some of the empirical evidence that can be useful to design the future of transportation.

REFERENCES

Balcombe, Richard, editor. 2004. *The demand for public transport: a practical guide*. TRL Report, TRL593.

Bento, Antonio M., Jonathan E. Hughes, and Daniel Kaffine. 2013. *Carpooling and driver responses to fuel price changes: Evidence from traffic flows in Los Angeles*. Journal of Urban Economics 77, 41-56.

Bento, Antonio, Daniel Kaffine, Kevin Roth and Matthew Zaragoza-Watkins. 2014. *The Effects of Regulation in the Presence of Multiple Unpriced Externalities: Evidence from the Transportation Sector*. American Economic Journal: Economic Policy, 6(3): 1-29.

Currie, Graham. 2005. *The demand performance of bus rapid transit.* Journal of Public Transportation 8: 41–55.

De Borger, Bruno, and Kurt Van Dender. 2003. *Transport tax reform, commuting, and endogenous values of time*. Journal of Urban Economics 53.3, 510-530.

Donna, Javier. 2015. *Estimating Switching Costs in Urban Travel Demand in Chicago*. Unpublished Manuscript.

Eiró, T., and L. Martínez. 2014. *Development of a discrete choice model to evaluate the impact of new transport alternatives*. Transportation Research Board 93rd Annual Meeting, Washington D. C.

Finger, Matthias and Künneke, Rolf. 2009. *Exploring socio technical governance regimes in liberalizing network industries.* Discussion paper series on the Coherence between institutions and technologies in infrastructures.

Finger, Matthias, Nadia Bert, and David Kupfer. 2015. *Mobility-as-a-Service: from the Helsinki experiment to a European model?* European Transport Regulation Observer.

Goodwin, P. B. 1992. *A Review of New Demand Elasticities with Special Reference to Short and Long Run Effects of Price Changes*. Journal of Transport Economics and Policy 26, no. 2, 155-69.

Graham, Daniel J., and Stephen Glaister. 2002. *The demand for automobile fuel: a survey of elasticities*. Journal of Transport Economics and Policy 36.1, 1-25.

Hensher, David A., Frank W. Milthorpe, and Nariida C. Smith. 1990. *The demand for vehicle use in the urban household sector: theory and empirical evidence*. Journal of Transport Economics and Policy, 119-137.

Holmgren, Johan. 2013. *An analysis of the determinants of local public transport demand focusing the effects of income changes*. European Transport Research Review 5.2, 101-107.

Hughes, Jonathan E., Christopher R. Knittel and Daniel Sperling. 2008. *Evidence of a Shift in the Short-Run Price Elasticity of Gasoline Demand*," The Energy Journal, vol. 29(1), pages 113-134.

ITF. 2015. *Urban Mobility System Upgrade: How shared self-driving cars could change city traffic, International Transport Forum Policy Papers*, No. 6, OECD Publishing, Paris.

Lam, Terence C., and Kenneth A. Small. 2001. *The value of time and reliability: measurement from a value pricing experiment*. Transportation Research Part E: Logistics and Transportation Review 37.2, 231-251.

Lucinda, Claudio R., Rodrigo Moita, Leandro Meyer and Bruno Ledo. 2015. *The Economics of Sub-optimal Policies for Traffic Congestion*. Unpublished Manuscript.

Martinez, L. M., and J. M. Viegas. 2009. *Activities, Transportation Networks and Land Prices as Key Factors of Location Choices: An Agent-Based Model for Lisbon Metropolitan Area (LMA)*. MIT-Portugal Working Paper Series.

McFadden, Daniel. 1974. *The measurement of urban travel demand*. Journal of Public economics 3.4, 303-328.

Moura, F., A. Duarte, and J. Viegas. 2007. *CAReFUL – Car fleet renewal as a key role for atmospheric emission reduction*. CESUR, Instituto Superior Técnico, Lisbon.

Parry, Ian WH, and Kenneth A. Small. 2005. *Does Britain or the United States have the right gasoline tax?* The American Economic Review 95.4, 1276-1289.

Paulley, Neil, Richard Balcombe, Roger Mackett, Helena Titheridge, John Preston, Mark Wardman, Jeremy Shires, and Peter White. 2006. *The demand for public transport: The effects of fares, quality of service, income and car ownership*. Transport Policy 13, no. 4, 295-306.

Santos, G. D., L. M. Martínez, J. M. Viegas, and D. Alves. 2011. *Design and deployment of an innovative mobility survey for the Lisbon metropolitan area oriented to assess the market potential and obtain the best configuration of new alternative transport modes*. European Transport Conference, Glasgow.

Small, Kenneth A. 1982. *The Scheduling of Consumer Activities: Work Trips*. The American Economic Review 72, no. 3, 467-79.

Small, Kenneth A., and Erik T. Verhoef. 2007. *The economics of urban transportation*. Chapter 2. Routledge.

Taylor, Brian D., and Camille NY Fink. 2003. *The factors influencing transit ridership: A review and analysis of the ridership literature*. University of California Transportation Center.

U. S. Department of Transportation. 2016. *History of Intelligent Transportation Systems*.

Viegas, J. M., and L. M. Martínez. 2010. *Generating the universe of urban trips from a mobility survey sample with minimum recourse to behavioural assumptions* in: Proceedings of the 12th World Conference on Transport Research, Lisbon.

Wardman, Mark. 2004. *Public transport values of time*. Transport policy 11.4, 363-377.

In: Infrastructure Investments
Editors: Gisele Ferreira Tiryaki et al.
ISBN: 978-1-53610-792-0

Chapter 7

USING OPTIMISATION AND MACHINE LEARNING TO VALIDATE THE VALUE OF INFRASTRUCTURE INVESTMENTS

Rodolfo García-Flores[1,]*, Soumya Banerjee[2] and George Mathews[2]
[1]CSIRO Data61, Australia
[2]CSIRO Data61, Australia

ABSTRACT

When stakeholders commit to building infrastructure as part of long-term planning, the final facilities are not normally amenable to modification after completion. Thus, users are forced to operate within the original specifications for, at least, as long as it takes to carry out major refurbishments, and even then, the constraints imposed by the original layout may be inescapable. Although the original infrastructure plans enhance (or limit) the users' ability to operate efficiently for years to come, managers and operators (that is, those living with the consequences of the strategic planning) do have some leeway to compensate for miscalculations by means of their operational planning.

*Corresponding author: Rodolfo.Garcia-Flores@csiro.au

In this chapter, we explore the use of quantitative techniques to amend bottlenecks and uncertainty in operations that affect the performance of infrastructure investments. To that end, we present a rail scheduling case study that combines Machine Learning and Operations Research methods to validate and maximise the value of already-existing infrastructure. We present a literature survey and a preliminary analysis of the system. We also anticipate that Machine Learning and optimisation techniques will be integrated to create tools that help stakeholders maximise the value of their current infrastructure, while meeting ever-changing demand.

Keywords: Infrastructure Value, Port Logistics, Infrastructure Planning, Rail Scheduling, Strategic Planning
AMS Subject Classification: 90-00, 90-02, 90B80.

Introduction

When governments or industry commit to building infrastructure, the resulting facilities are not normally amenable to modification after completion. Since these facilities are the product of strategic planning, they normally require very large investments and are meant to operate continuously for many years in the future, possibly with a few major refurbishments or retrofitting during the course of their lifetimes. This means that, despite having little flexibility to accommodate for changes in market conditions or in the natural environment, they must still face the constraints and bottlenecks produced by these changes, and which may not have been foreseen originally. This can produce unanticipated economical, social and political repercussions. A clear example is Castellón airport in Spain, which is currently in use but underutilised [1].

Fortunately, managers and operators have some margin to compensate for miscalculations by playing with their operational and tactical plans. Managers and operators must take advantage of the infrastructure available to them and "juggle" their resources in order to maximise their productivity, even in cases where the extant facilities represent more a

[1] `https://en.wikipedia.org/wiki/Castell%C3%B3n%E2%80%93Costa_Azahar_Airport`, accessed on the 11 of April 2016

limitation (or, in extreme situations, a liability) than an asset. In this chapter, we centre the discussion on quantitative approaches to assess the performance of large infrastructure investments when there are unforeseen changes in economic and environmental conditions, not taken into account at the strategical planning stage. These changes may be caused by:

1. *Variability in demand and supply.* Supply and demand are the most commonly used design variables for determining the capacity of new infrastructure investments, and their variability impacts directly on payback periods and facility utilisation.

2. *The economic environment.* In addition to supply and demand, other economic variables affect substantially the performance in time of large infrastructure facilities. These include fluctuations in market prices, labour costs, debts and subsidies, interest rates, insurance, and country risks.

3. *Social conditions.* Political will is not only the trigger to develop large construction projects, but often is also the engine that keeps them in operation, shuts them down, or carries out extensive refurbishment or auditing. Decision makers often know from the outset that some projects are not profitable, but they are realised nonetheless because they are indispensable to serve the population's needs. This is commonly the case with water networks and some transport infrastructure.

4. *The natural environment.* Increasingly, the environment is a crucial concern not only because of the need to preserve and maintain sustainable ecosystems, but also because of the massive changes that human activity is producing on global natural equilibria. Projects must now be designed to cope with increasing temperatures, rising sea levels, drought, floods, and higher population densities in most cities.

5. *Disruptions caused by market conditions.* In addition to the social and economic factors listed in items 1 to 3 above, less predictable

situations can also affect projected operation plans. These include economic crises, conflicts, shifts in public attitudes, bankruptcies and disruptions caused by new technologies.

6. *Accidents and natural disasters.* Accidents and natural disasters can seriously affect the integrity of physical infrastructure, but even in cases where the effects of these events on operations are short-term, they may derail strategic and tactic objectives and thus compromise the long-term viability of the facilities.

Often, these changes occur simultaneously and it becomes difficult to calculate the magnitude of their individual contributions.

Adopting a quantitative approach for infrastructure assessment has many advantages, besides the objective comparison of scenarios: by verbalising their problem, the stakeholders are forced to reflect on the rules and constraints that actually define their operations; they are compelled to adjust their expectations and articulate trade-offs explicitly; they can visualise alternatives of what is viable and what is not by examining exact or approximate solutions. Finally, delivering software to automatically solve the problem repeatedly allows operators to test what-if scenarios and can release valuable staff time for use in other business priorities.

The purpose of this chapter is to demonstrate that the combination of Operations Research (OR) and Machine Learning (ML) is an appropriate methodology to assess the value of existing infrastructure: if, on one hand, by using the best possible operation schedules calculated with plausible scenarios, the existing infrastructure and resources can cope with demand and avoid bottlenecks, the infrastructure is valuable and adequate. If, on the other hand, only uncertain forecasts and poor infrastructure are available, the calculation of a good schedule can still provide insights on where resources should be invested in the future and what aspects of the operation require improved collection of data.

Figure 1 is a schematic of the main quantitative techniques reported in the literature that have been used to validate the value of infrastructure investments. The focus of the reviewed projects, which use a variety of analytical tools (simulation, optimisation and various heuristics), is most frequently on assessing the economics, the value to society, or the effect of disruptions over the normal operations of the infrastructure.

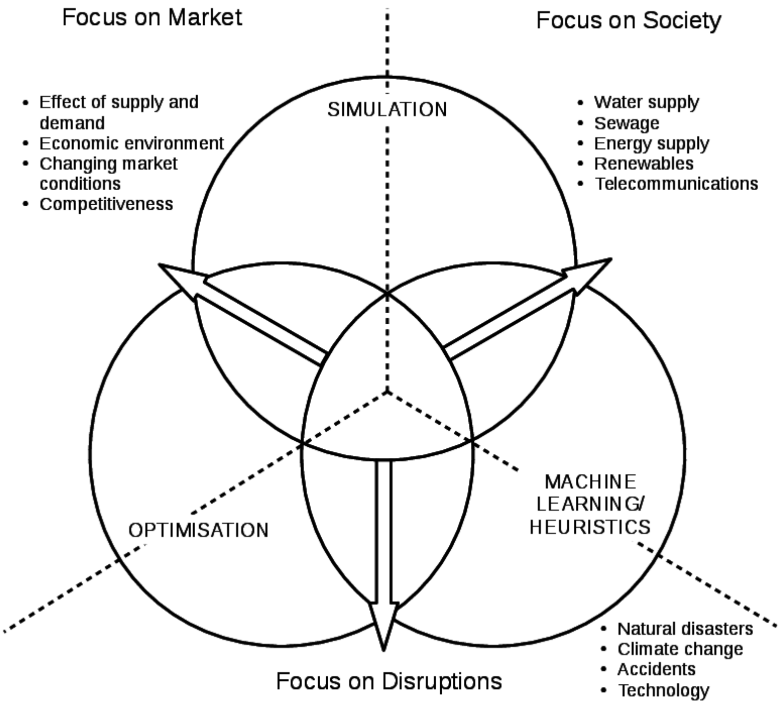


Figure 1. Quantitative techniques to validate the value of infrastructure investments.
The area of application of the validation tools is normally the economics, the value to society or the effect of disruptions over the normal operations of the infrastructure.

The structure of this chapter is as follows. Section 2. presents a non-exhaustive literature review of papers organised by the quantitative technique used, and grouped by area of application, using as a rough guide the schematic in Figure 1. Section 3. introduces the rail scheduling case study where we combine demand forecasting using ML techniques and formal OR methods to assess and maximise the value of already-existing infrastructure. Rail scheduling problems are defined not only by the system's own resources and constraints, but also by the accuracy of the parameters used, which are critical to correctly assess the value of the infrastructure. The discussion in this section and of the results in Section 4. will centre on the combined use of OR and ML techniques. Finally, Section 5. rounds up the discussion.

Literature Survey

The following review is organised around quantitative techniques to assess the value of infrastructure, and within the techniques, by the domain of application of the infrastructure being assessed. Many other concepts could have been used to organise the existing bibliography, as for example motivations (e.g., climate change mitigation, sector the infrastructure is used in, understanding economic cycles), cause of uncertainty (planned or unplanned), and solution approaches (empirical, quantitative, network-based, statistical, survey-based), all of which are equally valid. For a review on actions to mitigate climate change, we refer the reader to Trück et al. (2010), who presented an extensive study of the alternatives at the local level on the cost of mitigation and adaptation to climate change effects. For a complete review for managing disruptions in the abstract using OR techniques, please refer to Snyder et al. (2016), and for a similar review in the context of disaster management, to Galindo and Batta (2013).

Simulation

Simulation is the most popular approach to assess the performance of existing infrastructure. It works on the principle that the system's behaviour emerges from replicating its underlying structure, and has the advantage over analytical optimisation methods that it can describe large and complex systems. The following review illustrates major areas on which simulation has been used to validate the value of existing and planned infrastructure investments: services (such as water networks and energy), risk assessment (i.e., natural disasters and accidents), transportation and logistics, energy and industrial facilities, and is by no means exhaustive.

Simulation is widely used to validate investments in water infrastructure, as population grows and cities continue to develop while governments' resources become increasingly limited. Aleisa et al. (2011) presented a simulation model to make projections of future demand for Kuwait's water network, including the outputs of four existing waste water treatment plants, and concluded that only one of the plants will require a significant upgrade. Etchells and Malano (2005) focused on the increasing uncertainty in water supplies, analysed the sources of uncertainty in water allocation models, and reviewed a few existing models. They emphasised the current deficiencies of software models to validate existing infrastructure, given the uncertainty of all the factors involved. Harris et al. (2009) presented an extensive validation of a model that considers existing infrastructure at Cockburn Sound, Western Australia. The model took into account the capacity and influence of a desalination plant, the effect of harbour modifications and the development of quays, although the objective of this study was to simplify the environmental approval process, rather than assess performance of existing infrastructure. Smajgl et al. (2013) discussed an agent-based model to assess the impact of mainstream dams in the Mekong river area, land use changes, large-scale irrigation, and sea level rise. The model was useful to understand unintended side-effects of new investments (hydro-power, rubber plantations, irrigation schemes for food and energy crops, and mining), paying especial attention on their effect in achieving poverty alleviation targets. Tjandraatmadja et al. (2013) presented a modelling framework for Makassar City, Indonesia, that assesses the suitability of existing and proposed infrastructure to ensure fresh water supply. The framework considered the effects of population growth and climate change and proposes alternatives to enhance the city's water security.

Power load balance is a common theme of simulation projects dedicated to assess the value of electricity network infrastructure. To cite a few examples, Quezada et al. (2014) present an ambitious agent-based system to assess the effect of socio-technical factors that cause stress in the electric distribution network in southeast Queensland, including climate change and investment inertia, and propose adaptation strategies. Köpp et al. (2010) developed a demand and supply forecast method based on neural networks to help cope with the unavoidable intermittency of supply by solar and wind electricity generators. Additionally, they used a simulation to demonstrate that a smooth load balance is possible by using adequate control technology on both the consumer and the producer sides. Huang et al. (2015) combined simulation and optimisation to smooth the power load in electricity distribution networks. In Huang et al.'s (2015) paper, time-varying supply and demand information was used to obtain smooth schedules with minimal peak power and generation costs by using a non-linear model, and the results were validated via simulation.

Some simulation-based assessments of private industrial premises exist, although they are not as numerous as government-funded energy and water projects. This is partly because they tend to be not as capital-intensive, and partly because they are not open to public scrutiny, and thus they are not often reported in the literature. For example, Berends and Romme (2001) used simulation to understand the causes and effects that economic cycles have on capital-intensive industries, and in particular in the paper industry. Like many activities that require large investments, the paper industry is exposed to economic downturns, economies of scale, the need to "keep the machines running", the incentives to keep investing large amounts of capital in order to remain competitive (often in detriment of market prices), and the lag between investment decisions and the moment the new capacity is available for actual production[2]. Berends and Romme validated the model for price and capacity, and demonstrated that cyclicality is to a large extent endogenous (e.g., produced by adoption of new technology) and not only dependent on external market conditions.

Regarding the assessment of logistics and transportation infrastructure, we notice that much recent work uses newly collected data sets that were not available until recently, thanks to recent advances in on-site sensing and data collection. Thekdi and Lambert (2015) introduced the Corridor Trace Analysis tool for assessing the impact of developments on adjacent land on road transport networks. The tool prioritised corridor

[2]Like in many industries, these circumstances foster consolidation into a few major players.

segments that are vulnerable to adjacent land development. Thekdi and Lambert made use of a wealth of data not previously available, thanks to increased digitisation of transport operations, advances in satellite imagery, increased data storage capacities, and increased access to public data resources. Similarly, Higgins et al. (2013) and García-Flores et al. (2014) used recently collected data sets related to cattle movements to assess the robustness of road and cattle-producing infrastructure for the northern states of Australia, which is prone to disruption due to environmental change. The three models presented are strategic and operational, and cover simulation and optimisation aspects in order to provide recommendations on infrastructure repairs and new investments. Tsekeris (2014) used a database provided by the Greek government concerning all public investment projects funded by the European Commission. The database comprises road, railways, airports, seaports and urban public transport, and was used in an ambitious project to assess public expenditure inter-dependencies.

Some simulation studies are aimed specifically to understanding the disruptive effect of accidents and natural disasters. For example, Bruzzone et al. (2000) presented a simulation model of harbour and maritime environments with the aim of designing harbour and maritime infrastructures in order to determine the resources, structures and services needed to face possible emergencies. A case study of an oil spill was used to determine the amount of oil that could reach the coast of Genoa. Ferrario and Zio (2014) proposed an assessment framework to study the safety of a nuclear power plant in case of earthquake, and used Monte Carlo simulation to calculate the probability that the plant enters into an unsafe state. This model enable operators to determine the likelihood of certain parts of the plant recovering earlier than others.

Mathematical Programming

Simulation can model large and complex systems, but has the disadvantage of being only a tool for investigating a system's behaviour, unable by itself to suggest 'good' solutions; simulation is a *descriptive* modelling approach. By contrast, Mathematical optimisation aims at finding the 'best' possible solution. This solution is the 'best' in the sense that it is the most profitable set of actions a decision maker can possibly take, among a very large number of possible combinations of actions: optimisation is a *prescriptive* modelling approach. Because the number of possible combinations of actions is so large, optimisation models tend to be not as complex and detailed as simulation models. Uncertainty can be incorporated in these models by using different methodologies, which have been widely applied to infrastructure value assessment. Some of these include

1. *Stochastic programming.* Aims at producing a solution that, although may not be as good as a deterministic, optimal solution of a problem for which all the parameters are known with certainty, is far from being the worst given the possible uncertain scenarios (Kall and Wallace, 1994).

2. *Approximate dynamic programming.* Based on the idea that decisions should be made using estimates of the value of the states to which an action can take us, in contrast to 'myopic' policies that depend only on what is known at every time step (Powell, 2009).

3. *Robust optimisation.* Uses a measure of 'robustness' in the face of uncertainty, which is represented as deterministic variability in the value of the parameters (Bertsimas and Sim, 2004, Gabrel et al., 2014).

Detailed discussion of mathematical programming and related methodologies is beyond the scope of this chapter, and the interested reader is referred to the included references. We next review some of the applications of these techniques to infrastructure assessment projects. As in the case of simulation above, the following review covers industry, services, risk assessment and transportation, and is by no means exhaustive.

Regarding the assessment of logistics and transportation infrastructure, Burdett et al. (2015) considers the problem of earthwork planning, that is, the problem of strategically moving earth material from one place to another, a necessary task in any infrastructure project. Burdett et al. propose accurate mixed-integer program (MIP) strategic models for linear infrastructure projects that explicitly incorporates fuel consumption and terrain gradients. Mishalani and Koutsopoulos (2002) proposed a general methodology, based on dynamic programming, for modelling the spatial variation of causal variables (such as traffic, soil conditions, pavement design characteristics, weather) and the identification of regions of physical infrastructure that deteriorate uniformly over time. The methodology is useful to assign maintenance work to regions of similar deterioration. The model was validated satisfactorily using detailed data from three roadway facilities.

Regarding the design and operation of water networks, D'Ambrosio et al. (2015) presents a review on the use of mathematical programming techniques in fresh water supply and distribution. Among D'Ambrosio et al.'s (2015) findings is that design, operation, containment detection, and water quality management are the main areas of application of mathematical optimisation. Optimal operation of water networks is intrinsically related to giving existing infrastructure the best possible use. Projects of this type are normally difficult to solve and require simplifications, such as aggregating in the time dimension or linear approximation of nonlinear functions. The interested reader should also refer to Martin et al. (2012), which is a review that is broader in scope than D'Ambrosio et al.'s: it considers mostly mathematical programming models, but multi-objective optimisation and optimal control models are reviewed as well. Applications to optimise the operation of fresh water in industry are also common, such as Arzate et al. (2012), who presented a methodology to perform sensitivity analysis on the costs of investment required to upgrade treatment plants in water networks of refineries.

Regarding energy and electricity networks, a review is provided by Froger et al. (2016) in the context of maintenance scheduling, both corrective and preventative. Froger et al. note that there are differences between the needs of networks in regulated and unregulated markets: deregulated markets often present conflicts of interest between generation and transmission companies, whereas regulated markets focus mainly on reliability and costs. Uncertainty deserves a special mention in this review, and includes papers that use stochastic programming and heuristics.

Many OR papers dedicated to disaster management are related to the maintenance of system flow in emergency situations. For example, Matisziw and Murray (2009) proposed a novel constraint structure for network flow optimisation, and tested in Ohio's road network. The aim of the problem they solved was to identify network facilities most vital to network flow, more specifically, by identifying nodes and arcs associated to worst impact to system flow, given restrictions on the number of facilities damaged.

Approximate solutions for optimisation problems that are too large and difficult can be obtained by using *heuristics*. These are briefly reviewed in the next subsection.

Heuristics and Metaheuristics

Heuristics provide fast ways of solving problems approximately, but they do not provide any proof of optimality or give an estimation of the quality of a solution. Efficient heuristics for certain problems use information that is specific for that type of problem, so that the heuristic can take advantage of the mathematical structure of the problem's search space. This was proved in the *no free lunch* theorem (Wolpert and McReady, 1997). The implication for us as practitioners is that the knowledge about the class of problems that are better suited for the optimisation heuristic method of choice must also hold for the practical problem at hand we are trying to solve. Rothlauf (2011), distinguish three types of heuristic methods: heuristics (for construction or improvement of solutions), approximation algorithms and modern heuristics. Modern heuristics are general, problem-invariant and widely applicable search strategies, and are often called *metaheuristics*. Some applications related to infrastructure validation are reviewed next.

Won et al. (2012) introduce a heuristic based on genetic algorithms and rule extraction to, through future dividend policy, determine an optimal portfolio of investments; dividend policy is understood as the decisions about the relative proportion of dividends out of earnings over time. The proposed algorithm was used to evolve rules that represent the policy. It refines the multiple rules extracted through rule-based algorithms from dividend data sets using a genetic algorithm. Benchmarking on test data sets shows better results that when using rule extraction algorithms alone. Closely related to infrastructure, Zanakis and Becerra-Fernandez

(2005) focused on the use of data mining techniques to identify important factors associated with determining a country's competitiveness and the development of knowledge-based models to predict a country's competitiveness score. Many of the factors Zanakis and Becerra-Fernandez used to train neural networks and regression tree models were related to the countries' existing infrastructure, including telecommunications, water networks, and percentage of urban population.

Regarding transportation infrastructure, Chou (2009) presented a case-based reasoning expert prototype system to determine preliminary project and maintenance costs, using existing information from previous experience of pavement maintenance and construction. The system was used to assist decision makers in project screening and budget allocation, reusing existing project management information and reducing the impact of subsequent cost changes. In a similar vein, Deng et al. (2011) solved the problem of predicting passenger volume in Chinese highways to better estimate investment, management and maintenance decisions by using a combination of rough set theory and neural networks. Their combined rough set theory and neural network approach appeared to be more robust and stable than previously reported approaches. For risk assessment in port infrastructure, Mokhtari et al. (2012) present a decision support framework based in fuzzy set theory and evidential reasoning. The framework defined a hierarchy of risks and was applied to three Iranian ports. The results were tested using sensitivity analysis and the authors report that the methodology is being adapted to other engineering applications.

Regarding urban water networks, Marlow et al. (2015) noted that water distribution infrastructure is undervalued because it is buried and out of sight. This poses an additional challenge on managing the pipe networks, and Marlow et al. proposed a rule-based expert system to provide suggestions about the technical and economic risks related to pipeline repair and maintenance. The paper discussed at length the trade-offs between renovation, replacement and rehabilitation of water distribution infrastructure.

Machine Learning

Machine learning and statistical data analytics techniques play an important role in the valuation process by extracting from data the critical information needed by a simulation or mathematical programming model. This may be just a single nominal value, or a full probability distribution that defines the uncertainty, in a model parameter that specifies how the infrastructure may perform in the future. This includes two key problems: predicting the future demand for the infrastructure, and predicting the useful life left in the existing infrastructure.

The work of Li et al. (2014) developed an approach to failure prediction for pipes within underground water distribution networks. The approach was data driven and combined information regarding the pipe's age, diameter, depth, construction material, protective coating, internal water pressure, surrounding soil type, and others, with historical failure records to predict the probability of a pipe failing in the future. This is critical in determining when pipe renewal should occur and minimising the overall cost incurred by preventative and reactive maintenance.

Assessing damage in civil structures such as bridges is also an important application of machine learning. For instance the construction of bridges require significant upfront investment and any unplanned closures can cause significant disruption to the wider economy of the region. There are a number of approaches to detecting the early onset of damage to such structures that allow preventative maintenance to be scheduled in advance to avoid disruption at peak times. The work of Gul and Catbas (2009) provides an example of how anomaly detection and time series analysis techniques can be used to detect structural health. In Diez et al. (2016) similar unsupervised data driven techniques are being applied to detect the early onset of damage to the iconic Sydney harbour bridge.

Ports represent a critical piece of infrastructure for most economies. The ability to forecast port throughput enables stakeholders to make efficient decisions that not only covers infrastructure investments, but includes management of port development, operational restructuring and tariffs policy. There exist may studies that have developed forecasting methods and range from simple univariate techniques, for example autoregressive integrated moving average Kim et al. (2011) and exponential smoothing Abraham and Ledolter (2009) models, to more complex multivariate methods that model the interdependencies between a broader set of predictor variables such as socio-economic indicators, gross domestic products, commodity prices, etc. These methods include multivariable adaptive regression splines, dynamic factor models, vector autogreressive and auto-regressive integrated moving average with exogenous variables (Geng et al., 2015, Angelopoulos and Chlomoudis, 2015, Intihar et al., 2015). All of these methods are capable of extracting the models from data and capturing the uncertainty in how the demand will evolve in the future.

Hybrid and other Approaches

Combinations of solution methodologies are common and some examples in the literature have already been mentioned; a review with a taxonomy of hybrid approaches can be found in Talbi (2013). In this subsection we briefly review *matheuristics* and other methodologies that have been used to validate the value of infrastructure.

Matheuristics

Matheuristics are a special type of hybrid optimisation approach where mathematical programming algorithms interact with metaheuristics. For an algorithm to be considered a matheuristic, there must be a point in the algorithm where the solution strategy takes advantage of the mathematical structure of the problem or sub-problem being solved. These are also known as "model-based heuristics". One example in the assessment of energy networks is Fischetti et al. (2015), who present a new matheuristic that combines MIP optimisation and a greedy algorithm to solve the smart grid energy management problem. The proposed application addresses the demand side energy management problem by solving a scheduling problem involving multiple appliances with different operational constraints, user preferences, renewable energy sources, and batteries. Matheuristics are currently a very active area of research.

Other Approaches

Finally, some papers adopt other quantitative approaches to validate infrastructure. For example, Tsekeris (2014) presented an ambitious regression model to assess inter-dependencies among infrastructure investments, and assessed the effect of rail, road, air and sea transport investments on each other. Among other findings, the analysis demonstrated that increased relative growth share of maritime transport expenditure can stimulate the investment activity in other non-road transport expenditure categories. Also focusing on inter-dependencies, Ouyang (2016) presented a literature review of existing models for assessing the inter-dependencies of critical infrastructure systems, including empirical and agent-based approaches, which are beyond the scope of our discussion. Vandermeulen et al. (2011) used traditional cost-benefit analysis to assess the economic value of green infrastructure, which is normally not included in land use plans, but which has important social and environmental benefits.

Figure 2 shows some of the most relevant papers reviewed in this Section in the context of their focus and the techniques used.

Case Study: Strategic Rail Planning

The following case study considers containerised freight transport to and from Sydney's Port Botany. Containers can be transported directly between the premises of the freight owners and the port with trucks via the road network. Alternatively, they can be transported via an intermediate intermodal rail terminal. The case study considers the Greater Sydney metropolitan region that extends out 50km from the port. Critical to this study is an assessment of whether the existing network of railway lines and intermodal terminals is adequate to support future growth. This section addresses this issue by combining detailed demand forecasting methods with rail scheduling techniques to assess capacities and determine bottlenecks. The system comprises the following components:

1. **Production Areas**. These areas contain producers that send containers with goods for export and also contain the end customers for imports. The areas aggregate the total supply of containers for

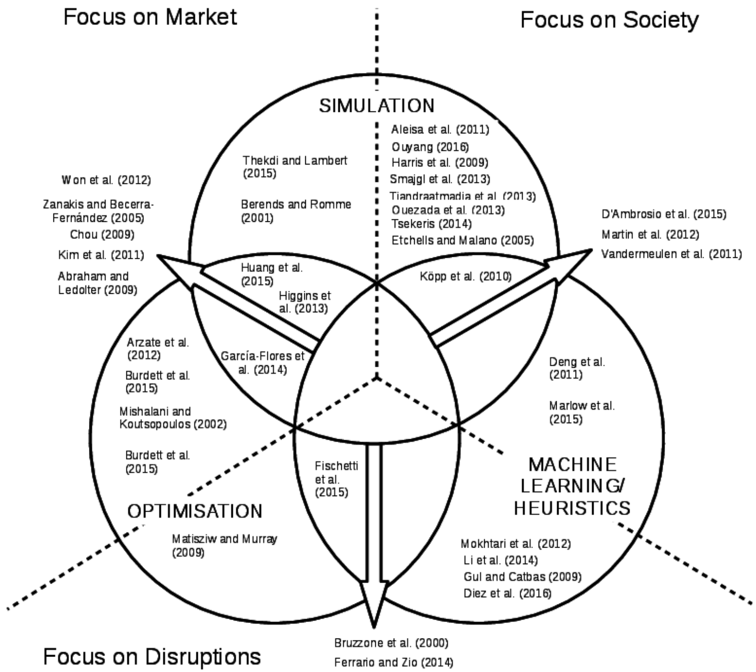


Figure 2. The (non-exhaustive) review of previous work organised by quantitative technique and application area.
Simulation is widely used to assess the value of existing infrastructure. All techniques are normally used in all three application areas.

export as well as demand of containers imported, to the suburb level. Customers within each suburb area determine their transport choice by selecting the service provider that best satisfies their price and service time requirements. The case study considers ten production areas.

2. **Inter-modal terminals (IMT)** An IMT is a transfer facility with road and rail access, and on-site warehousing. In the model presented here, each IMT runs its own rail assets (wagons/locomotives) and container handling assets (forklifts or cranes); in the real world, trains are operated by train companies that may or not own an IMT. For example, some logistics companies own trains and an IMT, but Cooks River IMT does not own any trains. On one hand, the effect of the assumption of IMT's resource ownership on the results of the simulation model is relevant, as the simulation is able to reproduce the competitive behaviour of IMTs and therefore the flow of containers of each company may be altered to suit individual IMT's interests. On the other hand, this assumption by itself does not affect the output of the optimisation model, as this model seeks the greater benefit of all participants in the system regardless of ownership. For the purpose of demonstrating the adequacy of existing infrastructure, this assumption is reasonable, as the costs of operating trains and other resources are the same regardless of ownership.

   Import containers change from trains to trucks at the IMT facilities on their way to the production areas, and export containers are placed into train consists from trucks on their way to the port terminals for shipping. The rate at which containers can be processed by the facility is determined by the number of handling assets, while the rate at which containers can be transported to and from the port is determined by the rail assets. Six IMTs are considered.

3. **Port terminals**. Terminals are export-shipping points and entry points for imports, and also have warehousing facilities available. Two terminals are included in the case study.

A simulation model of this system is described in Banerjee et al.

(2016), with further details found in Chi Thai (2011). It is worth remarking here that, unlike the simulation model or real life operations, the optimisation model presented in this chapter seeks the greater benefit of all participants in the system. The optimisation model suggests to producers the amounts and routes to use to transport their containers to maximise global benefit, and individual IMTs, who in real life compete with each other, have to follow these transportation plans, which may not strictly be reflected in reality. Other approaches to optimisation could be adopted, as for example Data Envelopment Analysis (DEA), which is normally used to compare the performance of (possibly competing) productive units of different sizes on equal grounds. Gundersen et al. (2009) illustrate the application of DEA in multimodal terminals. However, for the present project we consider that simulation is a better approach to model competing IMT behaviours in the context of the Grater Sydney case study. The aim of the optimisation model as presented is to assess and maximise the effectiveness of all existing resources in the supply chain.

A schematic of the system is depicted in Figure 3, and a map with the physical location of all the sites is shown in Figure 4. Producers have the option to send containers for export and receive imports directly from terminals by truck, or to use IMTs. If this is the case, containers are transferred from trains to trucks and vice versa in the IMTs for imports and exports, respectively. Storage facilities exist at terminals and warehouses.

The importance of understanding this system resides in the fact that, on one hand, IMTs reduce road congestion and exploit economies of scale by pooling demand from surrounding areas and using rail to transport containers to and from ports. Additionally, it may be in the interest of local government authorities to encourage rail transport in order to reduce congestion and extend the life of valuable road infrastructure. On the other hand, rail requires additional handling of containers (lift on and lift off), while direct transport from and to ports by truck may be more convenient and save the need for extra-handling in IMTs. In other words, there is a trade-off to be analysed in the attractiveness of truck versus rail that depends on a number of variables such as cost, total travel time,

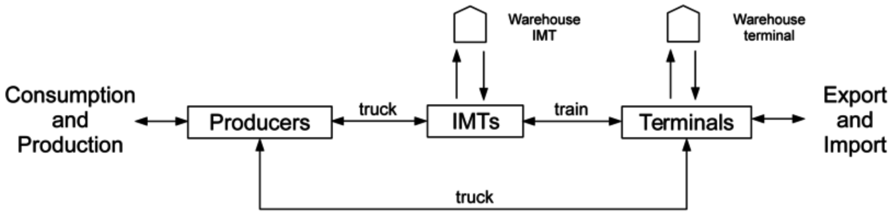


Figure 3. Schematic of the strategic bi-modal transportation tool. Containers are transported by truck between producers and IMTs and between producers and terminals. Transport by rail occurs only between IMTs and terminals. IMTs and terminals count with warehousing space to store containers.

frequency of services, risk, etc.

In Banerjee et al. (2016), we simulated this system with emphasis on IMTs by capturing costs, capacities and service times for different asset mixes, a demand forecast model, and competition from other offerings. By contrast, the present model, introduced in detail in Appendix A, assesses the value of existing infrastructure, namely the rail capacity, road usage and warehousing capacities at IMTs and terminals and provides a means for comparing quantitatively rail and road transportation modes by combining optimisation and the forecasting models developed in Banerjee et al. (2016). To be more precise, the aim of the present problem is to minimise the total costs of operating the entire supply chain, which comprise transportation costs, the costs of violating the soft inventory limits at warehouses, the costs of moving containers to and from warehouses, and the costs of hiring extra trains and container-handling equipment. The model is useful to assess the value of infrastructure because the magnitude of all these costs indicates if the available resources are sufficient or not.

We should point out that the optimisation model presented in Appendix A seeks the greater benefit for all participants in the system, although this may not be strictly true for individual IMTs, who compete

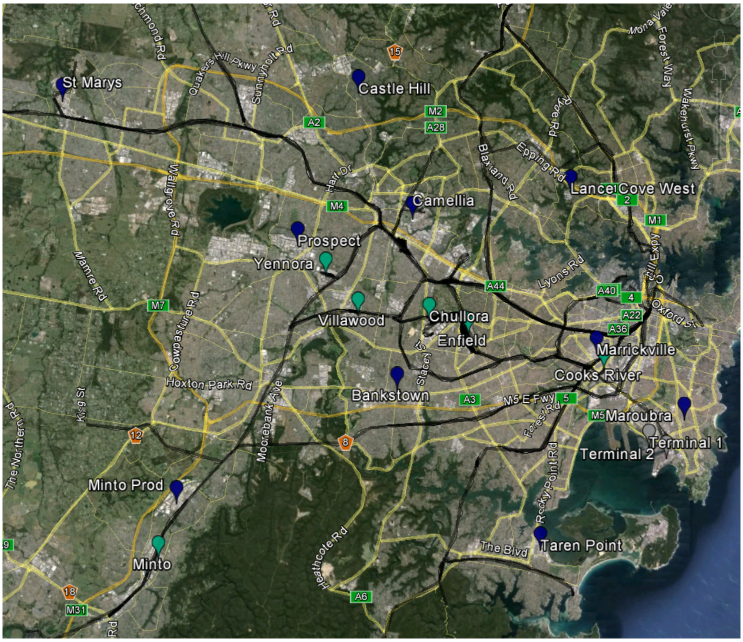


Figure 4. Location of all sites in the case study.
Producers are marked in blue, IMTs in green and terminals in gray. The model considers that all producers are connected to all terminals and to all IMTs by road (light green lines), and that all IMTs are connected to all terminals by rail (black lines); these links are not shown explicitly in the map.

with each other and try to maximise their own gains. However, the model is still useful to assess economic trade-offs and infrastructure value and use, and can be adapted to assess trade-offs for individual IMTs.

Facility Capacities

The basic set of parameters used for sites that hold inventories (IMTs and terminals) is shown in Table 1. The fields in this table include the type of each site as described above; the soft inventory limit, which represents the desirable upper limit on containers that can be stored in the facility; the hard inventory limit, which is the maximum container storage capacity of the facility; and the container handling capacity, which is the maximum number of containers that the facility can handle in a day. The other major set of parameters is the forecasted demand and supply, as shown in Figures 5 and 6.

We estimated that there is a fleet of 50 trains available, and we know that each train is able to carry 32 wagons. It is also known that, in turn, each wagon can accommodate three twenty-foot equivalent unit (TEU) slots, which is the same as three one-TEU containers, or two one-TEU container and one two-TEU container. The road transport cost is AUD \$360.00 plus AUD \$30.00 per additional kilometre, and the rail transport cost is fixed at AUD \$200.00. These values reflect the market conditions discussed in Piyatrapoom et al. (2006), Chi Thai (2011)[3].

Table 1. Parameters of IMTs and terminals in containers

Name	Type	Soft inventory limit	Hard inventory limit	Daily container handling capacity
Minto	IMT	4000	5000	2000
Yennora	IMT	10000	13000	3000
Chullora	IMT	5000	6000	2000
Enfield	IMT	4000	5000	2000
Cooks River	IMT	4000	5000	2000
Villawood	IMT	4000	5000	2000
Terminal 1	Terminal	80000	100000	4000
Terminal 2	Terminal	80000	100000	4000

Predicted Demand

The figures for supply (at producers/terminals) and consumption (at terminals/producers) of the exports/imports, respectively, are the main parameters we vary to assess the value of infrastructure. Figure 5 displays the production of export containers, specified in terms of a TEU, for a base case scenario. This corresponds to a representative 60 day window.

[3]The trends described in these reports for the port of Sydney remain valid, especially the figures regarding the growth areas and forecasts. the reason we claim this is true is that these trends are also reflected in the figures reported by the Australian Bureau of Statistics (Australian Government, 2016).

In addition to this base case, a more variable demand scenario is also considered. This scenario is displayed in Figure 6. Similar graphs for imports are not shown due to space limitations.

The two cases have been determined by analysing the statistical properties of throughputs of forecasted import and exports. The base case figures for container imports and exports in the Greater Sydney region were obtained from publicly available data sources (Australian Government, 2016) and are correct as aggregate figures for all IMT's. The proportion of containers that are sent to individual IMTs on each day was estimated using the percentages reported in New South Wales Government (2015). The variable demand scenario was calculated using the same mean but with a higher standard deviation than the observed data, representing higher future volatility in the throughput volumes. This represents an extreme but plausible scenario to assess the value and effectiveness of existing infrastructure. The two scenarios we analyse in Section 4. are the following: *a*) effect of reducing the number of total available trains, *b*) comparing the base case scenario with a peaky demand scenario.

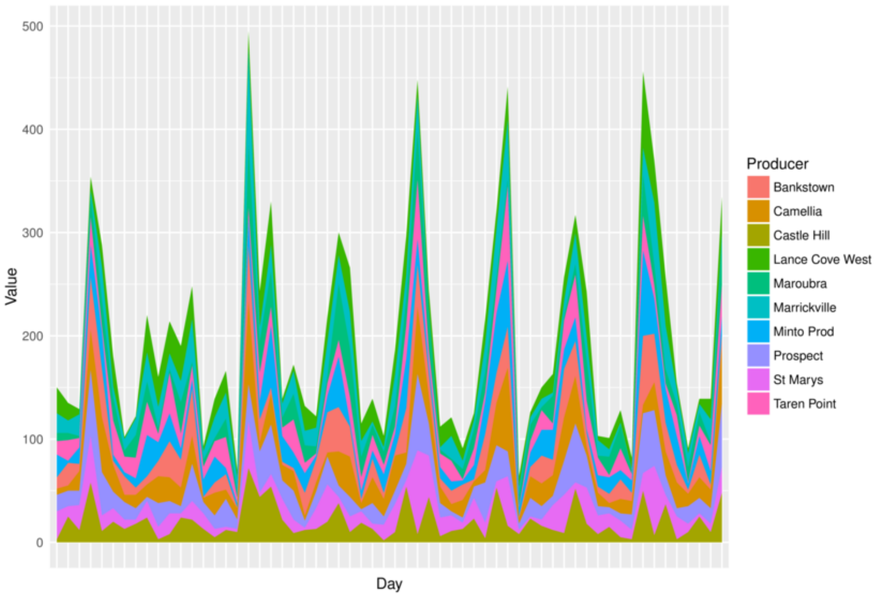


Figure 5. Base case scenario for production of one TEU containers. This Figure refers to exports by producer along the time horizon of 60 days.

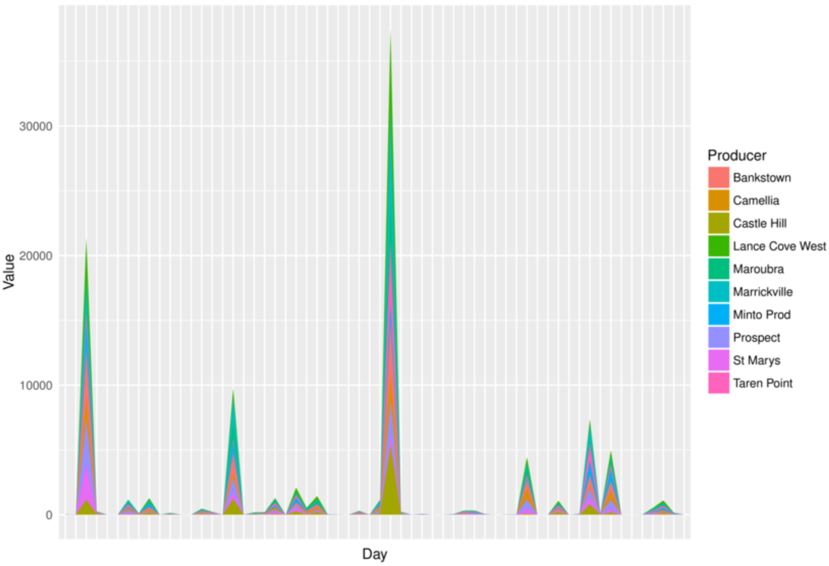


Figure 6. Variable scenario for production of one TEU containers. This Figure refers to exports by producer along the time horizon of 60 days.

Results

The problem was implemented in version 1.5 of the Clojure[4] language with an Excel interface. We obtained all the following results using version 12.4 of the CPLEX[5] optimiser in a 64-bit Intel Xeon CPU with two processors of eight cores (2.27 GHz) each and 48 GB of RAM. The problem has 32220 integer variables and 55352 constraints. A typical run takes around ten minutes to complete.

For the base case, the flow of one and two TEU containers for import per site, as calculated by the optimiser, is shown on Figure 7. Similar maps can be produced for exports, but these are not shown due to space restrictions. The actual amounts of exports and imports and percentage delivered directly to and from each source, or via an IMT are shown in Tables 2 and 3. The optimal solution for the base case does not recom-

[4]http://clojure.org/, accessed on the 8 of April 2013.

[5]http://www-01.ibm.com/software/commerce/optimization/cplex-optimizer/index.html, accessed on the 16 of May 2016.

mend transfer of containers in either direction at Villawood IMT, while Yennora takes the largest proportion of transported containers both for imports and exports, followed by Chullora and Cooks River. Most of the transport to terminals should be done by rail, as only Maroubra is close enough to send and receive containers exclusively by truck. Figure 8 shows the road infrastructure that actually get used at some moment of the 60-day time horizon. Increasing the cost of rail transport does little to modify this road layout, except for encouraging Taren Point to use roads directly to Terminal 2 instead of sending containers via Cooks River IMT (map not shown).

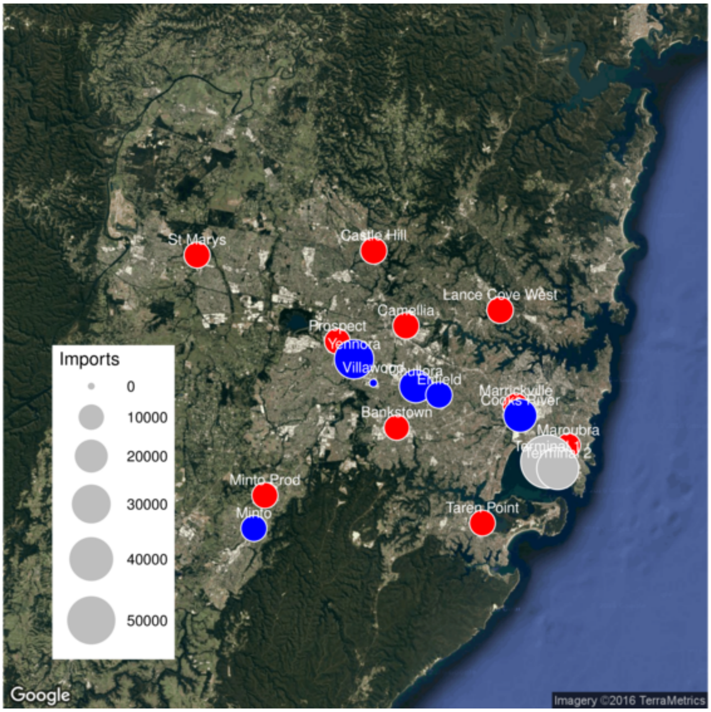


Figure 7. Total number of import containers handled by site. Producers are marked in red, IMTs in blue and terminals in gray. The size of the circle is proportional to the number of import containers handled. A similar map can be produced for exports (not shown).

Table 2. Total number of export containers and percentage received at terminals from IMTs or directly from producers for base case.

	Terminal 1	Terminal 2	% Terminal 1	% Terminal 2
Minto	819	977	7.72	14.89
Yennora	3266	1884	30.77	28.71
Chullora	2084	1584	19.64	24.14
Enfield	1066	704	10.04	10.73
Cooks River	1835	1414	17.29	21.55
Villawood	0	0	0.00	0.00
Lance Cove West	0	0	0.00	0.00
Marrickville	0	0	0.00	0.00
Maroubra	1543	0	14.54	0.00
Taren Point	0	0	0.00	0.00
Minto Prod	0	0	0.00	0.00
Bankstown	0	0	0.00	0.00
Camellia	0	0	0.00	0.00
Prospect	0	0	0.00	0.00
St Marys	0	0	0.00	0.00
Castle Hill	0	0	0.00	0.00
TOTAL:	10613	6563	100.00	100.00

Table 4 shows the effect of reducing the size of the fleet, scenario *a*). It is clear that the system collapses below a fleet size of thirty trains, as the number of trains needed and the number of days when these extra trains are needed also increase substantially. The variable demand scenario also wreaks havoc in the system, with 204 extra trains needed distributed in 34 days. Figure 9 show the result of having a more variable demand, scenario *b*). Under this scenario, some producers use multiple IMTs at different times to avoid the storage and handling bottlenecks that are present. For example, in the map we can observe that Camellia uses Yennora, Villawood and Chullora IMTs at different times during the 60-day horizon. Selecting these transport scheduling decisions is not trivial and requires a quantitative analysis tool like the model we presented. This analysis shows that the existing infrastructure can handle the two scenarios considered, provided that adequate decision support is available to the participants.

Table 3. Total number of import containers and percentage sent from terminals to IMTs or directly to producers for base case

	Terminal 1	Terminal 2	% Terminal 1	% Terminal 2
Minto	5018	4489	8.37	13.29
Yennora	15308	13643	25.54	40.39
Chullora	11014	7488	18.37	22.17
Enfield	6256	3882	10.44	11.49
Cooks River	14531	4274	24.24	12.65
Villawood	0	0	0.00	0.00
Lance Cove West	0	0	0.00	0.00
Marrickville	0	0	0.00	0.00
Maroubra	7818	0	13.04	0.00
Taren Point	0	0	0.00	0.00
Minto Prod	0	0	0.00	0.00
Bankstown	0	0	0.00	0.00
Camellia	0	0	0.00	0.00
Prospect	0	0	0.00	0.00
St Marys	0	0	0.00	0.00
Castle Hill	0	0	0.00	0.00
TOTAL:	59945	33776	100.00	100.00

CONCLUSION

When large infrastructure projects are built, they are not normally amenable to modification after completion, but instead they must operate continuously for many years in the future, possibly with only a few refurbishments during the course of their lifetimes. However, they must still face the constraints and bottlenecks produced by market changes and other disruptions, most of which may not have been foreseen originally and which can have unanticipated repercussions. Fortunately, managers and operators can compensate for miscalculations by playing with operational plans.

Table 4. Effect of fleet size. Between 20 and 30 trains the fleet cannot cope with demand. The cost of additional trains is set to AUD $ 100,000 per train.

Fleet size	Additional trains	Cost of additional trains	No. of days when trains are needed
50	2	200000.00	1
40	8	800000.00	2
30	6	600000.00	4
20	184	18400000.00	28

Figure 8. The optimised bi-modal supply chain of the base case.
In this map, producers are marked with a blue marker, IMTs are indicated with orange, terminals with green, road links are shown in orange and rail links in black.

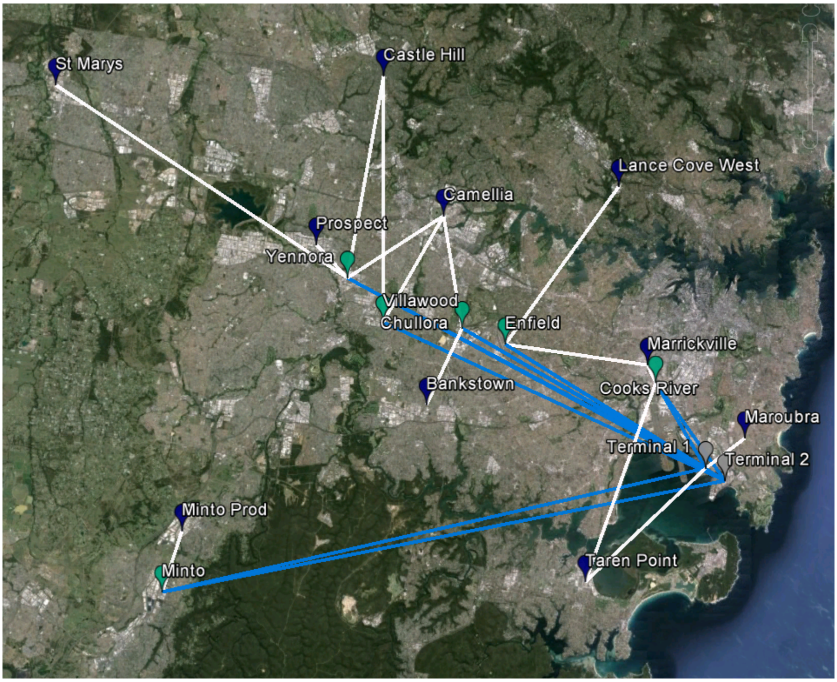


Figure 9. The optimised bi-modal supply chain of a variable demand and supply case scenario.
In this map, producers are marked with a blue marker, IMTs are indicated with light blue, terminals with gray, road links are shown in white and rail links in light blue.

In this chapter, we adopted the point of view of an operator who must make the most out of the existing infrastructure through optimal operation plans and accurate estimates of parameter data. The combination of OR and ML is an appropriate way to assess the value of existing infrastructure: if, on one hand, by using the best possible operation schedule calculated with plausible scenarios, the existing infrastructure and resources can cope with demand and avoid bottlenecks, the infrastructure is valuable and adequate. If, on the other hand, only uncertain forecasts and poor infrastructure are available, the calculation of a good schedule will provide insights of where resources should be invested in future.

Our literature review shows that simulation, optimisation and Machine Learning are widely used, alone or in combination, to study the economic and social value of existing infrastructure, and to analyse the impact of disruptions. The application areas span services (water and electricity networks), industry (mostly capital intensive, such as paper, oil and energy), and sudden, unforeseen events (e.g., climate change, technological innovations and disaster management). We also presented a case study based on the Greater Sydney rail system and studied a scenario with reduced number of trains, and a scenario with variable supply and demand. The results show that Villawood IMT may be an unnecessary facility for the base case from a global point of view, and that Yennora, Chullora and Cooks River take the largest proportion of transported containers both for imports and exports. We also demonstrated that a fleet size of less than 30 trains would struggle to cope with the volume of containers that need to be transported.

The quantitative model presented in Appendix A combines OR and ML, formalises operation rules in the supply chain, and facilitates comparison of scenarios. However, we should note that this model seeks the benefit of all participants in the system, although this may not strictly be in the interest of individual IMTs, who actually compete with each other. In any case, the model is still useful to assess economic trade-offs and infrastructure value and use, and can be modified to analyse the situation of individual participants in the suppy chain.

Acknowledgements

The authors wish to acknowledge Dr Thomas Vitsounis and Dr Yuriy Tyshetskiy for their valuable input, and Ms Blandine Vacher for her contribution in coding the problem.

A Mathematical Model

A1. Nomenclature

Sets

$\mathcal{K}^{\mathtt{EXP}}$ set of commodities (containers) that are exported

$\mathcal{K}^{\mathtt{IMP}}$ set of commodities (containers) that are imported

$\mathcal{K}$ set of commodities (containers)

$\mathcal{L}_P^{\mathtt{EXP}}$ set of valid links where containers for export are transported by rail (i.e., from IMTs to terminals)

$\mathcal{L}_R^{\mathtt{EXP}}$ set of valid links where containers for export are transported by truck (i.e., from producers to IMTs and from producers to terminals)

$\mathcal{L}_P^{\mathtt{IMP}}$ set of valid links where containers for import are transported by rail (i.e., from terminals to IMTs)

$\mathcal{L}_R^{\mathtt{IMP}}$ set of valid links where containers for import are transported by truck (i.e., from IMTs to producers and from terminals to producers)

$\mathcal{L}_P$ set of valid links where containers are transported by rail (i.e., from IMTs to terminals and vice versa)

$\mathcal{L}_R$ set of valid links where containers are transported by truck (i.e., from producers to IMTs and vice versa, and from producers to terminals and vice versa)

$\mathcal{P}$ set of terminals

$\mathcal{R}$ set of IMTs

$\mathcal{S}$ set of sources

$\mathcal{T}$ total number of planning periods (weeks) in the model

Decision variables

$\alpha^{\downarrow}_{ikt}$ number of containers of type $k \in \mathcal{K}$ short of the preferred lower inventory level at site $i \in \mathcal{R} \cup \mathcal{P}$ at time $t \in \mathcal{T}$

$\alpha^{\uparrow}_{ikt}$ number of containers of type $k \in \mathcal{K}$ in excess of the preferred upper inventory level at site $i \in \mathcal{R} \cup \mathcal{P}$ at time $t \in \mathcal{T}$

δ_{it} number of additional container-handling operations (by cranes or forklifts) at site $i \in \mathcal{R} \cup \mathcal{P}$ at time $t \in \mathcal{T}$

γ_t number of additional trains needed at time $t \in \mathcal{T}$

n_{ijt} number of trains running in rail track segment $(i,j) \in \mathcal{L}_P$ at time $t \in \mathcal{T}$

u^{+}_{ikt} number of containers of type $k \in \mathcal{K}$ that go into the warehouse of IMT $i \in \mathcal{R}$ at time $t \in \mathcal{T}$

u^{-}_{ikt} number of containers of type $k \in \mathcal{K}$ that go out of the warehouse of IMT $i \in \mathcal{R}$ at time $t \in \mathcal{T}$

v^{+}_{ikt} number of containers of type $k \in \mathcal{K}$ that go into the warehouse of terminal $i \in \mathcal{P}$ at time $t \in \mathcal{T}$

v^{-}_{ikt} number of containers of type $k \in \mathcal{K}$ that go out of the warehouse of terminal $i \in \mathcal{P}$ at time $t \in \mathcal{T}$

w_{ikt} number of containers of type $k \in \mathcal{K}$ that are stored in the warehouse of site $i \in \mathcal{R} \cup \mathcal{P}$ at time $t \in \mathcal{T}$

$x^{\mathtt{EXP}}_{ijkt}$ number of containers of type $k \in \mathcal{K}^{\mathtt{EXP}}$ that travel by truck through link $(i,j) \in \mathcal{L}^{\mathtt{EXP}}_R$ at time $t \in \mathcal{T}$

$x^{\mathtt{IMP}}_{ijkt}$ number of containers of type $k \in \mathcal{K}^{\mathtt{IMP}}$ that travel by truck through link $(i,j) \in \mathcal{L}^{\mathtt{IMP}}_R$ at time $t \in \mathcal{T}$

$y^{\mathtt{EXP}}_{ijkt}$ number of containers of type $k \in \mathcal{K}^{\mathtt{EXP}}$ that travel by rail through link $(i, j) \in \mathcal{L}^{\mathtt{EXP}}_{P}$ at time $t \in \mathcal{T}$

$y^{\mathtt{IMP}}_{ijkt}$ number of containers of type $k \in \mathcal{K}^{\mathtt{IMP}}$ that travel by rail through link $(i, j) \in \mathcal{L}^{\mathtt{IMP}}_{P}$ at time $t \in \mathcal{T}$

$z^{\mathtt{EXP}}_{ikt}$ number of containers of type $k \in \mathcal{K}$ that go to export from terminal $i \in \mathcal{P}$ at time $t \in \mathcal{T}$

$z^{\mathtt{IMP}}_{ikt}$ number of containers of type $k \in \mathcal{K}$ that come as import to producer $i \in \mathcal{S}$ at time $t \in \mathcal{T}$

Parameters

CC_i cost of additional container-handling equipment (cranes or fork-lifts) at site $i \in \mathcal{R} \cup \mathcal{P}$ at any time

CM_{ikt} cost of moving containers at site $i \in \mathcal{R} \cup \mathcal{P}$ of commodity $k \in \mathcal{K}$ at time $t \in \mathcal{T}$

$CONS_{ikt}$ number of containers of type $k \in \mathcal{K}$ that are imported to producer $i \in \mathcal{S}$ at time $t \in \mathcal{T}$

EXP_{ikt} number of containers of type $k \in \mathcal{K}$ that are exported from terminal $i \in \mathcal{P}$ at time $t \in \mathcal{T}$

FC_t cost of setting up additional trains at time $t \in \mathcal{T}$

$MAXNT$ maximum number of trains available at any moment

N number of wagons in a train (normally 32)

NS_k number of slots taken by container of type $k \in \mathcal{K}$

$PROD_{ikt}$ number of containers of type $k \in \mathcal{K}$ produced at $i \in \mathcal{S}$ at time $t \in \mathcal{T}$

PTC_{ijt} cost of moving a train at time $t \in \mathcal{T}$ from site $i \in \mathcal{R} \cup \mathcal{P}$ to site $j \in \mathcal{R} \cup \mathcal{P}, i \neq j$ by rail

Q_{ij} maximum capacity of road segment $(i, j) \in \mathcal{L}_R$

RTC_{ijkt} cost of transporting commodity $k \in \mathcal{K}$ at time $t \in \mathcal{T}$ from site $i \in \mathcal{S} \cup \mathcal{R}$ to site $j \in \mathcal{S} \cup \mathcal{P}, i \neq j$ by truck

SP_{kt} selling price of container of type $k \in \mathcal{K}$ at time $t \in \mathcal{T}$

SVC_{it} cost of violating the soft inventory limits of site $i \in \mathcal{R} \cup \mathcal{P}$ at time $t \in \mathcal{T}$

$W_i^{\mathtt{max}}$ maximum number of containers that cranes (or forklifts) at site $i \in \mathcal{R} \cup \mathcal{P}$ can move at any time

$WDES_i^{\mathtt{max}}$ desirable upper limit on storage capacity at site $i \in \mathcal{R} \cup \mathcal{P}$ at any time

$WDES_i^{\mathtt{min}}$ desirable lower limit on storage capacity at site $i \in \mathcal{R} \cup \mathcal{P}$ at any time

$WLIM_i$ maximum limit on storage capacity at site $i \in \mathcal{R} \cup \mathcal{P}$ at any time

Sub-indexes

i, j site

k commodity (i.e., type of container)

p terminal

t planning period

A2. Formulation

The aim of the problem is to maximise the profit, which is the sales of all containers minus the total costs of operating the entire supply chain. These costs comprise transportation costs, the costs of violating the soft inventory limits at warehouses, the costs of moving containers to and from warehouses, and the costs of hiring extra trains and container-handling equipment if the available resources are not sufficient.

Let us define the decision variables $x_{ijkt}^{\mathtt{EXP}}$ and $y_{ijkt}^{\mathtt{EXP}}$ as the containers for export that travel between sites i and j carrying commodity k at period t by truck and by rail, respectively; $x_{ijkt}^{\mathtt{IMP}}$ and $y_{ijkt}^{\mathtt{IMP}}$ as the containers for import that travel between sites i and j carrying commodity k at period t by truck and by rail, respectively; $z_{ikt}^{\mathtt{EXP}}$ and $z_{jkt}^{\mathtt{IMP}}$ the containers that go to export and that come as import at terminal i, respectively; u_{ikt}^{+} and u_{ikt}^{-} the containers that are moved into and out of the IMTs' warehouses, respectively; v_{ikt}^{+} and v_{ikt}^{-} the containers that are moved into and out of the terminals' warehouses, respectively; w_{ikt} and w_{ikt} the containers that are stored in the IMT and terminal warehouses, respectively; $\alpha_{ikt}^{\uparrow}$ and $\alpha_{ikt}^{\downarrow}$ the amount by which the minimum and maximum desired inventory levels at sites with warehouses (i.e., IMTs and terminals) are violated, respectively; γ_t the number of additional trains needed at time t and δ_{it} the number of additional container-handling equipment at IMT i at time t. We can now define the objective function as

$$
\begin{aligned}
\text{Maximise}\quad & \sum_{i\in\mathcal{P}} SP_{kt} \sum_{k\in\mathcal{K}^{\mathtt{IMP}}} \sum_{t\in\mathcal{T}} z_{ikt}^{\mathtt{IMP}} + \sum_{i\in\mathcal{P}} SP_{kt} \sum_{k\in\mathcal{K}^{\mathtt{EXP}}} \sum_{t\in\mathcal{T}} z_{ikt}^{\mathtt{EXP}} \\
& - \sum_{(i,j)\in\mathcal{L}_R^{\mathtt{EXP}}} \sum_{k\in\mathcal{K}^{\mathtt{EXP}}} \sum_{t\in\mathcal{T}} RTC_{ijkt}\ x_{ijkt}^{\mathtt{EXP}} - \sum_{(i,j)\in\mathcal{L}_R^{\mathtt{IMP}}} \sum_{k\in\mathcal{K}^{\mathtt{IMP}}} \sum_{t\in\mathcal{T}} RTC_{ijkt}\ x_{ijkt}^{\mathtt{IMP}} \\
& - \sum_{(i,j)\in\mathcal{L}_P} \sum_{t\in\mathcal{T}} PTC_{ijt}\ n_{ijt} - \sum_{i\in\mathcal{R}\cup\mathcal{P}} \sum_{t\in\mathcal{T}} SVC_{it} \sum_{k\in\mathcal{K}} (\alpha_{ikt}^{\uparrow} - \alpha_{ikt}^{\downarrow}) \\
& - \sum_{i\in\mathcal{R}} \sum_{k\in\mathcal{K}} \sum_{t\in\mathcal{T}} CM_{ikt} \left[u_{ikt}^{+} + u_{ikt}^{-} + \sum_{j\in\mathcal{S}} (x_{jikt}^{\mathtt{EXP}} + x_{ijkt}^{\mathtt{IMP}}) + \sum_{j\in\mathcal{P}} (y_{ijkt}^{\mathtt{EXP}} + y_{jikt}^{\mathtt{IMP}}) \right] \\
& - \sum_{i\in\mathcal{P}} \sum_{k\in\mathcal{K}} \sum_{t\in\mathcal{T}} CM_{ikt} \left[v_{ikt}^{+} + v_{ikt}^{-} + \sum_{j\in\mathcal{S}} (x_{jikt}^{\mathtt{EXP}} + x_{ijkt}^{\mathtt{IMP}}) + \sum_{j\in\mathcal{R}} (y_{jikt}^{\mathtt{EXP}} + y_{ijkt}^{\mathtt{IMP}}) + z_{ikt}^{\mathtt{EXP}} + z_{ikt}^{\mathtt{IMP}} \right] \\
& - \sum_{t\in\mathcal{T}} FC_t \gamma_t - \sum_{i\in\mathcal{R}} CC_i \sum_{t} \delta_{it}\ . \qquad (1)
\end{aligned}
$$

The problem is subject to the following constraints:

1. *Soft inventory capacity constraints.* There is a penalty for exceeding the desirable limits to the amount of containers that may be stored in the warehouses. For both IMTs and terminals:

$$
WDES_i^{\mathtt{min}} - \sum_{k\in\mathcal{K}} \alpha_{ikt}^{\downarrow} \le \sum_{k\in\mathcal{K}} w_{ikt} \le WDES_i^{\mathtt{max}} + \sum_{k\in\mathcal{K}} \alpha_{ikt}^{\uparrow} \qquad \forall i \in \mathcal{R}\cup\mathcal{P} \quad \forall t \in \mathcal{T}\ . \qquad (2)
$$

2. *Hard inventory capacity constraints.* There is a maximum number of containers that can be stored at sites with warehousing capability:

$$WDES_i^{\mathtt{max}} + \sum_{k \in \mathcal{K}} \alpha_{ikt}^{\uparrow} \leq WLIM_i \qquad \forall i \in \mathcal{R} \cup \mathcal{P} \quad \forall t \in \mathcal{T} \,. \tag{3}$$

3. *Crane capacity constraints.* There is a limit on how many containers the cranes (or forklifts) can move at any time period, and a penalty if the available container moving resources are not sufficient. For IMTs, containers must be moved in and out of the warehouses, as well as to trucks or trains if they are going to be brought into and sent out of the facility at the current planning period:

$$\begin{aligned}\sum_{k \in \mathcal{K}} (u_{ikt}^{+} + u_{ikt}^{-} + \sum_{j \in \mathcal{S}} x_{jikt}^{\mathtt{EXP}} + \sum_{j \in \mathcal{S}} x_{ijkt}^{\mathtt{IMP}} \\ + \sum_{j \in \mathcal{P}} y_{ijkt}^{\mathtt{EXP}} + \sum_{j \in \mathcal{P}} y_{jikt}^{\mathtt{IMP}}) - \delta_{it} \leq W_i^{\mathtt{max}} \quad \forall i \in \mathcal{R},\ \forall t \in \mathcal{T} \,.\end{aligned} \tag{4}$$

For terminals, in addition to all the containers moved into and out of the warehouses and the train and truck loading and unloading operations, we need to consider that the containers for export must be loaded into the ships, and containers for import must be unloaded from the ships:

$$\begin{aligned}\sum_{k \in \mathcal{K}} (v_{ikt}^{+} + v_{ikt}^{-} + \sum_{j \in \mathcal{S}} x_{jikt}^{\mathtt{EXP}} + \sum_{j \in \mathcal{S}} x_{ijkt}^{\mathtt{IMP}} + \sum_{j \in \mathcal{R}} y_{jikt}^{\mathtt{EXP}} + \sum_{j \in \mathcal{R}} y_{ijkt}^{\mathtt{IMP}} + \\ z_{ikt}^{\mathtt{EXP}} + z_{ikt}^{\mathtt{IMP}}) - \delta_{it} \leq W_i^{\mathtt{max}} \quad \forall i \in \mathcal{P},\ \forall t \in \mathcal{T} \,,\end{aligned} \tag{5}$$

where δ_{it} is the number of additional container-handling equipment needed at site i at time t and

$$z_{ikt}^{\mathtt{EXP}} = EXP_{ikt} \qquad \forall i \in \mathcal{P},\ \forall k \in \mathcal{K}^{\mathtt{EXP}},\ \forall t \in \mathcal{T} \,, \tag{6}$$

$$z_{ikt}^{\mathtt{IMP}} = IMP_{ikt} \qquad \forall i \in \mathcal{P},\ \forall k \in \mathcal{K}^{\mathtt{IMP}},\ \forall t \in \mathcal{T} \,. \tag{7}$$

4. *Conservation constraints at IMTs.* The expressions that define the

flow of containers at IMTs and their warehouses are, respectively,

$$\sum_{j\in\mathcal{S}} x^{\texttt{EXP}}_{jikt} + \sum_{j\in\mathcal{P}} y^{\texttt{IMP}}_{jikt} + u^{-}_{ikt} = \sum_{j'\in\mathcal{S}} x^{\texttt{IMP}}_{ij'kt} + \sum_{j'\in\mathcal{P}} y^{\texttt{EXP}}_{ij'kt} + u^{+}_{ikt}$$

$$\forall i \in \mathcal{R},\ \forall k \in \mathcal{K},\ \forall t \in \mathcal{T},\ \text{and} \tag{8}$$

$$w_{ik,t+1} = w_{ikt} + u^{+}_{ikt} - u^{-}_{ikt} \qquad \forall i \in \mathcal{R},\ \forall k \in \mathcal{K},\ \forall t \in \mathcal{T}\ . \tag{9}$$

We consider that the initial inventory level at all warehouses is zero.

5. *Conservation constraints at terminals.* The expressions that define the flow of containers at terminals and their warehouses are, respectively,

$$\sum_{j\in\mathcal{S}} x^{\texttt{EXP}}_{jikt} + \sum_{j\in\mathcal{R}} y^{\texttt{EXP}}_{jikt} + v^{-}_{ikt} + z^{\texttt{IMP}}_{ikt} = \sum_{j'\in\mathcal{S}} x^{\texttt{IMP}}_{ij'kt} + \sum_{j'\in\mathcal{R}} y^{\texttt{IMP}}_{ij'kt} + v^{+}_{ikt} + z^{\texttt{EXP}}_{ikt}$$

$$\forall i \in \mathcal{P},\ \forall k \in \mathcal{K},\ \forall t \in \mathcal{T},\ \text{and} \tag{10}$$

$$w_{ik,t+1} = w_{ikt} + v^{+}_{ikt} - v^{-}_{ikt} \qquad \forall i \in \mathcal{P},\ \forall k \in \mathcal{K},\ \forall t \in \mathcal{T}\ . \tag{11}$$

6. *Conservation constraints at producers.* It is assumed that the producers do not hold inventory and that all containers produced must enter the supply chain. The expressions that define the flow of containers in these sites are

$$\sum_{j\in\mathcal{R}\cup\mathcal{P}} x^{\texttt{IMP}}_{jikt} = CONS_{ikt} \qquad \forall i \in \mathcal{S},\ \forall k \in \mathcal{K}^{\texttt{IMP}},\ \forall t \in \mathcal{T}\ , \tag{12}$$

$$\sum_{j\in\mathcal{R}\cup\mathcal{P}} x^{\texttt{EXP}}_{ijkt} = PROD_{ikt} \qquad \forall i \in \mathcal{S},\ \forall k \in \mathcal{K}^{\texttt{EXP}},\ \forall t \in \mathcal{T}\ , \tag{13}$$

where $PROD_{ikt}$ and $CONS_{ikt}$ are the amounts of produced and consumed containers of export and import commodities, respectively, in producer i.

7. *Conservation of exports.* All commodities produced at producers must eventually go to export:

$$\sum_{i\in\mathcal{P}}\sum_{t\in\mathcal{T}} z^{\texttt{EXP}}_{ikt} = \sum_{j\in\mathcal{S}}\sum_{t\in\mathcal{T}} PROD_{jkt} \qquad \forall k \in \mathcal{K}^{\texttt{EXP}}\ . \tag{14}$$

8. *Constraints for consist assembly.* The space for containers in the trains is limited. We know that wagons have room for three slots; a slot is more properly known as a twenty-foot equivalent unit (TEU). We also know that containers can be 40 foot and occupy two TEUs, or 20-foot and occupy one TEU. For the time being, the model only distinguishes between containers for export or import. Thus, we have four commodities, $k = \{\texttt{20ft-EXP}, \texttt{40ft-EXP}, \texttt{20ft-IMP}, \texttt{40ft-IMP}\}$. Let n be the number of wagons in a consist (normally 32), n_{ijt} the number of trains that travel between sites i and j, and γ_t the number of additional trains needed in excess of the total available, $MAXNT$. To accommodate containers into trains, we need to define additional constraints. First, knapsack-like constraints so that the number of containers transported by rail in a road segment is not more than the capacity of all the trains travelling in that road segment at any given period:

$$\sum_{k \in \mathcal{K}^{\texttt{IMP}}} NS_k\ y^{\texttt{IMP}}_{ijkt} - 3\ N\ n_{ijt} \leq 0 \qquad \forall (i,j) \in \mathcal{L}^{\texttt{IMP}}_P,\ \forall t \in \mathcal{T}\ , \text{ and} \tag{15}$$

$$\sum_{k \in \mathcal{K}^{\texttt{EXP}}} NS_k\ y^{\texttt{EXP}}_{ijkt} - 3\ N\ n_{ijt} \leq 0 \qquad \forall (i,j) \in \mathcal{L}^{\texttt{EXP}}_P,\ \forall t \in \mathcal{T}\ , \tag{16}$$

where NS_k is the number of slots a container of type k takes. The total number of additional trains needed is

$$\sum_{(i,j) \in \mathcal{L}_P} n_{ijt} - \gamma_t \leq MAXNT \qquad \forall t \in \mathcal{T}\ . \tag{17}$$

The model assumes that every truck can only carry one container.

9. *Road capacity constraints.* We assume that roads have a fixed capacity:

$$\sum_{k \in \mathcal{K}} x^{\texttt{EXP}}_{ijkt} \leq Q_{ij} \qquad \forall (i,j) \in \mathcal{L}^{\texttt{EXP}}_R,\ \forall t \in \mathcal{T}\ , \tag{18}$$

$$\sum_{k \in \mathcal{K}} x^{\texttt{IMP}}_{ijkt} \leq Q_{ij} \qquad \forall (i,j) \in \mathcal{L}^{\texttt{IMP}}_R,\ \forall t \in \mathcal{T}\ , \tag{19}$$

where Q_{ij} is the capacity of the road segment (i,j).

10. *Upper bounds.* Finally, the upper bounds of the amounts transferred to and from warehouses are:

$$u^+_{ikt} \le W^{\mathtt{max}}_i \qquad \forall i \in \mathcal{R},\ \forall k \in \mathcal{K},\ \forall t \in \mathcal{T}\ , \tag{20}$$
$$u^-_{ikt} \le W^{\mathtt{max}}_i \qquad \forall i \in \mathcal{R},\ \forall k \in \mathcal{K},\ \forall t \in \mathcal{T}\ , \tag{21}$$
$$v^+_{ikt} \le W^{\mathtt{max}}_i \qquad \forall i \in \mathcal{P},\ \forall k \in \mathcal{K},\ \forall t \in \mathcal{T}\ , \tag{22}$$
$$v^-_{ikt} \le W^{\mathtt{max}}_i \qquad \forall i \in \mathcal{P},\ \forall k \in \mathcal{K},\ \forall t \in \mathcal{T}\ . \tag{23}$$

REFERENCES

Abraham, B. and J. Ledolter (2009). Statistical methods for forecasting. In *Statistical methods for forecasting.* John Wiley and Sons.

Aleisa, A., O. Alkassar, A. Al-Jadi, S. Al-Sabah, and R. Hishmi (2011, December). A simulation-based evaluation for a strategic residential wastewater master plan. In S. Jain, R. Creasey, J. Himmelspach, K. White, and M. Fu (Eds.), *Proceedings of the 2011 Winter Simulation Conference*, Winter Simulation Conference Proceedings, Phoenix, AZ, pp. 953–962. IEEE.

Angelopoulos, J. and K. Chlomoudis (2015, August). Towards a port services indicator – The case of US Ports. In *Proceedings of the Annual conference of the International Association of Maritime Economists (IAME)*, Hamburg, Germany.

Arzate, E., P. Huitzil, A. González, and B. Martínez (2012, June). Automated optimization model to perform sensitivity analysis on cost of investment required to upgrade treatment plants in water networks. In I. Bogle and M. Fairweather (Eds.), *Proceedings of the 22 European Symposium on Computer Aided Process Engineering (ESCAPE)*, Volume 30 of *Computer-Aided Chemical Engineering*, University College London, London, UK, pp. 1063–1067. Elsevier Science BV.

Australian Government (2016, September). Experimental statistics on international shipping container movements. Technical Report 5368.0.55.018, Australian Bureau of Statistics, ABS House 45

Benjamin Way, Belconnen ACT 2617.
http://www.abs.gov.au/AUSSTATS/abs@.nsf/Lookup/
5368.0.55.018Main+Features12009-10?OpenDocument.

Banerjee, S., G. Mathews, C. Aracena, B. Thorne, Y. Tyshetskiy, and T. Vitsounis (2016, October). Computationally simulating intermodal terminal attractiveness and demand. In *Proceedings of the 23 Intelligent Transport Systems World Congress*, Melbourne, Australia. Accepted.

Berends, P. and A. Romme (2001, December). Cyclicality of capital-intensive industries: a system dynamics simulation study of the paper industry. *Omega 29*(6), 543–552.

Bertsimas, D. and M. Sim (2004, January–February). The price of robustness. *Operations Research 52*(1), 35–53.

Bruzzone, A., R. Mosca, R. Revetria, and S. Rapallo (2000, Jun–Aug). Risk analysis in harbor environments using simulation. *Safety Science 35*(1–3), 75–86.

Burdett, R., E. Kozan, and R. Kenley (2015). Block models for improved earthwork allocation planning in linear infrastructure construction. *Engineering Optimization 47*(3), 347–369.

Chi Thai (2011). Metropolitan Intermodal Terminal Study 2011. Technical report, Shipping Australia Limited.

Chou, J. (2009, March). Web-based CBR system applied to early cost budgeting for pavement maintenance project. *Expert Systems with Applications 36*(2), 2947–2960.

D'Ambrosio, C., A. Lodi, S. Wiese, and C. Bragalli (2015, June). Mathematical programming techniques in water network optimization. *European Journal of Operational Research 243*(3), 774–778.

Deng, W., W. Li, and X. Yang (2011, April). A novel hybrid optimization algorithm of computational intelligence techniques for highway passenger volume prediction. *Expert Systems with Applications 38*(4), 4198–4205.

Diez, A., N. L. D. Khoa, M. Makki Alamdari, Y. Wang, F. Chen, and P. Runcie (2016). A clustering approach for structural health monitoring on bridges. *Journal of Civil Structural Health Monitoring*, 1–17.

Etchells, T. and H. Malano (2005, December). Identifying uncertainty in water allocation modelling. In A. Zerger and R. Argent (Eds.), *Proceedings of the International Congress on Modelling and Simulation (MODSIM05)*, Melbourne, Australia, pp. 2484–2490. Modelling and Simulation Society of Australia and New Zealand: University of Western Australia, Nedlands, WA, Australia.

Ferrario, E. and E. Zio (2014, May). Assessing nuclear power plant safety and recovery from earthquakes using a system-of-systems approach. *Reliability Engineering and System Safety 125*, 103–116.

Fischetti, M., G. Sartor, and A. Zanette (2015, January). MIP-and-refine matheuristic for smart grid energy management. *International Transactions in Operations Research 22*(1), 49–59.

Froger, A., M. Gendrau, J. Mendoza, and E. Pinson (2016, June). Maintenance scheduling in the electricity industry: A literature review. *European Journal of Operational Research 251*(3), 695–706.

Gabrel, V., C. Murat, and A. Thiele (2014, June). Recent advances in robust optimization: An overview. *European Journal of Operational Research 235*(3), 471–483.

Galindo, G. and R. Batta (2013, October). Review of recent developments in OR/MS research in disaster operations management. *European Journal of Operational Research 230*(2), 201–211.

García-Flores, R., A. Higgins, D. Prestwidge, and S. McFallan (2014, March). Optimal location of spelling yards for the northern Australian beef supply chain. *Computers and Electronics in Agriculture 102*, 134–145.

Geng, J., M. Li, Z. Dong, and Y. Liao (2015, June). Port throughput forecasting by mars-rsvr with chaotic simulated annealing particle swarm optimization algorithm. *Neurocomputing 147*, 239–250.

Gul, M. and F. N. Catbas (2009). Statistical pattern recognition for structural health monitoring using time series modeling: Theory and experimental verifications. *Mechanical Systems and Signal Processing 23*(7), 2192 – 2204.

Gundersen, E., E. Kaisar, and P. Scarlatos (2009, July). Assessing container terminal safety and security using Data Envelopment Analysis. In N. Mastorakis and M. Jha (Eds.), *Proceedings of the Second WSEAS International Conference on Urban Planning and Transportation*, Rhodes Islands, Greece, pp. 125–130. WSEAS: World Scientific And Engineering Academy And Society, Ag Loannou Theologou 17-23, 15773 Zographou, Athens, Greece.

Harris, J., J. Antenucci, P. Okely, R. Mills, and P. Yeates (2009, July). Three-dimensional hydrodynamic modelling of a coastal embayment for multiple use industrial discharge and infrastructure development. In R. Anderssen, R. Braddock, and L. Newham (Eds.), *Proceedings of the International Congress on Modelling and Simulation (MODSIM09)*, Cairns, Australia, pp. 4163–4169. Modelling and Simulation Society of Australia and New Zealand: University of Western Australia, Nedlands, WA, Australia.

Higgins, A., I. Watson, C. Chilcott, M. Zhou, R. García-Flores, S. Eady, S. McFallan, D. Prestwidge, and L. Laredo (2013). A framework for optimising capital investment and operations in livestock logistics. *The Rangelands Journal 35*(2), 181–191.

Huang, Y., S. Mao, and R. Nelms (2015, September). Smooth scheduling for electricity distribution in the smart grid. *IEEE Systems Journal 9*(3), 966–977.

Intihar, M., T. Krnberger, and D. Dragan (2015, August). The relationship between the economic indicators and the accuracy of container port throughput forecasting. In *Proceedings of the Annual* *conference of the International Association of Maritime Economists (IAME)*, Hamburg, Germany.

Kall, P. and S. Wallace (1994). *Stochastic programming.* New York: John Wiley and Sons.

Kim, J., J. Pak, Y. Wang, S. Park, and G. Yeo (2011, June). Study on forecasting container volume of port using SD and ARIMA. *Journal of Korean Navigation and Port Research 35*, 343–349.

Köpp, C., H. Mettenheim, M. Klages, and M. Breitner (2010, September). Analysis of electrical load balancing by simulation and neural network forecast. In B. Hu, K. Morasch, S. Pickl, and M. Siegle (Eds.), *Operations Research Proceedings 2010*, Operations Research Proceedings, Univ Bundeswehr Munchen, Neubiberg, Germany, pp. 519–524. German Operations Research Society (GOR): Springer-Verlag Berlin.

Li, Z., B. Zhang, Y. Wang, F. Chen, R. Taib, V. Whiffin, and Y. Wang (2014). Water pipe condition assessment: a hierarchical beta process approach for sparse incident data. *Machine Learning 95*(1), 11–26.

Marlow, D., S. Gould, and B. Lane (2015, December). An expert system for assessing the technical and economic risk of pipe rehabilitation options. *Expert Systems with Applications 42*(22), 8658–8668.

Martin, A., K. Klamroth, J. Lang, G. Leugering, A. Morsi, M. Oberlack, M. Ostrowski, and R. Rosen (Eds.) (2012). *Mathematical optimization of water networks*, Volume 162 of *International Series of Numerical Mathematics.* Basel: Springer.

Matisziw, T. and A. Murray (2009, January). Modeling s-t path availability to support disaster vulnerability assessment of network infrastructure. *Computers and Operations research 36*(1), 16–26.

Mishalani, R. and H. Koutsopoulos (2002, February). Modeling the spatial behavior of infrastructure condition. *Transportation Research Part B 36*(2), 171–194.

Mokhtari, K., J. Ren, C. Roberts, and J. Weng (2012, April). Decision support framework for risk management on sea ports and terminals using fuzzy set theory and evidential reasoning approach. *Expert Systems with Applications 39*(5), 5087–5103.

New South Wales Government (2015). Navigating the future: NSW Ports' 30 year master plan. Technical report, NSW Ports. http://www.nswports.com.au/assets/Publications/NSW-Ports-Master-Plan-2015.pdf.

Ouyang, M. (2016, January). Review on modeling and simulation of interdependent critical infrastructure systems. *Reliability Engineering and System Safety 121*, 43–60.

Piyatrapoom, N., J. M. Bunker, and L. Ferreira (2006). Freight intermodal terminal systems for port of Brisbane, Melbourne and Sydney. Technical report.

Powell, W. (2009, February). What you should know about approximate dynamic programming. *Naval Research Logistics 56*(3), 239–249.

Quezada, G., G. Grozev, S. Seo, and C. Wang (2014, April). The challenge of adapting centralised electricity systems: peak demand and maladaptation in South East Queensland, Australia. *Regional Environmental Change 14*(2), 463–473.

Rothlauf, F. (2011). *Design of Modern Heuristics: Principles and application.* Natural Computing Series. Springer Berlin Heidelberg.

Smajgl, A., J. Ward, and S. Egan (2013, December). Validating simulations of development outcomes in the Mekong region. In J. Piantadosi, R. Anderssen, and J. Bowland (Eds.), *Proceedings of the International Congress on Modelling and Simulation (MODSIM13)*, Adelaide, Australia, pp. 2096–2102. Modelling and Simulation Society of Australia and New Zealand: Modelling and Simulation Society of Australia and New Zealand Inc, MSSANZ, Christchurch, New Zealand.

Snyder, L., Z. Atan, P. Peng, Y. Rong, A. Schmitt, and B. Sinsoysal (2016, February). OR/MS models for supply chain disruptions: a review. *IIE Transactions 48*(2), 89–109.

Talbi, E. (2013, July). Combining metaheuristics with mathematical programming, constraint programming and machine learning. *4OR-A Quarterly Journal of Operations Research 11*(2), 101–150.

Thekdi, S. and J. Lambert (2015, June). Integrated risk management of safety and development on transportation corridors. *Reliability Engineering and System Safety 138*, 1–12.

Tjandraatmadja, G., D. Kirono, L. Neumann, S. Larson, S. Stone-Jovicich, R. Barkey, A. Amran, and M. Selintung (2013, December). Assessing urban water security and climate change adaptation in Makassar, Indonesia. In J. Piantadosi, R. Anderssen, and J. Bowland (Eds.), *Proceedings of the International Congress on Modelling and Simulation (MODSIM13)*, Adelaide, Australia, pp. 2235–2241. Modelling and Simulation Society of Australia and New Zealand: Modelling and Simulation Society of Australia and New Zealand Inc, MSSANZ, Christchurch, New Zealand.

Trück, S., W. Bradford, A. Henderson-Sellers, S. Matthew, J. Scott, M. Street, and R. Taplin (2010). *Climate alert: Climate change monitoring and strategy. Assessing climate change adaptation options for local governments*, pp. 362–399. Sydney, NSW, Australia: Sydney University Press.

Tsekeris, T. (2014, December). Multi-sectoral interdependencies of regional public infrastructure investments. *Socio-Economic Planning Sciences 48*(4), 263–272.

Vandermeulen, V., A. Verspecht, B. Vermeire, G. Huylenbroeck, and X. Gellynck (2011, November). The use of economic valuation to create public support for green infrastructure investments in urban areas. *Landscape and Urban Planning 103*(2), 198–206.

Wolpert, D. and W. McReady (1997, April). No free lunch theorems for optimisation. *IEEE Transactions on Evolutionary Computation 1*(1), 67–82.

Won, C., J. Kim, and J. Bae (2012, December). Using genetic algorithm based knowledge refinement model for dividend policy forecasting. *Expert Systems with Applications 39*(18), 13472–13479.

Zanakis, S. and I. Becerra-Fernandez (2005, October). Competitiveness of nations: A knowledge discovery examination. *European Journal of Operational Research 166*(1), 185–211.

In: Infrastructure Investments
Editors: Gisele Ferreira Tiryaki et al.
ISBN: 978-1-53610-792-0

Chapter 8

FUNDING AND FINANCING FOR WASTEWATER AND SANITATION INFRASTRUCTURE IN SOUTH AFRICA: PRICING, TARIFFS AND OPERATIONAL EFFICIENCY

Cornelius Ruiters* and Maselaganye P. Matji
Council for Scientific and Industrial Research (CSIR)
Pretoria South Africa

ABSTRACT

The aim of this chapter is to address the pricing and tariffing setting for the funding and financing of wastewater and sanitation provision infrastructure in South Africa. The existing public provision pricing and tariff setting models continue to characterise much of the publicly-provided wastewater and sanitation services in South Africa. Wastewater and sanitation infrastrcuture are mostly public-funded through grants and allocations from the National Revenue Fund (NRF), and full replacement costs are not presently used for pricing and tariffs setting for consumers/end-users. Many municipalities do not ring-fence their revenue and costs for water services, and hence are not in a position to inform the setting of tariffs. In general, there is inadequate and

* Corresponding Author: Cornelius Ruiters. E-mail: corneliusruiters07@gmail.com.

inconsistent provisioning for the full life-cycle cost of infrastructure. Maintenance is neglected, and there is mostly under-provisioning for depreciation of capital. In addition, lack of standardisation of reliable, appropriate wastewater treatment, sanitation technologies, and equipment contribute to maintenance challenges. Independent regulation of pricing and tariff setting is needed, as the current model is not effective.

Keywords: pricing, sanitation, tariffs, water services infrastructure, wastewater

INTRODUCTION

In South Africa, and in many other developing countries, urban and rural regions/areas face mainly three patterns of access to sanitation services; each requires appropriate policy, legislative, infrastructure and/or economic responses (Department of Water Affairs and Forestry (DWAF), 2001, 2003; World Bank (WB), 2010; National Treasury of South Africa (NT), 2013a, 2013b), i.e., 1) provision of basic sanitation services; 2) the upgrading to improved models; and 3) the challenge to provide access to improved sanitation to a larger, lower-income population. As population densities increase and water consumption rises, South Africa's burgeoning cities will eventually need to develop more extensive sewerage networks.

The Department of Water and Sanitation (DWS)[1], water utilities and municipalities are the primary players in the water services sector in South Africa, i.e., water supply and sanitation services, regional water services schemes, local water schemes, on-site sanitation and the collection and treatment of wastewater. The access to water services, including sanitation, is a constitutional basic right in South Africa and all people must have access to at least the basic level of service (Republic of South Africa (RSA), 1996). However, users have a reciprocal responsibility, namely, to use services responsibly and with due care, and to pay for services provided over and above the free basic sanitation allocations (DWAF, 1997a, 1997b, 2003). The South African government has prioritised not only the rollout of infrastructure necessary for the rendering of these services but also the provision of free basic services to poor households. The South African White Paper on Basic

[1] Department of Water Affairs and Forestry (known as DWAF) included forestry between 1994 and 2009. The Department of Water Affairs (DWA) was between 2009 and 2014; the research was conducted under DWA. Since May 2014, it has been renamed DWS (Department of Water and Sanitation).

Household Sanitation introduced the equitable share subsidy to allow the local government sector to overcome the burden of service delivery to the very poor (DWAF, 2001, 2003; NT, 2013c).

This chapter discusses the role of the pricing and tariff setting mechanisms in wastewater and sanitation services. It aims to address the pricing and tariffing setting for the funding and financing of wastewater and sanitation provision infrastructure in South Africa. Primary and secondary data was collected from the municipalities and water utilities regarding the wastewater and sanitation services tariffs per consumer type (residential, commercial, industrial and other or associated uses) in pre-defined blocks. A representative size of 269 municipalities, out of a total of 287 municipalities, was sampled, and the survey captured pricing and tariffs as applied by each municipality at present. Reports relating to infrastructure needs, funding and financing activities in wastewater and sanitation throughout South Africa were reviewed. Secondary data collection was obtained from the DWA (DWA, 2010, 2011, 2012), NT (NT, 2011a, 2011b, 2011c, 2011d, 2013a, 2013b, 2013c) Statistics South Africa (Stats SA) (Stats SA, 2012), water institutions, metropolitan municipalities, district and local municipalities (DWA, 2010, 2011, 2012; NT, 2011c, 2011d, 2013a, 2013b, 2013c) for wastewater and sanitation infrastructure. Reviews were done of revenue streams, local debt, expenditure restrictions, and other information relative to funding and financing of wastewater and sanitation infrastructure. The study addressed the following objectives:

- Pricing as a tool for signalling the value to address wastewater and sanitation needs and challenges and also to raise revenue for wastewater and sanitation provision infrastructure in South Africa (cf. Brineco-Garmendia et al., 2008; Organisation for Economic Cooperation and Development (OECD), 2010; WB, 2010; Ruiters, 2011, 2013; Matji and Ruiters, 2014, 2015a; Ruiters and Matji, 2015).
- Pricing and tariff setting given the socio-economic challenges facing South Africa (cf. Matthews, 2009; World Bank, 1994, 2010; Ruiters, 2011, 2013; Matji and Ruiters, 2014, 2015a; Ruiters and Matji, 2015).
- Due to population growth and increased urbanisation, the wastewater and sanitation infrastructure demands are increasing and economic tools (pricing and appropriate tariff setting) are important to demonstrate the true value of service and infrastructure provision (cf. Lamb, 1984; Lang and Merino, 1993; Eberhard, 1999, 2004; Palmer Development Group, 2000; Mander, 2002; Goodman and Hastak,

2006; OECD, 2010; WB, 1994, 2010; Vawda et al., 2011; Ruiters, 2011, 2013; Matji and Ruiters, 2014, 2015a; Ruiters and Matji, 2015b).

Besides this introduction, this chapter is structured into five other parts. The second section describes the wastewater and sanitation services value chain and institutional setup, while the third section focuses on the financing of wastewater and sanitation infrastructure. The fourth part addresses issues related to operational revenue and expenditure, and the fifth section conducts international comparisons of wastewater and sanitation services. The sixth section presents the concluding remarks.

WASTEWATER AND SANITATION SERVICES VALUE CHAIN AND INSTITUTIONAL SETUP

Water services infrastructure provision (including operations and maintenance) is hierarchical and developmental in nature in South Africa, based on administrative and/or political boundaries (DWAF, 1997a, 1997b, 1998, 2008, 2012b, 2013a; Eberhard, 1999, 2004; Hope and Garrod, 2005; Saleth and Dinar, 2005; Karar et al., 2011; Muller, 2014; Van Koppen and Schreiner, 2014). The hierarchy ranges from a national level to a local level with the responsibility for the implementation at each government level (sphere) varying within the administrative boundaries (Figures 1 and 2). The hierarchical impact associated with the value chain is improved accountability for water infrastructure, water security, operations and maintenance, quantity (availability) and quality, and better management of the supply and demand for water.

The under-investment in the water sector in South Africa was estimated at more than US$56 billion[2, 3] (DWA, 2011a; 2013a, 2013b). Whereas the capital investment needed in Africa based on minimum acceptable asset standards and accounting for both new infrastructure and rehabilitation of existing assets are conservatively estimated at $4.5 billion/year and maintenance requirements are $1.5 billion/year, obvious with inflation adjustments (WB, 2010). The scale of the water services infrastructure therefore does warrant immediate

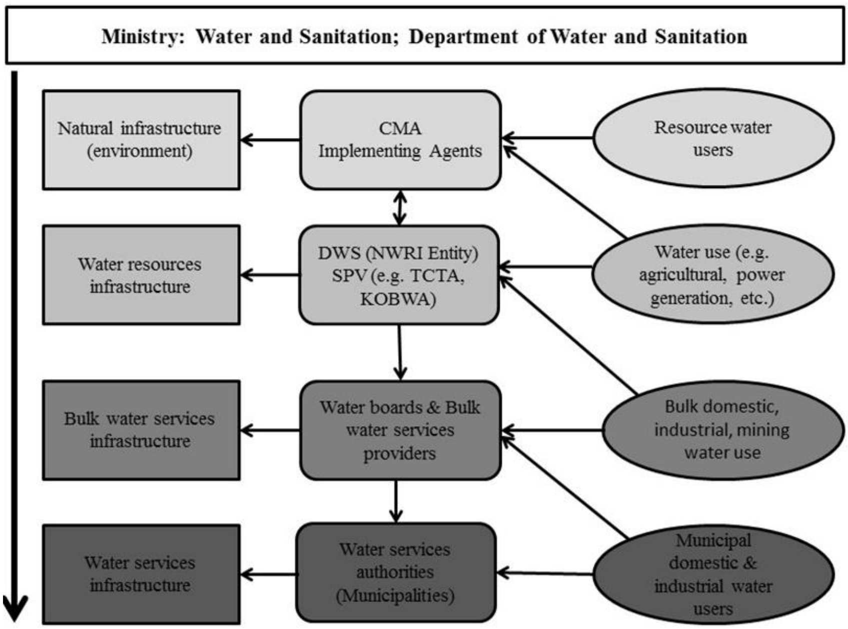
[2] The Rand (R or ZAR) is the South African monetary currency (R11 ≈ 1US$), where 100 cents = R1 (South African Rand).
[3] 1 Billion = 1×10^9.

attention for an appropriate pricing and tariff models for waste water and sanitation services. Thus, the principal drivers for the pricing and tariff setting of wastewater and sanitation services in the financing of water services infrastructure in South Africa (cf. Ruiters, 2011) are:

- ***Funding of water services infrastructure*** **-** National targets for eliminating the backlog in the provision of water and sanitation services have been set. These are challenging targets which will require increased investment, a rapid increase in capacity and much better use of existing capacity. Providing free basic water and sanitation services in a sustainable manner, particularly in rural areas, is also a significant challenge. Currently, the water and sanitation sector is seriously under-financed, revenue management is poor and there is limited or no investment in maintenance. These have led to the deterioration and the eventual collapse of infrastructure at municipal level (sphere).
- ***Maintenance and refurbishment of infrastructure*** **-** The inadequate maintenance and poor performance of water and sanitation services infrastructure pose a public health safety risk. This is mainly due to constraints on strategic budget planning and adequate investment for the rehabilitation, replacement and maintenance of water services infrastructure at local government (municipal) level (Matji and Ruiters, 2014).
- ***Financial management and cost-reflective tariffs*** **-** Water and sanitation users are calling for transparency in financial management and also in pricing and tariff determination process so that the sector can have cost reflective tariffs. Effective financial planning and pricing for the water sector require finding the right mix of revenues from water use charges, tariffs, grants and transfers, i.e., ‘sustainable cost recovery’ (Ruiters, 2011, 2013). Full cost recovery from tariffs, which may theoretically be the ideal solution, in practice remains a distant objective in South Africa. Increasing revenue from water use charges and tariffs requires a comprehensive approach, which includes reforming tariff levels and structures and increasing bill collection rates, but also increasing levels of service and putting in place social protection measures.
- ***Water institutions and governance*** **-** Effective and relevant fully functional water institutions and governance structures for funding and finance flow in the water sector value chain with assurance of

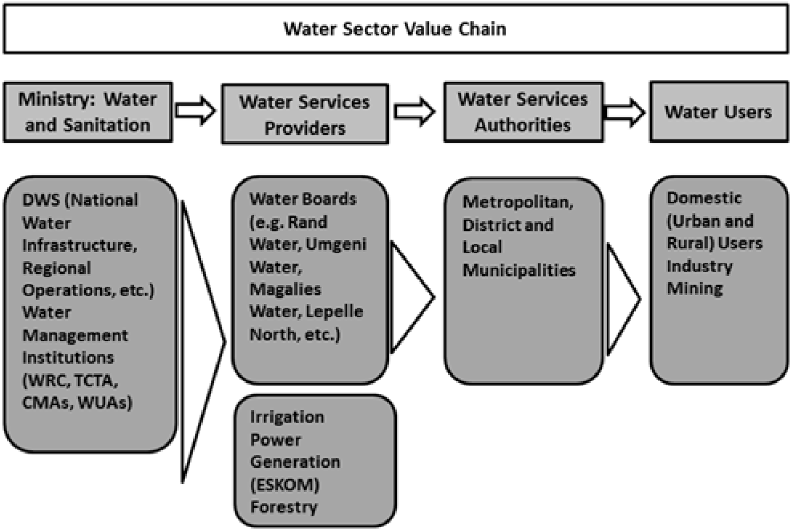
water services infrastructure delivery to all consumers/users (Figure 3; Ruiters and Matji, 2015). Consolidation, reform and modernisation of institutions and governance structures for water services infrastructure (water sector value chain) would provide a greater degree of certainty to the funders, financiers and developers.

Water supply is a key resource for effective wastewater (domestic, industrial and/or commercial) and sanitation (especially water-borne) services at an acceptable level of assurance/reliability, quality and accessibility. This is costly, both in terms of capital investment for infrastructure and operations and maintenance, and bulk collection and reticulation. To sustain both the water supply and sanitation services it is essential to recover costs and to maintain financial viability. Pricing for wastewater and sanitation services is a key principle and a mechanism to ensure that the water is used sustainably.



Source: Ruiters and Matji, 2015.

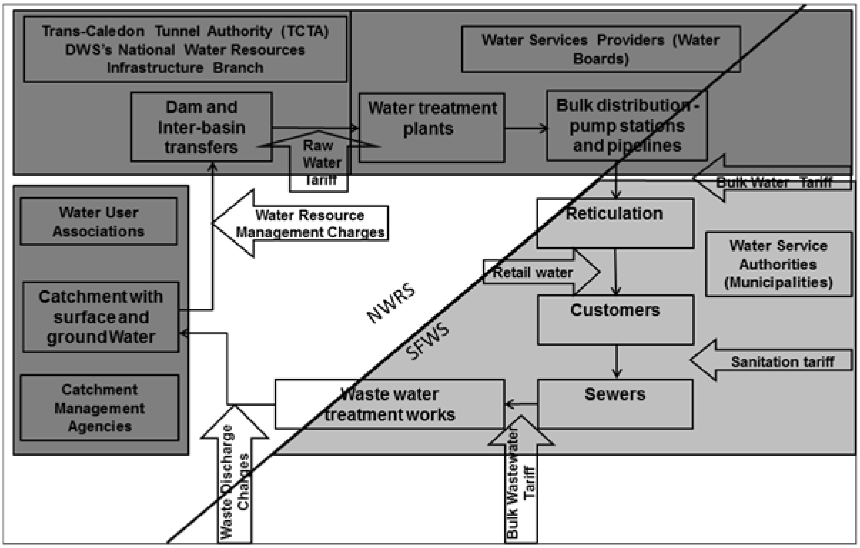
Figure 1. Financial flows and water infrastructure in South Africa.



Source: Ruiters, 2013; Ruiters and Matji, 2015.

Figure 2. Hierarchical institutional water sector value chain in South Africa.

To achieve the above principles, suitable cost reflective tariffs must be set by all relevant water services institutions involved in the water services cycle. The criteria for the setting of tariffs and the procedures are described in various policies and legislation (DWAF, 1997a, 1997b, 2003, 2007). These criteria include: 1) social equity; 2) infrastructure integrity, provision and sustainability; 3) financial sustainability; and 4) economic efficiency. In addition, there is a need for an independent regulator of tariffs and pricing across the whole value chain (Matji, 2013). Tariffs that are paid by consumers at the lower end of the value chain are directly affected by how water resources funding models are structured. The interconnectivity of water services and water resources infrastructure should be taken into account when setting tariffs for wastewater and sanitation services (Ruiters, 2011, 2013; Matji, 2013; Matji and Ruiters, 2014).



Source: Ruiters and Matji, 2015.

Figure 3. The water sector value chain in South Africa at the strategic and operational levels illustrating the water resources[4] and water services[5] components.

FINANCING OF WASTEWATER AND SANITATION INFRASTRUCTURE: PRICING AND TARIFFS

The pricing value chain and tariff setting for wastewater and sanitation services are integral components of the full water value chain (DWAF, 2003; OECD, 2010; WB, 2010). Appropriate pricing and tariff setting for wastewater and sanitation services have become an important issue for South Africa and has substantial impacts on demand and supply (Hollingworth, 1994; Hazelton, 1998; Eberhard, 1999; Palmer Development Group, 1998, 2000; Hosking, 2003; Still, 2003; Matthews, 2009; Ruiters, 2013; Matji and Ruiters, 2015a, 2015b). Different tariffs are employed in South Africa to ensure sufficient revenue to maintain and expand the wastewater and sanitation services infrastructure (cf. Figure 3). However, using standardised unit costs from South Africa's waste water and sanitation services sector budgets can

[4] NWRS = National Water Resources Strategy (DWA, 2013a.).
[5] SFWS = Strategic Framework for Water Services (DWAF, 2003.).

adequately provide for the investments in the different types of sanitation services per province (cf. Figures 4 and 5; NT, 2013a, 2013b).

The primary purpose of setting tariffs (rates) is to ensure that they cover both fixed and variable expenses, and to ensure the costs are spread equitably among customers/users (Eberhard, 1999, 2004; Vawda et al., 2011, Ruiters, 2013; Matji and Ruiters, 2014, 2015a). However, local conditions may require that other factors be considered in the pricing and tariff setting process. Examples of local considerations that may require modification of the tariff structure are:

(i) Where wastewater treatment is at, or very close to capacity, the water tariff structure may be modified to discourage excessive water use to avoid the expense of expanding the wastewater facility.
(ii) Legislative requirements mandate that certain requirements must be considered in setting tariffs for water systems.
(iii) If the systems were financed using loans or revenue bonds, there might be certain loan and/or bond requirements (covenants) that must be met.
(iv) The water entity (owner) may want to minimise the costs to low-income, elderly or fixed-income customers (poor communities) by shifting a greater portion of the costs to the larger (major) water users.

Municipal tariffs in South Africa are approved by local municipal councils in terms of a local tariff policy which must comply with nationally defined norms (DWAF, 1997a, 1997b, 2001, 2002, 2003, 2007; Eberhard, 1999, 2004; NT, 2003; Matji and Ruiters, 2014, 2015a). While tariffs have to comply with the principles and norms of the relevant policies and legislation, however, the interpretation and application vary, specifically at municipal level. Many municipalities do not ring-fence their revenue and costs for water services, hence are not in a position to inform the setting of tariffs. In general, there is inadequate and inconsistent provisioning for the full life-cycle cost of infrastructure. Maintenance is neglected and there is mostly under-provisioning for depreciation of capital. Factors such as inappropriate choice of imported technology that was installed without structured knowledge and skills transfer contribute to maintenance challenges. In addition to this, lack of standardisation of reliable, appropriate wastewater treatment, sanitation technologies, and equipment contribute to maintenance challenges.

Such service delivery is not sustainable and will rely on consistent support and cross-subsidisation from external sources. Financial legislation is addressing some of these deficiencies and compliance to minimum standards for asset management practices (NT, 2003; DPW, 2007). The legislation makes provision for the development and implementation of a framework for asset managent, i.e., 1) asset management plans; 2) asset register; 3) asset condition assessment; 4) asset norms and standards; 5) risk assessment and management; 6) monitoring of asset performance and; 7) asset maintenance management and planning. Future valuation of the tariffs will then be able to assess if adequate provision is made for return on assets, maintenance and operation of the infrastructure. This will require full knowledge of the current and depreciated replacement cost as well as the expected and remaining useful lives of the water services infrastructure (i.e., wastewater and sanitation). To manage the risk of failure, the sector urgently has to implement best practices on asset management and review tariffs on the full life-cycle costing to recover full costs for sustainable service provisioning.

A water services authority (municipality) must, when setting tariffs for wastewater and sanitation services provided to consumers and other users within its area of jurisdiction differentiate, where applicable, between at least the following categories:

(i) sanitation services to households (domestic sewerage and sanitation)
(ii) discharge of industrial effluent to a sewage treatment plant (industrial and commercial sewerage)
(iii) sanitation services other than those specified (i.e., on-site sanitation, latrines, etc.).

The results from this study indicate that the main components usually associated with wastewater and sanitation pricing and charges (tariff) setting at local government (municipal level) as per the water services (including wastewater and sanitation services) pricing policy and strategy (Figures 4 and 5; DWAF, 1997a, 1997b, 2003, 2007) include:

- Sewerage (wastewater) collection costs
- Wastewater treatment costs (including technology choice)
- Economic charges such as incentives to increase economic value and efficiency of water use for sanitation and wastewater services
- Management costs (overheads/administrative, cost of cost-recovery, customer service)

- Financing costs
- Location, distance and topography (bulk cost factors)
- Economy of scale
- Margins to achieve demand management objectives.

Based on the above, a three-category pricing and tariff structure for wastewater services (Box 1; Figures 4 and 5; Table 1) is applied by municipalities as prescribed by policy and legislation in South Africa (DWAF 1997a, 1997b, 2007; NT, 2003), i.e.,:

(i) Category 1 – a wastewater collection charge (includes the treatment costs)
(ii) Category 2 – a bulk wastewater tariff
(iii) Category 3 – a waste discharge tariff (polluter pays principle). However, this category is still under development.

The wastewater tariffs were measured across municipalities; however, comparisons were difficult except for metropolitan municipalities (Figures 4 and 5). From the results of the study, the following tariff types have been identified from the analysis (cf. Box 1; Table 1):

- Metred volume of effluent (US$/kl effluent)
- Tariff mark-up on the metred water use (US$/kl water used)
- Multiplier of the property value (US$/month/US$ of property value)
- Multiplier of the land size (US$/month/m^2 of land area)
- Multiplier of the building size (US$/month/m^2 of building size)
- Fixed tariff per stand size (US$/month for large, medium or small stands)
- Fixed tariff per household/dwelling (US$/house hold/month)

This research provides first-order comparisons (e.g., costs per household per month) to enable effective incorporation of the wastewater tariffs into the supply chain and add value to water services (Table 1). The categories include: 1) a flat rate amount for stands between 0 – 500 m^2, 500 – 1000 m^2; 2) 18% of all tariffs are linked to categories of metred water-use; 3) a free basic water supply per month for poor households for water-use between 0 - 6 kl and; 4) a flat rate for water-use between 6 – 15 kl, etc. The remaining portion (%) of the tariff structure is spread across the other alternative tariff units.

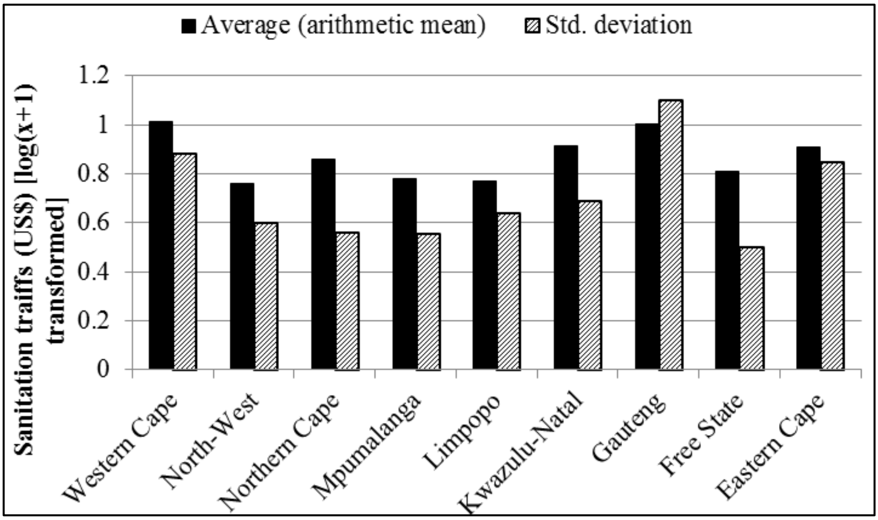


Figure 4. Mean (± SD) municipal sanitation tariffs (residential) per province in South Africa.

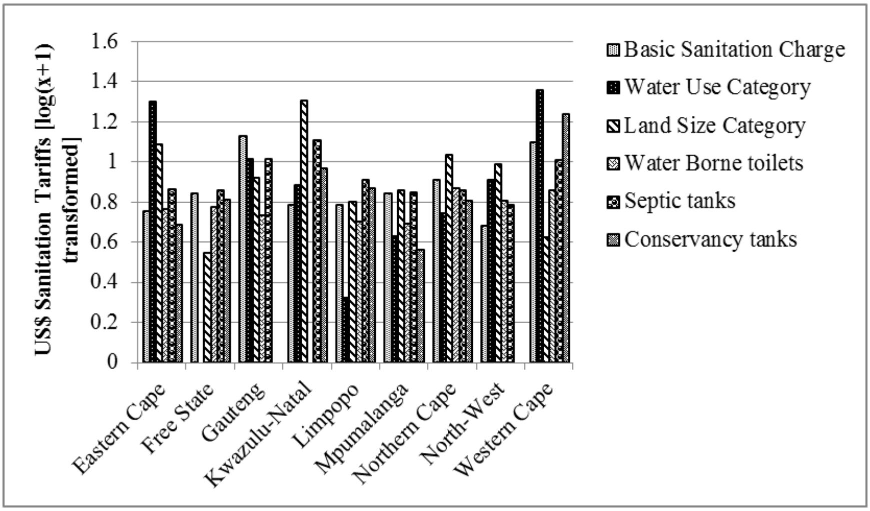


Figure 5. Mean municipal sanitation tariffs (residential) for the major sanitation services types charged per province in South Africa.

Table 1. The residential sanitation tariffs (incl. VAT[6]) in a common unit of US$/month

Type of sanitation tariff	Mean (±SD) tariff (US$/month)
Basic sanitation charge	7.90 ± 5.78
Water-use efficiency	8.33 ± 6.24
Land-size category	7.14 ± 4.66
Building-size category	10.50 ± 2.20
Water connection-size category	21.28 ± 8.69
Toilet-type category:	
Buckets	3.92 ± 1.77
Conservancy tanks	7.30 ± 3.64
Septic tanks	8.53 ± 4.08
Sewer connections	8.45 ± 8.69
VIP toilets	3.32 ± 4.10
WB toilets	4.90 ± 2.20
Chemical toilets	4.42 ± 3.62
Grand mean	7.45 ± 4.48

The mean residential sanitation tariffs result in a wastewater return of 20 kl/household/month (e.g., 60 - 70% of water used) (Table 1). However, unit rates for land value multipliers and building size multipliers are excluded from the consolidation. It is clear that a fixed rate per month (R/month) is most common. To bring the range of tariff units to a comparable base, the following assumptions have been made: for water use, related units change from US$/kl to US$/month assuming an average monthly water use of 30 kl/household/month, which typically reflected the tariff for most municipalities.

The results from the national mean sanitation tariffs per tariff type and service type (e.g., type of toilet) show the following (cf. Figures 4 and 5; Table 1):

- The typical fixed charge was US$7.90/month.
- Waterborne-related sewer connections were typically US$8.36/month with lower rates (e.g., US$4.90/month) for smaller households with one toilet only.

[6] VAT (Value Added Tax) = 14% in South Africa

- Septic tanks typically cost US$8.55/month, were more expensive than conservancy tanks, due to the higher density (loading), and related treatment intensity.
- Municipalities still have tariffs for buckets which cost on average US$3.90/month inclusive of repetitive collection (daily, bi-weekly or other). In comparison to previous years, the number of municipalities with bucket tariffs has decreased drastically, reflecting on the eradication of this sub-standard service.
- Ventilated Improved Pit (VIP) toilets were mostly considered to be free basic sanitation services and thus not included in the tariff structures.
- Indigent tariffs (where listed separately) were about 40% of the average residential tariffs (provinces vary between 32% and 65%).

Box 1: Types of Wastewater and Sanitation Tariff Structure for Supplying Services from Infrastructure in South Africa

Customer classifications: Dividing customers into group, i.e., residential, industrial and commercial. Many systems have different tariffs for each classification of customers/users.

***Wastewater and sanitation tariff structure*:**
Free basic sanitation services for poor communities
US$$x_1$ for basic sanitation charge
US$$x_2$ for water-use category (efficiency)
US$$x_3$ for land size category
US$$x_4$ for water connection size
US$$x_5$ for toilet type
US$$x_6$ for building-size category, with building size multiplier

Uniform single wastewater and sanitation tariff

Customers/users are charged a constant price per service, regardless of the amount used or discharged. This is often coupled to a minimum charge for having the service available.

Example:
US$$x_1$ minimum service availability charge per month for toilet type used.

Advantages: It is easy to administer and may encourage water conservation in water-scarce areas. The costs to the customer/user are in direct proportion to the toilet type used.

Disadvantages: It could encourage misuse/abuse in certain areas or communities.

Increasing wastewater and sanitation service block tariff

The price of wastewater and sanitation increases as the amount (volume) discharge or increases. Each succeeding consumption block is more expensive for the size. This structure is based on the assumption that tariffs should promote water efficiency.

Example:

US$ x_1 = Basic sanitation charge

US$ x_2 = Water-use category, i.e., 0-6 kl used (except for poor communities for which this quantity is free), 6-20 kl used, 20-60 kl used, >60 kl used

US$ x_3 = Land size category, with land size multiplier and land value multiplier

US$ x_4 = Water connection size category

US$ x_5 = Toilet type category

US$ x_6 =Building size category, with building size multiplier

Advantage: This promotes water efficiency, which is especially important in areas with limited water supplies or high wastewater treatment costs. Lower water use means less wastewater and smaller, less expensive wastewater treatment facilities. It provides a reasonable amount at a reasonable price and charges a premium for those using more.

Disadvantage: It is not easy to administer but could encourage water efficiency. This option is the most highly recommended model that is currently in operation for commercial and industrial users/customers.

The data show that wastewater services can represent a higher share of the water services billing than water supply (cf. Ruiters, 2011; NT, 2013b). However, tariff structures and prices vary in South Africa's provinces (Figure 5), including connection and fixed charges. This diversity of tariff structures reflects the degree of decentralisation of the tariff-setting process and a smaller number of municipalities use flat fees and decreasing block tariff structures. The trend in South Africa is the increasing use of fixed charges alongside volumetric components often combining a fixed and a variable element or the progressive increase in the weight of fixed charges in the overall bill (cf. Tables 2 and 3). This derives from differences in the cost of provision of the service and it also reflects policy choices as well (cf. DWAF, 1997a, 1997b, 2003, 2007; NT, 2003; OECD, 2010).

The above clearly exemplified a differential pricing strategy that the municipalities apply in South Africa to adjust pricing, tariff setting and costing based on various situations or circumstances. The price variations come in different forms for a particular group of people, i.e., indigent, senior or poor, and/or differentiated wastewater and/or sanitation services. The differential pricing strategy means that certain customers (i.e., vulnerable and marginal groups, low income groups, etc.) pay less for the same service than others. This option is typically well-received by others who do not qualify because the targeted groups are generally thought of as lower income and deserving of the discount. Knowledge of differential pricing allows municipalities to determine if this strategy is a possibility for its jurisdiction in terms of legislation and policy (DWAF, 1997a, 1997b, 2003, 2007; NT, 2003). The nature of differential pricing generally avoids the feeling of unfair treatment by those who do not qualify for the discount. The benefits of allowing differential pricing are for municipalities to expand to customers basic services that might not otherwise get it. The disadvantages of not applying full cost recovery will result in infrastructure operations and maintenance challenges (Matji and Ruiters, 2014, 2015a).

A comparison of residential sanitation tariffs and toilet type for urban and rural municipalities showed that basic sanitation charges are on average 23% more expensive in urban areas than in rural areas. Sanitation tariffs based on land size categories are shown to be more expensive in rural areas, which may be influenced by the generic assumptions made to translate the various tariff units to a generic unit of US$/month (Table 2). However, most other tariff types are very similar in rural and urban areas.

Larger municipalities have separate tariffs for commercial and industrial users (Table 3). The mean commercial tariff was US$38.27/month. Basic sanitation tariffs for commercial users were on average 62% higher than residential units whereas industrial units were on average 116% higher than residential units were, i.e., the principle of cross-subsidisation was applied. This may also relate to the higher effluent volumes and thus is to be expected. For water-use related tariffs, commercial and residential units were similarly charged (US$/kl water used).

OPERATIONAL REVENUE AND EXPENDITURE

All municipalities are expected to raise their own revenues in addition to the equitable share transfer they receive from national government (Figures 4 and 5). It creates a revenue-service link between the municipality and its customers, which empowers customers to hold the municipality directly accountable for the services it provides. A municipality's scope to raise own revenues depends on its fiscal capacity. The extent to which a municipality does raise own revenues in accordance with its fiscal capacity depends on its fiscal effort; this means the amount of attention it pays to ensuring effective revenue management. Figure 6 shows aggregate own revenue for different categories of municipalities for the financial period 2008/9 to 2013/14.

At an aggregate level, the most important sources of municipal revenue are service charges, transfers and property rates (NT, 2011d, 2013b; Stats SA, 2012). Figure 6 shows total operating revenue of municipalities in the new budget format prescribed by the Municipal Budget and Reporting Regulations. The figure shows that total operating revenue grew by 17.5% between 2008/09 and 2013/14. This was primarily driven by very rapid growth in 'transfers recognised – operational', and growth in service charge revenues. The rate of growth in service charge revenues increased by 18.6% over the period 2011/2012 to 2012/2013, reflecting the impact of higher electricity prices and the pressure on municipalities to implement cost-reflective tariffs for all trading services. Revenues from service charges were the largest source of municipal revenue. National transfers were the second largest source of revenue for local government. The very rapid growth in transfers results in this revenue source's share of total revenue increasing from 26.6% in 2008/09 to 32.6% in 2013/14. As a result, other revenue sources' shares of total revenue significantly declined. This trend changes significantly after 2009/10, with the share of revenues from service charges increasing from 41% in 2009/10 to 55% in 2013/14.

Municipalities' actual total operating expenditure increased in real terms by 11.8% annually from 2008/09 to 2013/14 and is estimated to grow by 6.4% over the medium term (Figures 6 and 7). However, operating budgeted expenditure per year per province varied significantly (Figure 8). There has been minor growth, although not significantly, in capital expenditure on wastewater and sanitation services per province (Figure 9). However, the green drop system for wastewater infrastructure management, implemented by the DWS, demonstrated clearly the extent of the operation and maintenance challenges in South Africa (DWA, 2012, 2013a, 2014).

There has been under-investment in the maintenance and refurbishment of infrastructure, which is evident in the number of service delivery failures across the country today (cf. Figure 4). In the water sector, water quality is an important indicator of the performance of a water treatment plant and/or sanitation services. If the quality of wastewater entering a water resource system is poor, it usually indicates that there are operational problems with the treatment plants, either in the forms of plant breakdowns, poor maintenance or delayed maintenance, and plants operating at above their built capacities.

Table 2. The average residential sanitation tariffs (incl. VAT) for rural and urban municipalities

	Rural	**Urban**
Type of sanitation tariff	**Mean (±SD) tariff (US$/month)**	**Mean (±SD) tariff (US$/month)**
Basic sanitation charge	6.98 ± 4.48	8.64 ± 12.66
Water-use category	7.60 ± 6.22	9.98 ± 5.69
Land-size category	9.00 ± 4.38	6.36 ± 5.17
Toilet-type category:		
Buckets	4.32 ± 1.79	4.00 ± 1.00
Conservancy tanks	7.34 ± 3.68	7.25 ± 1.00
Septic tanks	8.54 ± 4.12	8.70 ± 2.13
Sewer connections	5.81 ± 9.41	8.80 ± 1.00
Water-borne sewerage	4.84 ± 2.20	4.90 ± 2.92
Grand mean	**7.43** ± 4.59	**7.45** ± 4.48

Table 3. Sanitation tariffs (incl. VAT) for commercial and industrial-user categories

	Mean tariff (US$)		
Type of sanitation Tariff	**Unit**	**Commercial**	**Industrial**
Basic sanitation charge	US$/Month	12.80 ± 6.98	16.96 ± 8.64
Toilet- type category	US$/Month US$/Month/Toilet US$/Suction	7.70 ± 2.92 1.60 ± 4.10 23.30 ± 2.13	34.06 ± 12.66 - 6.94 ± 5.17
Water-use category	US$/kl	0.56 ± 0.69	0.28 ± 0.24
Land size category	US$/Month	37.30 ± 6.36	4.22 ± 5.17
Building size category	US$/Month	15.60 ± 2.20	-
Land size multiplier	US$/Month/ 100 m^2 land area	0.65 ± 0.17	-
Land value multiplier	US$/Month (multiplied with land area)	0.002 ± 0.001	-
Water connection size category	US$/Month	101.25 ± 21.28	-

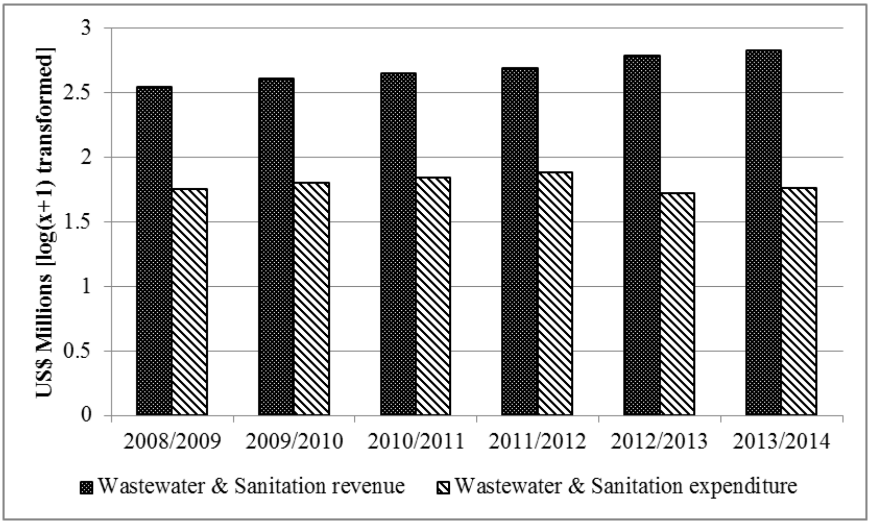


Figure 6. Wastewater and sanitation services operating budgeted expenditures and revenues for the financial period 2008/9 to 2013/2014 in South Africa.

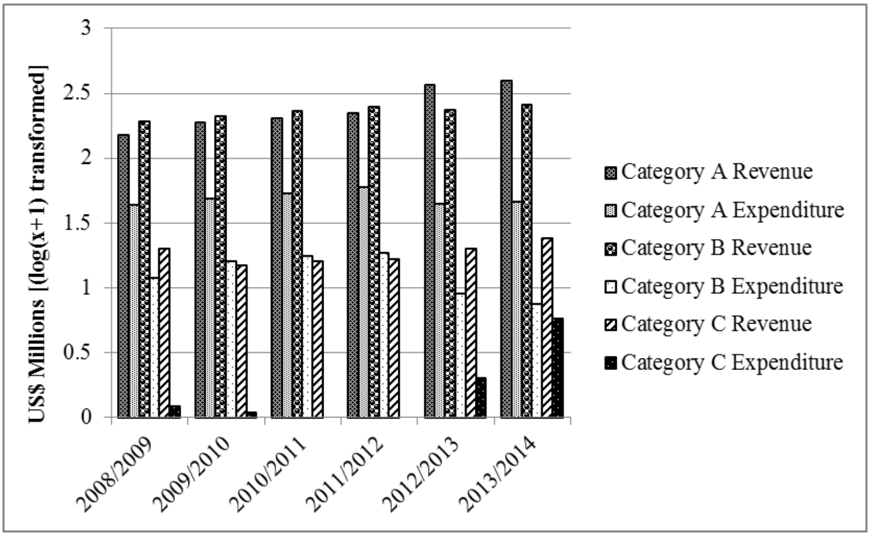


Figure 7. Wastewater and sanitation services operating budgeted expenditures and revenues for municipal categories for the financial period 2008/9 to 2013/14 in South Africa.

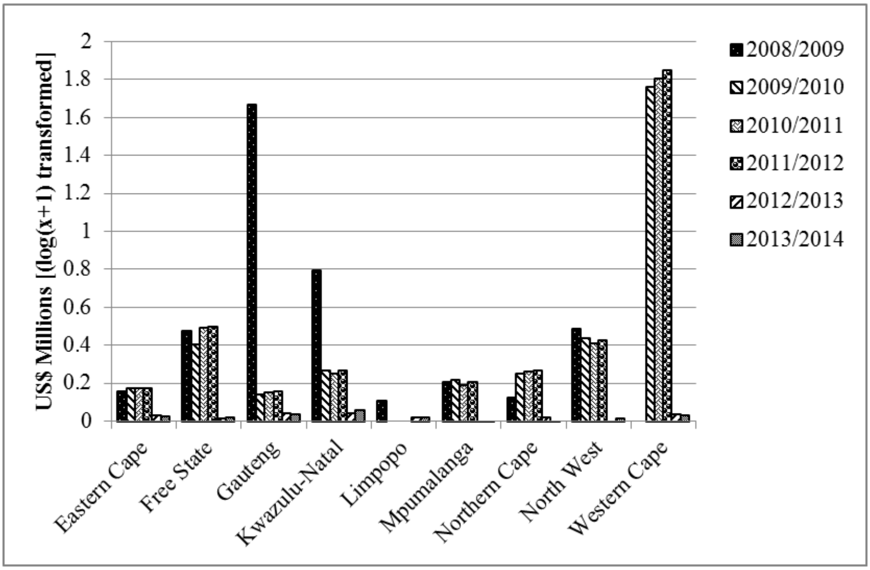


Figure 8. Wastewater and sanitation services operating budgeted expenditures per province for the financial period 2008/9 to 2013/14 in South Africa.

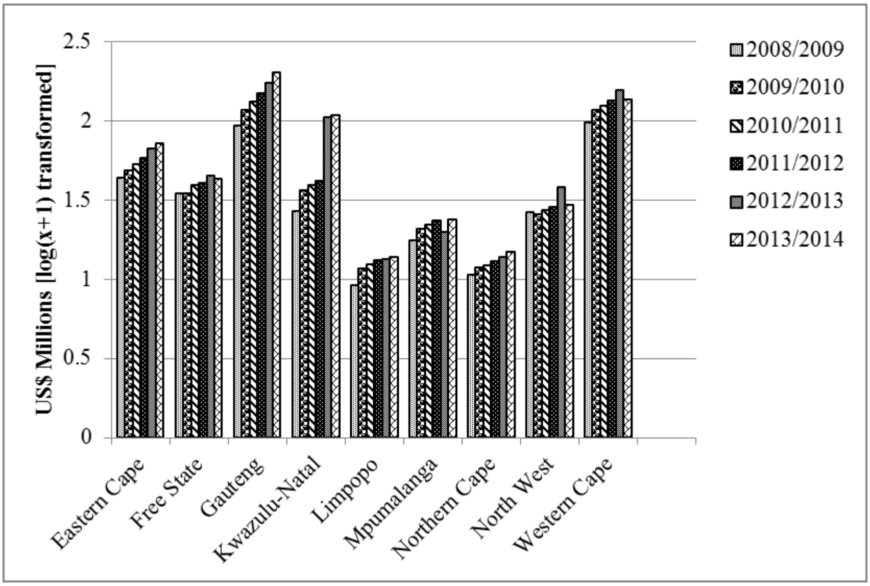


Figure 9. Wastewater and sanitation services operating budgeted revenues per province for the financial period 2008/9 to 2013/14 in South Africa.

This infrastructure problem is further compounded by the fact that many municipalities, especially the smaller and more rural municipalities, do not manage their assets strategically. They are often unaware of what assets they have, where those assets are located, how old those assets may be and what investments are required to extend the useful life of these assets. Without this information, it is almost impossible to determine the investment needs required. The development of an asset register is also a costly exercise, as many municipalities outsource this function because they do not have the requisite in-house capacity. As a result of funding constraints, this exercise is often deferred or completed through a phased approach. However, municipalities generally allocate approximately 5% to 12% of their annual operating budgets for repairs and maintenance (NT, 2011d, 2013b; Stats SA, 2012). Information on the actual repairs and maintenance spent by municipalities per asset class is currently not available.

In order to achieve the above, there is a need for improved economic regulation of bulk water prices. This should be some form of incentive-based regulation which is underpinned by a rate of return analysis. This should be more effective if undertaken by an independent regulator. Whilst prices for modest wastewater and sanitation users in major urban areas have been contained as a result of the introduction of free basic water policies and inclining block tariffs, the costs of wastewater and sanitation for larger users have increased sharply. There should be better regulatory oversight of retail pricing to ensure cost-reflectivity, transparency and consistency in the application of pricing to consumers (Eberhard, 1999, 2004; Ruiters, 2013; Matji, 2013; Matji and Ruiters, 2014, 2015a, 2015b). In the first step, this regulatory oversight should concentrate on the larger urban areas. The first step would be to require metropolitan municipalities to properly ring-fence there water and sanitation businesses and to significantly improve reporting. The second step would be to develop detailed guidelines for efficient pricing. The third step would be to regulate the pricing of water services in cities through independent regulation.

Tariff regulation frameworks and the appropriate regulations are needed to monitoring and measure:

- affordability of wastewater and sanitation services
- equitable access to wastewater and sanitation services
- profit/surplus margin of water institutions, i.e., water services authorities.

These actions are ultimately aimed at ensuring sustainable wastewater and sanitation services in South Africa. However, the role of the DWS as the water sector regulator has been severely criticised, including for wastewater and sanitation (Eberhard, 2004; Palmer Development Group, 2004a, 2004b; DWA, 2013a, 2014).

Compulsory national standards (the regulations) are prescribed that regulates the supply of wastewater and sanitation services to consumers (DWAF, 1997b, 2001, 2002; Eberhard, 1999, 2004; Palmer Development Group, 2002, 2004a, 2004b). The objectives are to protect consumers, municipalities and water services institutions and to ensure the application of sound management principles. A water services institution must, when determining its revenue requirements on which tariffs for water services are based, take into account at least the need to:

(a) recover the cost of wastewater and sanitation services provision
(b) recover overhead, operational and maintenance costs for sanitation services
(c) recover the cost of capital not financed through any grant, subsidy or donation
(d) provide for the replacement, refurbishment and extension of wastewater and sanitation services works
(e) ensure all households have access to basic sanitation.

International Comparisons of Wastewater and Sanitation Services

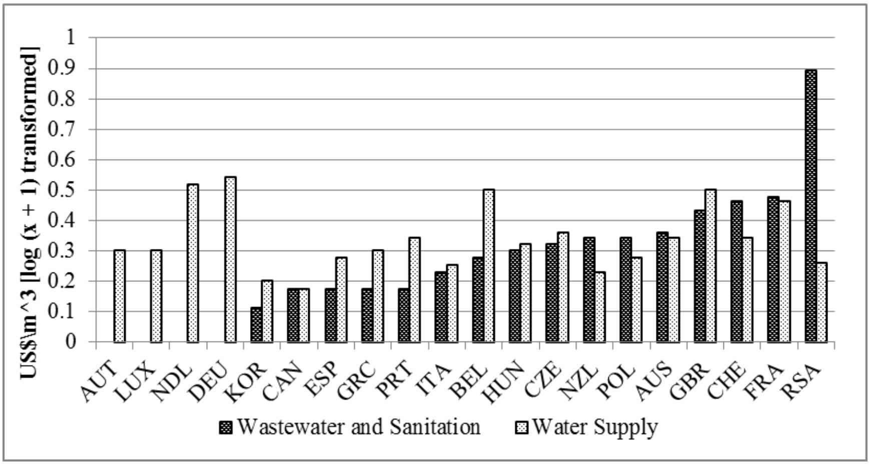
The international benchmarking and comparison of tariff levels need to take place but should be carried out very carefully. However, unimproved and improved (on site) latrines are the most prevalent sanitation service option in most developing countries, in particular Africa countries, and health benefits from their delivery are still questionable (WB, 2010). Waterborne sewerage (wastewater services) systems are rare in Africa and only half of the major cities operate a sewerage network at all. Only in Namibia, South Africa, and Senegal do the utilities cover the largest cities and/or towns provide universal access sewerage coverage (WB, 2010; DWA, 2011; NT, 2013a, 2013b). In most other countries on the African continent sewer networks barely reach

10% of the population within the service area of an urban utility (water services authority) (WB, 2010).

Most of the population in South Africa with piped water have water borne sanitation services connected to a sewer network in cities, secondary cities, larger and smaller towns. In the rest of Africa little than half of households with piped water also have water borne sanitation which are mainly connected to septic tanks rather than sewers (Mehta et al., 2005; WB, 2010). Furthermore, sanitation varies significantly across income groups in Africa with septic tanks and water borne sewerage in the middle and upper income groups, latrines most common amongst communities (50%) and open defecation confined to the lowest poorer population mostly in rural areas (Mehta et al., 2005; WB, 2010, DWA, 2011, Stats SA, 2012; NT, 2013a, 2013b).

In contrast, most OECD and developed non-OECD countries (including South Africa) levy separate charges for sewerage vs. wastewater treatment, although in most cases the basis for charging remains water consumption; only the size of the volumetric rate differs (DWAF, 2007; OECD, 2010). The latter is increasingly based on the pollution load of effluents, thus better reflecting actual treatment costs. Figure 10 presents separate average tariffs for water and wastewater services in US$ in a comparison of South Africa with OECD and non-OECD countries. The data shows that in half of the countries wastewater services can represent a higher share of the water bill than water supply. However, tariff structures vary within and across OECD countries and prices can vary from a factor 10 or more, as for variation in South Africa (cf. Figures 4 and 5), including connection and fixed tariffs. This diversity of tariff structures reflect the degree of decentralisation of the tariff-setting process and a smaller number of countries use flat tariffs and decreasing block tariff structures (OECD, 2010).

The trend in OECD countries, as is the case for South Africa, is the increasing use of fixed tariffs alongside volumetric components often combining a fixed and a variable element or the progressive increase in the weight of fixed tariffs in the overall bill. This derives from differences in the cost of provision of the service and reflects policy choices as well. In general, tariff levels in non-OECD countries tend to be much lower than in the OECD countries. Even in the water-scarce Middle-East/North Africa and Sub-Saharan Africa regions, no tariffs reach US$1, generally, the tariffs are below US$0.15/$m^3$, providing little incentive for water use efficiency and contributing little to cost recovery (OECD, 2010).



Source: OECD (2010.).

Figure 10. Separate average tariffs for water and for wastewater services in a comparison of South Africa with OECD and non-OECD countries.

CONCLUSION

While there have been substantial improvements in the rollout of water services infrastructure and the rendering of free basic water and sanitation, the sector does face some challenges going forward as implementation capacity remains a constraint. Furthermore, the sustainability of existing infrastructure cannot be neglected and is requiring more and more funding as infrastructure ages. In addition, the cost of extending the infrastructure network to outlying communities is not cost-effective or sustainable, which points to the need to explore alternative service-delivery options.

It is unlikely that wastewater and sanitation services prices will ever totally reflect the 'full cost' approach favoured by environmental economics, but some 'directionally correct' pricing structures are being designed. Strategic pricing of water and wastewater can play a greater role in meeting the investment needs of the country's wastewater and sanitation services (or slow growth rate of demand) and raise revenues to support critical capital investments. In the water sector, these imperatives are now greater than ever before. The following is recommended for future sanitation tariff analyses: Future assessments should track the annual changes (e.g., % increases) and trends between tariff types. It would be beneficial to the sector and to the

DWS, as the sector leader and regulator, to promote standardisation of tariff types so that comparisons can be made more efficiently and fairly between municipalities and user group.

This research only includes the sanitation tariffs charged by municipalities. Future assessments may consider including other charges in the wastewater discharge chain (e.g., bulk sanitation treatment costs) and effluent discharge levies specified in the effluent licenses. Financial analysis is possible only if both the tariffs and costs are available along the water and effluent discharge processes. Future research may want to collate typical sewage collection and treatment costs to enable a comparison between income from sanitation billing and the actual costs to provide the service. This will highlight whether wastewater and sanitation tariffs are fair and whether the services are run at a profit or a loss. Active management of the cost cycle of wastewater and sanitation services is a critical performance area for the DWS.

Tariff levels charged to households for wastewater and sanitation services vary greatly among municipalities in the provinces of South Africa, reflecting contrasted efforts to recover the costs of the services through pricing. Many municipalities' tariffs are inappropriately structured. The principal problem is that few municipalities understand how their various activities and services are being funded, and therefore what the tariffs needs to be to ensure financial sustainability. Other problems include:

- a failure to ensure that on average, service tariffs reflect the costs reasonably associated with rendering the service; i.e., that revenues and expenditures for the trading services break-even.
- limited use of inclining block tariffs, particularly for sanitation.
- overly generous rates rebates, exemptions and discounts, and a general movement towards the provision of free basic services (which is engendering a culture of dependency rather than adhering to the principle that everyone should make some payment for the municipal services they receive).

Municipal tariffs cannot be evaluated in isolation of the sequence of charges (i.e., tariff chain). The end user in effect pays for all up-stream and down-stream charges and this culminates in the municipal tariffs. A tariff monitoring, evaluation and regulation system must take cognisance of the average cost and value chain associated with water supply from source to user and then back to source, as well as the cost implications of wastewater management, treatment and release. Financial implications of the municipal tariffs can also not be evaluated comprehensively without knowing the volume of water used per tariff block (i.e., wastewater produced) and the success of billings recovered. The suitability and acceptability of a tariff must be measured by its financial impact on both the end user and the revenue it generates for the supplier. This also goes hand-in-hand with service level agreements to set the commitments upfront from both the supplier and the customer.

Currently, the DWS acts as a policy maker, regulator and implementer. This model should be reviewed to address conflicting interests. In order to ensure consistency, standardisation and equitable pricing and tariffs across the full value chain (i.e., source to end user to source), the process for establishing an independent water sector regulator should be fast-tracked. It would then be the responsibility of the independent water sector regulator to address significant variations in pricing and tariffs in the whole water sector value chain.

There is a general need for the strengthening of economic regulation of water throughout the water sector value chain. However, economic gains through investments that are more efficient and pricing are likely to be realised most cost-effectively by concentrating initially on the economic regulation of the major water services providers (i.e., water boards) and water services authorities serving the large or significant urban and industrial areas in South Africa. Where water services are provided through contracts, particular attention should be given to the financial component of these contracts to ensure that appropriate incentives for efficient pricing are built in. The commitments to strengthened economic regulation made in the water services white paper must be carried out. These commitments should be reflected in the regulation strategy which should address the regulatory issues in much more detail.

The unglamorous nature of wastewater and sanitation services put them at a disadvantage in the competition for the national revenue resources in South Africa, although huge strides have been made. The decentralisation and poor accounting for the sanitation sector expenditures can impede understanding of the public funds allocated; however, ring-fenced municipal grants have been introduced in the South African budgeting system. In South Africa, on-site sanitation services must be improved and cost-effective to ensure associated health benefits.

REFERENCES

Brineco-garmendia, Cecilia., Karlis, Smits. and Vivien, Foster. (2008). "Financing Infrastructure in Sub-Saharan Africa: Patterns, Issues, and Options." In AICD Background Paper 15, Africa Infrastructure Sector Diagnostic. Washington, DC: World Bank. (BRINECO-GARMENDIA et al. 2008).

Coldwell, David. and Francois, J. Herbst. (2004). Business Research. Cape Town: Juta. (Coldwell and Herbst 2004).

Cranston, Stephen. (2005). "Old Mutual: Closing the gap." Financial Mail, 28 January 2005. URL: http://www.businessday.co.za (Accessed 30 April 2014). (Cranston 2005).

Department of Cooperative Governance. (2010). Strategy for Accelerating and Improving Municipal Infrastructure Provision. Pretoria: Department of Cooperative Governance. (Department of Cooperative Governance 2010).

Department of Public Works. (2007). Government Immovable Management Act (Act No. 19 of 2007). Pretoria: Government Printer. (Department of Public Works 2007).

Department of Water Affairs and Forestry. (1997a). National Water Policy for South Africa. Pretoria: Government Printer. (Department of Water Affairs and Forestry 1997a).

Department of Water Affairs and Forestry. (1997b). Water Services Act. Pretoria: Government Printer. (Department of Water Affairs and Forestry 1997b).

Department of Water Affairs and Forestry. (1998.) National Water Act (Act No. 36 of 1998). Government Gazette 398 (No. 19182) of 26 August 1998. Cape Town: Government Printer. (Department of Water Affairs and Forestry 1998).

Department of Water Affairs and Forestry. (2001). White Paper on Basic Household Sanitation. Pretoria: Department of Water Affairs and Forestry. (Department of Water Affairs and Forestry 2001).

Department of Water Affairs and Forestry. (2002). Retail Water and Sanitation Guidelines. Pretoria: Department of Water Affairs and Forestry. (Department of Water Affairs and Forestry 2002).

Department of Water Affairs and Forestry. (2003). Water Services Strategic Framework. Pretoria: Department of Water Affairs and Forestry. (Department of Water Affairs and Forestry 2003).

Department of Water Affairs and Forestry. (2007). Raw Water Pricing Strategy of South Africa. Pretoria: Government Printer. (Department of Water Affairs and Forestry 2007).

Department of Water Affairs and Forestry. (2008). Emerging Institutional Models for Water in South Africa. Pretoria: Department of Water Affairs and Forestry. (Department of Water Affairs and Forestry 2008).

Department of Water Affairs. (2010). Water Services Tariffs 2009/2010. Pretoria: Department of Water Affairs. (Department of Water Affairs 2010).

Department of Water Affairs. (2011). Water Services Tariffs 2010/11. Pretoria: Department of Water Affairs. (Department of Water Affairs 2011).

Department of Water Affairs. (2012a). Water Services Tariffs 2011/12. Pretoria: Department of Water Affairs. (Department of Water Affairs 2012a).

Department of Water Affairs. (2012b). Institutional and Re-alignment: A Strategic Framework for Institutional Re-alignment and Reform. Pretoria: Department of Water Affairs. (Department of Water Affairs 2012b).

Department of Water Affairs. (2013a). National Water Resources Strategy: June 2013 2nd Edition. URL: http://www.dwa.gov.za/nwrs2013 (Accessed 3 October 2014). (Department of Water Affairs 2013a).

Department of Water Affairs. (2013b). Strategic Plan 2010/11-2012/13. URL: http://www.dwa.gov.za/documents (Accessed 3 October 2014). (Department of Water Affairs 2013b).

Department of Water Affairs. (2013c). Annual Report 2012-2013. URL: http://www.dwa.gov.za/Documents (Accessed 3 October 2014). (Department of Water Affairs 2013c).

Department of Water Affairs. (2014). Green Drop Report 2014. URL: http://www.dwa.gov.za/Documents (Accessed 3 October 2014). (Department of Water Affairs 2014).

Eberhard, Rolfe. (1999). Supply Pricing of Urban Water in South Africa. WRC Report No. 678/1/99, Pretoria: Water Research Commission. (Eberhard 1999).

Eberhard, Rolfe. (2004). Administered Prices: Water. A Report for National Treasury. http://www.treasury.gov.za/organisation/documents/default.htm (Accessed 1 November 2013). (Eberhard 2004).

Goodman, Alvin. S. and Makarand, Hastak. (2006). Infrastructure Planning Handbook: Planning, Engineering, and Economics. New York: McGraw-Hill. (Goodman and Hastak 2006).

Grove, Bennie. (2004). Development of Models for Economic Evaluation of the Integrated Management of the Quantity and Quality of Irrigation Water within River Catchments. WRC Report No. 1043/1/04, Pretoria: Water Research Commission. (Grove 2004).

Hazelton, Derek. (1998). Cost Recovery for Water Schemes to Developing Urban Communities. WRC Report No. 521/1/98, Pretoria: Water Research Commission. (Hazelton 1998).

Hollingworth, Brian. (1994). An Investigation into the Water Infrastructure Development Financial Allocation Pathways in Municipalities. WRC Report No. K5-1844, Pretoria: Water Research Commission.

(Hollingworth 1994). Hope, Rob. A. and Guy, D. Garrod. (2005). Evaluating changes in water institutions: methodological issues at the micro- and meso-levels. *Water Policy*, *7*, 21-34. (Hope and Garrod 2005).

Hosking, Stephen. G. (2003). Investigating the mechanism and processes used in setting water services tariffs. WRC Report No. K5/1871/3, Pretoria: Water Research Commission. (Hosking 2003).

Karar, Eiman., Gugu, Mazibuko., Thomas, Gyedu-Ababio. and Derek, Weston. (2011). "Catchment management agencies: A case study on institutional reform in South Africa." In Transforming water management in South Africa: Designing and implementing a new policy framework, edited by Barbara Schreiner and Rashid Hassan, 145–164. London: Springer. (Karar et al. 2011).

Kerdachi, Des. (2002). Review of industrial effluent tariff structures in SA. WRC Report No. 854/1/02, Pretoria: Water Research Commission. (Kerdachi 2002).

Lamb, Robert. A. (1984). Creative financing of the infrastructure. In Infrastructure Maintenance and Repair of Public Works, Annals, NY *Academy of Science*, Vol. *431*. (Lamb 1984).

Lang, Howard. J. and Daniel, N. Merino. (1993). The Selection Process for Capital Projects. New York: Wiley. (Lang and Merino 1993).

Mander, Myles. (2002). Incorporating Economic Considerations into Quantification, Allocation and Management of the Environmental Water Reserve. WRC Report No. 978/1/02, Pretoria: Water Research Commission. (Mander 2002).

Matji, Maselaganye, P. (2013). "Conceptual Framework for Public-Private Partnerships Model for Water Services in South Africa: Case Studies from Municipalities in the Limpopo and Gauteng Provinces." MBL diss., University of South Africa. (Matji 2013).

Matji, Maselaganye. P. and Cornelius, Ruiters. (2014). "Appropriate Solutions for Operation and Maintenance Challenges of Water Services Infrastructure in South Africa." Paper presented at the World Congress on Engineering Asset Management, Pretoria, South Africa, November 11-14. (Matji and Ruiters 2014).

Matji, Maselaganye. P. and Cornelius, Ruiters. (2015a). "Conceptual Framework for Public-Private Partnerships Model for Water Services Infrastructure Assets: Case Studies from Municipalities in the Limpopo and Gauteng Provinces." Paper presented at the AMPEAK Asset Management Conference, Sydney, Australia, May 19-23. (Matji and Ruiters 2015).

Matji, Maselaganye. P. and Cornelius, Ruiters. (2015b). Funding and Financing Mechanisms for Infrastructure Delivery: Multi-sector Analysis of Benchmarking of South Africa against Developed Countries. Paper presented at the AMPEAK Asset Management Conference, Sydney, Australia, May 19-23. (Matji and Ruiters 2015).

Marah, Lansana. (2003). Identifying Examples of Successful Cost Recovery Approaches in Low Income, Urban and Peri-urban Areas. WRC Report No. 1131/1/03, Pretoria: Water Research Commission. (Marah 2003).

Marah, Lansana. (2004). Effective cost recovery in a changing institutional and policy environment: Municipal demarcation, the 'free basic water' policy and financially sustainable service delivery. WRC Report No. 1384/1/04, Pretoria: Water Research Commission. (Marah 2004).

Matthews, Wendy. (2009). The Development of an Activity-based Costing Model to Quantify the Real Costs of Delivering Water Services in Rural Areas. WRC Report No. 1614/1/09, Pretoria: Water Research Commission. (Matthews 2009).

Mehta, Meera., Thomas, Fugelsnes. and Kamed, Virjee. (2005). Financing the Millennium Development Goals for Water Supply and Sanitation: What Will it Take? *Water Resources Development*, *21*(2), 239 -252. (Mehta et al. 2005).

Muller, Michael. A. (2014). Allocating Powers and Functions in a Federal Design: The Experience of South Africa. In Federal Rivers: Managing Water in Multi-layered Political Systems, edited by Dustin E. Garrick, George. R. M. Anderson, Daniel Connell, and Jamie Pittock, 179-194. Cheltenham UK: Edward Elgar. (Muller 2014).

National Treasury of South Africa. (2003). Municipal Finance Management Act (Act No. 56 of 2003). Pretoria, South Africa: Government Printer. (National Treasury of South Africa 2003).

National Treasury of South Africa. (2011a). Estimates of National Expenditure 2011: Vote 3. Cooperative Governance and Traditional Affairs. Pretoria: National Treasury. (National Treasury of South Africa 2011a).

National Treasury of South Africa. (2011b). Estimates of National Expenditure 2011: Vote 38. Water Affairs. Pretoria: National Treasury. (National Treasury of South Africa 2011b).

National Treasury of South Africa. (2011c). Division of Revenue Bill. Pretoria: Government Printer. (National Treasury of South Africa 2011c)

National Treasury of South Africa. (2011d). Local Government Budgets and Expenditure Review: 2006/07-2012/13. URL: http://www.treasury.gov.za/organisation/documents/default.htm (Accessed 1 November 2013). (National Treasury of South Africa 2011d).

National Treasury of South Africa. (2013a). Estimates of National Expenditure 2013. URL: http://www.treasury.gov.za/organisation/ nationalbudget/speech/default.htm (Accessed 1 November 2013). (National Treasury of South Africa 2013a).

National Treasury of South Africa. (2013b). The State of Local Government Finances and Financial Management as at 30 June 2013: Analysis Document. URL: http://www.treasury.gov.za/organisation/ppp/default.htm (Accessed 15 January 2014). (National Treasury of South Africa 2013b)

National Treasury of South Africa. (2013c). Minister of Finance Budget Speech 2013. URL: http://www.treasury.gov.za/organisation/nationalbudget/ene/default.htm (Accessed 1 November 2013). (National Treasury of South Africa 2013c).

Organisation of Economic Cooperation and Development. (2010). Pricing Water Resources and Water and Sanitation Services. Paris: OECD Publishing. (Organisation of Economic Cooperation and Development 2010)

Palmer Development Group. (1998). Water Supply Services Model Manual. WRC Report No. KV109/98, Pretoria: Water Research Commission. (Palmer Development Group 1998).

Palmer Development Group. (2000). Guidelines for Setting Water Tariffs with a Focus on Industrial, Commercial and Other Non-Residential Consumers. WRC Report No. 992/1/00, Pretoria: Water Research Commission. (Palmer Development Group 2000).

Palmer Development Group. (2002). Corporatisation of Municipal Services Providers. WRC Report No. TT 199/02, Pretoria: Water Research Commission. (Palmer Development Group 2002).

Palmer Development Group. (2004a). Economic Regulation of Water Services in South Africa. WRC Report No. 138/1/04, Pretoria: Water Research Commission. (Palmer Development Group 2004a).

Palmer Development Group. (2004b). Guidelines for Economic Regulation of Water Services in South Africa. WRC Report No. TT 229/04, Pretoria: Water Research Commission. (Palmer Development Group 2004b).

Republic of South Africa. (1996). Constitution of South Africa (Act no. 108). Statutes of the South Africa Constitutional Law. Cape Town: Office of the President. (Republic of South Africa 1996).

Ruiters, Cornelius. (2011). "Funding Models for the Financing of Water Infrastructure in South Africa: A Framework and Comparative Analysis of Alternatives." MBA diss., University of South Africa. (Ruiters 2011).

Ruiters, Cornelius. (2013). Funding Models for Financing Water Infrastructure in South Africa: A Framework and Critical Analysis of Alternatives. *Water SA*, *39*(2), 313-326. (Ruiters 2011).

Ruiters, Cornelius. and Maselaganye, P. Matji. (2015). Water Institutions and Governance Models for the Funding, Financing and Management of Water Infrastructure in South Africa. *Water SA*, *41*(5), 660-676. (Ruiters and Matji 2015).

Saleth, Rathinasamy. M. and Ariel, Dinar. (2005). Water Institutional Reforms: Theory and Practice. *Water Policy*, *7*, 1-20. (Saleth 2005).

Statistics South Africa. (2012). Financial Census of Municipalities for the Year Ended 30 June 2011. Pretoria: Statistics South Africa. (Statistics South Africa 2012).

Still, David. (2003). Cost and Tariff Model for Rural Water Supply Schemes. WRC Report No. 886/1/03, Pretoria: Water Research Commission. (Still 2003).

Tustin, Deon. R., Andre, A. Ligthelm., Johan, H. Martins. and Helgard, De J. Van. Wyk. (2005). Marketing Research in Practice. Pretoria: University of South Africa Press. (Tustin et al. 2005).

Van, Koppen. Barbara. and Barbara, Schreiner. (2014). Moving Beyond Integrated Water Resource Management: Developmental Water Management in South Africa. *Journal of Water Resources Development*, *30*(3), 543-558. (Van Koppen and Schreiner 2014).

Van, Vuuren. and David, S. (2003). Researching, Developing and Testing of Payment Strategies for the Lower Income Groups at Four Selected Communities in Order to Manage Charges for Water Use. WRC Report No. 202/1/03, Pretoria: Water Research Commission. (Van Vuuren 2003).

Van, Vuuren. and David, S. (2004). Payment Strategies and Price Elasticity of Demand for Water for Different Income Groups in Three Selected Urban Areas. WRC Report No. 1296/1/04, Pretoria: Water Research Commission. (Van Vuuren 2004).

Vawda, Mohamed., Nicola, King. and Michael, A. Muller. (2011). Pricing of Water for Cost Recovery, Economic Efficiency and Social Equity. In Transforming Water Management in South Africa: Designing and Implementing a New Policy Framework, edited by Babara Schreiner and Rashid Hassan, 181-201. Dordrecht, Germany: Springer. (Vawda et al. 2011).

World, Bank. (1994). World Bank Development Report: Infrastructure for Development. New York: Oxford University Press. (World Bank 1994).

World, Bank. (2010). Africa's Infrastructure: A Time for Transformation. Washington, DC: World Bank. (World Bank 2010).

About the Editors

Gisele Ferreira Tiryaki, PhD
Professor, Economics Department,
Universidade Federal da Bahia, Brazil

Gisele Ferreira Tiryaki holds a Master of Arts degree in economics from Northeastern University (Massachussets, USA) and a PhD degree in economics from George Mason University (Virginia, USA), granted in 1996 and 2001, respectively. Dr. Ferreira Tiryaki worked as a consultant at the World Bank headquarters in Washington D.C. for four years, researching about private participation in infrastructure projects in developing countries. After working as a professor for Anadolu University (Turkey) between 2002 and 2006, Dr. Ferreira Tiryaki moved to Brazil, where she was granted a research and teaching position at the Master program in Energy from Universidade Salvador (Bahia). Since 2009, she has worked as a professor at the Economics Department of Universidade Federal da Bahia, where she has devoted time researching about infrastructure regulation, industrial organization and applied macroeconomics.

André Luis Mota dos Santos, PhD
Professor, Economics Department,
Universidade Federal da Bahia, Brazil

André Luis Mota dos Santos was granted a Master of Arts and a PhD degrees in Economics from Universidade Estadual de Campinas (São Paulo, Brazil) in 2005 and 2010, respectively. Since 2009, Dr. Mota dos Santos has been working as a professor at the Economics Department of Universidade Federal da Bahia. His research focuses on issues related to economic development and institutions, with special attention to the impact of natural resources' abundance on an economy's performance.

INDEX

G

H

I

L

M

N

O

P

Q

R

S

T

U

V

W