European Commission

The impact of the clean energy transition on the determinants of economic growth

E3Modelling & Cambridge Econometrics
June 2022

Authors:

Kostas Fragkiadakis, Dimitris Fragkiadakis, Panagiotis Fragkos, Leonidas Paroussos (E3Modelling)

João Gorenstein Dedecca, Liliana Guevara Opinska, Foivos Petsinaris (Trinomics)

Boglarka Molnar, Richard Lewney, Bence Kiss-Dobronyi (Cambridge Econometrics)

EUROPEAN COMMISSION

Directorate-General for Energy
Directorate A — Energy policy: Strategy and Coordination
Unit A.4 — Chief Economist Unit

Contact: Energy Modelling Team

E-mail: *ENER-A4-MODELLING-TEAM@ec.europa.eu*

European Commission
B-1049 Brussels

The impact of the clean energy transition on the determinants of economic growth

Manuscript completed in March 2022

PDF	ISBN 978-92-76-52356-7	doi: 10.2833/56663	MJ-07-22-269-EN-N

Luxembourg: Publications Office of the European Union, 2022

Table of Contents

Executive Summary ... 7

Introduction ... 9

1. Determinants of economic growth ... 10

1.1. Literature review on the determinants of growth, the types of growth and sustainable development indicators ... 10

2. Energy as a factor of economic growth ... 14

2.1. Causality between energy and growth ... 14
2.1.1. Causality tests ... 15
2.1.2. Econometric results on energy-growth nexus – timeseries analysis 16
2.1.3. Econometric results on energy-growth nexus – panel data analysis 25

2.2. Energy substitution possibilities ... 32
2.2.1. Literature review on econometric methods used to estimate the energy substitution possibilities ... 33
2.2.2. Econometric results on the energy substitution possibilities ... 34
2.2.3. Summary and conclusion ... 66

3. Clean energy technologies and economic growth ... 67

3.1. R&D multiplier and leverage effects: Literature review for selected clean energy technologies ... 67
3.1.1. Introduction and summary of findings ... 67
3.1.2. Literature findings on the theoretical background of learning rates in clean energy technologies ... 69
3.1.3. Development of learning rates over time for selected clean energy technologies ... 73

3.2. Updating of industry and region spillover matrices based on patent data ... 84
3.2.1. Construction of spillover matrices – data and processing ... 84
3.2.2. Analysis of observed spillover effects ... 86

4. References ... 94

Appendix: Regions and economic sectors of the WIOD Database ... 98

Executive Summary

Findings regarding *the contribution of different factors to growth* differ across studies and depend on a multitude of aspects that characterise the socio-economic systems. These aspects mainly relate to the stage of development where a country is and the coverage and quality of its institutions and governance. That is why there are no universal best practice policies that would yield the same positive results irrespective of when and where are applied. However certain determinants like capital accumulation/investment and governance are considered to have a positive impact on growth in most studies considered. Much greater variability on findings exists for factors like public spending and education.

A similar conclusion rule applies in the case of *the role played by energy in growth*. Energy as an economic activity accounts for about 8% of global GDP, but energy is an essential resource for production and consumption. Depending on an economy's pattern of growth and stage of development, the economic system may become more or less energy intensive over time. The question as to whether and how energy drives economic growth has a different answer across countries and economic sectors. Regarding the causality between energy and GDP growth our findings are consistent with the neutrality hypothesis. Having said that, in the cases (countries and sectors) in which a cointegration relationship exists between energy and GVA or the series can be considered as stationary there is a greter tendency for the Growth hypothesis (energy drives GDP growth) to be accepted.

Regarding *the substitution possibilities of energy with capital and labour* a wide-range of econometric methods have been used. Both timeseries and panel data analysis strongly support the weak substitutability between energy and gross value added. Our statistical analysis identified the existence of structural breaks that indicate that the change in the substitutability between energy and value added is most likely driven by other factors than just changes in their prices and that a change in the level of the energy intensity to gross value added occurs as a shift at a point in time.

The estimates of all models used that allow to differentiate between *the short-run and long-run elasticity of substitution of energy and gross value added* do not provide statistical evidence that the long run elasticity can be considered as greater than the short run elasticity, contrary perhaps to what might have been expected. The elasticity of substitution between energy and gross value added remains stable across time. An indicative value of elasticities of substitution between energy and value added is around 0.5 for the different economic activities.

Our review of *learning-by-doing* rates suggests that different rates could apply for the period to 2030 compared with the longer run. Photovoltaics, onshore and offshore wind turbines, biomass and biofuels, and electric vehicles and batteries are the technologies that are expected to experience slower (although in some cases, still relatively high) learning rates for the period 2030-2050 than over 2015-2030.

The literature is much more spares for *learning-by-research* rates (that is, on the rate of decrease in unit costs for every doubling of research and development investment stock) for clean energy technologies, but some broad observations can be made. For the most part, the reported learning-by-research rates are as high as the learning-by-doing rates for the same technology and in some cases higher (for example in the case of three studies on wind turbines)

The following table presents the suggested learning-by-doing rates to use for the modelling of cost development of clean energy technologies over time.

Recommended learning-by-doing rates to use for the selected clean energy technologies

Clean energy technology	Recommended learning rate to use for the years 2015-2030	Recommended learning rate to use for the years 2031-2050
PV	20%	17%
Wind turbines - Onshore	7%	5%
Wind turbines – Offshore	11%	9%
Biomass and biofuels	7.8%	5%
Electric vehicles and batteries	18%	15%
Hydrogen production by electrolysis	7%	7%
Hydroelectricity	1.4%	1%
Synthetic fuels (power to X)	13%	13%
DAC (Direct Air Capture)	6%	6%
Biomass CCS and traditional CCS	7%	5%
Conventional technologies	3%	3%
Selected other sectors	18%	18%

Note: Learning rates are defined per doubling of cumulative output (learning-by-doing). Conventional technologies cover natural gas and nuclear. Selected other sectors cover chemicals and electronic goods.

Source: Reviewed by Cambridge Econometrics, 2019.

Our analysis of spillover effects evidenced by patent citations found that, unsurprisingly, the strongest linkages are within countries (a large share of citations were of sources from the same country). The US is an important citer of patents from other countries and an importance source of citations by other countries. Together Japan and the US are well ahead of the countries ranked 3rd and following and accounted for more than half of the (weighted and aggregated) citation shares. There is evidence of spillover clustering among certain regions. The EU15 countries cite each other heavily, but there are weaker links with and between the rest of the EU countries.. Similarly there is clustering among Asian countries again showing the impact of geographical/cultural proximity on knowledge dissemination. Citations across countries are mostly symmetric – the spillover effect generally seems to be two-way.

With regard to cross-sector spillover effects, again the most most important linkages are within-sector. Clustering of sectors is less obvious than it is for countries, but there are some examples, notably among industries covering activities of electronics manufacturing (mechanical engineering, electronics, electronic engineering and instruments).

While patent citations provide a valuable source of information about the diffusion of knowledge, there are obvious limitations, notably that they are a much better indicator of spillovers in manufacturing industry than in services.

1. Introduction

Empirical findings from the respective literature are inconclusive regarding the causal relationship between energy and growth (over 200 studies present mixed results). This study designs and performs new tests using the latest available statistics in order to empirically validate the role of energy to economic development. The new estimations and empirical results not only allow to better understand the link between the two but also are useful on their own right as they serve to update the respective key elasticities and reconsider substitution possibilities among production factors in both the large scale applied models GEM-E3 and E3ME.

The first section of this report provides a brief literature review on the key determinants of growth as these are considered in the main economic schools. The presentation of determinants of growth is important in order to illustrate the channels through which energy could contribute to growth. The second section is dedicated to the empirical findings on energy as a factor of economic growth. The analysis is based both on available estimates and on new estimations using the most recent econometric techniques. The final section is dedicated to the relation between R&D on energy and growth. This section discusses and presents the most recent findings on the impact of R&D on capital costs through accumulation of knowledge and the impact on costs through learning by repetition and economies of scale. The quantitative estimations and data collected are appended to the end of this report.

2. Determinants of economic growth

2.1. Literature review on the determinants of growth, the types of growth and sustainable development indicators

Economic growth is essential for countries to make progress towards sustainable development goals. The drivers of economic growth however differ by country and point in time. Future economic growth will largely be based on “new technologies that offer not only ‘catch-up’ potential but also ‘leapfrogging’ possibilities” while “economic growth will need to be environmentally sustainable (better management of natural resources with movement towards low carbon technologies)” OECD (2015).

This section provides a review on the determinants of economic growth according to different economic approaches and briefly discusses the empirical findings of the respective literature. The review includes both supply side (natural resources, capital goods, human resources and technology, institutions and governance) and demand side (public expenditure, efficiency, private spending) determinants of economic growth.

A recent review of the international literature (covering more than 30 studies, 100 countries and time period from 1960s to 2010) on the macroeconomic determinants of economic growth by Chirw et al (2016) identifies the following determinants of growth:

1. Investment and capital accumulation
2. Real exchange rate
3. Human capital and education
4. Budget surplus
5. Inflation
6. Public expenditure
7. Domestic savings
8. Population growth
9. Terms of trade
10. Institutional quality and rule of law
11. Trade openness
12. Technical progress and labour productivity
13. Money supply, interest rates and credit to the private sector
14. Natural resources

The evidence regarding the contribution of the above-mentioned factors to growth is not consistent across studies and partly depends on the stage of development that each country is. For example capital accumulation and investment is considered to have a positive impact on growth in most studies considered. Whereas public spending and education depends on whether the country is already developed or developing. Non-economic variables such as the rule of law and governance are found to be important growth determinants and their variability across studies is small.

Table 1: A review of the determinants of growth (Developing Countries)

Author(s)	Countries and Sample Period	Determinants with a Significant Positive (+) or Negative (-) Impact on Growth
	Developing Countries	
Dollar 1992	1976-1985 95 developing economies	– Investment (+) – Real exchange rate variability (-)
Fischer 1992	Sub-Saharan Africa and Latin American and Caribbean countries; 1970-1985	– Human capital, investment, and budget surplus (+) – Inflation (-)
Knight, Loayza, and Villanueva 1993	98 countries; 76 developing countries; 1960-1985	– Physical and human capital(+) – Public expenditure (+) – Tariff rates (-)
Most and Vann de Berg 1996	11 Sub-Saharan Africa countries	– Foreign aid (+) – Domestic savings (-) (+) depending on the country – Foreign Direct Investment (-)(+) – Population growth (-)
Hamilton and Monteagudo 1998	98 non-oil countries; 75 intermediate countries; 22 OECD 1960-1985	– Investment (+) – Population (-) – Education (-)
Barro 1999	100 countries (Asia, Latin America, OECD, Sub-Saharan Africa; 1960-1995	– Investment share, growth rate of terms of trade, years of schooling, rule of law index, democracy index and trade openness (+) – Government consumption, total fertility rate, and inflation (-)
Burnside and Dollar 2000	56 countries (16 middle-income, and 40 low-income countries); 1970-1993	– Foreign aid (+) – Budget surplus, institutional quality, trade openness (+)
Chen and Feng 2000	China (29 provinces); 1978-1989	– Education and trade (+) – Inflation and state-owned enterprises (-)
Radelet, Sachs, and Whang Lee 2001	14 Asian countries; 1965-1990	– Education, government savings, trade openness, quality of institutions, life expectancy, growth of working age population (+) – Labour productivity, natural resource (-)
Barro 2003	87 countries 1965-1995	– Average years of school attainment, investment, rule of law, democracy, trade openness, and terms of trade (+) – Initial level of per capita GDP, life expectancy, fertility rate, government consumption, inflation rate and landlockedness (-) nomic growth.
Bhaskara-Rao and Hassan 2011	Bangladesh; 1970-2007	– FDI, money supply, and trade openness (+) – Government expenditure and inflation (-)
Chang and Mendy 2012	36 African countries; 1980-2009	– Foreign aid, Trade, (+) – Foreign direct investment, domestic investment, and gross domestic savings (-) - Foreign aid (-) (+)
Anyanwu 2014	53 African countries; 1996-2010	– Africa: investment, education, resource, (governance), (+) – China: domestic investment, and trade openness (+) – China: Population growth, inflation, credit to the private sector,(-)

Source: Adjusted from Chirw et al (2016).

Table 2: A review of the determinants of growth (Developed Countries)

Author(s)	Countries and Sample Period	Determinants with a Significant Positive (+) or Negative (-) Impact on Growth
	Developed Countries	
Bleaney, Gemmell, and Kneller 2001	22 OECD countries; 1970-1995	– Investment (+) – Taxation (-)
Freire-Seren 2002	Spain (Spanish Regions); 1964-1991	– Human capital and investment (+)
Anaman 2004	Brunei Darussalam; 1971-2001	– Exports and investment (+) – High government size(-)
Acikgoz and Mert 2005	3 Asian countries; Time series data (Taiwan, 1951-2007; Korea, 1953-2007; and Hong Kong, 1960-2007)	Investment (+)
Bayraktar 2006	Turkey; 1968-1998	– Investment and human capital (+) – Inflation (-)
Asheghian 2009	Japan; 1971-2006	– TFP (+) - Investment (+)
Checherita Westphal and Rother 2012	12 Euro countries; 1970-2008	– Government balance, government debt, private savings, and trade openness (+) – Population growth, the square of government debt and real interest rates (-)

Source: Adjusted from Chirw et al (2016)

Table 3: A review of the determinants of growth (South-East and Central European Countries)

Author(s)	Countries and Sample Period	Determinants with a Significant Positive (+) or Negative (-) Impact on Growth
	South-East and Central European Countries	
Botric and Slijepcevic 2008	6 South-Eastern European countries; 1995-2005	– Inflation and interest rate spread (-) – General government balance as a share of GDP (+)
De Grauwe and Schnabl 2008	18 South-Eastern and Central European countries; 1994-2004	– Exchange rate stability (+) – Growth rate of dollar exports and budget defcit (-)
Prochniak 2011	10 Central and Eastern European countries; 1993-2009	– Investment , human capital development, fnancial sector development, high services share in GDP, high share of working age population, development of information, communication and technology (ICT), high private sector share in GDP, economic freedom, and progress in market and structural reforms (+) – Budget defcits, public debt, interest rates, and inflation (-)
Fetahi-Vehapi, Sadiku, and Petkovski 2015	10 South East European countries; 1996-2012	– Trade openness, human capital development,investment, FDI(+) – Population (-)

Source: Adjusted from Chirw et al (2016)

Barro (2003) finds empirical evidence that indicates that income is positively affected by better governance, low government consumption low inflation, high life expectancy, high level of education and improving terms of trade. Prince (2019) applied cointegration analysis on data from 1980-2016 to examine the effect of infrastructure on economic growth in the United States and found a positive and significant effect in the long run but not in the short run.

As OECD (2015) indicates there are no best-practice policies that will always yield the same positive result, "*For example, creating stronger incentives for private investment may require improving the security of property rights in one country, but enhancing the financial sector in another. Technological catch-up may call for stronger or weaker patent protection, depending on the level of development*".

3. Energy as a factor of economic growth

3.1. Causality between energy and growth

This section focus on the energy-growth nexus and empirically examines the role of energy in economic growth. The causal relationship between energy and economic growth has attracted attention over the last three decades. More than 200 research papers and reviews have been published to support one of the four possible directions of the causal relationship between energy and economic growth. Kalimeris et al. (2014) carried out a meta-analysis investigation of the following four directions of the energy-GDP causal relationship:

- No causality between energy use and economic growth (**neutrality hypothesis**)
- Unidirectional causality from GDP growth to energy use (**conservation hypothesis**)
- Unidirectional causality from energy use to GDP growth (**growth hypothesis**)
- Bi-directional causality between energy use and GDP growth (**feedback hypothesis**)

The existing literature is inconclusive as to whether energy causes economic growth or vice versa. Kalimeris et al. (2014) in their extensive meta-analysis found that from 686 causality tests that are derived by 158 research papers, 22.6% of the cases examined support the no causality hypothesis, 23.8% the conservation hypothesis, 28.1% the growth hypothesis and 25.5% the feedback hypothesis. These proportions vary when we focus on the acceptance of each hypothesis conditionally to a specific methodology (i.e. papers that applied the Toda – Yamamoto causality test, the Granger causality test etc.).

These mixed econometric results raise important questions regarding the appropriateness of the type of methodology, the selected set of data and variables that are used to examined such a hypothesis (Beaudreau, 2010).

Our purpose is to further investigate the link between energy use and gross value added by using multi-method econometric analysis and by focusing at the sectoral level. Time series and panel analysis are carried on the data derived from the WIOD database. The WIOD database contains data for 39 countries and 35 economic sectors[1]. To investigate causality, Granger causality (1987), Toda-Yamamoto (1995) and cointegration tests are applied for time-series data. Pedroni (1999) and Kao (1999) methods are used for panel data.

In the next subsection a short description of the methods that are commonly used in the literature to examine the causality between energy and economic growth are presented. In the subsection 3.1.2 the results of our econometric analysis are presented focusing at the sectoral level examining the causality between energy-use and gross value added, by using a multi-method approach, for 39 countries and 35 economic sectors.

[1] See Table 0-55 and Table 0-56 in the Appendix for more details about the regions and economic sectors used in the econometric analysis Table 0-56.

3.1.1. Causality tests

Several approaches have been used in the literature to empirically examine the direction of the causality between energy use and economic growth. Beaudreau (2010) classified the energy-GDP causality tests into three classes: early tests, cointegration tests and post-cointegration tests.

Early tests

In the early tests the causality between two variables is examined by empirically testing if the past values of the causal variable give additional information to predict the future values of the explanatory variable. The main caveat of these tests is that in order for the inference to be valid the variables under examination should be stationary[2]. The early tests[3] of Granger (1969) and Sims (1972) are characterised as invalid to be used in empirical economic analysis as it was shown, in early 1980s, that many economic variables, among which energy-use and GDP, are not stationary. In particular, it is found that many economic variables are not stationary at their levels but they become stationary if first differences are taken).

Based on these findings, it is proposed to estimate causality by using the Granger causality test on the first difference of the timeseries, which in most cases were found to be stationary.

Cointegration tests

In the late 1970s the first unit root econometric tests that test for stationarity appeared in the literature. Unit root tests such as ADF-test (Dickey and Fuller (1979, 1981)) and PP-test (Phillips and Perron (1988)) are widely applied in econometric studies and have become a pre-requisite test even in modern econometric analysis. In the presence of a unit root in the variables the Ordinary Least Square (OLS) estimator results in misleading statistical evidence of a linear relationship between independent non-stationary variables, the so-called "spurious regression[4]".

The unit root tests have also been used to identify the order of integration[5] of a variable that is necessary information when searching for a cointegration relationship[6]. In the early 1980s the first cointegration tests for whether a cointegrating relationship between two variables exists appeared in the literature (see Engle and Granger (1987), Johansen and Juselius (1990)).

Granger (1988) argued that a cointegration relationship implies causation but without identifying the direction of causation. Granger (1988) suggested to use an error correction model (ECM) which not only can identifies the direction of causation but also allows a

---

[2] We can say that a variable is stationary when its statistical measures such as mean, variance etc., do not change over time.

[3] See section: Causality tests.

[4] The term "spurious" is used when the causal effect of a variable to one other variable comes from a third variable that affect these two variables. In such case, a correlation of the first two variables is appeared even when neither has a causal effect on the other.

[5] We say that a variable is integrated of order p if the p order differences of the variable do not have a unit root, which means that it is stationary.

[6] We say that a cointegration exists if each variable under examination is integrated of order p but a linear combination of the variables is integrated of order lower than p. This linear combination of the variables is called cointegration relationship.

distinction between short run and long run causality. ECM can be applied in cases where the variables under examination have the same order of integration and are cointegrated.

Time series analysis often suffers from low explanatory power when the time series are short. In such cases a panel data analysis, that offers more observations and hence degrees of freedom, may result in more efficient estimates.

Kao (1999) transferred the Engle and Granger (1987) idea for cointegration in panel data analysis and developed a panel data cointegration test with homogenous individual coefficient.

Pedroni (1999) noted that when the maintained hypothesis of homogeneity of the coefficients of the regression is false then the null hypothesis of no cointegration may not be rejected even if the variables are, in fact, cointegrated. Pedroni (1999) extended the Kao (1999) panel data cointegration test by including a deterministic trend and heterogeneity across individuals in the cointegration relationship.

Post-cointegration tests

The number of tests that are required (unit root test, cointegration test and inference in the ECM) to identify causality may lead to counterintuitive results. To overcome this problem other tests, that are based on a single model are developed.

Toda and Yamamoto (1995) developed a methodology for identifying causality regardless of whether the variables examined have the same order of integration or are co-integrated. By using a vector autoregressive (VAR) model and appropriate lag lengths of the variables, they showed that a Wald type statistic is valid for testing causality.

Bound tests that have better small sample properties have been developed as an alternative to cointegration tests. Pesaran (2001) used an autoregressive distributed lag (ARDL) model and proposed standard F-statistic bound tests for cointegration regardless of whether the variables have the same order of integration.

A large number of research papers have used one or more of the above methodologies for identifying the existence of a cointegrating and causal relationship between energy use and economic growth. For an extensive literature review of econometric methods and results see Kalimeris, Richardson and Bithas (2014) and Ozturk (2010).

3.1.2. Econometric results on energy-growth nexus – timeseries analysis

For the causality analysis data for energy volumes and the quantity index of gross value added have been drawn from the WIOD database for 39 countries and 35 economic sectors. In particular, energy volumes have been derived from the WIOD environmental accounts[7] and the gross value added from the socio-economic accounts[8]. All data used are available in the Causality WIOD Data excel file that is attached to this report.

To apply causality tests for energy use and gross value added, the order of integration of the timeseries needs to be identified. The ADF unit root tests[9] for the energy and gross value added timeseries by each sector are presented in Table 0-4 and

[7] http://www.wiod.org/database/eas13

[8] http://www.wiod.org/database/seas13

[9] The code of the timeseries unit root tests can be found in the Appendix.

Table 0-5, respectively.

Table 0-4: ADF unit root tests for energy-use by sector

	I(0)	I(1)	n/a	> I(1)		I(0)	I(1)	n/a	> I(1)
Total industries	7	24	0	8	Construction	5	31	0	3
Agriculture, Hunting, Forestry and Fishing	12	24	0	3	Sale, Maintenance and Repair of Motor Vehicles and Motorcycles; Retail Sale of Fuel	11	24	2	2
Mining and Quarrying	13	24	0	2	Wholesale Trade and Commission Trade, Except of Motor Vehicles and Motorcycles	9	27	0	3
Food, Beverages and Tobacco	10	28	0	1	Retail Trade, Except of Motor Vehicles and Motorcycles; Repair of Household Goods	9	28	0	2
Textiles and Textile Products	22	16	0	1	Hotels and Restaurants	8	28	0	3
Leather, Leather and Footwear	16	20	2	1	Inland Transport	5	29	0	5
Wood and Products of Wood and Cork	6	27	0	6	Water Transport	5	25	3	6
Pulp, Paper, Paper , Printing and Publishing	9	27	0	3	Air Transport	15	23	1	0
Coke, Refined Petroleum and Nuclear Fuel	10	20	4	5	Other Supporting and Auxiliary Transport Activities; Activities of Travel Agencies	8	26	0	5
Chemicals and Chemical Products	8	29	0	2	Post and Telecommunications	10	26	0	3
Rubber and Plastics	12	22	0	5	Financial Intermediation	10	27	0	2
Other Non-Metallic Mineral	12	22	0	5	Real Estate Activities	11	26	0	2
Basic Metals and Fabricated Metal	5	24	0	10	Renting of M&Eq and Other Business Activities	7	25	0	7
Machinery, Nec	14	22	0	3	Public Admin and Defence; Compulsory Social Security	7	28	0	4
Electrical and Optical Equipment	11	23	0	5	Education	10	25	0	4
Transport Equipment	9	25	0	5	Health and Social Work	7	28	0	4
Manufacturing, Nec; Recycling	13	25	0	1	Other Community, Social and Personal Services	6	26	0	7
Electricity, Gas and Water Supply	7	25	0	7	sum total	339	879	12	135

Table 0-5: ADF unit root tests for gross value added by sector

	I(0)	I(1)	n/a	> I(1)		I(0)	I(1)	n/a	> I(1)
TOT	11	3	0	25	**F**	13	14	0	12
AtB	15	20	0	4	**50**	8	17	2	12
C	9	24	0	6	**51**	6	18	0	15
15t16	10	21	0	8	**52**	7	20	0	12
17t18	13	20	0	6	**H**	11	17	0	11
19	15	19	2	3	**60**	8	19	0	12
20	9	17	0	13	**61**	9	19	3	8
21t22	13	15	0	11	**62**	10	24	1	4
23	12	18	4	5	**63**	9	19	0	11
24	8	21	0	10	**64**	3	18	0	18
25	7	16	0	16	**J**	6	23	0	10
26	11	14	0	14	**70**	8	23	0	8
27t28	5	20	0	14	**71t74**	5	18	0	16
29	8	19	0	12	**L**	3	25	0	11
30t33	4	19	0	16	**M**	11	16	0	12
34t35	10	20	0	9	**N**	10	18	0	11
36t37	6	22	0	11	**O**	14	15	0	10
E	10	24	0	5	**sum total**	317	655	12	381

Most of the timeseries are found to be integrated of order one.

Table 0-6 combines the order of integration of energy and gross value added. Four cases are considered. Column (1) corresponds to the case that both timeseries (energy and gross value added) are stationary; column (2) corresponds to the case that both timeseries are not stationary but are integrated of order (1); column (3) corresponds to the cases that one timeseries is integrated of order zero and the other of order one; and finally column (4) corresponds to all other cases.

Table 0-6: Combined results of integration of the timeseries energy-use and gross value added

	(1)*	(2)*	(3)*	(4)*		(1)*	(2)*	(3)*	(4)*
TOT	1	0	11	27	**F**	2	12	12	13
AtB	6	12	14	7	**50**	2	9	13	15
C	3	14	14	8	**51**	1	14	7	17
15t16	3	15	12	9	**52**	1	13	13	12
17t18	6	7	19	7	**H**	3	11	12	13
19	7	10	16	6	**60**	2	13	9	15
20	5	14	4	16	**61**	0	13	13	13
21t22	4	12	11	12	**62**	4	15	15	5
23	2	10	15	12	**63**	4	15	7	13
24	1	14	12	12	**64**	1	14	6	18
25	2	8	11	18	**J**	1	18	9	11
26	4	9	10	16	**70**	4	17	9	9
27t28	1	12	6	20	**71t74**	1	13	7	18
29	5	11	9	14	**L**	1	19	5	14
30t33	1	10	9	19	**M**	4	12	10	13
34t35	3	13	10	13	**N**	2	14	10	13
36t37	3	15	9	12	**O**	3	12	12	12
E	3	17	8	11	**sum total**	96	437	369	463

*Columns: (1) both timeseries are I(0), (2) both are I(1), (3) one I(0) and one (1), (4) other.

Of the total number of 1365 cases, only in 96 were both series stationary I(0) allowing the Granger causality test to be performed without any further econometric analysis. In all other cases the test of Granger causality may be invalid. There are 451 cases (463 – 12 cases for which data are missing either for energy-use or for gross value added) that are non-stationary even if we take first differences.

By excluding the cases characterised by strong non-stationarity, the timeseries analysis focuses on the causality test at the remaining 902 cases (excluding all the cases that belong to the column (4)) by using an appropriate econometric technique.

Table 0-7 presents the results for the cases in which both energy and gross value added are stationary. The results indicate that the neutrality hypothesis has the highest rate of acceptance in these cases.

Table 0-7: Causality test on cases that both energy and gross value added are stationary, I(0)

	Energy does not Granger cause GVA	Energy Granger cause GVA	Total
GVA does not Granger cause Energy	Neutrality 43 (45%)	Growth 20 (21%)	63 (66%)
GVA Granger cause Energy	Conservation 23 (24%)	Feedback 8 (27%)	33 (34%)
Total	66 (69%)	30 (31%)	96

For the cases in which both energy-use and gross value added are stationary after taking first differences, I(1) (437 cases), the Engle and Granger (1987) and the Johansen and Juselius (1990) tests for cointegration are performed.

Of the total number of 437 cases, 130 were found to have stationary residuals if the Engle and Granger cointegration test is used and 295 cases if the Johansen and Juselius cointegration test is used. Our results strongly support Bilgili (1998) that concludes that Johansen methodology dominates the Engle-Granger methodology in cointegration analyses. Table 0-8 presents detailed results on cointegration tests by sector.

Table 0-8: Cointegration tests by sector in cases that both energy and gross value added are found to be integrated of order *one.*

	Both I(1)	EG*	JJ*		Both I(1)	EG*	JJ*
TOT	0	0	0	**F**	12	2	8
AtB	12	4	6	**50**	9	6	5
C	14	3	10	**51**	14	6	9
15t16	15	4	13	**52**	13	7	9
17t18	7	1	5	**H**	11	7	7
19	10	2	5	**60**	13	4	10
20	14	1	9	**61**	13	4	11
21t22	12	1	7	**62**	15	4	11
23	10	3	7	**63**	15	5	8
24	14	3	8	**64**	14	3	12
25	8	2	7	**J**	18	4	12
26	9	4	8	**70**	17	6	12
27t28	12	5	6	**71t74**	13	3	11
29	11	4	6	**L**	19	5	11
30t33	10	4	5	**M**	12	4	6
34t35	13	4	9	**N**	14	8	11
36t37	15	3	8	**O**	12	1	11
E	17	3	12	**sum total**	437	130	295

*EG: Engle and Granger cointegration test, JJ: Johansen and Juselius cointegration test

Table 0-9 and Table 0-10 present the results for the cases in which both energy and gross value added are integrated of order one and cointegrated, based on Engle Granger and Johansen and Juselius cointegration tests, respectively. The direction of the causality is identified by using a Vector Error Correction Model of the two variables that allows to identify if a short-run or/and a long-run causality exists.

The results indicate that the neutrality hypothesis has the highest rate of acceptance in these cases.

Table 0-9: Causality test on cases that both energy and gross value added are integrated of order one, I(1) and cointegrated based on Engle and Granger cointegration test

	Energy does not Granger cause GVA	Energy Granger cause GVA	Total
GVA does not Granger cause Energy	Neutrality 61 (47%)	Growth 34 (26%)	95 (73%)
GVA Granger cause Energy	Conservation 23 (18%)	Feedback 12 (9%)	35 (27%)
Total	84 (65%)	46 (35%)	130

Table 0-10: Causality test on cases that both energy and gross value added are integrated of order one, I(1) and cointegrated based on Johansen and Juselius cointegration test

	Energy does not Granger cause GVA	Energy Granger cause GVA	Total
GVA does not Granger cause Energy	Neutrality 166 (56%)	Growth 63 (21%)	229 (78%)
GVA Granger cause Energy	Conservation 47 (16%)	Feedback 19 (6%)	66 (22%)
Total	213 (72%)	82 (28%)	295

In the cases in which both energy use and gross value added are I(1) and a cointegration vector does not exist, the Granger causality test is invalid to be performed in the levels of the timeseries and is performed in their first differences. Table 0-11 and Table 0-12 present the results on these cases. It is found that the neutrality hypothesis has the highest rate of acceptance when no cointegration exists between energy and gross value added.

Table 0-11: Causality test on cases that both energy and gross value added are integrated of order one, I(1) and not cointegrated based on Engle and Granger cointegration test

	Energy does not Granger cause GVA	Energy Granger cause GVA	Total
GVA does not Granger cause Energy	Neutrality 221 (72%)	Growth 36 (12%)	257 (84%)
GVA Granger cause Energy	Conservation 43 (14%)	Feedback 7 (2%)	50 (16%)
Total	264 (87%)	43 (13%)	307

Table 0-12: Causality test on cases that both energy and gross value added are integrated of order one, I(1) and not cointegrated based on Johansen and Juselius cointegration test

	Energy does not Granger cause GVA	Energy Granger cause GVA	Total
GVA does not Granger cause Energy	Neutrality 106 (75%)	Growth 14 (10%)	103 (85%)
GVA Granger cause Energy	Conservation 18 (13%)	Feedback 4 (3%)	39 (15%)
Total	124 (87%)	28 (13%)	142

Table 0-13 presents the results of the post-cointegration test of Toda and Yamamoto (1995) that allows to examine the causality between energy and gross value added regardless of the integration and cointegration of the time series. Due to lack of degrees of freedom we consider only the cases for which both timeseries have an order of integration lower than two (see: (1), (2) and (3) of

Table 0-6).

Table 0-13: Causality test for cases in which energy and gross value added are integrated of order lower than two based on the Toda and Yamamoto test

	Energy does not cause GVA	Energy cause GVA	Total
GVA does not cause Energy	Neutrality 652 (72%)	Growth 102 (11%)	754 (84%)
GVA cause Energy	Conservation 123 (14%)	Feedback 25 (3%)	148 (16%)
Total	775 (86%)	127 (14%)	902

In conclusion, timeseries analysis indicates that the neutrality hypothesis is supported as the hypothesis with the highest rate of acceptance across all the tests reported in this section. The existence of a cointegration relationship between the energy use and gross value strongly affects the causality tests: in the cases for which a cointegration relationship is found or the series can be considered as stationary there is a greater number of cases in which the Growth hypothesis is accepted.

3.1.3. Econometric results on energy-growth nexus – panel data analysis

The test for causality between energy use and gross value added in a panel context is conducted in three steps. First, the order of integration in the two timeseries is tested. Second, after having established the order of integration in the series, panel cointegration tests are used to examine the long-run relationships between the variables in question. Then, in cases in which the both series are found to be stationary or integrated of order one but cointegrated, dynamic panel causality tests can be used to evaluate the short-run and log-run direction of causality between energy and gross value added.

Panel causality tests allow more accurate and efficient estimates to made of causality between energy and gross value added than those of timeseries analysis since they permit a larger number of observations and hence degrees of freedom.

Table 0-14 and

Table 0-15 present the p-values of two panel data unit root tests by sector for energy and gross value added, respectively. The LLC unit root test proposed by Levin, Lin, Chu (2002) for common unit root across individuals and the IPS unit root test proposed by Im, Pesaran, Shin (2003) for heterogenous unit roots have been used to identify the order of integration of energy and gross value added.

Table 0-14: p-values of LLC and IPS panel unit roots on Energy

	LLC[1]	IPS[1]	LLC[2]	IPS[2]		LLC[1]	IPS[1]	LLC[2]	IPS[2]
TOT	0.015	0.057	0.195	0.986	**F**	0.067	0.176	0.791	0.701
AtB	0.000	0.360	0.000	0.000	**50**	0.000	0.008	0.000	0.000
C	0.001	0.099	0.000	0.000	**51**	0.079	0.971	0.000	0.018
15t16	0.134	0.293	0.000	0.000	**52**	0.000	0.133	0.000	0.008
17t18	0.000	0.203	0.000	0.000	**H**	0.043	0.155	0.088	0.394
19	0.000	0.228	0.000	0.000	**60**	0.002	0.794	0.000	0.061
20	0.223	0.842	0.000	0.166	**61**	0.000	0.006	0.035	0.260
21t22	0.332	0.471	0.000	0.646	**62**	0.000	0.037	0.000	0.000
23	0.028	0.086	0.000	0.002	**63**	0.182	0.921	0.000	0.001
24	0.000	0.045	0.000	0.003	**64**	0.174	0.332	0.000	0.000
25	0.070	0.115	0.000	0.000	**J**	0.000	0.033	0.000	0.004
26	0.000	0.025	0.002	0.079	**70**	0.000	0.002	0.000	0.004
27t28	0.250	0.283	0.844	0.436	**71t74**	0.046	0.828	0.000	0.000
29	0.000	0.001	0.001	0.004	**L**	0.104	0.130	0.005	0.004
30t33	0.000	0.036	0.000	0.004	**M**	0.094	0.813	0.032	0.442
34t35	0.119	0.286	0.189	0.570	**N**	0.588	0.970	0.000	0.029
36t37	0.000	0.000	0.000	0.000	**O**	0.223	0.981	0.495	0.307
E	0.003	0.244	0.104	0.592					

[1] Unit root test with a deterministic constant
[2] Unit root test with a deterministic constant and trend

Table 0-15: p-values of LLC and IPS panel unit roots on gross value added

	LLC[1]	IPS[1]	LLC[2]	IPS[2]		LLC[1]	IPS[1]	LLC[2]	IPS[2]
TOT	0.868	0.999	0.000	0.083	**F**	0.001	0.703	0.732	0.198
AtB	0.997	0.335	0.000	0.000	**50**	0.027	0.453	0.018	0.012
C	0.000	0.060	0.016	0.641	**51**	0.336	0.989	0.952	0.694
15t16	0.069	0.991	0.005	0.163	**52**	0.969	1.000	0.247	0.082
17t18	0.987	1.000	0.638	0.840	**H**	0.000	0.085	0.129	0.034
19	0.387	0.989	0.000	0.000	**60**	0.162	0.599	0.439	0.024
20	0.016	0.030	1.000	0.864	**61**	0.003	0.192	0.000	0.088
21t22	0.623	0.973	0.424	0.035	**62**	0.507	0.046	0.000	0.000
23	0.000	0.407	0.000	0.000	**63**	0.000	0.491	0.981	0.053
24	0.025	0.781	1.000	0.994	**64**	1.000	1.000	1.000	1.000
25	0.041	0.524	1.000	0.755	**J**	0.988	1.000	0.762	0.876
26	0.267	0.160	0.969	0.000	**70**	0.005	1.000	0.002	0.016
27t28	0.141	0.641	1.000	0.474	**71t74**	0.998	1.000	0.000	0.000
29	0.845	0.910	0.317	0.005	**L**	0.301	1.000	0.449	0.736
30t33	0.252	0.790	0.368	0.280	**M**	0.998	1.000	0.042	0.128
34t35	0.000	0.204	0.000	0.014	**N**	1.000	1.000	0.285	0.344
36t37	0.352	0.959	0.922	0.299	**O**	0.000	0.998	0.059	0.138
E	0.000	0.000	0.000	0.000					

[1] Unit root test with a deterministic constant
[2] Unit root test with a deterministic constant and trend

We discuss the results using the LLC and IPS panel unit roots tests for the cases in which both energy use and gross value added are integrated at the same order at the 5% level of significance.[10].

For the cases in which both energy and gross value added are stationary, I(0), the Granger causality test is applied by using the OLS estimator. By testing the cross-section dependence on the residuals with the Breusch-Pagan LM test, the Pesaran scaled LM test and the Pesaran cross dependence test, it is found that in all cases the residuals are cross dependent. On the presence of cross dependency, the variance-covariance matrix with panel corrected standard errors (OLS PCSE) is used. In addition, the FMOLS developed by Pedroni (2000) and the Dynamic OLS (DOLS) panel data estimators are used for comparison purposes. Both DOLS and FMOLS are preferred to the OLS estimator because they take care of small sample bias and endogeneity bias by taking the leads and lags of the first-differenced regressors.

[10] The sectors: 17t18, 21t22, 30t33, 34t35, 36t37, 51, 52, 60, 64, J, L and N are excluded from the panel data analysis.

Table 0-16, Table 0-17 and Table 0-18 present the panel causality tests in the 20 cases[11] in which either the LLC or the IPS unit root test suggests that both energy and gross value added are integrated of order zero by using the variance-covariance matrix with panel corrected standard errors, the FMOLS and the DOLS estimators, respectively.

[11] Sectors: TOT, AtB, C, 15t16, 19, 20, 23, 24, 25, 26, 29, E, 50, H, 61, 62, 63, 70, 71t74, M

Table 0-16: Panel causality test with panel corrected standard errors on cases that energy and gross value added are integrated of order zero

	Energy does not cause GVA	Energy cause GVA	Total
GVA does not cause Energy	Neutrality 5 (25%)	Growth 5 (25%)	10 (50%)
GVA cause Energy	Conservation 7 (35%)	Feedback 3 (15%)	10 (50%)
Total	12 (60%)	8 (40%)	20

Table 0-17: FMOLS panel causality test on cases that energy and gross value added are integrated of order zero

	Energy does not cause GVA	Energy cause GVA	Total
GVA does not cause Energy	Neutrality 3 (15%)	Growth 6 (30%)	14 (45%)
GVA cause Energy	Conservation 5 (25%)	Feedback 6 (30%)	6 (55%)
Total	8 (40%)	12 (60%)	20

Table 0-18: DOLS panel causality test on cases that energy and gross value added are integrated of order zero

	Energy does not cause GVA	Energy cause GVA	Total
GVA does not cause Energy	Neutrality 9 (45%)	Growth 7 (35%)	16 (80%)
GVA cause Energy	Conservation 4 (20%)	Feedback 0 (0%)	4 (20%)
Total	13 (65%)	7 (35%)	20

The panel data causality tests result in a higher acceptance rates of the Growth hypothesis as compared to the timeseries causality tests. The DOLS estimator, that is influenced more by the time dimension, has more similar results to the timeseries causality test as compared to the FMOLS and OLS estimators.

Table 0-19 presents the causality tests by sector and by estimator. The results for Chemicals, Rubber and Plastics, Coke, Refined Petroleum and Nuclear Fuel, and Electricity, Gas and Water Supply support the Growth hypothesis. These sectors are all highly dependent on energy and the Growth hypothesis suggests that by knowing the change in energy-use for these sectors we can forecast the future values of their gross value added.

Table 0-19: Panel causality tests by sector and by estimator

	OLS (PCSE)	FMOLS	DOLS
AtB	Neutrality	Neutrality	Neutrality
C	Conservation	Conservation	Conservation
15t16	Conservation	Conservation	Neutrality
19	Conservation	Conservation	Neutrality
20	Conservation	Conservation	Neutrality
23	Growth	Feedback	Growth
24	Growth	Growth	Growth
25	Growth	Growth	Growth
26	Feedback	Feedback	Conservation
29	Conservation	Growth	Conservation
E	Growth	Growth	Growth
50	Neutrality	Neutrality	Neutrality
H	Neutrality	Growth	Growth
61	Conservation	Feedback	Neutrality
62	Neutrality	Neutrality	Neutrality
63	Neutrality	Growth	Neutrality
70	Growth	Feedback	Growth
71t74	Conservation	Conservation	Neutrality
M	Feedback	Feedback	Conservation
TOT	Feedback	Feedback	Growth

For the cases in which both energy and gross value added are integrated of order one[12] the Pedroni (1999) and Kao (1999) panel cointegration tests have been performed and a vector error correction model applied in cases in which a cointegration relationship was found to exist.

Table 0-20: Pedroni and Kao cointegration tests

	27t28	F	M	O	TOT
Panel v	0.257	0.000	0.823	0.002	0.686
Panel rho	0.984	0.155	0.758	0.000	1.000
Panel PP	0.994	0.001	0.166	0.000	1.000
Panel ADF	0.910	0.000	0.031	0.000	1.000
Group rho	0.995	0.976	0.993	0.884	1.000
Group PP	0.295	0.259	0.151	0.001	1.000
Group ADF	0.000	0.000	0.053	0.000	0.565
Kao	0.000	0.000	0.019	0.082	0.000

[12] Only 5 cases: the sectors 27t28, F, M, O and TOT.

The Kao cointegration test, which assumes homogeneity across countries by sector, identifies a cointegration relationship in the four sectors (27t28, F, M and TOT). Pedroni cointegration tests, which assume heterogeneity across countries by sector, provide mixed outcomes depending on the cointegration test used. Since in each sector there is at least one cointegration test that supports that a cointegration relationship exist we proceed to the Granger causality test via an Error Correction model.

Table 0-21: Panel ECM causality test

27t28	F	M	O	TOT
Feedback	Feedback	Conservation	Feedback	Feedback

3.2. Energy substitution possibilities

This section focus on the energy substitution possibilities and the work previously carried out for DG ENER[13]. Our analysis has been extended[14] to cover more countries including all the 28 EU member states and selected non-EU countries by utilising the WIOD[15] database. Our analysis has also been extended so as to employ recent advancments in econometric techniques. Time series analysis with and without structural breaks have been applied, linear and non-linear cointegration relationships have been examined and panel data techniques have been used.

Our analysis focuses on the size of the substitution elasticity and whether this varies across countries, sectors and over time. This section presents estimates of the size of the elasticity of substitution through various econometric methods. Firstly, time series analysis are applied using the ADF unit root test (Dickey Fuller (1979)) and units root test that allow for one-time structural break proposed by Zivot and Andrews (1992). Secondly, linear and non-linear cointegration are examined by using methodologies that allow for a possible structural break. For linear cointegration the methodology of Engle and Granger (1987) and Gregory and Hansen (1996) that allows for one-time structural break at an unknown time are used. For non-linear cointegration the method of Shin and Yu (2014) that includes asymmetries from a positive and a negative change in an ARDL model is applied.

Furthermore, the panel unit root tests of Levin, Lin, Chu (2002) and Im, Pesaran, Shin (2003) have been used to identify the order of integration of the panels. Cointegration tests developed by Pedroni (1999) and Kao (1999) have been used to identify a cointegration relationship. Finally, to estimate the elasticity of substitution of energy and gross value added the OLS estimator with panel corrected standard errors has been used.

[13] European Commission (2017), 'Study on the Macroeconomics of Energy and Climate Policies', available from https://ec.europa.eu/energy/en/data-analysis/energy-modelling/macroeconomic-modelling

[14] In the previous analysis only 12 EU member states are covered.

[15] http://www.wiod.org/home

3.2.1. Literature review on econometric methods used to estimate the energy substitution possibilities

Van der Werf (2008) and Kemfert (1998) mention that the (KL)E nesting structure offers the best fit to historical data and their findings reject the null hypothesis that the elasticity of substitution between capital-labour and energy is equal to one in favour of the hypothesis that the elasticity is lower than one. Using panel data sets it is found that the sizes of the elasticities vary over countries and industries. Kemfert (1998) used three approaches for the CES production function to estimate the elasticity of substitution among capital, labour and energy and found the elasticity of substitution between capital-labour and energy to be 0.458.

In addition, Koesler and Schymura (2015) used non-linear least squares estimation in a three-level nested KLEM on data taken from the WIOD database to estimate the elasticities of substitution between capital and labour, capital-labour and energy, and production factors with intermediate factors. The elasticities of substitution between capital-labour and energy vary between 0 and 7.86. Antoszewski (2019) used data from the WIOD database and applied panel data techniques, finding that the elasticity of substitution between capital-labour and energy was lower than one.

The econometric estimation of the elasticity of substitution can be based on various econometric methods, some of them already presented in the previous section. In this subsection, we therefore present only the methods that consider a possible structural break in the timeseries or a non-linear relationship.

Times series methods that consider structural break

Perron (1989) developed a unit root test that considers a possible structural break in the timeseries and claimed that if a structural break exists then the ADF unit root test of Dickey and Fuller (1979) and the PP unit root test of Phillips and Perron (1988) can be invalid. If a variable is stationary and a structural break exists the ADF and PP unit root test may fail to reject the null hypothesis of non-stationarity.

Zivot-Andrews (1992) proposed a unit root test that the structural break point is endogenously selected. They proposed three different econometric models: a model with a structural break on the constant, a model with a structural break on the deterministic trend and a model with a structural break both on the constant and the deterministic trend. Each model is estimated using the OLS estimator and by assuming that the possible structure break belongs on the time interval $[2\,, T-1]$. Among all the subsequent values the point that achieves the lowest t-statistic on the null hypothesis of the presence of unit root is selected. Although this approach ensures that the unit root test is consistent with the existence of a structure break in the timeseries, it does not identify if a structural break exist.

Gregory and Hansen (1996) developed a cointegration test that allows for a structural break in the cointegration relationship. The classical ADF unit root test may suffer from low power if a structural break exists but has not been included in the cointegration relationship. They proposed three different models depending on the nature of the structural break: (i) a structural break on the constant, (ii) a structural break on the constant and trend and (iii) a structural break on the constant and the elasticity/slope.

Times series methods that consider non-linearity

Shin and Yu (2014) note that in cases in which an asymmetric relationship between the variables of examination exists, the cointegration results of a symmetric linear relationship may not be valid. Ndoricimpa (2017) used the non-linear ARDL model of Shin and Yu (2014) for South Africa and identified asymmetries in the relationship among the energy-use, pollution emissions and real output.

3.2.2. Econometric results on the energy substitution possibilities

The WIOD release 2013 database has been used. A new version of this database, release 2016, which covers 56 sectors and 43 regions of the world economy within the time span 2000–2014 is available but it does not include an updated version of the WIOD energy data and therefore it was not possible to use this 2016 update for the purpose of this analysis.

The indicators that have been used for the estimation of the elasticity of substitution between gross value added and energy are presented in Table 0-22. The database contains balanced data for all variables included in the analysis for the period 1995 – 2009.

Table 0-22: Variables that are used on the econometric model for the estimation of the elasticity of substitution

Variable	Unit	Description
QE	Total energy in TJ	Total energy in TJ is derived from the WIOD Environmental Accounts, available at: http://www.wiod.org/database/eas13
QKL	Gross value added (in million $, 1995)	Gross value added at national currencies and constant prices 1995 are derived from the Socio-Economic Accounts available at: http://www.wiod.org/database/seas13 Exchange rates that convert the national currencies into US$ are derived from World Input-Output Tables available at: http://www.wiod.org/database/wiots13
PE	Energy price index	The cost of energy use by sector is derived from World Input-Output Tables available at: http://www.wiod.org/database/wiots13 The energy price index was computed as the ratio of the cost of energy use by sector divided by the total energy use in TJ by sector.
PKL	Gross value-added price index	The gross value-added price index is derived from the Socio-Economic Accounts available at: http://www.wiod.org/database/seas13

The estimation of the elasticity of substitution between gross value added and energy is based on the first order conditions of the producers' profit maximization problem[16] of a CES production function. The econometric model that is used is:

$$\ln\frac{QE_t}{QKL_t} = a + \varphi \cdot t - \sigma \cdot \ln\frac{PE_t}{PKL_t} + u_t \qquad [1]$$

[16] For more details see: Energy Resilience and Vulnerability in the EU and Other Global Regions. 'Study on the Macroeconomics of Energy and Climate Policies', available from ttps://ec.europa.eu/energy/sites/ener/files/documents/macro_energy_resilience_and_vulnerability_0.pd

where: σ the Hicksian elasticity of substitution (Hicks, 1932).

The Hicksian elasticity of substitution (HES):

$$\sigma = \frac{\partial\left(\frac{QE_t}{QKL_t}\right)/\left(\frac{QE_t}{QKL_t}\right)}{\partial\left(\frac{PKL_t}{PE_t}\right)/\left(\frac{PKL_t}{PE_t}\right)}$$

is selected as a consistent with the CES production function measure. The selected econometric model is an indirect or an economic approach of the elasticity of substitution estimation. Henningsen et al. (2019) applied a direct technical approach for the estimation of the elasticity of substitution, that is a direct non-linear estimation of the CES production function. Such an approach does not require any price data to provide estimates of the elasticity of substitution. However, this represents the major disadvantage of this technique. Price data may contain important information that is essential for the estimation of the Hicksian elasticity of substitution (Broadstock et al. (2007) and Antoszewski (2019)).

Henningsen et al. (2019) used non-linear goodness-of-fit testing to identify the CES production function that best fits with the data. They conclude that: "*none of our estimated elasticities—neither based on the data in Kemfert (1998), nor based on our newly created dataset—are reliable and can be used in policy modelling. Our findings, and those of previous studies following the same approach, are sobering and indicate severe inherent problems with estimating nested CES production functions of the KLE type through a direct approach, in particular when using short time series*". Henningsen et al. (2019) results indicate that, due to the multi-set of parameters that a typical CES function has, it is unlikely to identify among different nesting schemes.

Antoszewski (2019) used panel data analysis and the economic approach to estimate the elasticity of substitution of a nested-CES production function by utilising the WIOD database. He provides a wide-range estimates of the elasticity of substitution by sector that most of then lie in the interval [0,1]. Antoszewski (2019) also performed a wide-range F-test to conclude that there is a strong argument against the arbitrary use of Leontief and/or Cobb-Douglas specifications in multi-sector CGE models as in many sectors the hypothesis of a Leontief type (σ=0) or a Cobb-Douglas type (σ=1) elasticity of substitution is rejected. These results are very similar to those reported in the previous study for DG ENER[17] that identifies weak substitutability between energy and gross value added.

In this report, a wide range of tests and econometric approaches have been used to estimate the elasticity of substitution. All the estimates have been made by using the Eviews-9.5[18] econometric program.

Time-series analysis

Unit root tests for the timeseries Q-ratio ($\ln\frac{QE_t}{QKL_t}$) and P-ratio ($\ln\frac{PE_t}{PKL_t}$) have been implemented. The standard ADF unit root test and the formula[19] of McKinnon (2010), which provides critical values consistent with the small size of the sample, have been

[17] European Commission (2017), 'Study on the Macroeconomics of Energy and Climate Policies', available from https://ec.europa.eu/energy/en/data-analysis/energy-modelling/macroeconomic-modelling

[18] https://www.eviews.com/home.html

[19] This is: $\beta_\infty + \beta_1/T + \beta_2/T^2 + \beta_3/T^3$ where β_∞, β_1, β_2 and β_3 are derived from McKinnon (2010) and T the number of observations in the unit root test.

used. The Zivot – Andrews (1992) test that considers possible structural break in the timeseries has also implemented.

In both variables the unit root tests have been applied on all sectors and countries timeseries, 1326[20] cases, by assuming an autoregressive scheme with or without a deterministic part (constant or constant and trend). The lags of the autoregressive scheme have been based on the Schwarz criterion. In the Zivot – Andrews test, a structural break either at the deterministic constant or at the deterministic constant and trend, has been examined.

Table 0-23 presents a summary of the unit roots tests. In a large number of cases (562 cases) it is found that both Qratio and Pratio series are stationary if a structural break is considered. A different order of integration between the two ratios, zero or one, has been found in 350 cases. Strong non-stationarity for either of the two timeseries has been found to only 39 cases. Finally, in 360 cases have be found that the Qratio and the Pratio are non-stationary at their levels but they are stationary at theirs' first differences.

Table 0-23: ADF and Zivot-Andrews unit root tests

		ADF				Zivot - Andrews			
		Qratio				Qratio			
		I(0)	I(1)	>I(1)	Total	I(0)	I(1)	>I(1)	Total
Pratio	I(0)	151	193	15	359	562	140	0	702
	I(1)	253	590	23	866	210	360	8	578
	>I(1)	22	53	11	86	0	9	22	31
	Total	426	836	49	1311	772	509	30	1311

Based on Zivot – Andrews results regarding the order of integration of the Qratio and the Pratio the econometric models applied in each case are the following:

Case 1: Both timeseries are integrated of order zero according. The OLS estimator in the linear regression model has been used.

Case 2: Timeseries have a different order of integration (zero or one). The OLS estimator applied in the linear regression model at the first differences of the timeseries (OLS – FD).

Case 3: Both timeseries are integrated of order one. A cointegration relationship is examined.

1. Case 3.1: A cointegration relationship exists. The OLS estimator is applied on the error-correction model (ECM) by using the Engle and Granger 2-step approach.
2. Case 3.2: A cointegration under the presence of structural break exists. The OLS estimator is applied on the error-correction model (ECM – GH) by using a Gregory Hansen 2-step approach.
3. Case 3.3: A cointegration relationship does not exist. The OLS estimator applied in the linear regression model at the first differences of the timeseries (OLS – FD).

[20] 15 cases were identified without data available either on price ratio or on volume ratio.

4. Case 4: Timeseries are integrated of order higher than one (strong non-stationarity behaviour). These cases are excluded[21] from the analysis.

In addition, Pesaran (2001) proposed to estimate an autoregressive distributed lag (ARDL) model at a one step, regardless of whether the timeseries are stationary or not. Although the ARDL model involves potentially non-stationary variables, a bound test for cointegration has been proposed that is valid for inference even if the order of integration of the variables is different, zero or one.

Shin and Yu (2014) proposed a bound test procedure developed by Pesaran (2001) in a non-linear ARDL framework for testing non-linear cointegration. The non-linear ARDL (NARDL) model allows to capture possible asymmetries in the adjustment process. The speed of adjustment to the long-run equilibrium may differs if there is a positive or a negative change on the ratio of prices of energy to value added.

The different models that have been applied in each case for the estimation of the elasticity of substitution are summarised in Table 0-24.

Table 0-24: Model selection depending on the unit roots tests

<table>
<tr><td colspan="2" rowspan="2"></td><th colspan="4">Qratio</th></tr>
<tr><th>I(0)</th><th colspan="3">I(1)</th></tr>
<tr><th rowspan="3">Pratio</th><th>I(0)</th><td>OLS
OLS – FD
ARDL
NARDL</td><td colspan="3">OLS – FD
ARDL
NARDL</td></tr>
<tr><th rowspan="2">I(1)</th><td rowspan="2">OLS – FD
ARDL
NARDL</td><th rowspan="2">Cointegration</th><th>no</th><td>OLS – FD
ARDL
NARDL</td></tr>
<tr><th>yes</th><td>ECM
ECM – GH
OLS – FD
ARDL
NARDL</td></tr>
</table>

The OLS estimator is used for the linear regression model [1] for the 562 cases in which the Qratio and the Pratio are both stationary.

[21] One can try to transform the model by taking such order of differences in the timeseries as the order of integration.

Table 0-25 presents an overview of the statistical significance of the estimated elasticity of substitution and compares with the results of the previous study. In many cases the estimated elasticity of substitution is positive and statistically significant (44%). Only in a few cases (7%) was a statistically significant estimate found that had an opposite sign from that expected from economic theory. The results do not differ substantially from those of the EC (2017) study.

Table 0-25: Estimation of the elasticity of substitution between energy and gross value added on the cases that both timeseries are stationary by using the ADF test

	new	previous study	
	OLS	**OLS (FD)**	**FMOLS**
Estimated equations	562	407	407
Cases with positive sign on σ and significant results	44%	40%	31%
Cases with positive sign on σ and insignificant results	30%	46%	41%
Cases with negative sign on σ and insignificant results	19%	12%	19%
Cases with negative sign on σ and significant results	7%	2%	9%

Residual based tests have been used to check the validity of the results: (i) the Breusch-Godfrey serial correlation LM test has been used to test for serial correlation on the residuals, (ii) the White heteroskedasticity test has been used to test for heteroskedasticity on the residuals and (iii) the Jarque – Bera test has been used to test if the residuals follow the normal distribution.

Table 0-26 presents for each case presented in

Table 0-25 the percentage of cases that do not suffer for serial correlation (LM-test), for heteroskedasticity (White), for non-normality (JB) and for any of the three (All) problem in the residuals. The 77% of the cases with positive sign and significant estimate on σ pass all the residual based test implemented. In the cases that either serial correlation or heteroskedasticity have been found the feasible generalized least square (FGLS) estimator is used. The FGLS uses the OLS estimated variance-covariance matrix to correct for the serial correlation or/and the heteroskedasticity identified in the residuals.

Table 0-26: Residual based tests on the OLS estimations

	LM-test	White	JB	All
Cases with positive sign on σ and significant results	87%	88%	97%	77%
Cases with positive sign on σ and insignificant results	89%	85%	95%	72%
Cases with negative sign on σ and insignificant results	91%	89%	94%	75%
Cases with negative sign on σ and significant results	95%	87%	97%	82%

Cointegration tests that examine for a possible cointegration relationship between the Qratio and the Pratio have been implemented at the 360 cases that both timeseries have been found to be integrated at order one by the Zivot – Andrews unit root tests.

Table 0-27 presents the results of the Johansen and Juselius (1990) and the Engle and Granger (1987) cointegration tests. Among the 360 cases, that both Qratio and Pratio are integrated of order one, the 149 cases support a cointegration relationship by both cointegration tests. In 317 cases either the Engle – Granger cointegration test or the Johansen and Juselius cointegration test supports the existence of the cointegration relationship between the Qratio and the Pratio.

Table 0-27: Cointegration tests on cases that both Qratio and Pratio are integrated of order one

		Johansen and Juselius cointegration test		
		No	**Yes**	**Total**
Engle and Granger cointegration test	**No**	43	117	**160**
	Yes	51	149	**200**
	Total	**94**	**266**	**360**

Gregory, Nason and Watt (1994) demonstrate that the power of the ADF test falls when a structural break is present, so procedures for cointegration tests that include structural breaks need to be examined. Three models are considered:

1. Model A: a cointegration relationship with a deterministic constant and a shift in the constant
2. Model B: a cointegration relationship with a deterministic constant and trend and a shift in the constant
3. Model C: a cointegration relationship with a deterministic constant and a shift in constant and the elasticity.

Table 0-28 presents the number of cases by sector that have found a cointegration relationship between the Qratio and Pratio based on the models A, B and C of Gregory and Hansen (1996) cointegration test.

By allowing for a possible structural break of the form of the model A, model B or model C in the cointegration relationship the number of cases that support the existence of a cointegration relationship is 214.

Table 0-28: Cointegration test with possible structural break based on Gregory Hansen

	Number of cases that a cointegration exist by using either				Number of cases that Qratio and Pratio are I(1)
	A	B	C	A or B or C	
Agriculture, Hunting, Forestry and Fishing	3	5	3	6	11
Mining and Quarrying	3	3	3	4	10
Food, Beverages and Tobacco	4	4	2	5	8
Textiles and Textile Products	2	5	2	6	11
Leather, Leather and Footwear	3	5	3	7	11
Wood and Products of Wood and Cork	3	4	2	5	12
Pulp, Paper, Paper , Printing and Publishing	5	5	4	9	15
Coke, Refined Petroleum and Nuclear Fuel	2	4	5	7	9
Chemicals and Chemical Products	1	5	1	5	10
Rubber and Plastics	0	3	1	4	10
Other Non-Metallic Mineral	5	7	6	8	12
Basic Metals and Fabricated Metal	5	7	8	10	16
Machinery, Nec	5	3	4	7	14
Electrical and Optical Equipment	7	4	4	8	10
Transport Equipment	4	4	4	5	11
Manufacturing, Nec; Recycling	3	6	1	6	11
Electricity, Gas and Water Supply	7	11	10	14	16
Construction	4	6	5	7	11
Sale, Maintenance and Repair of Motor Vehicles and Motorcycles; Retail Sale of Fuel	3	3	3	5	6
Wholesale Trade and Commission Trade, Except of Motor Vehicles and Motorcycles	5	3	5	6	9
Retail Trade, Except of Motor Vehicles and Motorcycles; Repair of Household Goods	2	4	4	6	11
Hotels and Restaurants	4	7	4	8	9
Inland Transport	2	5	4	6	11
Water Transport	1	1	1	2	5
Air Transport	7	10	6	11	18
Other Supporting and Auxiliary Transport Activities; Activities of Travel Agencies	1	4	3	5	6
Post and Telecommunications	1	1	1	3	5
Financial Intermediation	3	3	2	4	10
Real Estate Activities	3	1	1	4	9
Renting of M&Eq and Other Business Activities	4	4	3	6	10
Public Admin and Defence; Compulsory Social Security	3	6	2	7	11
Education	6	4	3	7	10
Health and Social Work	4	6	3	7	13
Other Community, Social and Personal Services	2	2	2	4	9
ALL	117	155	115	214	360

Based on the cointegration tests (

Table 0-27 and

Table 0-28), the error correction model (ECM) and the error-correction model with a structural break in the cointegration relationship (ECM – GH) were applied to these 214 cases. In only 34 cases did we find evidence that the elasticity of substitution varies across time. In all other cases a structural break was for the constant term of the linear regression. This result indicates that the elasticity of substitution is stable across time but at some points in time there is a structural break that affects the level of the Qratio without affecting the level of the elasticity of substitution. Further analysis is needed to identify which factors could explain such structural breaks.

Table 0-29: Long-run elasticities of substitution by using the ECM model (sectoral overview)

	Max (1)	Min (1)	Median (1)	% of total (1)	% of total (2)	% of total (3)	# of cases
AtB	0.59	0.33	0.41	66.7	16.7	16.7	6
C	0.00	0.00	0.00	0.0	25.0	75.0	4
15t16	0.41	0.22	0.31	80.0	20.0	0.0	5
17t18	0.56	0.33	0.51	50.0	33.3	16.7	6
19	0.55	0.22	0.40	71.4	0.0	28.6	7
20	0.63	0.28	0.63	60.0	0.0	40.0	5
21t22	0.75	0.31	0.42	55.6	11.1	33.3	9
23	1.21	0.18	0.51	85.7	0.0	14.3	7
24	0.22	0.17	0.20	40.0	40.0	20.0	5
25	0.43	0.27	0.39	75.0	25.0	0.0	4
26	1.25	0.22	0.38	37.5	37.5	25.0	8
27t28	0.67	0.26	0.43	40.0	10.0	50.0	10
29	0.75	0.38	0.66	42.9	14.3	42.9	7
30t33	0.97	0.16	0.33	37.5	12.5	50.0	8
34t35	0.60	0.25	0.36	60.0	40.0	0.0	5
36t37	0.43	0.25	0.31	66.7	33.3	0.0	6
E	0.79	0.08	0.16	42.9	21.4	35.7	14
F	0.48	0.18	0.23	71.4	0.0	28.6	7
50	0.58	0.38	0.40	60.0	40.0	0.0	5
51	0.46	0.46	0.46	16.7	16.7	66.7	6
52	0.15	0.15	0.15	16.7	16.7	66.7	6
H	0.83	0.21	0.46	50.0	12.5	37.5	8
60	0.00	0.00	0.00	0.0	0.0	100.0	6
61	0.00	0.00	0.00	0.0	0.0	100.0	2
62	0.49	0.22	0.31	27.3	45.5	27.3	11
63	0.54	0.34	0.44	80.0	20.0	0.0	5
64	0.00	0.00	0.00	0.0	0.0	100.0	3
J	0.35	0.35	0.35	25.0	25.0	50.0	4
70	0.00	0.00	0.00	0.0	50.0	50.0	4
71t74	0.63	0.17	0.40	33.3	16.7	50.0	6
L	0.36	0.16	0.16	42.9	14.3	42.9	7
M	0.23	0.23	0.23	14.3	14.3	71.4	7
N	0.55	0.55	0.55	14.3	42.9	42.9	7
O	0.32	0.32	0.32	25.0	25.0	50.0	4
ALL	1.25	0.08	0.38	42.5	20.1	37.4	214

(1): The cases with a negative σ in equation [1]

(2): The cases with a positive but statistically insignificant σ in equation [1] at a 5% level of significance

(3): The cases with a positive and statistically significant σ in equation [1] at a 5% level of significance

Table 0-30: Long-run elasticities of substitution by using the ECM-GH model (sectoral overview)

	Max (1)	Min (1)	Median (1)	% of total (1)	% of total (2)	% of total (3)	# of cases
AtB	0.53	0.24	0.31	83.3	16.7	0.0	6
C	0.45	0.18	0.31	50.0	0.0	50.0	4
15t16	0.37	0.13	0.37	60.0	40.0	0.0	5
17t18	0.68	0.17	0.31	83.3	16.7	0.0	6
19	1.63	0.25	0.53	85.7	14.3	0.0	7
20	0.91	0.27	0.51	80.0	20.0	0.0	5
21t22	0.77	0.23	0.31	55.6	22.2	22.2	9
23	1.25	0.13	0.49	71.4	0.0	28.6	7
24	0.45	0.28	0.37	40.0	60.0	0.0	5
25	0.94	0.27	0.50	100.0	0.0	0.0	4
26	1.31	0.24	0.27	50.0	37.5	12.5	8
27t28	0.46	0.36	0.41	20.0	30.0	50.0	10
29	0.91	0.20	0.43	100.0	0.0	0.0	7
30t33	0.98	0.29	0.72	50.0	25.0	25.0	8
34t35	0.48	0.19	0.35	80.0	20.0	0.0	5
36t37	0.69	0.28	0.37	66.7	16.7	16.7	6
E	0.34	0.12	0.17	28.6	50.0	21.4	14
F	0.22	0.16	0.19	57.1	42.9	0.0	7
50	0.54	0.15	0.54	60.0	40.0	0.0	5
51	0.84	0.25	0.25	50.0	33.3	16.7	6
52	0.42	0.42	0.42	16.7	16.7	66.7	6
H	1.40	0.24	0.72	37.5	37.5	25.0	8
60	1.31	0.16	0.42	50.0	0.0	50.0	6
61	0.00	0.00	0.00	0.0	0.0	100.0	2
62	0.55	0.36	0.44	45.5	54.5	0.0	11
63	0.87	0.21	0.40	80.0	20.0	0.0	5
64	0.71	0.71	0.71	33.3	0.0	66.7	3
J	0.74	0.74	0.74	25.0	25.0	50.0	4
70	0.63	0.27	0.45	50.0	0.0	50.0	4
71t74	0.56	0.56	0.56	16.7	33.3	50.0	6
L	0.43	0.13	0.21	57.1	28.6	14.3	7
M	0.26	0.11	0.18	28.6	28.6	42.9	7
N	0.61	0.61	0.61	14.3	14.3	71.4	7
O	0.00	0.00	0.00	0.0	25.0	75.0	4
ALL	1.63	0.11	0.39	50.5	25.7	23.8	214

(1): The cases with a negative σ in equation [1]

(2): The cases with a positive but statistically insignificant σ in equation [1] at a 5% level of significance

(3): The cases with a positive and statistically significant σ in equation [1] at a 5% level of significance

In the timeseries analysis the OLS - FD, the ARDL and the NARDL models have been used in all the 1272 cases regardless the order of integration of the timeseries or if a cointegration relationship exists.

An overview by region and by sector for the OLS estimator using first differences of the timeseries is presented in

Table 0-31 and Table 0-32. In more than 80% of cases the elasticity of substitution has a positive sign which is consistent with economic theory. The results strongly support the hypothesis of weak substitutability between energy and the composite good of capital and labour. The results of the OLS-FD model should be treated with caution as the elasticity of substitution is estimated by a model which includes the first differences of the timeseries and do not correspond exactly to the initial model [1].

Table 0-31: Elasticity of substitution by using the OLS-FD model (region overview)

	Max (1)	Min (1)	Median (1)	% of total (1)	% of total (2)	% of total (3)
AUT	0.87	0.24	0.56	61.8	35.3	2.9
BEL	0.71	0.10	0.26	55.9	32.4	11.8
DEU	0.99	0.13	0.57	55.9	38.2	5.9
ESP	0.90	0.08	0.37	55.9	38.2	5.9
FRA	0.64	0.11	0.55	17.6	47.1	35.3
GBR	0.72	0.13	0.29	50.0	38.2	11.8
GRC	1.19	0.18	0.59	32.4	29.4	38.2
ITA	0.83	0.21	0.34	29.4	44.1	26.5
NLD	0.70	0.17	0.45	47.1	44.1	8.8
POL	0.70	0.11	0.62	20.6	61.8	17.6
PRT	0.77	0.23	0.39	58.8	38.2	2.9
ROU	0.85	0.10	0.41	67.6	29.4	2.9
AUS	0.88	0.13	0.50	61.8	35.3	2.9
BGR	0.43	0.08	0.14	20.6	29.4	50.0
BRA	0.82	0.15	0.19	23.5	64.7	11.8
CAN	1.07	0.14	0.58	26.5	61.8	11.8
CHN	0.86	0.27	0.57	67.6	26.5	5.9
CYP	0.99	0.32	0.54	38.2	50.0	11.8
CZE	0.82	0.28	0.49	26.5	50.0	23.5
DNK	0.92	0.15	0.29	35.3	35.3	29.4
EST	0.83	0.16	0.48	58.8	35.3	5.9
FIN	1.34	0.16	0.57	35.3	44.1	20.6
HUN	0.96	0.22	0.50	52.9	41.2	5.9
IDN	0.78	0.07	0.23	35.3	47.1	17.6
IND	1.03	0.19	0.64	64.7	20.6	14.7
IRL	0.75	0.17	0.33	29.4	47.1	23.5
JPN	0.85	0.10	0.33	41.2	47.1	11.8
KOR	0.85	0.08	0.36	52.9	29.4	17.6
LTU	0.91	0.12	0.49	32.4	50.0	17.6
LUX	1.00	0.21	0.49	44.1	35.3	20.6
LVA	1.21	0.17	0.48	35.3	38.2	26.5
MEX	0.56	0.08	0.23	41.2	52.9	5.9
MLT	2.32	0.44	0.91	73.5	11.8	14.7
RUS	0.22	0.13	0.17	14.7	55.9	29.4
SVK	0.65	0.20	0.45	29.4	52.9	17.6
SVN	1.02	0.07	0.75	85.3	11.8	2.9
SWE	0.56	0.11	0.33	32.4	55.9	11.8
TUR	0.84	0.08	0.41	38.2	44.1	17.6
USA	0.78	0.23	0.38	23.5	52.9	23.5
ALL	2.32	0.07	0.46	42.9	41.1	16.0

(1): The cases with a negative σ in equation [1]

(2): The cases with a positive but statistically insignificant σ in equation [1] at a 5% level of significance

(3): The cases with a positive and statistically significant σ in equation [1] at a 5% level of significance

Table 0-32: Elasticity of substitution by using the OLS-FD model (sectoral overview)

	Max (1)	Min (1)	Median (1)	% of total (1)	% of total (2)	% of total (3)
AtB	0.72	0.07	0.38	38.5	46.2	15.4
C	0.96	0.18	0.54	56.4	30.8	12.8
15t16	0.66	0.15	0.44	46.2	43.6	10.3
17t18	1.06	0.20	0.46	56.4	38.5	5.1
19	1.34	0.19	0.58	61.5	28.2	10.3
20	1.00	0.20	0.54	74.4	23.1	2.6
21t22	0.84	0.11	0.42	53.8	38.5	7.7
23	1.02	0.08	0.24	35.9	35.9	28.2
24	2.32	0.17	0.41	41.0	38.5	20.5
25	0.92	0.15	0.62	69.2	30.8	0.0
26	0.86	0.13	0.34	38.5	53.8	7.7
27t28	0.77	0.11	0.32	33.3	48.7	17.9
29	0.97	0.15	0.47	59.0	30.8	10.3
30t33	1.10	0.20	0.44	46.2	38.5	15.4
34t35	0.97	0.18	0.49	59.0	30.8	10.3
36t37	1.21	0.22	0.62	64.1	30.8	5.1
E	0.65	0.07	0.25	25.6	43.6	30.8
F	0.81	0.08	0.29	64.1	30.8	5.1
50	0.91	0.13	0.47	35.9	48.7	15.4
51	1.01	0.11	0.34	51.3	30.8	17.9
52	0.76	0.17	0.37	30.8	56.4	12.8
H	0.96	0.22	0.56	25.6	51.3	23.1
60	1.09	0.10	0.25	28.2	59.0	12.8
61	0.96	0.08	0.30	30.8	35.9	33.3
62	1.02	0.21	0.56	61.5	30.8	7.7
63	1.00	0.21	0.42	23.1	59.0	17.9
64	0.91	0.27	0.47	33.3	35.9	30.8
J	1.03	0.18	0.43	33.3	46.2	20.5
70	1.08	0.22	0.70	20.5	41.0	38.5
71t74	1.04	0.17	0.60	28.2	51.3	20.5
L	0.80	0.11	0.34	38.5	41.0	20.5
M	1.05	0.08	0.40	35.9	35.9	28.2
N	1.15	0.21	0.54	30.8	43.6	25.6
O	0.75	0.11	0.60	28.2	69.2	2.6
ALL	2.32	0.07	0.46	42.9	41.1	16.0

(1): The cases with a negative σ in equation [1]

(2): The cases with a positive but statistically insignificant σ in equation [1] at a 5% level of significance

(3): The cases with a positive and statistically significant σ in equation [1] at a 5% level of significance

The ARDL model is an extension of the model [1] with short-term dynamics. By including either one-year or two-year lag dynamics the model becomes:

$$\ln\frac{QE_t}{QKL_t} = a + \varphi \cdot t + \beta \ \ln\frac{PE_t}{PKL_t} + \sum_{j=1}^{m} \gamma_j \ln\frac{QE_{t-j}}{QKL_{t-j}} + \sum_{j=1}^{m} \delta_j \ln\frac{PE_{t-j}}{PKL_{t-j}} + u_t \qquad [2]$$

which can be written as

$$\Delta\left(\ln\frac{QE_t}{QKL_t}\right) = a + \varphi \cdot t + b \ \ln\frac{PE_{t-1}}{PKL_{t-1}} + c \ \ln\frac{QE_{t-1}}{QKL_{t-1}} + \sum_{j=1}^{m-1} c_j \, \Delta\left(\ln\frac{QE_{t-j}}{QKL_{t-j}}\right) + \sum_{j=1}^{m-1} d_j \ln\left(\frac{PE_{t-j}}{PKL_{t-j}}\right) + u_t \qquad [3]$$

Pesaran (2001) proposed a bound test to test if the estimated model [3] is valid for inference. Comparison of the critical values of Pesaran (2001) bound tests with the F-statistic of the null hypothesis: $H_0\colon b = c = 0$ identifies whether cointegration exists. Based on the ARDL, the Banjeree, Dolado and Mestre (1998) t-statistic bound test, provide even more statistical evidence that the selected model is valid for inference. That is a bound test on the t-statistic of the null hypothesis: $H_0\colon b = 0$.

From the 1272 cases examined, only 196 cases reject the null hypothesis in both bound tests.

Table 0-33 presents an overview of the estimates of the elasticity of substitution by using the ARDL model.

Table 0-33: Long-run elasticities of substitution by using the ARDL model (sectoral overview)

	Max (1)	Min (1)	Median (1)	% of total (1)	% of total (2)	% of total (3)	% of total (4)	# of cases
AtB	0.43	0.27	0.34	10.5	7.9	21.1	60.5	38
C	1.06	0.21	0.46	10.3	10.3	15.4	64.1	39
15t16	0.40	0.28	0.34	5.3	13.2	23.7	57.9	38
17t18	1.00	0.28	0.59	13.5	5.4	21.6	59.5	37
19	1.25	0.30	0.74	16.2	10.8	8.1	64.9	37
20	1.36	0.38	0.60	8.3	13.9	11.1	66.7	36
21t22	0.95	0.12	0.50	13.2	10.5	23.7	52.6	38
23	1.26	0.17	0.69	15.2	18.2	18.2	48.5	33
24	0.23	0.08	0.16	5.3	7.9	18.4	68.4	38
25	2.13	0.45	1.15	10.8	16.2	32.4	40.5	37
26	1.28	0.29	0.43	8.1	32.4	21.6	37.8	37
27t28	0.09	0.09	0.09	2.6	10.3	20.5	66.7	39
29	0.00	0.00	0.00	0.0	16.2	18.9	64.9	37
30t33	0.85	0.38	0.62	7.9	2.6	21.1	68.4	38
34t35	0.87	0.40	0.58	10.3	28.2	12.8	48.7	39
36t37	0.69	0.23	0.46	5.1	17.9	12.8	64.1	39
E	0.00	0.00	0.00	0.0	10.5	26.3	63.2	38
F	0.64	0.17	0.25	7.9	10.5	15.8	65.8	38
50	0.60	0.24	0.30	10.8	18.9	13.5	56.8	37
51	1.63	0.51	0.72	10.8	13.5	10.8	64.9	37
52	0.72	0.27	0.48	10.3	5.1	20.5	64.1	39
H	0.28	0.28	0.28	2.6	15.4	38.5	43.6	39
60	0.77	0.36	0.57	5.6	8.3	25.0	61.1	36
61	1.17	1.12	1.15	6.1	15.2	21.2	57.6	33
62	1.17	0.84	1.00	5.4	10.8	13.5	70.3	37
63	0.75	0.11	0.43	5.3	7.9	23.7	63.2	38
64	0.59	0.29	0.47	10.5	7.9	13.2	68.4	38
J	0.57	0.57	0.57	2.8	13.9	16.7	66.7	36
70	0.73	0.32	0.68	7.7	12.8	25.6	53.8	39
71t74	0.81	0.17	0.35	11.1	16.7	22.2	50.0	36
L	0.43	0.22	0.40	11.1	11.1	16.7	61.1	36
M	0.00	0.00	0.00	0.0	10.5	36.8	52.6	38
N	0.26	0.26	0.26	2.6	10.5	23.7	63.2	38
O	1.37	0.26	0.47	10.3	7.7	41.0	41.0	39
ALL	2.13	0.08	0.49	7.7	12.6	20.8	58.9	1272

(1): The cases with a negative b/c in equation [3]

(2): The cases with a positive but statistically insignificant b/c in equation [3] at a 5% level of significance

(3): The cases with a positive and statistically significant b/c in equation [3] at a 5% level of significance

(4): The cases that no cointegration is found based on bound tests of Pesaran (2001) and Banjeree (1998).

Table 0-34: Short-run elasticities of substitution by using the ARDL model (sectoral overview)

	Max (1)	Min (1)	Median (1)	% of total (1)	% of total (2)	% of total (3)	% of total (4)	# of cases
AtB	0.71	0.12	0.28	13.2	10.5	15.8	60.5	38
C	0.69	0.34	0.40	15.4	7.7	12.8	64.1	39
15t16	0.38	0.15	0.33	7.9	10.5	23.7	57.9	38
17t18	0.95	0.44	0.62	10.8	10.8	18.9	59.5	37
19	1.66	0.18	0.72	24.3	5.4	5.4	64.9	37
20	1.12	0.37	0.47	22.2	5.6	5.6	66.7	36
21t22	0.58	0.13	0.45	18.4	10.5	18.4	52.6	38
23	1.08	0.21	0.48	21.2	9.1	21.2	48.5	33
24	0.40	0.17	0.19	7.9	5.3	18.4	68.4	38
25	1.17	0.42	0.71	24.3	10.8	24.3	40.5	37
26	1.18	0.20	0.38	18.9	18.9	24.3	37.8	37
27t28	0.45	0.29	0.33	7.7	5.1	20.5	66.7	39
29	0.62	0.19	0.46	13.5	10.8	10.8	64.9	37
30t33	0.75	0.56	0.60	7.9	5.3	18.4	68.4	38
34t35	0.84	0.18	0.51	33.3	5.1	12.8	48.7	39
36t37	0.76	0.34	0.66	20.5	2.6	12.8	64.1	39
E	0.22	0.22	0.22	2.6	10.5	23.7	63.2	38
F	0.84	0.12	0.24	15.8	10.5	7.9	65.8	38
50	0.58	0.23	0.24	13.5	21.6	8.1	56.8	37
51	1.61	0.19	0.50	16.2	5.4	13.5	64.9	37
52	0.51	0.17	0.29	12.8	5.1	17.9	64.1	39
H	0.48	0.23	0.33	12.8	10.3	33.3	43.6	39
60	0.57	0.13	0.39	13.9	2.8	22.2	61.1	36
61	0.90	0.16	0.27	9.1	12.1	21.2	57.6	33
62	0.93	0.32	0.81	10.8	2.7	16.2	70.3	37
63	0.87	0.25	0.67	10.5	7.9	18.4	63.2	38
64	0.48	0.46	0.47	5.3	15.8	10.5	68.4	38
J	0.59	0.35	0.42	11.1	8.3	13.9	66.7	36
70	0.69	0.23	0.46	5.1	10.3	30.8	53.8	39
71t74	0.42	0.18	0.35	11.1	16.7	22.2	50.0	36
L	0.40	0.16	0.39	8.3	8.3	22.2	61.1	36
M	0.73	0.19	0.51	7.9	7.9	31.6	52.6	38
N	0.75	0.75	0.75	2.6	13.2	21.1	63.2	38
O	0.68	0.49	0.63	7.7	17.9	33.3	41.0	39
ALL	1.66	0.12	0.46	13.1	9.4	18.6	58.9	1272

(1): The cases with a negative b/c in equation [3]

(2): The cases with a positive but statistically insignificant b/c in equation [3] at a 5% level of significance

(3): The cases with a positive and statistically significant b/c in equation [3] at a 5% level of significance

(4): The cases that no cointegration is found based on bound tests of Pesaran (2001) and Banjeree (1998).

The process of convergence to the long-run equilibrium relationship may follow a different adjustment rate in cases in which the short-term dynamics lead to positive or negative change from the long-run equilibrium relationship. Shin and Yu (2014) proposed a bound test procedure developed by Pesaran (2001) in a non-linear ARDL framework in order to test for cointegration with asymmetric adjustment in the long-run equilibrium relationship. By using the NARDL model only 115 cases accepted the null hypothesis that an asymmetric non-linear cointegration relationship between the Qratio and Pratio exists.

Table 0-35 presents the 93 cases in which either the short run or long run elasticity of substitution of energy and gross value added from a positive or negative change in prices is statistically significance at a 5% level.

It is found that in many cases the elasticity of substitution is higher in periods when there is a decrease in the relative ratio of prices between energy and gross value added as compared to periods in which the relative ratio of prices was increasing.

Table 0-35: Short-run and long-run substitution of elasticity by using an NARDL model

Region	Sector	long-run positive	long-run negative	short-run positive	short-run negative
ITA	**15t16**	0.23	1.20	stat	1.13
RUS	**15t16**	0.44	sign	0.49	0.14
TUR	**15t16**	0.61	0.13	stat	0.41
FRA	**17t18**	0.69	0.35	0.74	sign
BGR	**17t18**	0.88	stat	stat	stat
BGR	**19**	0.89	stat	stat	stat
RUS	**19**	0.79	stat	stat	0.38
CHN	**20**	0.56	1.80	0.73	stat
LUX	**20**	1.02	1.42	1.00	1.13
BEL	**21t22**	stat	0.20	sign	0.39
ESP	**21t22**	stat	0.77	stat	1.16
TUR	**21t22**	0.62	stat	stat	stat
ESP	**23**	stat	0.62	sign	stat
CZE	**23**	0.54	0.90	0.33	0.70
IND	**24**	0.68	stat	stat	sign
LUX	**24**	0.89	2.68	stat	stat
DNK	**25**	0.65	1.02	0.47	0.90
PRT	**26**	0.16	1.06	stat	sign
BGR	**26**	0.61	sign	0.55	sign
IDN	**26**	0.58	0.39	1.04	stat
KOR	**26**	0.12	stat	0.44	sign
IRL	**27t28**	stat	sign	0.38	sign
RUS	**27t28**	0.31	stat	stat	0.23
FRA	**29**	0.24	sign	sign	sign
BRA	**29**	sign	stat	sign	0.27

Region	Sector	long-run positive	long-run negative	short-run positive	short-run negative
JPN	29	0.93	1.64	stat	4.20
SVK	29	0.56	stat	stat	1.05
SVN	29	stat	stat	0.82	sign
FRA	30t33	0.29	sign	sign	sign
ITA	30t33	0.24	0.94	sign	1.11
POL	34t35	0.52	sign	sign	0.36
ROU	34t35	1.16	0.26	1.19	stat
CZE	34t35	0.57	stat	0.57	stat
IRL	34t35	0.44	stat	stat	stat
SWE	34t35	0.31	sign	stat	1.45
TUR	34t35	0.94	0.31	0.71	0.39
AUT	36t37	stat	0.78	stat	0.82
ITA	36t37	stat	0.58	stat	0.65
SVK	36t37	0.44	sign	stat	sign
BEL	50	0.20	0.27	0.47	stat
AUS	50	stat	sign	0.44	sign
BGR	50	0.37	1.11	stat	stat
IDN	51	1.01	0.80	1.04	stat
SVN	51	stat	stat	0.86	stat
BRA	52	0.26	0.20	0.31	stat
IDN	52	0.97	0.78	1.06	stat
GBR	60	0.57	stat	stat	stat
IDN	60	0.17	0.41	stat	0.19
KOR	60	0.15	sign	sign	0.34
ITA	61	stat	0.52	sign	stat
DNK	61	stat	2.71	stat	1.50
IND	61	sign	stat	0.74	stat
USA	61	0.93	1.90	stat	stat
BEL	62	0.31	stat	0.31	sign
CZE	62	0.60	1.22	sign	1.50
TUR	62	1.00	0.28	stat	stat
POL	63	stat	0.58	stat	stat
JPN	63	stat	0.75	0.64	sign
MLT	63	0.85	1.32	0.95	1.47
EST	64	0.27	sign	stat	sign
LTU	64	0.23	stat	stat	sign
LVA	64	stat	4.18	sign	4.31
GRC	70	stat	stat	sign	2.35
NLD	70	stat	stat	0.61	sign
HUN	70	stat	1.17	1.27	sign
BRA	71t74	0.09	sign	0.10	sign
SVK	71t74	0.38	stat	stat	sign
FIN	AtB	0.23	0.55	sign	stat

Region	Sector	long-run positive	long-run negative	short-run positive	short-run negative
CZE	**C**	0.79	1.13	stat	1.38
BGR	**F**	stat	0.98	0.90	0.85
FIN	**F**	0.14	0.41	sign	0.42
IDN	**F**	0.40	0.33	0.53	stat
JPN	**F**	0.14	stat	0.19	sign
BGR	**H**	0.49	stat	stat	stat
DNK	**H**	sign	0.34	stat	sign
PRT	**J**	0.30	sign	0.52	sign
BRA	**J**	0.54	stat	stat	stat
DNK	**J**	0.64	stat	0.94	stat
FIN	**J**	1.33	3.07	sign	1.18
SVN	**J**	0.75	sign	0.68	sign
IDN	**L**	0.79	0.46	0.90	sign
LUX	**L**	0.41	1.32	sign	0.95
SVN	**L**	0.48	stat	1.14	sign
SWE	**L**	0.31	0.72	stat	0.44
BGR	**M**	stat	1.10	stat	0.74
IDN	**M**	sign	sign	sign	0.36
IRL	**M**	stat	0.46	stat	sign
JPN	**M**	0.80	0.52	stat	stat
LUX	**M**	0.30	1.28	stat	stat
DNK	**N**	stat	sign	0.60	sign
LUX	**N**	0.49	0.70	stat	0.70
IRL	**O**	stat	0.61	sign	sign
TUR	**O**	stat	0.27	sign	stat

*sign: The cases that the elasticity of substitution has a negative sign

*stat: The cases that the elasticity of substitution has a positive but insignificant sign

Panel data analysis

This section discusses the application of panel data analysis for the estimation of the elasticity of substitution between energy and value added. The unit roots test of Levin, Lin, Chu (2002) and Im, Pesaran, Shin (2003), apply only in cases in which the panels are balanced[22], and these are presented in

[22] Sectors 23, 61, 62 and L have unbalanced data

Table 0-36 and Table 0-37. Since both Qratio and Pratio can be considered as stationary variables the OLS estimator is used.

Table 0-36: p-values of the panel unit root tests on Qratio

	LLC[1]	IPS[1]	LLC[2]	IPS[2]		LLC[1]	IPS[1]	LLC[2]	IPS[2]
AtB	0.000	0.014	0.000	0.000	**F**	0.002	0.033	0.000	0.001
C	0.002	0.095	0.000	0.036	**50**	0.000	0.002	0.000	0.000
15t16	0.001	0.084	0.000	0.000	**51**	0.000	0.624	0.000	0.000
17t18	0.004	0.893	0.002	0.114	**52**	0.000	0.000	0.000	0.000
19	0.014	0.555	0.000	0.074	**H**	0.000	0.000	0.000	0.000
20	0.005	0.362	0.001	0.036	**60**	0.036	0.590	0.000	0.005
21t22	0.124	0.510	0.000	0.006	**61**	n/a	n/a	n/a	n/a
23	n/a	n/a	n/a	n/a	**62**	n/a	n/a	n/a	n/a
24	0.000	0.176	0.000	0.000	**63**	0.000	0.153	0.000	0.013
25	0.001	0.043	0.000	0.009	**64**	0.012	0.813	0.000	0.000
26	0.000	0.060	0.000	0.000	**J**	0.001	0.523	0.007	0.026
27t28	0.266	0.738	0.004	0.150	**70**	0.000	0.004	0.000	0.000
29	0.000	0.293	0.000	0.016	**71t74**	0.000	0.064	0.000	0.000
30t33	0.000	0.359	0.000	0.003	**L**	n/a	n/a	n/a	n/a
34t35	0.000	0.023	0.018	0.017	**M**	0.000	0.001	0.000	0.030
36t37	0.000	0.027	0.000	0.000	**N**	0.002	0.135	0.000	0.033
E	0.000	0.000	0.000	0.299	**O**	0.003	0.001	0.000	0.000

[1] Unit root test with a deterministic constant
[2] Unit root test with a deterministic constant and trend

Table 0-37: p-values of the panel unit root tests on Pratio

	LLC[1]	IPS[1]	LLC[2]	IPS[2]		LLC[1]	IPS[1]	LLC[2]	IPS[2]
AtB	0.111	0.733	0.000	0.000	**F**	0.000	0.002	0.000	0.011
C	0.000	0.117	0.001	0.018	**50**	0.000	0.182	0.000	0.012
15t16	0.000	0.306	0.000	0.000	**51**	0.009	0.439	0.000	0.000
17t18	0.014	0.584	0.000	0.029	**52**	0.000	0.117	0.000	0.060
19	0.158	0.370	0.000	0.005	**H**	0.000	0.000	0.000	0.694
20	0.000	0.056	0.000	0.037	**60**	0.008	0.445	0.004	0.046
21t22	0.863	0.832	0.000	0.128	**61**	n/a	n/a	n/a	n/a
23	n/a	n/a	n/a	n/a	**62**	n/a	n/a	n/a	n/a
24	0.004	0.990	0.000	0.001	**63**	0.267	0.467	0.000	0.000
25	0.000	0.630	0.000	0.000	**64**	0.317	0.998	0.000	0.000
26	0.002	0.241	0.000	0.019	**J**	0.000	0.013	0.000	0.017
27t28	0.000	0.736	0.000	0.145	**70**	0.000	0.001	0.000	0.032
29	0.004	0.786	0.000	0.000	**71t74**	0.000	0.000	0.000	0.013
30t33	0.009	0.888	0.000	0.004	**L**	n/a	n/a	n/a	n/a
34t35	0.000	0.440	0.000	0.044	**M**	0.000	0.000	0.000	0.029
36t37	0.000	0.071	0.000	0.000	**N**	0.000	0.000	0.000	0.001
E	0.086	0.391	0.000	0.000	**O**	0.000	0.000	0.000	0.039

[1] Unit root test with a deterministic constant
[2] Unit root test with a deterministic constant and trend

The Breusch-Pagan LM test, the Pesaran scaled LM test and the Pesaran cross dependence test have been used to test for cross-section dependence on the residuals of the OLS estimations. It is found that in all cases the residuals are cross dependent and therefore the covariance matrix with panel corrected standard errors is more valid to be used for inference.

To estimate the elasticity of substitution of energy and gross value added the following steps have been followed:

Step 1: Estimate the model with panel least squares.

Step 2: Test by using a LM test if the random effects should be included.

Step 3: Estimate the model with fixed effect approach that allows the intercept to vary across countries or/and time period.

Step 4: Test by using an F-test if the fixed effects should be excluded by the model.

Step 5: If the null hypothesis of Step 2 and Step 4 is rejected, the Hausman test is used to select between the fixed or random effects. If the null hypothesis of the Hausman test is rejected then the model with random effects is selected, otherwise the model with fixed effects is used.

Table 0-38 presents three different estimates of the substitution of elasticity between energy and gross value added by using the OLS (column 2), random effects (column 3) and fixed effects (column 4). The econometric results are compared with the corresponding estimates of Antoszewski (2019) (column 1). By using the p-value of the Hausman test (column 5) the model selected among the three different specifications is presented with bold numbers in

Table 0-38. All the estimates are statistically signigicant at 1% level and support the hypothesis of weak substitutability between energy and gross value added.

Table 0-38: Panel data estimates by sector of the elasticity of substitution between energy and gross value added

	(1)	(2)	(3)	(4)	HT
AtB	0.212	0.217	0.152	0.155	0.687
C	0.211	0.359	0.322	0.325	0.709
15t16	0.390	0.433	0.190	0.201	0.225
17t18	0.498	0.448	0.272	0.288	0.291
19	0.318	0.476	0.252	0.275	0.145
20	0.468	0.586	0.380	0.400	1.000
21t22	0.318	0.601	0.200	0.221	1.000
23	0.493	n/a	n/a	n/a	n/a
24	0.252	0.663	0.252	0.265	0.193
25	0.638	0.481	0.415	0.421	0.716
26	0.522	0.536	0.256	0.272	0.205
27t28	0.346	0.720	0.218	0.260	0.000
29	0.724	0.557	0.512	0.515	0.897
30t33	0.786	0.655	0.531	0.543	0.693
34t35	0.458	0.399	0.439	0.436	0.858
36t37	0.627	0.504	0.427	0.434	0.338
E	0.084	0.343	0.070	0.081	1.000
F	0.343	0.382	0.097	0.107	1.000
50	0.378	0.493	0.297	0.308	0.362
51	0.396	0.349	0.301	0.305	0.804
52	0.359	0.284	0.127	0.137	1.000
H	0.430	0.481	0.117	0.137	0.007
60	0.036	0.350	0.076	0.087	1.000
61	0.433	n/a	n/a	n/a	n/a
62	0.248	n/a	n/a	n/a	n/a
63	0.258	0.351	0.095	0.112	1.000
64	0.573	0.527	0.352	0.368	0.429
J	0.398	0.581	0.181	0.207	0.142
70	0.181	0.543	0.088	0.097	0.069
71t74	0.189	0.429	0.177	0.187	0.370
L	0.262	n/a	n/a	n/a	n/a
M	−0.2250	0.475	0.139	0.146	0.125
N	0.303	0.549	0.188	0.200	0.144
O	0.306	0.488	0.185	0.195	1.000

(1): Antoszewski (2019) σ(kle) estimates

(2): Estimates by using the OLS estimator with panel corrected standard errors

(3): Estimates by using the OLS random effects estimator with panel corrected standard errors

(4): Estimates by using the OLS fixed effects estimator with panel corrected standard errors

(5): Hausman test

3.2.3. Summary and conclusion

A wide-range of econometric methods have been used to estimate the elasticity of substitution between energy and gross value added by region and by sector. Both timeseries and panel data analysis strongly support the hypothesis of weak substitutability between energy and gross value added.

By applying the Zivot – Andrews test and the Gregory and Hansen cointegration test it is found that in most cases there is statistical evidences to support the existence of a structural break for the constant term of the econometric equation [1]. This result indicates that possibly there are other factors, not only prices, that affect the ratio of energy to gross value added and that a consequent change in the level of the energy intensity to gross value added occurs at a point in time.

The estimates of the error correction and the ARDL models that allow to differentiate between the short-run and long-run elasticity of substitution of energy and gross value added do not provide statistical evidence that the long run elasticity can be considered to be greater than the short run elasticity. The elasticity of substitution between energy and gross value added remains stable across time.

Finally, by using the NARDL model that allows to examine for non-linearities in the elasticity of substitution between the energy and gross value added, it is found that in some cases there is an asymmetric adjustment to the long-run equilibrium which is more intense in periods in which there is a decrease in the relative ratio or prices of energy to gross value added.

4. Clean energy technologies and economic growth

4.1. R&D multiplier and leverage effects: Literature review for selected clean energy technologies

4.1.1. Introduction and summary of findings

This report reviews the empirical literature on learning rates for clean energy technologies and recommends rates to be used in updated versions of E3ME and GEM-E3.

We review the relevant, evidence-based literature on the estimation of the learning-by-doing and learning-by-research rates and report the latest available learning rates for a selection of clean energy technologies: photovoltaics, wind turbines, biomass and biofuels, electric vehicles and batteries, hydrogen production by electrolysis, hydroelectricity, synthetic fuels, direct air capture technologies, biomass and traditional carbon capture and storage technologies; as well as for a few conventional technologies in energy (natural gas and nuclear) and some further key basic industries (chemicals and electronic goods). Based on this review of academic papers and those published by industry experts or associations, we suaggest a set of learning-by-doing rates to use in updated versions of E3ME and GEM-E3, in particular to inform the modelling carried out in the 'Study on the Macroeconomics of the Energy Union' for the European Commission Directorate-General for Energy.

There are a number of key dimensions influencing how we arrive from the rates cited in the literature to those suggested to be used in the modelling exercise for the time period 2015-2050. The first one is to identify the most recently-cited rates for the selected technologies to make sure that the rates adopted for modelling are up to date. The second is to examine whether there is any trend in estimated rates over time (which could be interpreted as meaning that the technology is reaching a more mature phase in which the pace of cost reduction might be decelerating, or which could raise the rate of cost reduction in the light of more recent experience). The third is to review the experience of selected older technologies to see whether anything can be learned with regard to the time profile of learning rates for more mature technologies. We summarize the key findings by the identified dimensions below.

Updating of the estimated rates

An important implication of the review is that for some of the clean energy technologies, more recent literature that has been reviewed seem to project slightly higher learning-by-doing rates than those suggested by (International Energy Agency, 2015) (especially regarding shorter-run projections). Since (International Energy Agency, 2015) was a key source for the learning rates currently used in the models, the implication is a small increase in the rates used in the models, at least for the first part of the period to 2050. Notably, these technologies were the following: photovoltaics, offshore wind turbines, biomass and biofuels, and electric vehicles and batteries - for these technologies the application of the relatively higher learning rate assumptions is suggested.

Trend in estimated rates over time

Importantly, a key goal with including conventional technologies (such as natural gas or nuclear) was to illustrate that while learning-by-doing rates are projected to decrease with the development of certain technologies, they can also be seen to persist over a relatively long time period. This implies that the same can apply for relatively newer clean energy technologies too and has been carried forward into the rates recommended for use in the models.

Slower learning rates in the longer run

For the cases where it has been evidenced, different learning-by-doing rates are suggested for the first part of the period to 2030 and for the longer run. Photovoltaics, onshore and offshore wind turbines, biomass and biofuels, and electric vehicles and batteries are the technologies that are expected to experience slower (although in some cases, still relatively high) learning rates for the period 2030-2050 than over 2015-2030.

Compared to the learning-by-doing rates, the literature review has provided much thinner evidence on learning-by-research rates (that is, on the rate of decrease in unit costs for every doubling of research and development investment stock) of the selected clean energy technologies. This does not provide a strong basis for specific recommendations on the learning-by-research rate to use for each technology, but some broad observations can be made. For the most part, the reported learning-by-research rates are as high as the learning-by-doing rates for the same technology and in some cases higher: notably, three studies on wind turbines found higher learning-by-research rates than the learning-by-doing rates for that technology. A similar outcome was found for hydroelectricity and natural gas combined cycle turbines, and also, in the case of one research article, for electric vehicles.

Table 0-39 presents the suggested learning-by-doing rates to use for the modelling of cost development of clean energy technologies over time.

Table 0-39 Recommended learning-by-doing rates to use for the selected clean energy technologies

Clean energy technology	Recommended learning rate to use for the years 2015-2030	Recommended learning rate to use for the years 2031-2050
PV	20%	17%
Wind turbines - Onshore	7%	5%
Wind turbines – Offshore	11%	9%
Biomass and biofuels	7.8%	5%
Electric vehicles and batteries	18%	15%
Hydrogen production by electrolysis	7%	7%
Hydroelectricity	1.4%	1%
Synthetic fuels (power to X)	13%	13%
DAC (Direct Air Capture)	6%	6%
Biomass CCS and traditional CCS	7%	5%
Conventional technologies	3%	3%
Selected other sectors	18%	18%

Note: Learning rates are defined per doubling of cumulative output (learning-by-doing). Conventional technologies cover natural gas and nuclear. Selected other sectors cover chemicals and electronic goods.

Source: Reviewed by Cambridge Econometrics, 2019.

4.1.2. Literature findings on the theoretical background of learning rates in clean energy technologies

The distinction between learning-by-research and learning-by-doing

In the area of technological development research, learning curve theory refers to the assumption of an exponential correlation between the unit cost and the cumulative production of a specific technology (Heuberger, Rubin, Staffell, Shah, & Dowell, 2018)[23]. The simplest one-factor learning curve model's core assumption is that for each doubling of the cumulative capacity installed (reflecting a doubling of 'production experience'), the unit cost is reduced by a factor x (often denoted by LR: the learning rate). As an extension to this one-factor model, other learning effects (e.g. learning attributed to research efforts) can be included in the equation, resulting in a so-called two-factor learning curve model[24].

Although the two-factor learning curve model is likely to offer greater accuracy and a more explicit representation of R&D-driven technological advances, its applicability is often limited by data availability (Heuberger, Rubin, Staffell, Shah, & Dowell, 2018). Furthermore, modelling problems stem from the uncertainty associated with R&D processes (high risks regarding their outcomes), and from the potential to develop new products which the 2FLC model cannot capture (i.e., some disruptive innovation products, the learning process of which is often not traceable and/or the learning curve of which is most likely non-linear).

We distinguish two learning rates in this review: the learning-by-doing rate and learning-by-research rate, as per the definition given in (E3-Modelling - Cambridge Econometrics, 2017):

- **Learning-by-doing rate**: the rate of decrease of unit costs for every doubling of cumulative output;
- **Learning-by-research rate**: the rate of decrease of unit costs for every doubling of R&D (research and development) stock.

Albeit some literature suggests considering further specifications within the typology of learning dynamics[25], in the context of this paper learning-by-doing is considered to encompass all relevant cost-influencing factors[26], such as changes in the specific components of the production process (e.g. technical innovation, up-sizing or increase in labour productivity), in the product itself (e.g. a re-design of product with technology remaining the same) or in input prices of labour or materials.

Learning effects in the form of learning-by-research or learning-by-doing are apparent in all technologies, but the relevance and outcomes of learning-by-research vs. learning-by-doing, and the relative importance of public vs. private sector involvement in R&D, vary greatly across technologies, depending on: how mature the technology is, what industry the technology is predominantly associated with, and what is the geographical coverage of investigation. The role of public support for R&D is closely related to the state of the product cycle: (Popp, 2019) synthesized earlier research and concluded that direct public R&D or public support for the private sector research are less important for more mature

[23] The notion of learning-by-doing and the approach to estimate the rate were first explored in the pioneering work of (Arrow, 1962).

[24] For one of the first notable works done on the two-factor learning curve (2FLCs) see: (Kouvaritakis, Soria, & Isoard, 2000).

[25] Besides learning-by-doing and learning-by-research, the notable works of (Junginger, et al., 2006) or (Kahouli-Brahmi, 2008) suggest rates for learning-by-using (induced by customer feedback), learning-by-interacting (spill-over effects across technologies) and learning-by-economies of scale (mostly induced by supply chain improvements).

[26] Listed factors are based on (European Commission, 2018).

technologies, but are needed to foster development of emerging, less mature and more advanced technologies.

Impact of public and private R&D in the innovation cycle

The various motivations for increasing the deployment of viable clean energy technologies[27] include: reducing greenhouse gas emissions, mitigating the negative impacts of climate change, and reducing (potentially politically costly) dependence on foreign sources of energy (Boulatoff & Boyer, 2016). The question of how these energy systems are designed and financed is a crucial one to address if one seeks to build a sustainable pathway in support of clean technology financing. While governments and the corporate sphere have different motivations in financing clean projects and investing in clean energy R&D processes, clean energy technology firms are expected to be more successful on average in countries where intense R&D is undertaken by both businesses and the government (Boulatoff & Boyer, 2016, p. 148). The authors examine how clean technology industries performed over a 10-year period between 2004 and 2014, and assess how government involvement (in terms of R&D investments) in specific clean industries affects the cost of doing research in clean technology and the performance of businesses within the industry. The authors investigate whether public R&D functions as a *complement* to private sector R&D or as a *substitute* to it (crowding out private R&D). The authors refer to the findings of (Zúñiga-Vicente, Alonso-Borrego, Forcadell, & Galán, 2012) and conclude that considerable complementarity seems to exist between public and private R&D (although the estimates vary depending on the country considered and the methodology used).

Public and private R&D processes differ in their key channels of diffusion. Private R&D typically yields patents that can form the basis of substantial royalty revenues (depending on the maturity of market and its structure, these are likely to be monopoly rents) and so provides the innovator with the prospect of sustained early-mover advantages. Hence, the key channels of diffusion for the innovative product / technology are largely market-based. This is much less true in the case of public R&D, where the benefits resulting from the R&D investments are much more diffuse, and in the majority of cases, accrue to the users of the arising products and technological innovations (and their customers).

(Miremadi, Saboohi, & Arasti, 2019) discuss the role of public R&D and knowledge spillovers in the development of renewable energy technologies. The authors model renewable energy technology knowledge flows as a function of public R&D expenditure. The core argument as to why public R&D investment plays a crucial role in clean technologies development is as follows. First, support for the *absorption of international knowledge spillovers* is key in driving the clean energy transition in the domestic market, which requires the expansion of domestic absorptive capacity: existence and operation of scientific bodies, laboratories, scale and application of public R&D or industrial policies. Second, the development of a new technology is a lengthy and uncertain process and, at least in the early stages of development, the uncertainty (and the associated level of risk) may be so high as to discourage most private sector investors. A key role for government is, therefore, to overcome this barrier in the process of resource allocation in clean energy innovations and to fill the *finance gap* (Miremadi, Saboohi, & Arasti, 2019, p. 2).

From a conventional innovation cycle point of view, private R&D spending is assumed to be closer to the commercialisation stage of technological development, suggesting that

[27] In accordance with a broad and conventionally accepted definition, clean technology firms are considered to include those providing products and services that harness renewable materials and renewable energy sources, reduce the use of natural resources, or eliminate or cut GHG emissions and waste (Boulatoff & Boyer, 2016, p. 148).

private R&D spending will show a larger reduction in unit cost per euro spent than public R&D spending. (Veugelers, 2016) reviewed the evidence on the impact of public R&D spending on rates of return in R&D projects. This review makes the case that, particularly in the area of clean energy innovations, market mechanisms will not provide the socially optimal amount of innovation due to a combination of two market failures: environmental externalities (the social benefits of reduced pollution are much larger than any private benefits) and knowledge externalities (in cases where at least some of the knowledge is non-rival and non-excludable). Partly for this reason, and supported by the findings of research on patent citations in the SIMPATIC project[28] (Veugelers, 2016, p. 4), the suggestion is that R&D for clean technologies should receive higher public support (subsidies to the private sector) in preference to R&D in technologies that do not have an associated environmental benefit.

Estimating historical learning rates and making future projections

This section discusses the main kinds of approaches that have been followed to estimate historical learning rates and to make projections into the future.

Learning/experience curve approach to estimate historical learning rates

The key assumption of the conventionally accepted and widely used learning-by-doing theory builds on the observation that historical cost reductions in certain technologies have been correlated with the increase in their cumulative production or installed capacity (European Commission, 2018, p. 4). This learning-by-doing (or experience curve) approach is widely used (European Commission, 2018) (Louwen, Junginger, & Krishnan, 2018) in research investigating various aspects of the energy transition, and while the basic steps of the approach are relatively easy to develop, the robustness of historical learning estimations will largely depend on the quality of empirical data used.

The most widely-used formulation of the one-factor learning-by doing approach (the experience curve approach) describes a log-linear relationship between the unit cost of the technology and its cumulative output or installed capacity. This relationship can be transformed into the fraction of cost reduction resulting from a doubling of cumulative capacity (or production), and in its simplest form, can be expressed in the following logarithmic form:

$$\log C(cum) = \log C_1 + b * \log cum$$

Where $C(cum)$ is the cost C of the product at cumulative production *cum*, C_1 is the cost of producing the first unit, and b is the so-called experience parameter (Louwen, Junginger, & Krishnan, 2018). The experience curve parameter b represents the slope of the linear representation of the experience curve (which can be plotted in a double-logarithmic graph). The slope of this line indicates the rate at which the cost of a technology decreases. Using the experience parameter b, the learning rate (LR) can be expressed as:

$$LR = 1 - 2^b$$

The interpretation is that, for example, at a learning rate of 10% every doubling of *cumulative* production decreases the cost of an additional unit of product by 10%.

[28] SIMPATIC was a European Commission research project focusing on the impact assessment of research and innovation policies across Europe. For more information visit: https://cordis.europa.eu/project/rcn/102503/factsheet/en

The REFLEX project[29] that was part of the Horizon 2020 Work Programme "Secure, clean and efficient energy" of the EU and which ended in May 2019, aimed at modelling the dynamics and impacts of the transition of the European energy system. The learning rate estimation approach in the project proceeded as follows (Junginger, 2018, p. 4):

- collecting empirical data on installed capacity and the evolution of costs for the specific technology;
- revising/updating experience curves (which, as described above, are normally expressed as a log linear equation describing the reduction of total product cost as a function of cumulative production of that product), for the specific technology;
- incorporating experience curves into an energy model to allow for exogenous modelling of technological development and cost reductions;
- taking into account statistical uncertainty of the updated learning rates;
- decomposing the learning curve by cost parameters (depending on data availability): to reflect to what extent lower costs are driven by, e.g. certain input prices.

The rate of learning-by-doing for an entire technology may actually be dominated by learning for one or some of the individual components only. This creates the rationale for a component-based learning curve approach: as different components of a more complex technology, such as a coal-fired plant, are at different stages of maturity, further growth in cumulative production capacity will likely to result in a more rapid decrease in the cost of a newer component (in case of a coal-fired plant, for example, a carbon capture unit) than in the cost of a more mature component with the same initial learning rate but with a much larger base of initial installed capacity (Rubin, Azevedo, Jaramillo, & Yeh, 2015). (Ferioli, Schoots, & van der Zwaan, 2009) investigated the possibility of combining historical learning curves for single product (technology) components to derive an aggregate learning curve for the total product. This aggregated cost reduction impact of individual technology components, often used in the case of power stations or wind turbines, is also called as compound learning (Heuberger, Rubin, Staffell, Shah, & Dowell, 2018).

Learning/experience curve approach to make future projections

The most common form of estimating future cost reductions for particular clean technologies is to take the historically derived learning rates and combine those with future growth projections for that technology (European Commission, 2018). The European Commission's 2018 publication provided projections for the cost evolution of selected low-carbon energy technologies based on a range of reviewed studies that researched historical learning rates.

The UK Energy and Research Centre's 2013 report (Gross, et al., 2013) suggests there are two principal (and complementary) approaches for the estimation of future cost trends of clean technologies. The first one is engineering assessment (including parametric modelling) and the second is the learning-by-doing/experience curve approach. The experience curve approach is hard to apply in the absence of historical cost and deployment data, thus, the engineering assessment approach is considered to be more appropriate for early-stage technologies with limited market exposure (Gross, et al., 2013, p. 2). Another advantage of the engineering assessment approach, in the case where the drivers of cost reduction are considered 'parametrically' (by breaking down the costs of a technology into a set of sub-parameters), is that it is possible to test for the sensitivity of

[29] http://reflex-project.eu/

certain key cost parts. The report highlights some key complexities associated with the learning(-by-doing) approach, including: whether learning rates vary through time and as technologies mature; whether there is likely to be a ‘cost floor’ in the future development of any particular clean technology; and whether there is too much uncertainty to project the future deployment (and hence the cost) of a specific clean technology.

A 2016 review of wind energy costs and cost drivers (Wiser, et al., 2016) highlights a general critique of the use of learning curves in assessing future cost trajectories, namely that they do not adequately capture the many causal mechanisms that can bring about cost reduction. In particular, the authors note that in case of wind energy, very few published studies focus on the most relevant metric of wind energy costs (the levelized cost of electricity - LCOE) focusing instead on just one element of the LCOE, the upfront capital costs of deployment. The authors conclude that the combination of the learning curve approach with the engineering assessment approach, ideally with detailed modelling of specific possible technology advances, can provide technologically sophisticated information on both cost and performance. The authors also recommend that additional expert knowledge be gathered to inform estimates of future cost trajectories. Such knowledge can be developed through formal expert elicitation procedures (whereby uncertain quantities are estimated based on careful assessment of the knowledge and beliefs of experts about those quantities), interviews and workshops.

4.1.3. Development of learning rates over time for selected clean energy technologies

In light of the reviewed literature, a key observation is that while a wide set of factors can influence the learning rates of certain clean technologies, these factors can be grouped under two broad categories: *time* and *industry specificities.*

The *time factor* is important in three ways. Estimates of learning rates vary depending on:

1) when the reviewed piece of work / research has been done (is it relatively old or more recent?);
2) the length of the time period of historical data used in the estimation of the learning rate; and
3) (related to the first point), whether the cost projection is for an already-mature technology or an early-stage one.

Industry specificities are also important, such as the existence/ or absence of certain technology policy measures, industry expectations about returns and innovation benefits, and the nature of competition in the industry.

Early stage technologies are often characterized by higher learning rates, but at the same time they are surrounded by higher uncertainties about the future evolution of costs and profits, thereby increasing the risk of making no profit or even losses on the R&D investment.

In this chapter we summarize findings from a literature review of learning rates (both learning-by-doing and learning-by-research rates) that includes academic research and industry reports and publications.

Photovoltaics

In the case of solar PVs, the total system capital cost is the sum of the PV module cost and the balance-of-system (BOS) cost (including electrical installation, inverters, wiring and power electronics) (Rubin, Azevedo, Jaramillo, & Yeh, 2015). Most of the learning curve literature reviewed and presented below focuses on the PV module cost, but there are some examples of estimates for the cost trajectories of BOS or equipment parts of the solar PV system.

Table 0-40 Learning rates for Photovoltaics based on reviewed literature

Clean energy technology	Type of learning rate (by doing / by research)	Estimated decrease in unit cost (estimated learning rates)	Years covered in study / across the studies	Assumptions applied for global installed capacity	Source
PV	by doing	20%	various	2020: 424GW; 2030: 643GW; 2050: 1464GW	European Commission (2018)
PV	by doing	18.6% ± 1%	various	2017: 401 GW, 2030: 2346 GW	Louwen et al. (2018)
PV	by doing	17%	various	n.a.	E3-Modelling - CE (2017)
PV	by doing	22.5%	various	n.a.	ITRPV - VDMA (2017)
PV	by doing	23%	1980-2015	n.a.	Fraunhofer (2016)
PV	by doing	14% // 18% // 32% (min // mean // max)	1959–2011	n.a.	Rubin et al. (2015)
PV	by doing	19% - 23% (long-term); 10% (short-term)	various	10% until reaching 5000 GW global capacity, 19-23% afterwards	Agora Energiewende (2015)
PV	by doing	21%	various	n.a.	Nemet - Husmann (2012)
PV	by doing	29.2% // 19.6%	2006-2012 // 1980-2012	n.a.	Fraunhofer Institute (2012)
PV	by doing	19%	various	n.a.	Swanson (2006)
PV	by doing	18.4%	various	n.a.	Kobos et al. (2006)
PV	by doing	20%	various	n.a.	Criqui et al. (2000)
PV	by research	12%	various	n.a.	E3-Modelling - CE (2017)
PV	by research	10% // 12% // 14.3% (min // mean // max)	1959–2011	n.a.	Rubin et al. (2015)
PV	by research	14.3%	various	n.a.	Kobos et al. (2006)
PV	by research	10%	various	n.a.	Criqui et al. (2000)
PV BOS	by doing	12.9% ± 1.7%	various	2017: 401 GW, 2030: 2346 GW	Louwen et al. (2018)
PV modules	by doing	21.4% ± 0.8%	various	2017: 401 GW, 2030: 2346 GW	Louwen et al. (2018)

Note: Learning rates are defined per doubling of cumulative output (learning-by-doing) or per doubling of the knowledge stock (learning-by-research). Where available, estimation error intervals are indicated for the estimated learning rates, denoted by ± x%, or rate ranges and mean, denoted by (min/mean/max). "Various" years covered: where the referred literature derived and indicated a ***best estimate*** *of the learning rate based on various studies and time periods covered. The references are sorted by type of technology, type of learning rate and then date of publication (most recent first).*

Source: Reviewed by Cambridge Econometrics, 2019.

Wind turbines (onshore, offshore, non-categorized)

Recent research (Ryberg, et al., 2019) suggests that an increased rollout of onshore wind turbines across Europe could technically provide the continent with more than 10 times its current electricity needs[30]. Offshore wind farms have also been increasingly deployed in northern Europe, and projections again point to continued growth (Rubin, Azevedo, Jaramillo, & Yeh, 2015). In the following we present learning rates suggested by the literature. Wherever it was well-specified in the reviewed literature, we indicate the *type* of wind turbine (*onshore, offshore*); the others are listed as *undefined*.

Table 0-41 Learning rates for Wind turbine (onshore) based on reviewed literature

Clean energy technology	Type of learning rate (by doing / by research)	Estimated decrease in unit cost (estimated learning rates)	Years covered in study / across the studies	Assumptions applied for global installed capacity	Source
Wind turbine (onshore)	by doing	5%	various	2020: 606GW; 2030: 927GW; 2050: 1521GW	EC (2018)
Wind turbine (onshore)	by doing	5.9%	various	n.a.	Louwen et al. (2018)
Wind turbine (onshore)	by doing	7%	various	n.a.	International Energy Agency (2015)
Wind turbine (onshore)	by doing	3.1% / 9.6% / 13.1% (min/mean/max)	1979–2010	n.a.	Rubin et al. (2015)
Wind turbine (onshore)	by research	16.5%	1979–2010	n.a.	Rubin et al. (2015)

Note: Learning rates are defined per doubling of cumulative output (learning-by-doing) or per doubling of the knowledge stock (learning-by-research). Where available, estimation error intervals are indicated for the estimated learning rates, denoted by ± x%, or rate ranges and mean, denoted by (min/mean/max). "Various" years covered: where the referred literature derived and indicated a ***best estimate*** *of the learning rate based on various studies and time periods covered. The references are sorted by type of technology, type of learning rate and then date of publication (most recent first).*

Source: Reviewed by Cambridge Econometrics, 2019.

[30] Cited by: https://www.carbonbrief.org/europe-could-get-10-times-its-electricity-needs-from-onshore-wind-study-says

Table 0-42 Learning rates for Wind turbine (offshore) based on reviewed literature

Clean energy technology	Type of learning rate (by doing / by research)	Estimated decrease in unit cost (estimated learning rates)	Years covered in study / across the studies	Assumptions applied for global installed capacity	Source
Wind turbine (offshore)	by doing	11% (in 2020, decreasing to 5% by 2030)	various	2020: 30GW; 2030: 47GW; 2050: 108GW	EC (2018)
Wind turbine (offshore)	by doing	10.3%	various	n.a.	Louwen et al. (2018)
Wind turbine (offshore)	by doing	5% / 12% / 19% (min/mean/max)	1985–2001	n.a.	Rubin et al. (2015)
Wind turbine (offshore)	by research	4.9%	1985–2001	n.a.	Rubin et al. (2015)

Note: Learning rates are defined per doubling of cumulative output (learning-by-doing) or per doubling of the Note: Learning rates are defined per doubling of cumulative output (learning-by-doing) or per doubling of the knowledge stock (learning-by-research). Where available, estimation error intervals are indicated for the estimated learning rates, denoted by ± x%, or rate ranges and mean, denoted by (min/mean/max). "Various" years covered: where the referred literature derived and indicated a ***best estimate*** *of the learning rate based on various studies and time periods covered. The references are sorted by type of technology, type of learning rate and then date of publication (most recent first).*

Source: Reviewed by Cambridge Econometrics, 2019.

Table 0-43 Learning rates for Wind power (undefined) based on reviewed literature

Clean energy technology	Type of learning rate (by doing / by research)	Estimated decrease in unit cost (estimated learning rates)	Years covered in study / across the studies	Assumptions applied for global installed capacity	Source
Wind power	by doing	7%	various	n.a.	E3-Modelling - CE (2017)
Wind power	by doing	4.1%	2003-2007	n.a.	Qiu - Anadon (2012)
Wind power	by doing	10-20%	various	n.a.	Nemet (2012)
Wind power	by doing	3.8%	various	n.a.	Söderholm et al. (2007)
Wind power	by doing	5.4%	various	n.a.	Klaasen et al. (2005)
Wind power	by research	10.5%	various	n.a.	E3-Modelling - CE (2017)
Wind power	by research	4.3%	2003-2007	n.a.	Qiu - Anadon (2012)
Wind power	by research	16.4%	various	n.a.	Söderholm et al. (2007)
Wind power	by research	12.6%	various	n.a.	Klaassen et al. (2005)

Note: Learning rates are defined per doubling of cumulative output (learning-by-doing) or per doubling of the knowledge stock (learning-by-research). Where available, estimation error intervals are indicated for the estimated learning rates, denoted by ± x%, or rate ranges and mean, denoted by (min/mean/max). "Various" years covered: where the referred literature derived and indicated a ***best estimate*** *of the learning rate based on various studies and time periods covered. The references are sorted by type of technology, type of learning rate and then date of publication (most recent first).*

Source: Reviewed by Cambridge Econometrics, 2019.

Biomass and biofuels

Biomass technologies cover a wide range of process types, but for the current review we consider biomass to include all sub-technologies for combined heat and power generation (European Commission, 2018). Most learning curve estimates in the literature reviewed for biomass-based power generation has focused on fluidized bed combustion for combined heat and power (CHP) and on the production of biogas.

Table 0-44 Learning rates for Biomass and biofuels based on reviewed literature

Clean energy technology	Type of learning rate (by doing / by research)	Estimated decrease in unit cost (estimated learning rates)	Years covered in study / across the studies	Assumptions applied for global installed capacity	Source
Biomass	by doing	11% (2016-2020); 7.8% (2021-2030); 4.4% (2031-2040); 1.8% (2041-2050) ('Medium' learning rate scenario)	2016–2050 (forecast)	n.a.	Handayani et al. (2019)
Biomass	by doing	0% / 11% / 24% (min/mean/max)	1976-2005	n.a.	Rubin et al. (2015)
Biogas production	by doing	Of total cost: 24% (1984-1997); 15% (1984-1991); 0% (1991-2001); Of op. and maintenance costs: 2%	various	102.5 MWe	Junginger et al. (2006)
CHP from biomass	by doing	5%	various	2020: 182GW; 2030: 226GW; 2050: 350GW	EC (2018)
Fluidized bed combustion for CHP	by doing	7-10% (of total capital costs)	1976-2005	n.a.	Koornneef et al. (2007)
Fluidized bed combustion for CHP	by doing	23% (of total capital costs)	1990-2002	102.5 MWe	Junginger et al. (2006)

Note: Learning rates are defined per doubling of cumulative output (learning-by-doing) or per doubling of the knowledge stock (learning-by-research). Where available, estimation error intervals are indicated for the estimated learning rates, denoted by ± x%, or rate ranges and mean, denoted by (min/mean/max). "Various" years covered: where the referred literature derived and indicated a ***best estimate*** *of the learning rate based on various studies and time periods covered. The references are sorted by type of technology, type of learning rate and then date of publication (most recent first).*

Source: Reviewed by Cambridge Econometrics, 2019.

Electric vehicles and batteries

The future cost trajectories for electric cars will be of major importance for modelling the mobility transition. In our review we differentiate three specific vehicle types: BEV (battery electric vehicle), HEV (hybrid electric vehicle), and, where it was not indicated by the reviewed sources, we refer to EV (electric vehicles technology) in general.

The uptake dynamics and cost projections of electric vehicles will largely depend on the introduction of progressive measures aiming to foster the deployment of electric mobility solutions. The latest EV outlook report of the IEA (International Energy Agency, 2019) suggests that governments should allocate funds to accelerate research and innovation in the advanced lithium-ion and solid state battery technologies. A further recommendation is that a stepping-up of funding for battery manufacture should be accompanied by requirements regarding the sustainability of battery cell manufacture.

Table 0-45 Learning rates for Electric vehicles and batteries based on reviewed literature

Clean energy technology	Type of learning rate (by doing / by research)	Estimated decrease in unit cost (estimated learning rates)	Years covered in study / across the studies	Assumptions applied for global installed capacity	Source
EV	by doing	8%	various	n.a.	E3-Modelling - CE (2017)
EV	by doing	6-9% (of li-on battery cost)	various	n.a.	Nykvist - Nilsson (2015)
EV	by doing	9.5% (of battery cost)	various	n.a.	International Energy Agency (2013)
EV	by doing	8% (of li-on battery cost)	various	n.a.	Mayer et al. (2012)
EV	by doing	17% (of li-on battery cost)	various	n.a.	Nagelhout - Ros (2009)
EV	by research	15%	various	n.a.	E3-Modelling - CE (2017)
EV	by research	27%	various	n.a.	Mayer et al. (2012)
BEV battery packs	by doing	15.2% ± 2.9%	various	n.a.	Louwen et al. (2018)
FCEV fuel cell stacks	by doing	18% ± 1.7%	various	n.a.	Louwen et al. (2018)
HEV	by doing	7% (of total cost)	1999-2010	n.a.	Weiss et al. (2012)
HEV battery packs	by doing	10.3% ± 3.3%	various	n.a.	Louwen et al. (2018)
PEV	by doing	20% (of battery cost)	various	n.a.	International Energy Agency (2017)

Note: EV stands for electric vehicle. HEV stands for hybrid electric vehicle. BEV stands for battery electric vehicle. FCEV stands for fuel cell electric vehicle. Learning rates are defined per doubling of cumulative output (learning-by-doing) or per doubling of the knowledge stock (learning-by-research). Where available, estimation error intervals are indicated for the estimated learning rates, denoted by ± x%, or rate ranges and mean, denoted by (min/mean/max). "Various" years covered: where the referred literature derived and indicated a ***best estimate*** *of the learning rate based on various studies and time periods covered. The references are sorted by type of technology, type of learning rate and then date of publication (most recent first).*

Source: Reviewed by Cambridge Econometrics, 2019.

Hydrogen production by electrolysis

Hydrogen electrolysis is a traditional production technique, the learning rates for which are still quite high. The technical specificities of the technology lead to a more and more increasing requirement for hydrogen purity that will expectedly increase production costs as well (Element Energy, 2018), while an accelerated take-up of cleaner production techniques within hydrogen electrolysis (e.g. *blue hydrogen* where carbon capture technology is included in the process of electrolysis, or *green hydrogen* where hydrogen is produced from entirely renewable energy sources) will likely result in unit cost decrease in the medium term. This literature review was unable to identify learning rates specific to these latter (less mature) technologies, however.

Table 0-46 Learning rates for Hydrogen production by electrolysis based on reviewed literature

Clean energy technology	Type of learning rate (by doing / by research)	Estimated decrease in unit cost (estimated learning rates)	Years covered in study / across the studies	Assumptions applied for global installed capacity	Source
Alkaline water electrolyser (AWE)	by doing	18% ± 6% (of capital costs)	various	n.a.	Schmidt et al. (2017)
Proton Exchange Membrane electrolyser (PEM)	by doing	7%	various	Total installed capacity is currently less than 50MW	Element Energy (2018)

Note: Learning rates are defined per doubling of cumulative output (learning-by-doing) or per doubling of the knowledge stock (learning-by-research). Where available, estimation error intervals are indicated for the estimated learning rates, denoted by ± x%, or rate ranges and mean, denoted by (min/mean/max). "Various" years covered: where the referred literature derived and indicated a ***best estimate*** *of the learning rate based on various studies and time periods covered. The references are sorted by type of technology, type of learning rate and then date of publication (most recent first).*

Source: Reviewed by Cambridge Econometrics, 2019.

Hydroelectricity

Hydropower is a key source of electricity in many parts of the world (Rubin, Azevedo, Jaramillo, & Yeh, 2015) with further huge growth potential in certain economies. In the table below, learning rates for hydroelectric power plants are presented. Being a highly mature technology, learning rates for hydroelectric power plants are now considered to be slow.

Table 0-47 Learning rates for Hydroelectricity based on reviewed literature

Clean energy technology	Type of learning rate (by doing / by research)	Estimated decrease in unit cost (estimated learning rates)	Years covered in study / across the studies	Assumptions applied for global installed capacity	Source
Hydroelectric power plant	by doing	1%	various	n.a.	EC (2018)
Hydroelectric power plant	by doing	1.4%	1975-1990	n.a.	Kouvaritakis et al. (2000)
Large hydro power plant	by doing	1.96%	1980-2001	n.a.	Jamasb (2007)
Large hydro power plant	by research	2.63%	1980-2001	n.a.	Jamasb (2007)
Small hydro power plant	by doing	0.48%	1988-2001	n.a.	Jamasb (2007)
Small hydro power plant	by research	20.6%	1988-2001	n.a.	Jamasb (2007)

Note: Learning rates are defined per doubling of cumulative output (learning-by-doing) or per doubling of the knowledge stock (learning-by-research). Where available, estimation error intervals are indicated for the estimated learning rates, denoted by ± x%, or rate ranges and mean, denoted by (min/mean/max). "Various" years covered: where the referred literature derived and indicated a ***best estimate*** *of the learning rate based on various studies and time periods covered. The references are sorted by type of technology, type of learning rate and then date of publication (most recent first).*

Source: Reviewed by Cambridge Econometrics, 2019.

Synthetic fuels (power to X)

In our current review, we have considered methane or liquid fuels as synthetic fuels (derived from power-to-gas, PtG or power-to-liquid, PtL processes), which will play a key role in decarbonising chemicals, the industrial sector, and parts of the transport sector (Agora Energiewende, 2018).

Interest in synthetic fuels has picked up relatively recently and so the literature is relatively sparse. We identified one sound and relevant source (Agora Energiewende, 2018) that provides current estimates and projected learning rates for power-to-gas (PtG) facilities (based on a German example).

Table 0-48 Learning rates for Synthetic fuels (power to X) based on reviewed literature

Clean energy technology	**Type of learning rate (by doing / by research)**	**Estimated decrease in unit cost (estimated learning rates)**	**Years covered in study / across the studies**	**Assumptions applied for global installed capacity**	**Source**
Power-to-Gas (PtG)	by doing	13%	2020-2050 (forecast, based on 2014 learning rate estimation)	0.03 GW of power-to-gas facilities (in Germany)	Agora Energiewende (2018)

Note: Learning rates are defined per doubling of cumulative output (learning-by-doing) or per doubling of the knowledge stock (learning-by-research). Where available, estimation error intervals are indicated for the estimated learning rates, *denoted by ± x%, or rate ranges and mean, denoted by (min/mean/max). "Various" years covered: where the referred literature derived and indicated a **best estimate** of the learning rate based on various studies and time periods covered. The references are sorted by type of technology, type of learning rate and then date of publication (most recent first).*

Source: Reviewed by Cambridge Econometrics, 2019.

DAC (Direct air capture)

The overall feasibility and adaptability of large-scale biological CO_2-removal technologies to achieve ambitious climate targets is still unclear, but Direct Air Carbon Capture and Storage (DACCS) could offer an option to reduce climate mitigation costs (Realmonte, et al., 2019). While the learning rates for this technology are encouragingly high based on the reviewed research outputs below, (Realmonte, et al., 2019) as of today the key factor limiting the deployment of direct air carbon capture technology is the rate at which it can be scaled up.

Table 0-49 Learning rates for Direct air capture based on reviewed literature

Clean energy technology	Type of learning rate (by doing / by research)	Estimated decrease in unit cost (estimated learning rates)	Years covered in study / across the studies	Assumptions applied for global installed capacity	Source
DACCS (Direct Air Carbon Capture and Storage)	by doing	6%	various	n.a.	Realmonte et al. (2019)
DACCS (Direct Air Carbon Capture and Storage)	by doing	15%	2020-2050	Global initial DAC capacity in 2020: 1.5 MtCO2/a	Fasihi et al. (2019)
DAC / capital costs	by doing	10.1%	various	n.a.	Nemet - Brandt (2012)
DAC / energy costs	by doing	13.5%	various	n.a.	Nemet - Brandt (2012)
DAC / operation and maintenance costs	by doing	13.5%	various	n.a.	Nemet - Brandt (2012)

*Note: Learning rates are defined per doubling of cumulative output (learning-by-doing) or per doubling of the knowledge stock (learning-by-research). Where available, estimation error intervals are indicated for the estimated learning rates, denoted by ± x%, or rate ranges and mean, denoted by (min/mean/max). "Various" years covered: where the referred literature derived and indicated a **best estimate** of the learning rate based on various studies and time periods covered. The references are sorted by type of technology, type of learning rate and then date of publication (most recent first).*

Source: Reviewed by Cambridge Econometrics, 2019.

Biomass CCS and Traditional CCS

The literature distinguishes the following carbon capture technologies: in coal-based power generation, gas-based power generation and in biomass-based power generation. The table below, mainly building on (European Commission, 2018), lists observed and relevant learning-by-doing rates for specific types.

Table 0-50 Learning rates for Biomass CCS and Traditional CCS based on reviewed literature

Clean energy technology	Type of learning rate (by doing / by research)	Estimated decrease in unit cost (estimated learning rates)	Assumptions applied for global installed capacity	Source
Pulverised coal supercritical, CCS post-combustion	by doing	2.1%	2020: 0 GW; 2030: 1 GW; 2050: 11 GW	EC (2018)
Pulverised coal supercritical, CCS oxyfuel	by doing	2.8%	2020: 0 GW; 2030: 1 GW; 2050: 11 GW	EC (2018)
Lignite integrated gasification and combined cycle, CCS precombustion	by doing	5%	2020: 0 GW; 2030: 1 GW; 2050: 11 GW	EC (2018)
Coal integrated gasification and combined cycle, CCS precombustion	by doing	5%	2020: 0 GW; 2030: 1 GW; 2050: 11 GW	EC (2018)
Natural gas combined cycle, CCS post-combustion	by doing	2.2%	n.a.	EC (2018)
Biomass integrated gasification and combined cycle, CCS precombustion	by doing	5%	n.a.	EC (2018)
Equipment for CCS technologies	by doing	7%	n.a.	E3-Modelling - CE (2017)
Equipment for CCS technologies	by research	7%	n.a.	E3-Modelling - CE (2017)

Note: Learning rates are defined per doubling of cumulative output (learning-by-doing) or per doubling of the knowledge stock (learning-by-research). Where available, estimation error intervals are indicated for the estimated learning rates, denoted by ± x%, or rate ranges and mean, denoted by (min/mean/max). The references are sorted by type of technology, type of learning rate and then date of publication (most recent first). Years covered are not indicated, therefore not presented in the table.

Source: Reviewed by Cambridge Econometrics, 2019.

Conventional technologies

For comparison with the learning rates estimated for clean-energy technologies (whether mature or emerging), this section reviews the learning rates associated with more conventional technologies, such as: ICE (internal combustion engine), CCG (combined-cycle gas) and nuclear power.

Table 0-51 Learning rates for Conventional technologies (ICE, CCG, Nuclear) based on reviewed literature

Technology	Type of learning rate (by doing / by research)	Estimated decrease in unit cost (estimated learning rates)	Years covered in study / across the studies	Assumptions applied for global installed capacity	Source
Coal (IGCC)	by doing	2.5%-7.6%	projection (as of 2013)	n.a.	Azevedo et al. (2013)
ICE (internal combustion engine)	by doing	2%	-	n.a.	Graham et al. (2018)
Natural gas (combined-cycle gas)	by doing	25%	1991-1997	70.000 MW	Colpier – Cornland (2002)
Natural gas (gas-fired combustion turbine)	by doing	0.65%-5.3%	1980-1998	n.a.	Azevedo et al. (2013)
Natural gas (gas-fired combustion turbine)	by research	2.4%-17.7%	1980-1998	n.a.	Azevedo et al. (2013)
Nuclear	by doing	3%	-	n.a.	Graham et al. (2018)
Nuclear	by doing	1-6%	1975–1993	n.a.	Azevedo et al. (2013)

Note: Learning rates are defined per doubling of cumulative output (learning-by-doing) or per doubling of the knowledge stock (learning-by-research). Where available, estimation error intervals are indicated for the estimated learning rates, denoted by ± x%, or rate ranges and mean, denoted by (min/mean/max). "Various" years covered: where the referred literature derived and indicated a ***best estimate*** *of the learning rate based on various studies and time periods covered. The references are sorted by type of technology, type of learning rate and then date of publication (most recent first).*

Source: Reviewed by Cambridge Econometrics, 2019.

Selected other sectors

Finally, this review turns to the learning rates reported for selected other of the key non-clean energy sectors, namely chemicals and semiconductors.

The period covered by the studies refer back to an era when the products were relatively new or at least not yet mature, making them a broadly relevant comparator for some of the clean energy technologies. The rates shown in the table are towards the upper end of the range reported above for clean energy technologies, suggesting that assumptions for the modelling that rely on the rates reported above at not unusually high compared with the experience of other sectors.

Table 0-52 Learning rates for selected non-energy sectors (chemicals, semiconductors) based on reviewed literature

Technology	Type of learning rate (by doing / by research)	Estimated decrease in unit cost (estimated learning rates)	Years covered in study / across the studies	Assumptions applied for global installed capacity	Source
Chemicals - Weighted total of 37 chemicals	by doing	28%	1950/1960 - 1972	n.a.	Lieberman (1984)
Chemicals - Cellophane	by doing	16%	1924-1950	n.a.	Junginger et al. (2008)
Chemicals - Cellophane	by doing	14%	1925-1950	n.a.	Junginger et al. (2008)
Chemicals - Polyester	by doing	13%	1973-1983	n.a.	Junginger et al. (2008)
Semiconductors (DRAM - dynamic random-access memory)	by doing	20%	1974-1992	n.a.	Irwin - Klenow (1994)

Note: Learning rates are defined per doubling of cumulative output (learning-by-doing) or per doubling of the knowledge stock (learning-by-research). Where available, estimation error intervals are indicated for the estimated learning rates, denoted by ± x%, or rate ranges and mean, denoted by (min/mean/max). "Various" years covered: where the referred literature derived and indicated a ***best estimate*** *of the learning rate based on various studies and time periods covered. The references are sorted by type of technology, type of learning rate and then date of publication (most recent first).*

Source: Reviewed by Cambridge Econometrics, 2019.

4.2. Updating of industry and region spillover matrices based on patent data

Both E3ME and GEM-E3 make use of estimates of spillover matrices to capture the extent to which the benefits of R&D in one industry and region are transmitted to other industries and regions that did not undertake the R&D but which benefit from the diffusion of the knowledge. Such spillovers can in principle occur through trade and direct investment links, in which trading partners or subsidiaries acquire knowledge from the source industry-region, or through technological similarities and cultural or institutional proximity. The approach adopted here and described in the following section is motivated by the linkages implied by the second of these processes and draws evidence from patent citations data recorded at a detailed product level.

In both models the impact of spillovers is to increase the stock of knowledge in the recipient industries and regions in the same way as if they had undertaken the R&D themselves. This in turn contributes to economic performance in the way in which each model represents the role of knowledge: raising productivity, lowering costs and improving the quality of products and hence strengthening trade competitiveness.

4.2.1. Construction of spillover matrices – data and processing

We have built on the methodology presented by Meijers & Verspagen (Meijers & Verspagen, 2010) and used by Cambridge Econometrics in the MONROE project (Cambridge Econometrics, 2019) to model effects of technology spillovers across countries and across industries. Data from the United States Patent and Trademark Office

(USPTO) have been obtained and processed[31] for the years 2012-2016: data for all these years are combined to provide results that are effectively an average across the period.

Processing the data involved

- the classification of patents by sector (E3ME and GEM-E3 relevant sectors) and region
- the weighting of patent citations according to the total number of citations per patent, to adjust for the fact that some patents are more highly cited

The weights are calculated using the number of cited patents per citing patent, so that the weighted shares that represent the relative importance of spillover effects from any single patent to other industries/countries in the matrices used in modelling are calculated as follows for any given patent ($patentN$) (see also Figure 0.1, table (1) below):

$$weight\_share_{patentN} = \frac{1}{number\_of\_cited_{patentN}}$$

In the subsequent calculations and aggregations we use this $weight\_share$ as representing the spillover effect from one patent to another, and therefore on an aggregated level, between industries and countries / regions.

The classification of patents by sector has been done based on a two-stage process for each of the E3ME and GEM-E3 industry classifications (following a very similar approach used in the MONROE project (Cambridge Econometrics, 2019):

1. A concordance table published by Eurostat (Patent Statistics: Concordance IPC V8 – NACE REV.2 (2014)) was used to classify patents into NACE Rev. 2 sectors.
2. In the case of E3ME, only small adjustments were needed to go from the NACE Rev.2 sectors to the model's industry classification. The outputs of the process are matrices with rows indicating citing industries and columns indicating cited industries.
3. In the case of GEM-E3, to aggregate up to the model's sectoral classification, we calculate the shares of the different NACE2 sectors in a given GEM-E3 industry and apply these shares as weights to construct 'average' spillover coefficients. The output of the process is a long-format database of the aggregated patent metrics by all relevant dimensions.

[31] This has been done with the support of Prof. Dieter Kogler and Changjun Lee from University College Dublin.

Figure 0.1 - Data structure in different steps of the processing

(1) Processed USPTO data

app_y	citing_country	citing_NACE2	citing_patent	cited_patent	cited_NACE2	cited_country	weight_share
2011	US	26.2	7921252	6526468	26.2	CA	0.0068
2011	US	26.2	7921252	6595929	32.5	US	0.2877
2011	US	26.2	7921252	6820160	26.2	US	0.0068
2011	US	26.2	7921252	7289553	26.1	JP	0.1370
2011	US	26.2	7921252	7289553	26.3	JP	0.4795
2011	US	26.2	7921252	7330544	26.3	FR	0.0068
2011	US	26.2	7921252	7365454	26.2	US	0.0137
2011	US	26.2	7921252	7610422	26.3	US	0.0068
2011	US	26.2	7921252	7689724	26.2	US	0.0137
2011	US	26.2	7921252	7765344	26.2	US	0.0137
2011	US	26.2	7921252	7840742	26.2	US	0.0274

$\Sigma = 1$

Weights sum to 1 for every *citing* patent

...

(2) Constructed E3ME spillover matrix

citing / cited	1	2	3	4	5	6	7	8	...
1	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	
2	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	
3	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	
4	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	
5	0.0000	0.0000	0.0000	0.0000	47.6374	0.0499	0.8281	0.0000	
6	0.0000	0.0000	0.0000	0.0000	0.0394	44.3531	0.3461	0.3787	
7	0.0000	0.0000	0.0000	0.0000	0.8784	2.3820	22.7917	0.3978	
8	0.0000	0.0000	0.0000	0.0000	0.0005	0.5187	0.0284	4.8189	
...									

(3) Constructed data for GEM-E3

citing_country	cited_country	citing_GEM	cited_GEM	value
AE	SG	Market Services	Market Services	0.0488
AE	US	Consumer Goods Industries	Consumer Goods Industries	0.0027
AE	US	Consumer Goods Industries	Electronic Goods	0.0109
AE	US	Consumer Goods Industries	Other Equipment Goods	0.0058
AE	US	Consumer Goods Industries	Transport equipment	0.0091
AE	US	Electronic Goods	Consumer Goods Industries	0.0109
AE	US	Electronic Goods	Electronic Goods	0.0437
AE	US	Electronic Goods	Other Equipment Goods	0.0231
AE	US	Electronic Goods	Transport equipment	0.0366
AE	US	Market Services	Consumer Goods Industries	0.0008

Note: These are only data excerpts for illustrating the transformation steps of the processing. Presented content of the tables are for illustrative purposes only.

Source: Cambridge Econometrics, 2019.

For aggregating the patents to geographical categories (regions in E3ME) an analogous procedure was followed. Figure 0.1 illustrates the data structure at the beginning and end of the process.

4.2.2. Analysis of observed spillover effects

This section presents a short descriptive analysis of the resulting estimates of spillover effects. Since the processed spillover data are to be used to update E3ME and GEM-E3's assumptions for spillover relationships, the relationships observed here will be reflected in the two models' properties for the transmission of knowledge. The analysis is presented in two parts: region-region linkages and industry-industry linkages.

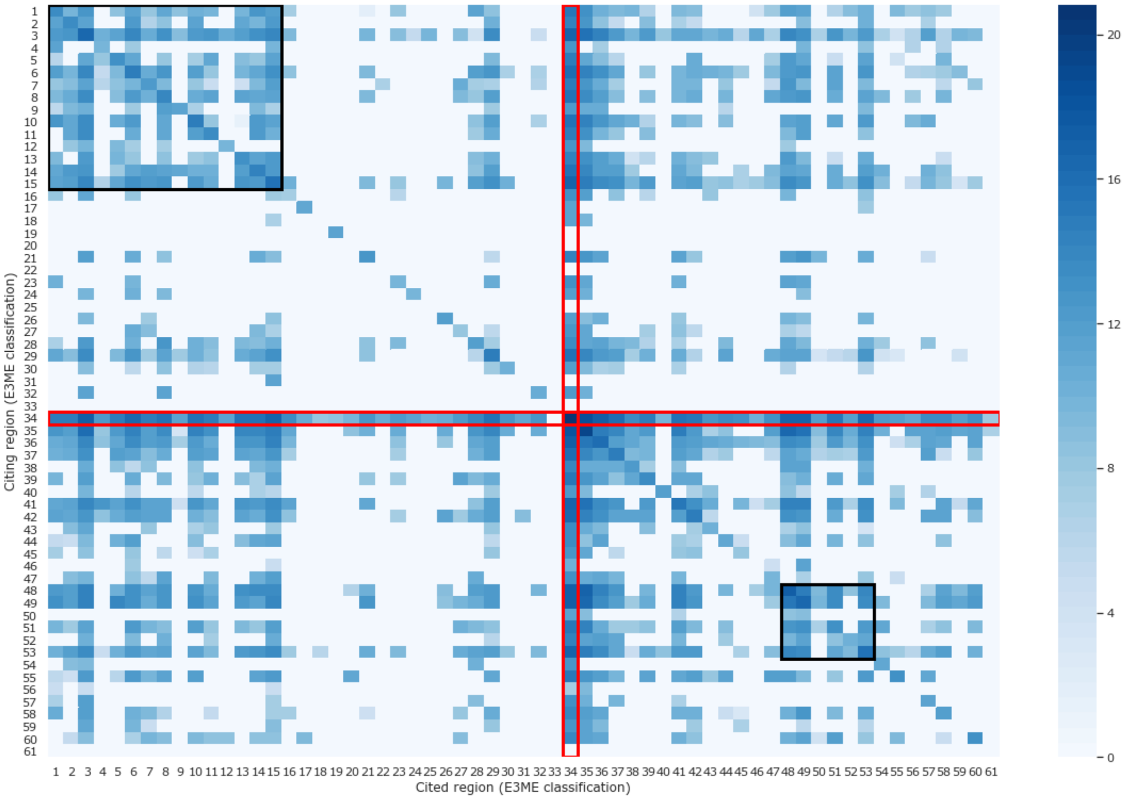
Region-to-region spillover effects

The relative importance of region to region spillover effects is shown in Figure 0.2. The heatmap here shows how the citing regions (on the y-axis) are linked through the patents with the cited regions (on the x-axis). For the purposes of this visualisation, the data shown are for the 61 regions defined in E3ME, which can be found in

Table 0.53.

The elements on the main diagonal show the tendency for own-country citations of patents (from one patent to another).

Figure 0.2 – Region-to-region spillover linkages



Note: Darker colours reflect stronger linkages. For visualisation purposes, the coefficients are transformed through multiplication by 100,000 followed by taking logarithms.

Source: Cambridge Econometrics, 2019.

The US (region code 34, highlighted with red in both dimensions) is an important citer of patents from other regions and an importance source of citations by other regions. This reflects its overall size, its advanced technological state, its well-developed system of patenting and the fact that the database we are using is the USPTO's data. The heatmap shows clustering among certain regions, as illustrated by the groupings highlighted with a black border. The upper left-hand corner of the heatmap shows that the EU15 countries cite each other heavily, in marked contrast to the weaker links with and between the rest of the EU countries (region codes 16-27 and 31). The clustering among regions 48-53 relates mostly to the Asian countries (Korea, Taiwan, and ASEAN) and again shows the impact of geographical/cultural proximity on knowledge dissemination.

The heat map shows a mostly symmetric matrix, showing that citations are symmetric – the spillover effect generally seems to be two-way.

While the database is dominated by the US, Japan too is an important source and user of citations. Table 0.54 shows the top 5 regions cited and citing. Japan and the US are well ahead of the countries ranked 3rd and following. Together they account for more than half of the (weighted and aggregated) citation shares.

While South Korea and Germany appear in both tables, Canada and Taiwan appear once on each side: Canada is cited more than it cites, while Taiwan cites more than it is cited.

Table 0.53 E3ME 61 region classification

E3ME 61 region classification			
1 Belgium	16 Czech Republic	31 Croatia	46 Colombia
2 Denmark	17 Estonia	32 Turkey	47 Rest of Latin America
3 Germany	18 Cyprus	33 Macedonia	48 Korea
4 Greece	19 Latvia	34 USA	49 Taiwan
5 Spain	20 Lithuania	35 Japan	50 Indonesia
6 France	21 Hungary	36 Canada	51 Rest of ASEAN
7 Ireland	22 Malta	37 Australia	52 Rest of OPEC
8 Italy	23 Poland	38 New Zealand	53 Rest of world
9 Luxembourg	24 Slovenia	39 Russian Federation	54 Ukraine
10 Netherlands	25 Slovakia	40 Rest of Annex I	55 Saudi Arabia
11 Austria	26 Bulgaria	41 China	56 Nigeria
12 Portugal	27 Romania	42 India	57 South Africa
13 Finland	28 Norway	43 Mexico	58 Rest of Africa
14 Sweden	29 Switzerland	44 Brazil	59 Africa OPEC
15 UK	30 Iceland	45 Argentina	60 Malaysia
			61 Kazakhstan

Source: Cambridge Econometrics, 2019.

Table 0.54 Top 5 regions cited in the database (left), Top 5 regions citing in the database (right)

Cited region	Share of total weighted citations
United States of America	38.15%
Japan	19.63%
Germany	6.3%
South Korea	4.55%
Canada	4.5%

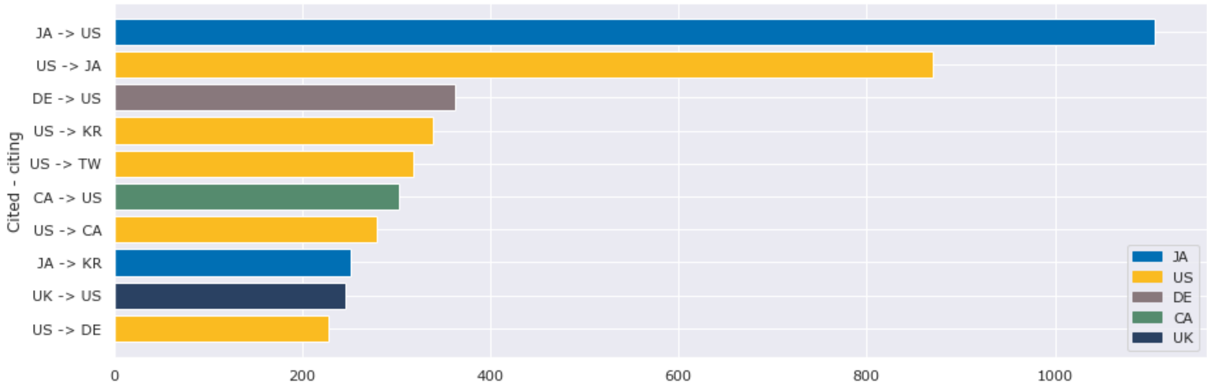
Citing region	Share of total weighted citations
United States of America	36.61%
Japan	14.34%
South Korea	7.8%
Taiwan	6.4%
Germany	4.21%

Note: Values are presented as a share of total weighted citations.

Source: Cambridge Econometrics, 2019.

Figure 0.3 shows the most important bilateral linkages. The US appears as one of the partners in 9 out of 10 cases, the remaining case being the strong link between Japan and South Korea. South Korea cites Japanese patents even more than the US cites UK

Figure 0.3 – Top region-to-region spillover linkages



Note: Measure of the displayed data is the value of the calculated weight_share coefficients. Colouring is based on the cited region.

Source: Cambridge Econometrics, 2019.

patents.

Industry-to-industry spillover effects

Investigating the calculated industry-to-industry matrices based on the E3ME 44-sector classification shows that the classification structure is not fully covered by the database: patents in the database are classified to 19 sectors – various manufacturing industries (e.g.: food, drink & tobacco; electronics; chemicals; etc.), construction and computing services. It appears that there are no patents classified into sectors such as agriculture, transport, education or other services.

Linkages are once stronger along the diagonal – indicating that spillover effect is important internally (i.e.: inside the sector). “Clustering” here appears less significantly, but still visible – on good example is sectors 16-19 (highlighted on Figure 0.4). These are industries covering activities of electronics manufacturing (mechanical engineering, electronics, electronic engineering and instruments), therefore the cross-industry linkage seems logical.

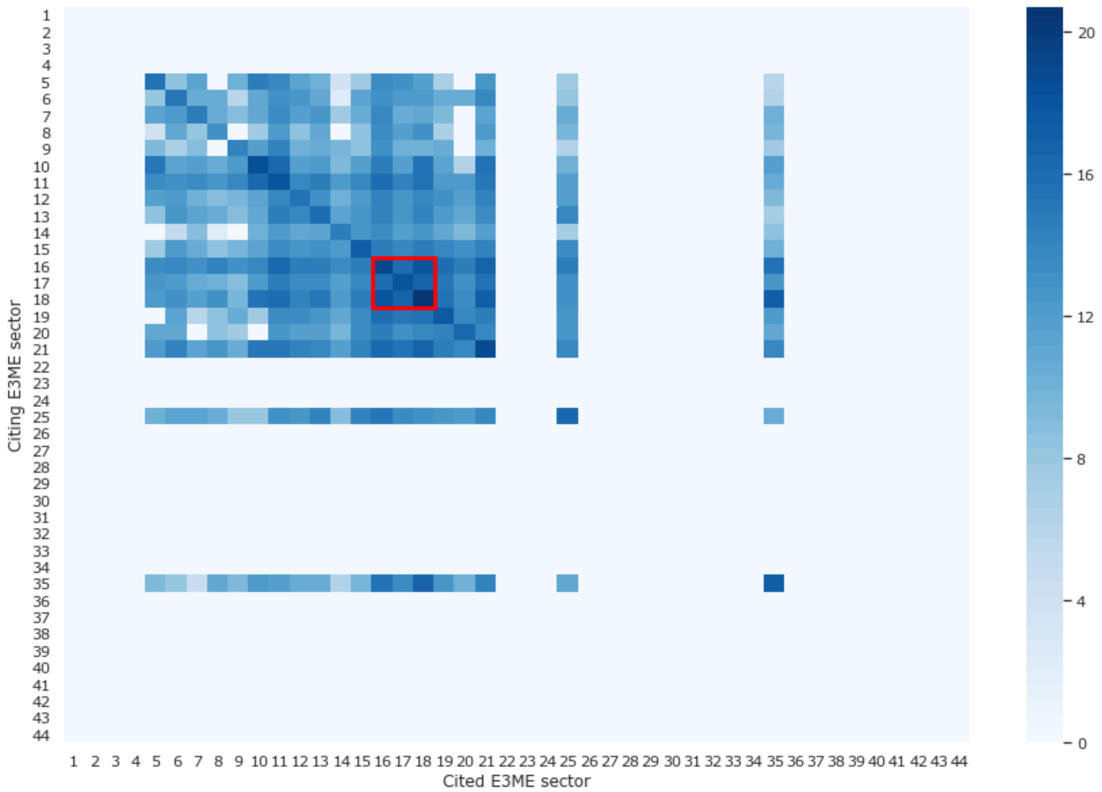


Figure 0.4 – Industry-to-industry spillover linkages (E3ME 44 sector class.)

Note: Darker colours reflect stronger linkages. For visualisation purposes, the coefficients are transformed through multiplication by 100,000 followed by taking logarithms.

Source: Cambridge Econometrics, 2019.

Plotting out the distribution of the “linkage scores” (Figure 0.5) we find that the scores follow a distribution that can be described as log-normal, indicating that there are certain connections which are much stronger than the rest. Therefore, it is important to analyse these stronger linkages. In Figure 0.6 the top 10 industry-industry linkages are presented: the “cluster” that was presented earlier between the manufacturing sectors of electronics is dominant (4 out of 10) and most other linkages are related to these sectors too.

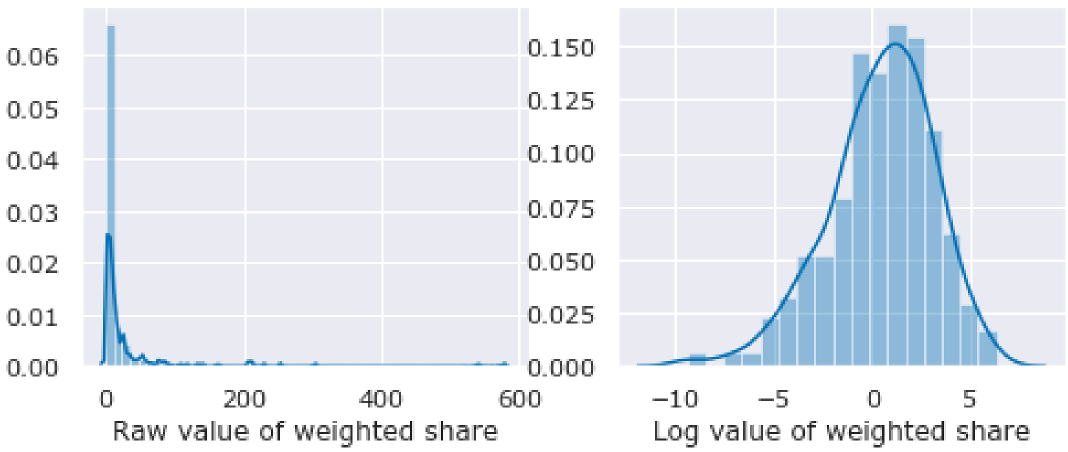


Figure 0.5 – Distribution of industry-industry spillover scores

Note: Distribution of weighted share across industry-industry relations.

Source: Cambridge Econometrics, 2019.

Running a similar analysis on the database using GEM-E3's broader classifications yields somewhat different results, but the main findings are consistent. Classifying to GEM-E3's classification system was done using NACE Rev2 sectors. The USPTO database, as it was described before, consist information on the NACE classification of the patents, while GEM-E3's broad sectors can be translated to several NACE Rev2 groups and classes[32].

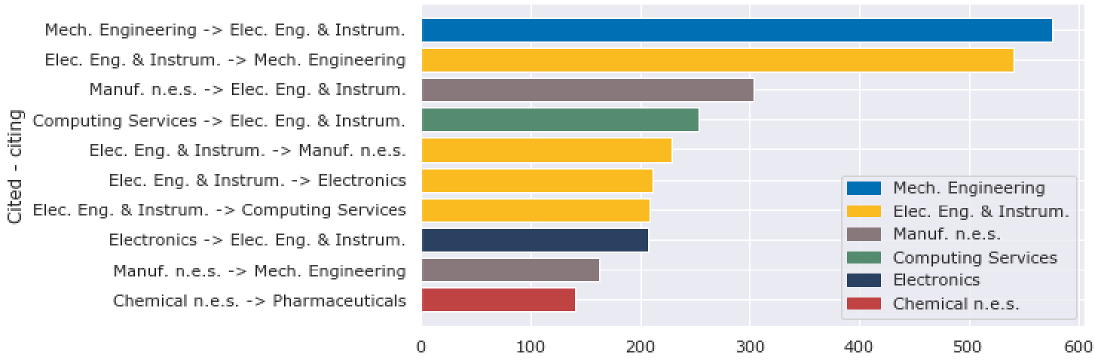


Figure 0.6 – Top industry-industry linkages

Note: Measure of the displayed data is the value of the calculated weight_share coefficients. Colouring is based on the cited industry.
Source: Cambridge Econometrics, 2019.

Therefore, a simple method was used for translating the values from NACE Rev2 classification to GEM-E3: the classes and groups that were provided as parts of a GEM-E3 broad category were aggregated and shares were calculated based on them creating a coefficient for sharing out patent value with a certain NACE code between the GEM-E3 categories. As an example, if in the database there was a patent with NACE code 12 (Manuf. of tobacco products) and with a weight_share of 1 (meaning only 1 citation) it was shared out between the broad GEM-E3 sector "Agriculture" and "Consumer Goods Industries" with values 0.5-0.5.

Figure 0.7 is the result of this process. It presents and overview of the spillover linkages in the database using the GEM-E3 broad sector classification. There are similarities, like the strong diagonal, but due to the different classification structure there are obvious differences – because of the method of "sharing out" certain patent values across sectors there are sectors which were presented as "empty" when using the E3ME classifications: the Agriculture is one of such sectors.

The previously analysed strong linkage between different industries of electronics manufacturing here is represented by the significant connections between "Electronic goods" and "Other Equipment Goods". With the later consisting NACE classes such as "Manufacturing of rubber products" but also "Electronic components" and "Communication equipment".

In the GEM-E3 classification services are grouped together to one category "Market Services". While with the E3ME classification there was no linkages presented towards services, again due to the "sharing out" process services (while having relatively weak links) act as both source and target sectors for the spillover effects.

[32] NACE Rev2 consists of three levels: Divisions (highest level), groups and classes. GEM-E3 to NACE Rev2 correspondence tables were available as linkages between groups/classes and broad GEM-E3 sectors.

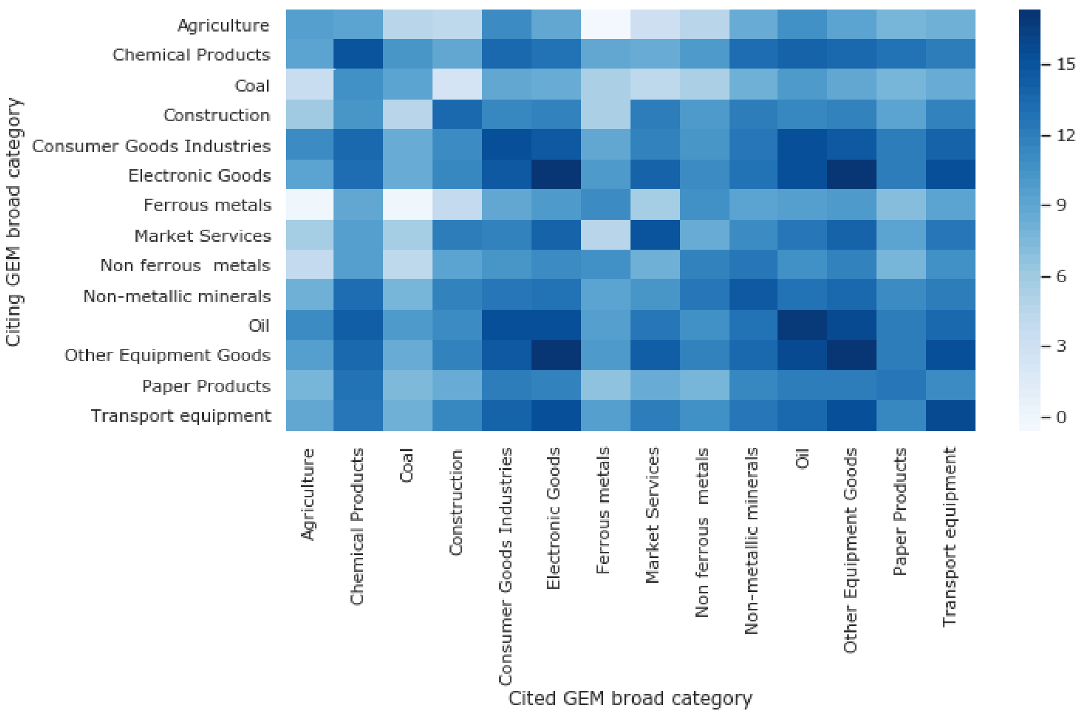


Figure 0.7 – Top industry-industry spillover linkages (GEM-E3 broad class.)

Note: Darker colours reflect stronger linkages. For visualisation purposes, the coefficients are transformed through multiplication by 100,000 followed by taking logarithms.
Source: Cambridge Econometrics, 2019.

5. References

Agora Energiewende. (2018). The Future Cost of Electricity-Based Synthetic Fuels. Retrieved Sep 6, 2019, from https://www.agora-energiewende.de/fileadmin2/Projekte/2017/SynKost_2050/Agora_SynKost_Study_EN_WEB.pdf

Antoszewski Michał (2019). Wide-range estimation of various substitution elasticities for CES production functions at the sectoral level. Energy Economics 83, 272–289.

Arrow, K. J. (1962). The economic implications of learning by doing. The Review of Economic Studies, 29(3), 155-173.

Banerjee A, Dolado J, Mestre R. (1998). Error-correction mechanism tests for cointegration in single-equation framework. J Time Ser Analysis;19:267-83.

Beaudreau Bernard C. (2010). On the methodology of energy-GDP Granger causality tests. Energy 35,3535-3539.

Bilgili, Faik, (1998). Stationarity and cointegration tests: Comparison of Engle - Granger and Johansen methodologies, MPRA Paper 75967, University Library of Munich, Germany.

Boulatoff, C., & Boyer, C. M. (2016). What is the impact of private and public R&D on clean technology firms' performance? An international perspective. Journal of Sustainable Finance & Investment, 7(2). Retrieved Sep 2, 2019, from https://www.tandfonline.com/doi/abs/10.1080/20430795.2016.1251813

Broadstock, D., Hunt, L., Sorrell, S., 2007. UKERC review of evidence for the rebound effect. Technical report 3: elasticity of substitution studies. Working Paper – REF UKERC/WP/TPA/2007/011. Centre, UK Energy Research.

Cambridge Econometrics. (2019). Technical description of the R&I module of the E3ME model. Deliverable D.4.4.2 within the MONROE project.

Dickey David A. and Wayne A. Fuller (1979). Distribution of the Estimators for Autoregressive Time Series With a Unit Root. Journal of the American Statistical Association, Vol. 74, No. 366, 427-431.

Dickey David A. and Wayne A. Fuller (1981). Likelihood ratio statistics for autoregressive time series with a unit root. Econometrica, Vol. 49, No. 4.

E3-Modelling - Cambridge Econometrics. (2017). A technical case study on R&D and technology spillovers of clean energy technologies. European Commission, DG Energy.

Element Energy. (2018). Hydrogen supply chain evidence base. Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/760479/H2_supply_chain_evidence_-_publication_version.pdf

European Commission. (2018). Cost development of low-carbon energy technologies: Scenario-based cost trajectories to 2050, 2017 Edition. European Commission. Retrieved Aug 26, 2019, from http://publications.jrc.ec.europa.eu/repository/bitstream/JRC109894/cost_development_of_low_carbon_energy_technologies_v2.2_final_online.pdf

Ferioli, F., Schoots, K., & van der Zwaan, B. C. (2009). Use and limitations of learning curves for energy technology policy: A component-learning hypothesis. Energy Policy, 37(7), 2525-2535. Retrieved Sep 3, 2019, from https://www.sciencedirect.com/science/article/pii/S0301421508006101

Granger C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods. Econometrica, Vol. 37, No. 3.

Gregory Allan W., Hansen Bruce E. (1996). Residual-based tests for cointegration in models with regime shifts. Journal of Econometrics 70, 99- 126.

Gregory Allan W., Nason James M., Watt David G (1996) Testing for structural breaks in cointegrated relationships. Journal of Econometrics 71, 321-341.

Gross, R., Heptonstall, P., Greenacre, P., Candelise, C., Jones, F., & Castillo, A. (2013). Presenting the future: An assessment of future costs estimation methodologies in the electricity generation sector. UK Energy Research Centre. Retrieved Aug 30, 2019, from http://www.ukerc.ac.uk/asset/3009A2A7-BE27-4DE6-B22B4D12BFA98D95

Henningsen, A., Henningsen, G., van der Werf, E. (2019). Capital-labour-energy substitution in a nested CES framework: a replication and update of Kemfert (1998). Energy Economics 82, 16-25.

Heuberger, C. F., Rubin, E. S., Staffell, I., Shah, N., & Dowell, N. M. (2018, Jun 15). Power capacity expansion planning considering endogenous technology cost learning. Applied Energy, 974. Retrieved Sep 3, 2019, from https://www.sciencedirect.com/science/article/pii/S0306261917309479

Im Kyung So, Pesaran M. Hashem, Shin Yongcheol (2003). Testing for unit roots in heterogeneous panels. Journal of Econometrics 115, 53 – 74.

International Energy Agency. (2015). Projected Costs of Generating Electricity - 2015 edition.

International Energy Agency. (2019). Global EV Outlook 2019. Retrieved Sep 3, 2019, from https://webstore.iea.org/download/direct/2807?fileName=Global_EV_Outlook_2019.pdf

Johansen Søren, Juselius Katarina (1990). Maximum likelihood estimation and inference on cointegration – with applications to the demand for money. Oxford Bulletin of Economics and Statistics, 52, 2.

Junginger, M. (2018). Riding down the experience curves: How technological learning may shape the energy transition. EMP-E conference 2018 'Modelling Clean Energy Pathways' presentation material. Brussels, Belgium. Retrieved Aug 26, 2019, from http://www.energymodellingplatform.eu/uploads/1/8/5/0/18504136/2_junginger_emp-e_presentation_25.9.2018.pdf

Kalimeris Panos, Richardson Clive, Bithas Kostas (2014). A meta-analysis investigation of the direction of the energy-GDP causal relationship: implications for the growth-degrowth dialogue. Journal of Cleaner Production 67, 1-13.

Kao, C. (1999). Spurious Regression and Residual Based Tests for Cointegration in Panel Data. Journal of Econometrics, 90, 1-44.

Kemfert, C. (1998). Estimated substitution elasticities of a nested CES production function approach for Germany. Energy Econ. 20 (3), 249–264.

Koesler, S., Schymura, M. (2015). Substitution elasticities in a constant elasticity of substitution framework – empirical estimates using nonlinear least squares. Econ. Syst. Res. 27 (1), 101–121.

Kouvaritakis, N., Soria, A., & Isoard, S. (2000). Modelling energy technology dynamics: methodology for adaptative expectations model with learning-by-doing and learning-by-searching. International Journal of Global Energy Issues, 14(1-4), 104-115.

Levin Andrew, Lin Chien-Fu, Chu Chia-Shang James (2002). Unit root tests in panel data: asymptotic and finite-sample properties. Journal of Econometrics 108, 1–24.

Louwen, A., Junginger, M., & Krishnan, A. (2018). Technological Learning in Energy Modelling: Experience Curves - Policy Brief. REFLEX project. Retrieved Aug 26, 2019, from http://reflex-project.eu/wp-content/uploads/2018/12/REFLEX_policy_brief_Experience_curves_12_2018.pdf

MacKinnon James G. (2010). Critical Values for Cointegration Tests. Queen's Economics Department Working Paper No. 1227.

Meijers, H., & Verspagen, B. (2010). Construction of technology flow matrices. Athens: Deliverable 1.2b within the Demeter project.

Miremadi, I., Saboohi, Y., & Arasti, M. (2019). The influence of public R&D and knowledge spillovers on the development of renewable energy sources: The case of Nordic countries. Technological Forecasting and Social Change (under publication). Retrieved Sep 2, 2019, from https://www.sciencedirect.com/science/article/pii/S0040162518309132

Ndoricimpa Arcade (2017). Analysis of asymmetries in the nexus among energy use, pollution emissions and real output in South Africa. Energy 125, 543-551.

Ozturk, I., 2010. A literature survey on energyegrowth nexus. Energ. Policy 38 (1), 340-349.

Pedroni, P., (1999). Critical values for cointegration tests in heterogeneous panels with multiple regressors. Oxford Bulletin of Economics and Statistics, special issue 0305-9049.

Pedroni, P., (2000). Fully modified OLS for the heterogeneous cointegrated panels. Advances in Econometrics 15, 93–130.

Perron, P. (1989). The great crash, the oil price shock, and the unit root hypothesis. Econometrica 57, 1361-1401.

Pesaran M. Hashem, Shin Yongcheol and Smith Richard J. (2001). Bounds testing approaches to the analysis of level relationships. Journal of applied econometrics *J. Appl. Econ.* 16: 289–326 (2001) DOI: 10.1002/jae.616.

Phillips Peter C. B. and Perron Pierre (1988). Testing for a unit root in time series regression. Biometrika (1988), 75, 2, pp. 335-46.

Popp, D. (2019, Mar). Environmental policy and innovation: a decade of research. Retrieved Sep 6, 2019, from https://www.ifo.de/DocDL/cesifo1_wp7544.pdf

Realmonte, G., Drouet, L., Gambhir, A., Glynn, J., Hawkes, A., & Köberle, A. C. (2019). An inter-model assessment of the role of direct air capture in deep mitigation pathways. Nature Communications. Retrieved Sep 6, 2019, from https://www.nature.com/articles/s41467-019-10842-5

Robert F. Engle and C. W. J. Granger (1987). Co-Integration and Error Correction: Representation, Estimation, and Testing. Econometrica, 55, 2, 251 – 276.

Rubin, E. S., Azevedo, I. M., Jaramillo, P., & Yeh, S. (2015). A review of learning rates for electricity supply technologies. Energy Policy, 86, 198-218.

Ryberg, D. S., Caglayan, D. G., Schmitt, S., Linssen, J., Stolten, D., & Robinius, M. (2019). The future of European onshore wind energy potential: Detailed distribution and simulation of advanced turbine designs. Energy. Retrieved Sep 3, 2019, from https://www.sciencedirect.com/science/article/abs/pii/S0360544219311818?dgcid=rss_sd_all

Shin Y, Yu B, Greenwood-Nimmo M. (2014) Modelling Asymmetric Cointegrationand Dynamic Multipliers in a Nonlinear ARDL Framework. Ch.9 of R.C. Sickles and W.C. Horrace (eds.), Festschrift in Honor of Peter Schmidt: Econometric Methods and Applications, DOI 10.1007/978-1-4899-8008-3__9, Springer, New York.

Sims Christopher A. (1972). Money, Income and Causality. *The American Economic Review*, Vol. 62, No. 4., pp. 540-552.

Toda Hiro Y., Taku Yamamoto (1995). Statistical inference in vector autoregressions with possibly integrated processes. Journal of Econometrics, 66, 225-250.

van der Werf, E. (2008). Production functions for climate policy modeling: an empirical analysis. Energy Econ. 30 (6), 2964–2979.

Veugelers, R. (2016). Getting the most from public R&D spending in times of austerity: Some insights from simpatic analysis. 2016/01. Retrieved Aug 27, 2019, from https://bruegel.org/wp-content/uploads/2016/02/WP-2016_01-1.pdf

Wiser, R., Jenni, K., Seel, J., Baker, E., Hand, M., Lantz, E., & Smith, A. (2016, Jun). Forecasting Wind Energy Costs and Cost Drivers: The Views of the World's Leading Experts. IEA Wind. Retrieved Aug 30, 2019, from https://emp.lbl.gov/sites/all/files/lbnl-1005717.pdf

Zivot, E. and Andrews, D. W. K. (1992), Further Evidence on the Great Crash, the Oil-Price Shock, and the Unit-Root Hypothesis. Journal of Business & Economic Statistics, Vol. 10, No. 3, pp. 251-270.

Zúñiga-Vicente, J. Á., Alonso-Borrego, C., Forcadell, F., & Galán, J. I. (2012). Journal of Economic Surveys, 28(1), 36-67. Retrieved Sep 2, 2019, from https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-6419.2012.00738.x

Appendix: Regions and economic sectors of the WIOD Database

Table 0-55: Sector definitions for abbreviations used in econometric analysis

Abbr.	Description
TOT	Total industries
AtB	Agriculture, Hunting, Forestry and Fishing
C	Mining and Quarrying
15t16	Food, Beverages and Tobacco
17t18	Textiles and Textile Products
19	Leather, Leather and Footwear
20	Wood and Products of Wood and Cork
21t22	Pulp, Paper, Paper , Printing and Publishing
23	Coke, Refined Petroleum and Nuclear Fuel
24	Chemicals and Chemical Products
25	Rubber and Plastics
26	Other Non-Metallic Mineral
27t28	Basic Metals and Fabricated Metal
29	Machinery, Nec
30t33	Electrical and Optical Equipment
34t35	Transport Equipment
36t37	Manufacturing, Nec; Recycling
E	Electricity, Gas and Water Supply
F	Construction
50	Sale, Maintenance and Repair of Motor Vehicles and Motorcycles; Retail Sale of Fuel
51	Wholesale Trade and Commission Trade, Except of Motor Vehicles and Motorcycles
52	Retail Trade, Except of Motor Vehicles and Motorcycles; Repair of Household Goods
H	Hotels and Restaurants
60	Inland Transport
61	Water Transport
62	Air Transport
63	Other Supporting and Auxiliary Transport Activities; Activities of Travel Agencies
64	Post and Telecommunications
J	Financial Intermediation
70	Real Estate Activities
71t74	Renting of M&Eq and Other Business Activities
L	Public Admin and Defence; Compulsory Social Security
M	Education
N	Health and Social Work
O	Other Community, Social and Personal Services

Table 0-56: Region definitions for abbreviations used in econometric analysis

Abbr.	Description
AUS	Australia
AUT	Austria
BEL	Belgium
BRA	Brazil
BGR	Bulgaria
CAN	Canada
CHN	China
CYP	Cyprus
CZE	Czech Republic
DNK	Denmark
EST	Estonia
FIN	Finland
FRA	France
DEU	Germany
GRC	Greece
HUN	Hungary
IND	India
IDN	Indonesia
IRL	Ireland
ITA	Italy
JPN	Japan
KOR	Korea, Republic of
LVA	Latvia
LTU	Lithuania
LUX	Luxembourg
MLT	Malta
MEX	Mexico
NLD	Netherlands
POL	Poland
PRT	Portugal
ROU	Romania
RUS	Russia
SVK	Slovak Republic
SVN	Slovenia
ESP	Spain
SWE	Sweden
TUR	Turkey
GBR	United Kingdom
USA	United States

Augmented Dickey – Fuller (ADF) Unit Root test

```
' ADF Unit Root Test for Energy Use
!j=0
table(1500,5) adf_DEN_A
table(1500,5) adf_DEN_B
table(1500,5) adf_DEN_C

for %reg  AUT_ BEL_ DEU_ ESP_ FRA_ GBR_ GRC_ ITA_ NLD_ POL_ PRT_ ROU_ AUS_ BGR_ BRA_ CAN_ CHN_ CYP_
CZE_ DNK_ EST_ FIN_ HUN_ IDN_ IND_ IRL_ JPN_ KOR_ LTU_ LUX_ LVA_ MEX_ MLT_ RUS_ SVK_ SVN_ SWE_ TUR_
USA_

for  %sec AtB_ C_ 15t16_ 17t18_ 19_ 20_ 21t22_ 23_ 24_ 25_ 26_ 27t28_ 29_ 30t33_ 34t35_  36t37_  E_ F_  50_ 51_ 52_
H_ 60_ 61_ 62_ 63_ 64_ J_ 70_ 71t74_ L_ M_ N_ O_ tot_

%name = %reg + %sec
%var0 = "va_q"
%var1 = "en_q"

%GDP = %name  + %var0
%EN    = %name  + %var1

if ((@min({%EN}) <= 0) or (@min({%GDP}) <= 0)) then
!j = !j +1
    adf_DEN_A(!j,1) = %reg
    adf_DEN_A(!j,2) = %sec
    adf_DEN_A(!j,3) = "n/a"
    adf_DEN_A(!j,4) = "n/a"
    adf_DEN_A(!j,5) = "n/a"

    adf_DEN_B(!j,1) = %reg
    adf_DEN_B(!j,2) = %sec
    adf_DEN_B(!j,3) = "n/a"
    adf_DEN_B(!j,4) = "n/a"
    adf_DEN_B(!j,5) = "n/a"

    adf_DEN_C(!j,1) = %reg
    adf_DEN_C(!j,2) = %sec
    adf_DEN_C(!j,3) = "n/a"
    adf_DEN_C(!j,4) = "n/a"
    adf_DEN_C(!j,5) = "n/a"
else
!j=!j+1

call adf({%EN},1,3)
call adf({%EN},2,3)
call adf({%EN},3,3)
endif

next
next
' ---------------------------------------------------------------------------------------------------
' Arguments
'----------------------------------------------------------------------------------------------------
'series Y              'variable to perform the unit root test
'scalar Model          '1 = None, 2 = Constant, 3 = Constand and Trend
'scalar Maxlag         'Maximum number of lags for unit root testing
'string Criterion      'Selection criteria for unit root testing (i.e. aic / sic / hqc)
' ---------------------------------------------------------------------------------------------------

subroutine adf(series Y, scalar Model, scalar Maxlag)
smpl @all

!maxlag=Maxlag

'level(3,1) : ADF statistic
'level(2,1) : Lag selected (based on Schwarz criterion)
'level(4,1):  P-value of the ADF statistic based on McKinnon critical values

"MODEL 1 : Unit Root test without deterministic constant or trend
if Model=1 then
  uroot(adf, none, maxlag=!maxlag, save=level, dif = 1)  Y
  adf_DEN_A(!j,1) = %reg
  adf_DEN_A(!j,2) = %sec
  adf_DEN_A(!j,3) =  level(3,1)
```

```
  adf_DEN_A(!j,4) =  level(2,1)
  adf_DEN_A(!j,5) =  level(4,1)
endif

"MODEL 2 : Unit Root test with deterministic constant
if Model=2 then
  uroot(adf, const, maxlag=!maxlag, save=level, dif = 1) Y
  adf_DEN_B(!j,1) = %reg
  adf_DEN_B(!j,2) = %sec
  adf_DEN_B(!j,3) = level(3,1)
  adf_DEN_B(!j,4) = level(2,1)
  adf_DEN_B(!j,5) = level(4,1)
endif

"MODEL 3 : Unit Root test with deterministic constant and trend
if Model = 3 then
  uroot(adf, trend, maxlag=!maxlag, save=level, dif = 1) Y
  adf_DEN_C(!j,1) = %reg
  adf_DEN_C(!j,2) = %sec
  adf_DEN_C(!j,3) = level(3,1)
  adf_DEN_C(!j,4) = level(2,1)
  adf_DEN_C(!j,5) = level(4,1)
endif

delete level
endsub
```

Engle and Granger Cointegration test

```
!j=0
table  coint_const
table  coint_trend

scalar mc_binf
scalar mc_b1
scalar mc_b2
scalar mc_b3

for %reg   AUT_ BEL_ DEU_ ESP_ FRA_ GBR_ GRC_ ITA_ NLD_ POL_ PRT_ ROU_ AUS_ BGR_ BRA_ CAN_ CHN_ CYP_
CZE_ DNK_ EST_ FIN_ HUN_ IDN_ IND_ IRL_ JPN_ KOR_ LTU_ LUX_ LVA_ MEX_ MLT_ RUS_ SVK_ SVN_ SWE_ TUR_
USA_

for  %sec AtB_ C_ 15t16_ 17t18_ 19_ 20_ 21t22_ 23_ 24_ 25_ 26_ 27t28_ 29_ 30t33_ 34t35_  36t37_  E_ F_  50_ 51_ 52_
H_ 60_ 61_ 62_ 63_ 64_ J_ 70_ 71t74_ L_ M_ N_ O_ tot_

%name = %reg + %sec
%var0 = "va_q"
%var1 = "en_q"

%DEP1= %name + %var0
%IND1= %name + %var1

if  ((@min({%IND1}) <= 0) or (@min({%DEP1})) <= 0) then
!j = !j +1
    coint_const(!j,1) = %reg
    coint_const(!j,2) = %sec
    coint_const(!j,3) = "n/a"
    coint_const(!j,4) = "n/a"
    coint_const(!j,5) = "n/a"
    coint_const(!j,6) = "n/a"
    coint_const(!j,7) = "n/a"

    coint_trend(!j,1) = %reg
    coint_trend(!j,2) = %sec
    coint_trend(!j,3) = "n/a"
    coint_trend(!j,4) = "n/a"
    coint_trend(!j,5) = "n/a"
    coint_trend(!j,6) = "n/a"
    coint_trend(!j,7) = "n/a"
else

!j=!j+1
```

```
call cointegration({%DEP1},{%IND1},1,3)
call cointegration({%DEP1},{%IND1},2,3)

endif
next
next
' -----------------------------------------------------------------------------------------------------
' Arguments
'------------------------------------------------------------------------------------------------------
'series Y            ' dependent variable
'Series G            '  independent variable
'scalar Model        ' 1 = Consant, 2 = Consant and Trend (in the cointegration relationship)
'scalar Maxlag       ' Maximum number of lags for unit root testing
' -----------------------------------------------------------------------------------------------------

subroutine cointegration(series Y, series G, scalar Model,scalar Maxlag)
smpl @all

!maxlag=Maxlag

equation gr

if Model=1 then
'MODEL 1: cointegration relationship with constant

'McKinnon formula for critical values at 5% level of significance
mc_binf = -2.862
mc_b1 = -2.890
mc_b2 = -4.234
mc_b3 = -40.040

   gr.ls Y c G
   gr.makeresid res
   uroot(adf, none, maxlag=!maxlag, save=level) res
   coint_const(!j,1) = %reg
   coint_const(!j,2) = %sec
   coint_const(!j,3) = level(3,1)
   coint_const(!j,4) = level(2,1)
   coint_const(!j,5) = level(4,1)
   coint_const(!j,6) = level(1,1)
   coint_const(!j,7) = mc_binf + mc_b1/coint_const(!j,6) + mc_b2/(coint_const(!j,6)^2) + mc_b3/(coint_const(!j,6)^3)

else if Model=2 then
'MODEL 2 : cointegration relationship with constant and trend
mc_binf = -3.410
mc_b1 = -4.3904
mc_b2 = -9.036
mc_b3 = -45.37

    gr.ls Y c  G @trend
    gr.makeresid res
    uroot(adf, none, maxlag=!maxlag, save=level) res
    coint_trend(!j,1) = %reg
    coint_trend(!j,2) = %sec
    coint_trend(!j,3) = level(3,1)
    coint_trend(!j,4) = level(2,1)
    coint_trend(!j,5) = level(4,1)
    coint_trend(!j,6) =  level(1,1)
    coint_trend(!j,7) = mc_binf + mc_b1/coint_trend(!j,6) + mc_b2/(coint_trend(!j,6)^2) + mc_b3/(coint_trend(!j,6)^3)

endif
endif
  delete level gr
endsub
```

Johansen and Juselius Cointegration test

```
!j=0
table Johansen_const
table Johansen_trend
```

```
for %reg  AUT_ BEL_ DEU_ ESP_ FRA_ GBR_ GRC_ ITA_ NLD_ POL_ PRT_ ROU_ AUS_ BGR_ BRA_ CAN_ CHN_ CYP_
CZE_ DNK_ EST_ FIN_ HUN_ IDN_ IND_ IRL_ JPN_ KOR_ LTU_ LUX_ LVA_ MEX_ MLT_ RUS_ SVK_ SVN_ SWE_ TUR_
USA_

for  %sec AtB_ C_ 15t16_ 17t18_ 19_ 20_ 21t22_ 23_ 24_ 25_ 26_ 27t28_ 29_ 30t33_ 34t35_  36t37_  E_ F_  50_ 51_ 52_
H_ 60_ 61_ 62_ 63_ 64_ J_ 70_ 71t74_ L_ M_ N_ O_ tot_

%name = %reg + %sec
%var0 = "va_q"
%var1 = "en_q"

%DEP1= %name + %var0
%IND1= %name + %var1

if  ((@min({%IND1}) <= 0) or (@min({%DEP1}) <= 0)) then
!j = !j +1
Johansen_const(!j,1) = %reg
Johansen_const(!j,2) = %sec
Johansen_const(!j,3) = "n/a"
Johansen_const(!j,4) = "n/a"
Johansen_const(!j,5) = "n/a"
Johansen_const(!j,6) = "n/a"

Johansen_trend(!j,1) = %reg
Johansen_trend(!j,2) = %sec
Johansen_trend(!j,3) = "n/a"
Johansen_trend(!j,4) = "n/a"
Johansen_trend(!j,5) = "n/a"
Johansen_trend(!j,6) = "n/a"
else
!j=!j+1
group aa {%dep1} {%ind1}
'MODEL d: cointegration relationship with constant
aa.coint(d,2)
freeze(johc) aa.coint(d,2)
Johansen_const(!j,1) = %reg
Johansen_const(!j,2) = %sec
Johansen_const(!j,3) = johc(13,5)
Johansen_const(!j,4) = johc(14,5)
Johansen_const(!j,5) = johc(25,5)
Johansen_const(!j,6) = johc(26,5)
'MODEL b: cointegration relationship with constant and trend
aa.coint(b,2)
freeze(joht) aa.coint(b,2)
Johansen_trend(!j,1) = %reg
Johansen_trend(!j,2) = %sec
Johansen_trend(!j,3) = joht(13,5)
Johansen_trend(!j,4) = joht(14,5)
Johansen_trend(!j,5) = joht(25,5)
Johansen_trend(!j,6) = joht(26,5)

delete johc joht aa
endif
next
next
```

Error Correction Model and causality tests

```
table(3000,14) ecm_ENtoVA

!i=0
for %reg  AUT_ BEL_ DEU_ ESP_ FRA_ GBR_ GRC_ ITA_ NLD_ POL_ PRT_ ROU_ AUS_ BGR_ BRA_ CAN_ CHN_ CYP_
CZE_ DNK_ EST_ FIN_ HUN_ IDN_ IND_ IRL_ JPN_ KOR_ LTU_ LUX_ LVA_ MEX_ MLT_ RUS_ SVK_ SVN_ SWE_ TUR_
USA_

for  %sec AtB_ C_ 15t16_ 17t18_ 19_ 20_ 21t22_ 23_ 24_ 25_ 26_ 27t28_ 29_ 30t33_ 34t35_  36t37_  E_ F_  50_ 51_ 52_
H_ 60_ 61_ 62_ 63_ 64_ J_ 70_ 71t74_ L_ M_ N_ O_ tot_

  %name = %reg + %sec
   %var0 = "VA_Q"
   %var1 = "EN_Q"

%DEP1= %name + %var0
```

```
%IND1= %name + %var1

if ((@min({%IND1}) <= 0) or (@min({%DEP1}) <= 0)) then
!i = !i +1
ecm_ENtoVA(!i,1) = %reg
ecm_ENtoVA(!i,2) = %sec
ecm_ENtoVA(!i,3) = "n/a"
ecm_ENtoVA(!i,4) = "n/a"
ecm_ENtoVA(!i,5) = "n/a"
ecm_ENtoVA(!i,6) = "n/a"
ecm_ENtoVA(!i,7) = "n/a"
ecm_ENtoVA(!i,8) = "n/a"
ecm_ENtoVA(!i,9) = "n/a"
ecm_ENtoVA(!i,10) = "n/a"
ecm_ENtoVA(!i,11) = "n/a"
        else
equation coint.ls {%dep1} c {%ind1}
coint.makeresid res

equation eq1{%name}.ls D({%dep1}) C res(-1)  D({%dep1}(-1))  D({%ind1}(-1))
equation eq2{%name}.ls  D({%dep1}) C res(-1) D({%dep1}(-1)) D({%dep1}(-2)) D({%ind1}(-1)) D({%ind1}(-2))
equation eq3{%name}.ls  D({%dep1}) C res(-1) D({%dep1}(-1)) D({%dep1}(-2)) D({%dep1}(-3))  D({%ind1}(-1)) D({%ind1}(-2))
D({%ind1}(-3))

freeze(STR1)   eq1{%name}.wald C(2)=0, C(4)=0
freeze(SR2)    eq2{%name}.wald C(5)=0, C(6)=0
freeze(STR2)   eq2{%name}.wald C(2)=0, C(5)=0, C(6)=0
freeze(SR3)    eq3{%name}.wald C(6)=0, C(7)=0, C(8)=0
freeze(STR3)   eq3{%name}.wald C(2)=0, C(6)=0, C(7)=0 , C(8)=0

!i = !i + 1

ecm_ENtoVA(!i,1) = %reg
ecm_ENtoVA(!i,2) = %sec
ecm_ENtoVA(!i,3) = eq1{%name}.@pvals(2)
ecm_ENtoVA(!i,4) = eq1{%name}.@pvals(4)
ecm_ENtoVA(!i,5) = STR1(6,4)
ecm_ENtoVA(!i,6) = eq2{%name}.@pvals(2)
ecm_ENtoVA(!i,7) = SR2(6,4)
ecm_ENtoVA(!i,8) =  STR2(6,4)
ecm_ENtoVA(!i,9) = eq3{%name}.@pvals(2)
ecm_ENtoVA(!i,10) = SR3(6,4)
ecm_ENtoVA(!i,11) =  STR3(6,4)

delete eq2{%name} res str1 sr2 str2 sr3 str3 coint
delete eq1{%name} eq3{%name}

endif
NEXT
NEXT
```

Granger causality test at the first differences of the timeseries

```
table(1500,14)  cs_dif_entova
!i=0
for %reg  AUT_ BEL_ DEU_ ESP_ FRA_ GBR_ GRC_ ITA_ NLD_ POL_ PRT_ ROU_ AUS_ BGR_ BRA_ CAN_ CHN_ CYP_
CZE_ DNK_ EST_ FIN_ HUN_ IDN_ IND_ IRL_ JPN_ KOR_ LTU_ LUX_ LVA_ MEX_ MLT_ RUS_ SVK_ SVN_ SWE_ TUR_
USA_

for %sec  AtB_ C_ 15t16_ 17t18_ 19_ 20_ 21t22_ 23_ 24_ 25_ 26_ 27t28_ 29_ 30t33_ 34t35_  36t37_  E_ F_  50_ 51_ 52_
H_ 60_ 61_ 62_ 63_ 64_ J_ 70_ 71t74_ L_ M_ N_ O_ tot_

  %name = %reg + %sec
   %var0 = "VA_Q"
   %var1 = "EN_Q"

%DEP1= %name + %var0
%IND1= %name + %var1

if ((@min({%IND1}) <= 0) or (@min({%DEP1}) <= 0)) then
!i = !i +1
cs_dif_entova(!i,1) = %reg
cs_dif_entova(!i,2) = %sec
```

```
cs_dif_entova(!i,3) = "n/a"
cs_dif_entova(!i,4) = "n/a"
cs_dif_entova(!i,5) = "n/a"
cs_dif_entova(!i,6) = "n/a"
else

equation eq1r{%name}.ls  D({%DEP1}) C D({%DEP1}(-1))
equation eq1u{%name}.ls D({%DEP1}) C D({%DEP1}(-1)) D({%IND1}(-1))

equation eq2r{%name}.ls  D({%DEP1}) C D({%DEP1}(-1)) D({%DEP1}(-2))
equation eq2u{%name}.ls D({%DEP1}) C D({%DEP1}(-1)) D({%DEP1}(-2))  D({%IND1}(-1))  D({%IND1}(-2))

equation eq3r{%name}.ls  D({%DEP1}) C D({%DEP1}(-1)) D({%DEP1}(-2)) D({%DEP1}(-3))
equation eq3u{%name}.ls D({%DEP1}) C D({%DEP1}(-1)) D({%DEP1}(-2)) D({%DEP1}(-3)) D({%IND1}(-1)) D({%IND1}(-2))
D({%IND1}(-3))

!i = !i + 1

cs_dif_entova(!i,1) = %reg
cs_dif_entova(!i,2) = %sec
' Causality with lag =1
cs_dif_entova(!i,3) = ((eq1r{%name}.@ssr - eq1u{%name}.@ssr) * eq1u{%name}.@df) / (eq1u{%name}.@ssr)
cs_dif_entova(!i,4) = eq1u{%name}.@df
cs_dif_entova(!i,5) = @qfdist(0.95,1,eq1u{%name}.@df)
if  (((eq1r{%name}.@ssr - eq1u{%name}.@ssr) * eq1u{%name}.@df) / (eq1u{%name}.@ssr) >@qfdist(0.95,1,eq1u{%name}.@df))
then
cs_dif_entova(!i,6) = "yes"
else
cs_dif_entova(!i,6) ="no"
endif
' Causality with lag =2
cs_dif_entova(!i,7) = ((eq2r{%name}.@ssr - eq2u{%name}.@ssr) * eq2u{%name}.@df) / (2 * eq2u{%name}.@ssr)
cs_dif_entova(!i,8) = eq2u{%name}.@df
cs_dif_entova(!i,9) = @qfdist(0.95,2,eq2u{%name}.@df)
if  (((eq2r{%name}.@ssr - eq2u{%name}.@ssr) * eq2u{%name}.@df) / (2 * eq2u{%name}.@ssr)
>@qfdist(0.95,2,eq2u{%name}.@df))  then
cs_dif_entova(!i,10) = "yes"
else
cs_dif_entova(!i,10) ="no"
endif
' Causality with lag =3
cs_dif_entova(!i,11) = ((eq3r{%name}.@ssr - eq3u{%name}.@ssr) * eq3u{%name}.@df) / (3 * eq3u{%name}.@ssr)
cs_dif_entova(!i,12) = eq3u{%name}.@df
cs_dif_entova(!i,13) = @qfdist(0.95,3,eq3u{%name}.@df)
if  (((eq3r{%name}.@ssr - eq3u{%name}.@ssr) * eq3u{%name}.@df) / (3 * eq3u{%name}.@ssr)
>@qfdist(0.95,3,eq3u{%name}.@df))  then
cs_dif_entova(!i,14) = "yes"
else
cs_dif_entova(!i,14) ="no"
endif

delete eq1r{%name}
delete eq1u{%name}
delete eq2r{%name}
delete eq2u{%name}
delete eq3r{%name}
delete eq3u{%name}

endif
NEXT
NEXT
```

Toda – Yamamoto causality test

```
table(1500,14) toda_entogdp
!i=0
for %reg  AUT_ BEL_ DEU_ ESP_ FRA_ GBR_ GRC_ ITA_ NLD_ POL_ PRT_ ROU_ AUS_ BGR_ BRA_ CAN_ CHN_ CYP_
CZE_ DNK_ EST_ FIN_ HUN_ IDN_ IND_ IRL_ JPN_ KOR_ LTU_ LUX_ LVA_ MEX_ MLT_ RUS_ SVK_ SVN_ SWE_ TUR_
USA_

for  %sec AtB_ C_ 15t16_ 17t18_ 19_ 20_ 21t22_ 23_ 24_ 25_ 26_ 27t28_ 29_ 30t33_ 34t35_  36t37_  E_ F_  50_ 51_ 52_
H_ 60_ 61_ 62_ 63_ 64_ J_ 70_ 71t74_ L_ M_ N_ O_ tot_
```

```
%name = %reg + %sec
%var0 = "VA_Q"
%var1 = "EN_Q"

%DEP1= %name + %var0
%IND1= %name + %var1

if  ((@min({%IND1}) <= 0) or (@min({%DEP1}) <= 0)) then
!i = !i +1
toda_entogdp(!i,1) = %reg
toda_entogdp(!i,2) = %sec
toda_entogdp(!i,3) = "n/a"
toda_entogdp(!i,4) = "n/a"
toda_entogdp(!i,5) = "n/a"
toda_entogdp(!i,6) = "n/a"
toda_entogdp(!i,7) = "n/a"
else

!i = !i + 1
'ADF is a parameter that includes the order of integration of each the series
' First column includes the cases that both series are I(0)
if adf_max_int(!i,1) = 1 then
' Model that one of the series are I(0)
equation equ1{%name}.ls   {%DEP1} C {%DEP1}(-1) {%IND1}(-1)
' Model that one of the series are I(0) with one extra lag for autocorrelation
equation equ2{%name}.ls   {%DEP1} C {%DEP1}(-1) {%DEP1}(-2)  {%IND1}(-1)   {%IND1}(-2)
' Model that one of the series are both I(0) with two extra lag for autocorrelation
equation equ3{%name}.ls   {%DEP1} C {%DEP1}(-1) {%DEP1}(-2)   {%DEP1}(-3)  {%IND1}(-1)   {%IND1}(-2)   {%IND1}(-3)

toda_entogdp(!i,1) = %reg
toda_entogdp(!i,2) = %sec
toda_entogdp(!i,3) = equ1{%name}.@tstats(3)
toda_entogdp(!i,4) = equ1{%name}.@df
toda_entogdp(!i,5) = equ1{%name}.@pvals(3)

freeze(toda) equ2{%name}.wald c(4)=0 , c(5) =0
toda_entogdp(!i,6) = toda(6,4)
delete toda

freeze(toda) equ3{%name}.wald c(5)=0 , c(6) =0,  c(7) = 0
toda_entogdp(!i,7) = toda(6,4)
delete toda

delete equ1{%name}
delete equ2{%name}
delete equ3{%name}

else
'ADF is a parameter that includes the order of integration of each the series
' Second column includes the cases that both series are I(1)
' Third column includes the cases that one of the series is I(1) and the other I(0)
if ((adf_max_int(!i,2) = 1)  or (adf_max_int(!i,3) = 1)) then
equation eqI1u1{%name}.ls {%DEP1} C {%DEP1}(-1) {%DEP1}(-2)  {%IND1}(-1)  {%IND1}(-2)
equation eqI1u2{%name}.ls {%DEP1} C {%DEP1}(-1) {%DEP1}(-2)  {%DEP1}(-3) {%IND1}(-1)  {%IND1}(-2)  {%IND1}(-3)

toda_entogdp(!i,1) = %reg
toda_entogdp(!i,2) = %sec
toda_entogdp(!i,3) = eqI1u1{%name}.@tstats(4)
toda_entogdp(!i,4) = eqI1u1{%name}.@df
toda_entogdp(!i,5) = eqI1u1{%name}.@pvals(4)

freeze(toda)  eqI1u2{%name}.wald  c(5)=0, c(6) =0
toda_entogdp(!i,6) = toda(6,4)
delete toda

delete eqI1u1{%name}
delete eqI1u2{%name}

else
toda_entogdp(!i,1) = %reg
toda_entogdp(!i,2) = %sec
toda_entogdp(!i,3) = "n/a"
toda_entogdp(!i,4) = "n/a"
toda_entogdp(!i,5) = "n/a"
toda_entogdp(!i,6) = "n/a"
```

```
toda_entogdp(!i,7) = "n/a"
endif
endif

endif

NEXT
NEXT
```

Levin, Lin, Chu and Im, Pesaran, Shin panel unit root tests

```
!j=0
Table LLC_const
Table LLC_trend
Table IPS_const
Table IPS_trend

for %sec  AtB  C  15t16  17t18  19  20  21t22  23  24  25  26  27t28  29  30t33  34t35   36t37   E  F   50  51 52 H  60 61 62  63  64  J 70
71t74  L  M N O  tot

!j = !j +1
%var0 = "va_"
%var1 = "en_"

%dep = %var0 + %sec
%ind = %var1 + %sec

Uroot(llc,maxlag=3,const,dif=0,save=llcconstdep) {%dep}
Uroot(llc,maxlag=3,const,dif=0,save=llcconstind) {%ind}
Uroot(llc,maxlag=3,trend,dif=0,save=llctrenddep) {%dep}
Uroot(llc,maxlag=3,trend,dif=0,save=llctrendind) {%ind}
Uroot(ips,maxlag=3,const,dif=0,save=ipsconstdep) {%dep}
Uroot(ips,maxlag=3,const,dif=0,save=ipsconstind) {%ind}
Uroot(ips,maxlag=3,trend,dif=0,save=ipstrenddep) {%dep}
Uroot(ips,maxlag=3,trend,dif=0,save=ipstrendind){%ind}

LLC_const(!j,1) = %sec
LLC_const(!j,2) = llcconstdep(41,2)
LLC_const(!j,3) = llcconstind(41,2)

LLC_trend(!j,1) = %sec
LLC_trend(!j,2) = llctrenddep(41,2)
LLC_trend(!j,3) = llctrendind(41,2)

 IPS_const(!j,1) = %sec
 IPS_const(!j,2) = ipsconstdep(41,2)
 IPS_const(!j,3) = ipsconstind(41,2)

IPS_trend(!j,1) = %sec
IPS_trend(!j,2) = ipstrenddep(41,2)
IPS_trend(!j,3) = ipstrendind(41,2)

delete llcconstdep llcconstind llctrenddep llctrendind ipsconstdep ipsconstind ipstrenddep ipstrendind
NEXT
```

Pedroni and Kao cointegration tests

```
!j=0
Table Pedroni_none
Table Pedroni_const
Table Pedroni_trend

for %sec  AtB  C  15t16  17t18  19  20  21t22  23  24  25  26  27t28  29  30t33  34t35   36t37   E  F   50  51 52 H  60 61 62  63  64  J 70
71t74  L  M N O  tot
!j = !j +1
%var0 = "va_"
%var1 = "en_"

%dep = %var0 + %sec
%ind = %var1 + %sec
```

```
group ped {%dep} {%ind}
'Pedroni cointegration test with no deterministic part
freeze(savenone) ped.coint(pedroni,none,maxlag=3)
'Pedroni cointegration test with a constant in the deterministic part
freeze(saveconst) ped.coint(pedroni,const,maxlag=3)
'Pedron cointegration test with a constant and tred in the deterministic part
freeze(savetrend) ped.coint(pedroni,trend,maxlag=3)
'Kao cointegration test
freeze(savekao) ped.coint(Kao, maxlag=3)

if savenone(15,4)="Prob." then
'In cases that the panels are inbalanced, an additional row is required for reporting these cases
Pedroni_none(!j,1) = %sec
' p- values of the Pedroni cointegration tests with no deterministic part
Pedroni_none(!j,2) = savenone(16,4)
Pedroni_none(!j,3) = savenone(17,4)
Pedroni_none(!j,4) = savenone(18,4)
Pedroni_none(!j,5) = savenone(19,4)
Pedroni_none(!j,6) = savenone(24,4)
Pedroni_none(!j,7) = savenone(25,4)
Pedroni_none(!j,8) = savenone(26,4)
' p- values of the Kao cointegration tests with no deterministic part
Pedroni_none(!j,9) = savekao(12,5)
else
Pedroni_none(!j,1) = %sec
' p- values of the Pedroni cointegration tests with no deterministic part
Pedroni_none(!j,2) = savenone(15,4)
Pedroni_none(!j,3) = savenone(16,4)
Pedroni_none(!j,4) = savenone(17,4)
Pedroni_none(!j,5) = savenone(18,4)
Pedroni_none(!j,6) = savenone(23,4)
Pedroni_none(!j,7) = savenone(24,4)
Pedroni_none(!j,8) = savenone(25,4)
' p- values of the Kao cointegration tests with no deterministic part
Pedroni_none(!j,9) = savekao(12,5)
endif

if savenone(15,4)="Prob." then
Pedroni_const(!j,1) = %sec
' p- values of the Pedroni cointegration test with a constant in the deterministic part
Pedroni_const(!j,2) = saveconst(16,4)
Pedroni_const(!j,3) = saveconst(17,4)
Pedroni_const(!j,4) = saveconst(18,4)
Pedroni_const(!j,5) = saveconst(19,4)
Pedroni_const(!j,6) = saveconst(24,4)
Pedroni_const(!j,7) = saveconst(25,4)
Pedroni_const(!j,8) = saveconst(26,4)
' p- values of the Kao cointegration test with a constant in the deterministic part
Pedroni_const(!j,9) = savekao(12,5)
else
Pedroni_const(!j,1) = %sec
' p- values of the Pedroni cointegration test with a constant in the deterministic part
Pedroni_const(!j,2) = saveconst(15,4)
Pedroni_const(!j,3) = saveconst(16,4)
Pedroni_const(!j,4) = saveconst(17,4)
Pedroni_const(!j,5) = saveconst(18,4)
Pedroni_const(!j,6) = saveconst(23,4)
Pedroni_const(!j,7) = saveconst(24,4)
Pedroni_const(!j,8) = saveconst(25,4)
' p- values of the Kao cointegration test with a constant in the deterministic part
Pedroni_const(!j,9) =  savekao(12,5)
endif

if savenone(15,4)="Prob." then
Pedroni_trend(!j,1) = %sec
' p- values of the Pedroni cointegration test with a constant and trend in the deterministic part
Pedroni_trend(!j,2) = savetrend(16,4)
Pedroni_trend(!j,3) = savetrend(17,4)
Pedroni_trend(!j,4) = savetrend(18,4)
Pedroni_trend(!j,5) = savetrend(19,4)
Pedroni_trend(!j,6) = savetrend(24,4)
Pedroni_trend(!j,7) = savetrend(25,4)
Pedroni_trend(!j,8) = savetrend(26,4)
' p- values of the Kao cointegration test with a constant and trend in the deterministic part
```

```
Pedroni_trend(!j,9) = savekao(12,5)
else
Pedroni_trend(!j,1) = %sec
' p- values of the Pedroni cointegration test with a constant and trend in the deterministic part
Pedroni_trend(!j,2) = savetrend(15,4)
Pedroni_trend(!j,3) = savetrend(16,4)
Pedroni_trend(!j,4) = savetrend(17,4)
Pedroni_trend(!j,5) = savetrend(18,4)
Pedroni_trend(!j,6) = savetrend(23,4)
Pedroni_trend(!j,7) = savetrend(24,4)
Pedroni_trend(!j,8) = savetrend(25,4)
' p- values of the Kao cointegration test with a constant and trend in the deterministic part
Pedroni_trend(!j,9) =  savekao(12,5)
endif

delete savenone saveconst savetrend savekao ped
next
```

Panel causality test with panel corrected standard errors

```
!j=0
Table Estimation_PCSE

for %sec  AtB  C  15t16  17t18  19  20  21t22  23  24  25  26  27t28  29  30t33  34t35   36t37   E  F   50  51 52 H  60 61 62  63  64  J 70
71t74  L  M N O  tot
!j = !j +1
%var0 = "va_"
%var1 = "en_"

%dep = %var0 + %sec
%ind = %var1 + %sec

equation eq{%dep}.ls(cx=f, cov=cxsur) {%dep} {%ind}(-1) {%dep}(-1) c

equation eq{%ind}.ls(cx=f, cov=cxsur)   {%ind}  {%dep}(-1) {%ind}(-1) c

equation eq2d.ls(cx=f, cov=cxsur) {%dep} {%ind}(-1) {%ind}(-2) {%dep}(-1) {%dep}(-2) c

equation eq2i.ls(cx=f, cov=cxsur)   {%ind}  {%dep}(-1) {%dep}(-2)  {%ind}(-1) {%ind}(-2) c

freeze(csdep) eq2d.wald c(1)=0 , c(2)=0
freeze(csind)  eq2i.wald c(1)=0 , c(2)=0

Estimation_PCSE(!j,1) = %sec
Estimation_PCSE(!j,2) = eq{%dep}.@pvals(1)
Estimation_PCSE(!j,3) = eq{%ind}.@pvals(1)
Estimation_PCSE(!j,4) = csdep(6,4)
Estimation_PCSE(!j,5) = csind(6,4)


delete eq{%dep} eq{%ind} eq2d eq2i csdep csind
next
```

FMOLS estimator for panel causality test

```
!j=0
Table fmols

for %sec  AtB  C  15t16  17t18  19  20  21t22  23  24  25  26  27t28  29  30t33  34t35   36t37   E  F   50  51 52 H  60 61 62  63  64  J 70
71t74  L  M N O  tot
!j = !j +1
%var0 = "va_"
%var1 = "en_"

%dep = %var0 + %sec
%ind = %var1 + %sec

equation eq{%dep}.cointreg(method=fmols ,trend = const , lltype = sic , pancov = sandwich) {%dep} {%ind}(-1) {%dep}(-1)
equation eq{%ind}.cointreg(method=fmols ,trend = const , lltype = sic ,  pancov = sandwich) {%ind}  {%dep}(-1) {%ind}(-1)
```

```
equation eq2d.cointreg(method=fmols ,trend = const , lltype = sic , pancov = sandwich) {%dep} {%ind}(-1) {%ind}(-2) {%dep}(-1) {%dep}(-2)
equation eq2i.cointreg(method=fmols ,trend = const , lltype = sic ,  pancov = sandwich)  {%ind}  {%dep}(-1) {%dep}(-2)  {%ind}(-1) {%ind}(-2)

freeze(csdep) eq2d.wald c(1)=0 , c(2)=0
freeze(csind)  eq2i.wald c(1)=0 , c(2)=0


fmols(!j,1) = %sec
fmols(!j,2) = eq{%dep}.@pvals(1)
fmols(!j,3) = eq{%ind}.@pvals(1)
fmols(!j,4) = csdep(6,4)
fmols(!j,5) = csind(6,4)

delete eq{%dep} eq{%ind} eq2d eq2i csdep csind
next
```

DOLS estimator for panel causality test

```
!j=0
Table DOLS

for %sec  AtB  C  15t16  17t18  19  20  21t22  23  24  25  26  27t28  29  30t33  34t35   36t37   E  F   50  51 52 H  60 61 62  63  64  J 70 71t74  L  M N O  tot
!j = !j +1
%var0 = "va_"
%var1 = "en_"

%dep = %var0 + %sec
%ind = %var1 + %sec

equation eq{%dep}.cointreg(method=DOLS ,trend = const , lltype = sic , pancov = sandwich) {%dep} {%ind}(-1) {%dep}(-1)
equation eq{%ind}.cointreg(method=DOLS ,trend = const , lltype = sic ,  pancov = sandwich) {%ind}  {%dep}(-1) {%ind}(-1)

DOLS(!j,1) = %sec
DOLS(!j,2) = eq{%dep}.@pvals(1)
DOLS(!j,3) = eq{%ind}.@pvals(1)

delete eq{%dep} eq{%ind}
next
```

Panel ECM causality test

```
!j=0
Table ECM

for %sec  AtB  C  15t16  17t18  19  20  21t22  23  24  25  26  27t28  29  30t33  34t35   36t37   E  F   50  51 52 H  60 61 62  63  64  J 70 71t74  L  M N O  tot

!j = !j +1
%var0 = "va_"
%var1 = "en_"
%dep = %var0 + %sec
%ind = %var1 + %sec

equation eqcoint.cointreg(method=fmols ,trend = const , lltype = sic , pancov = sandwich) {%dep}  {%ind}
eqcoint.makeresid res

equation ecm1.ls D({%dep}) c res(-1) D({%ind}(-1)) D({%dep}(-1))
equation ecm2.ls D({%ind}) c res(-1) D({%dep}(-1)) D({%ind}(-1))

equation ecm3.ls D({%dep}) c res(-1) D({%ind}(-1)) D({%ind}(-2)) D({%dep}(-1)) D({%dep}(-2))
equation ecm4.ls D({%ind}) c res(-1) D({%dep}(-1)) D({%dep}(-2)) D({%ind}(-1)) D({%ind}(-2))

freeze(str1) ecm1.wald c(2)=0 , c(3)=0
freeze(str2) ecm2.wald c(2)=0 , c(3)=0
freeze(sr3)  ecm3.wald c(3)=0 , c(4)=0
freeze(str3) ecm3.wald c(2)=0 , c(3)=0 , c(4)=0
freeze(sr4)  ecm4.wald c(3)=0 , c(4)=0
```

```
freeze(str4) ecm4.wald c(2)=0 , c(3)=0 , c(4)=0

ECM(!j,1) = %sec
ECM(!j,2) = ecm1.@pvals(2)
ECM(!j,3) = ecm1.@pvals(3)
ECM(!j,4) = str1(6,4)
ECM(!j,5) = ecm2.@pvals(2)
ECM(!j,6) = ecm2.@pvals(3)
ECM(!j,7) = str2(6,4)
ECM(!j,8) = ecm3.@pvals(2)
ECM(!j,9) = sr3(6,4)
ECM(!j,10) = str3(6,4)
ECM(!j,11) = ecm4.@pvals(2)
ECM(!j,12) = sr4(6,4)
ECM(!j,13) = str4(6,4)

delete eqcoint res str1 str2 str3 str4 sr3 sr4 ecm1 ecm2 ecm3 ecm4
next
```

Zivot-Andrews Unit Root Test

```
'Zivot-Andrews Unit Root Test
table(3000,14) zivot_c_q_3lag

!j=0
for %reg  AUT_ BEL_ DEU_ ESP_ FRA_ GBR_ GRC_ ITA_ NLD_ POL_ PRT_ ROU_ AUS_ BGR_ BRA_ CAN_ CHN_ CYP_
CZE_ DNK_ EST_ FIN_ HUN_ IDN_ IND_ IRL_ JPN_ KOR_ LTU_ LUX_ LVA_ MEX_ MLT_ RUS_ SVK_ SVN_ SWE_ TUR_
USA_

for %sec AtB_ C_ 15t16_ 17t18_ 19_ 20_ 21t22_ 23_ 24_ 25_ 26_ 27t28_ 29_ 30t33_ 34t35_ 36t37_ E_ F_ 50_ 51_ 52_
H_ 60_ 61_ 62_ 63_ 64_ J_ 70_ 71t74_ L_ M_ N_ O_

table(7,2) ZAZ
  %name = %reg + %sec
   %var0 = "q"
   %var1 = "p"

%DEP1= %name + %var0
%IND1= %name + %var1

if ((@min({%IND1}) <= 0) or (@min({%DEP1}) <= 0)) then
!j = !j +1
zivot_c_q_3lag(!j,1) = %reg
zivot_c_q_3lag(!j,2) = %sec
zivot_c_q_3lag(!j,3) = "n/a"
zivot_c_q_3lag(!j,4) = "n/a"
zivot_c_q_3lag(!j,5) = "n/a"
zivot_c_q_3lag(!j,6) = "n/a"
zivot_c_q_3lag(!j,7) = "n/a"
else
call zivot(log({%DEP1}),"C",3)
!j = !j +1
zivot_c_q_3lag(!j,1) = %reg
zivot_c_q_3lag(!j,2) = %sec
zivot_c_q_3lag(!j,3) = ZAZ(3,2)
zivot_c_q_3lag(!j,4) = ZAZ(4,2)
zivot_c_q_3lag(!j,5) = ZAZ(5,2)
zivot_c_q_3lag(!j,6) = ZAZ(6,2)
zivot_c_q_3lag(!j,7) = ZAZ(7,2)
endif
delete zaz
next
next

'Zivot Andrews subroutine test
'Reference: Zivot, E. and Andrews, D. W. K. (1992), "Further Evidence on the Great Crash, the Oil-Price Shock, and the Unit-Root
Hypothesis", Journal of Business & Economic Statistics, Vol. 10, No. 3, pp. 251-270.

subroutine zivot(series y,string %Model,scalar maxlag)
!trim = 0.15 'Trimming parameter
series DY = D(Y)

!nobs = @obs(y)-1
```

```
equation temp.ls dy c @trend y(-1)
!aic0 = log(temp.@ssr/!nobs)+2*(temp.@ncoef/!nobs)
!bic0 = log(temp.@ssr/!nobs)+ log(!nobs)*(temp.@ncoef/!nobs)
!min_aic = !aic0
!min_bic = !bic0
!nobs = @obs(y)-1-maxlag
for !lag=maxlag to 1 step -1
  equation temp.ls dy y(-1) c @trend dy(-1 to -!lag)
  !aic = log(temp.@ssr/!nobs)+2*(temp.@ncoef/!nobs)
  !bic = log(temp.@ssr/!nobs)+log(!nobs)*(temp.@ncoef/!nobs)
      if !bic < !min_bic then
         !min_aic = !aic
         !min_bic = !bic
         !best_lag = !lag
         else if !min_bic = !bic0 then
         !best_lag =0
         endif
     endif
next

!znobs = @obs(y) - !best_lag
!lower = 1+!best_lag+@round(!znobs*!trim)
!upper =  @obs(y)-@round(!znobs*!trim)

vector(!upper-!lower+1) results

for !i = !lower+1 to !upper
  if !best_lag=0 and %Model = "A" then
     equation temp.ls DY Y(-1) C @trend (@trend>!i-2)
     else if !best_lag=0 and %Model = "B" then
     equation temp.ls DY Y(-1) C @trend (@trend>!i-2)*(@trend-(!i-2))
     else if !best_lag=0 and %Model = "C" then
     equation temp.ls DY Y(-1) C @trend (@trend>!i-2) (@trend>!i-2)*(@trend-(!i-2))
     else if !best_lag>0 and %Model = "A" then
     equation temp.ls DY Y(-1) C @trend (@trend>!i-2) DY(-1 to -!best_lag)
     else if !best_lag>0 and %Model = "B" then
     equation temp.ls DY Y(-1) C @trend (@trend>!i-2)*(@trend-(!i-2)) DY(-1 to -!best_lag)
     else if !best_lag>0 and %Model = "C" then
     equation temp.ls DY Y(-1) C @trend (@trend>!i-2) (@trend>!i-2)*(@trend-(!i-2)) DY(-1 to -!best_lag)
     endif
     endif
     endif
     endif
     endif
  endif
  results(!i-!lower) = temp.@tstats(1)
next

'The minimum from the results vector stored in t_min(1)
vector t_min =@cmin(results)
!t_min = t_min(1)
vector break = @cimin(results)+!lower
'# of obsrevation with the minimum break(1)
!break = break(1)

series DT = (@trend>!break-2)*(@trend-(!break-2))
if %Model = "A" or %Model="C" then
   series DU = @trend> !break-2
endif

'Select the model equation
if !best_lag=0 and %Model="A" then
   equation ZA.ls DY Y(-1) C @trend DU
   else if !best_lag=0 and %Model="B" then
   equation ZA.ls DY Y(-1) C @trend DT
   else if !best_lag=0 and %Model="C" then
   equation ZA.ls DY Y(-1) C @trend DU DT
   else if !best_lag>0 and %Model = "A" then
   equation ZA.ls DY Y(-1) C @trend DU DY(-1 to -!best_lag)
   else if !best_lag>0 and %Model = "B" then
   equation ZA.ls DY Y(-1) C @trend DT DY(-1 to -!best_lag)
   else if !best_lag>0 and %Model = "C" then
   equation ZA.ls DY Y(-1) C @trend DU DT DY(-1 to -!best_lag)
   endif
```

```
   endif
   endif
   endif
   endif
endif

'Store values for reporting
ZAZ(1,1) = "Variable(s)"
ZAZ(3,1) = "t-stat(s)"
ZAZ(4,1) = "Lag(s)"
ZAZ(5,1) = "Break"
ZAZ(6,1) = "DU1 p-value"
ZAZ(1,2) = y.@name
ZAZ(3,2) = !t_min
ZAZ(4,2) = !best_lag
if @datestr(@now,"F") = "?" then
  ZAZ(5,2) = !break-1
else
  ZAZ(5,2) = @otod(!break-1)
endif
ZAZ(6,2) = @tdist(za.@tstat(4),za.@regobs-za.@ncoef)
zaz(7,2)  = za.@pvals(1)
setline(ZAZ, 2)
smpl @all

delete temp break results t_min DY
endsub
```

OLS estimation of the elasticity of substitution and diagnostic test

```
!j=0
table OLS

for %reg   AUT_ BEL_ DEU_ ESP_ FRA_ GBR_ GRC_ ITA_ NLD_ POL_ PRT_ ROU_ AUS_ BGR_ BRA_ CAN_ CHN_ CYP_
CZE_ DNK_ EST_ FIN_ HUN_ IDN_ IND_ IRL_ JPN_ KOR_ LTU_ LUX_ LVA_ MEX_ MLT_ RUS_ SVK_ SVN_ SWE_ TUR_
USA_

for  %sec AtB_  C_  15t16_  17t18_  19_  20_  21t22_  23_  24_  25_  26_  27t28_  29_  30t33_  34t35_   36t37_   E_  F_   50_  51_  52_
H_  60_ 61_ 62_  63_  64_  J_ 70_  71t74_  L_  M_ N_ O_

  %name = %reg + %sec
    %var0 = "q"
    %var1 = "p"

%DEP1= %name + %var0
%IND1= %name + %var1

if  ((@min({%IND1}) <= 0) or (@min({%DEP1})) <= 0) then
!j = !j +1
OLS(!j,1) = %reg
OLS(!j,2) = %sec
OLS(!j,3) ="n/a"
OLS(!j,4) = "n/a"
OLS(!j,5) = "n/a"
OLS(!j,6) = "n/a"
OLS(!j,7) = "n/a"
OLS(!j,8) = "n/a"
OLS(!j,9) = "n/a"
OLS(!j,10) = "n/a"
OLS(!j,11) = "n/a"
OLS(!j,12) = "n/a"
OLS(!j,13) = "n/a"
else
!j=!j+1
equation  estim.ls(cov="HAC") log({%dep1}) c log({%ind1})
var          jb.ls 0 0 log({%dep1}) @ c log({%ind1})
equation  estimtrend.ls(cov="HAC") log({%dep1}) c log({%ind1}) @trend
var          jbtrend.ls 0 0 log({%dep1}) @ c log({%ind1}) @trend

freeze(lm)      estim.auto(3)
freeze(white) estim.white
freeze(norm) jb.jbera(factor=chol)
freeze(lmtrend)      estimtrend.auto(3)
```

```
freeze(whitetrend) estimtrend.white
freeze(normtrend) jbtrend.jbera(factor=chol)

OLS(!j,1) = %reg
OLS(!j,2) = %sec
OLS(!j,3) = estim.c(2)
OLS(!j,4) = estim.@pvals(2)
OLS(!j,5) = estimtrend.c(2)
OLS(!j,6) = estimtrend.@pvals(2)
OLS(!j,7) = estimtrend.@pvals(3)
OLS(!j,8) = lm(3,5)
OLS(!j,9) = white(3,5)
OLS(!j,10) = norm(25,4)
OLS(!j,11) = lmtrend(3,5)
OLS(!j,12) = whitetrend(3,5)
OLS(!j,13) = normtrend(25,4)

delete estim jb lm white norm lmtrend whitetrend normtrend estimtrend jbtrend

endif
next
next
```

OLS estimation of the elasticity of substitution at the first differences

```
table(2000,14) old_fd_estimates

!i = 0

for %reg  AUT_ BEL_ DEU_ ESP_ FRA_ GBR_ GRC_ ITA_ NLD_ POL_ PRT_ ROU_ AUS_ BGR_ BRA_ CAN_ CHN_ CYP_
CZE_ DNK_ EST_ FIN_ HUN_ IDN_ IND_ IRL_ JPN_ KOR_ LTU_ LUX_ LVA_ MEX_ MLT_ RUS_ SVK_ SVN_ SWE_ TUR_
USA_

for %sec AtB_  C_  15t16_  17t18_  19_  20_  21t22_  23_  24_  25_  26_  27t28_  29_  30t33_  34t35_  36t37_  E_  F_  50_  51_ 52_
H_  60_ 61_ 62_  63_  64_  J_ 70_  71t74_  L_  M_ N_ O_

' Define names of the variables dynamically
%name = %reg + %sec
%var0    = "q"
%var1    = "p"

' Define the dependent and independent variables of the regression
%DEP1 = %name + %var0
%IND1  =  %name + %var1

if  ((@min({%IND1}) <= 0) or (@min({%DEP1}) <= 0)) then
!i = !i + 1
      old_fd_estimates(!i, 1) = %reg
      old_fd_estimates(!i, 2) = %sec
      old_fd_estimates(!i, 3) = "n/a"
      old_fd_estimates(!i, 4) = "n/a"
      old_fd_estimates(!i, 5) = "n/a"
      old_fd_estimates(!i, 6) = "n/a"
      old_fd_estimates(!i, 7) = "n/a"
      old_fd_estimates(!i, 8) = "n/a"
      old_fd_estimates(!i, 9) = "n/a"
else
      equation eq{%name}.ls(cov = "hac")  D(log({%DEP1})) D(log({%IND1})) c
' Store the results of each regression i=1 for the first equation, i=2 for the second etc.
      !i = !i + 1
     old_fd_estimates(!i, 1) = %reg
      old_fd_estimates(!i, 2) = %sec
' Store the estimation for the coefficient C(1) from the regression [substitution of elasticity]
      old_fd_estimates(!i, 3) = eq{%name}.c(1)
' Store the standard error
      old_fd_estimates(!i, 4) = eq{%name}.@tstats(1)
' Store the p-value
      old_fd_estimates(!i, 5) = eq{%name}.@pvals(1)
' Store the estimation for the coefficient C(2) from the regression
      old_fd_estimates(!i, 6) = eq{%name}.c(2)
' Store the t-statistic
      old_fd_estimates(!i, 7) = eq{%name}.@pvals(2)
' Store the R-squared
```

```
    old_fd_estimates(!i, 8) = eq{%name}.@r2
'Store the degrees of freedom
    old_fd_estimates(!i,9) = eq{%name}.@df
delete eq{%name}
endif
next
next
```

ECM estimation of the elasticity of substitution

```
!j=0
table  eg_none
table  eg_const
table  eg_trend
table eg_estim_none
table eg_estim_const
table eg_estim_trend

scalar mc_binf
scalar mc_b1
scalar mc_b2
scalar mc_b3

for %reg  AUT_ BEL_ DEU_ ESP_ FRA_ GBR_ GRC_ ITA_ NLD_ POL_ PRT_ ROU_ AUS_ BGR_ BRA_ CAN_ CHN_ CYP_
CZE_ DNK_ EST_ FIN_ HUN_ IDN_ IND_ IRL_ JPN_ KOR_ LTU_ LUX_ LVA_ MEX_ MLT_ RUS_ SVK_ SVN_ SWE_ TUR_
USA_

for  %sec AtB_  C_  15t16_  17t18_  19_  20_  21t22_  23_  24_  25_  26_  27t28_  29_  30t33_  34t35_   36t37_   E_  F_   50_  51_ 52_
H_  60_ 61_ 62_  63_  64_  J_ 70_  71t74_  L_  M_ N_ O_

  %name = %reg + %sec
  %var0 = "q"
   %var1 = "p"

%DEP1= %name + %var0
%IND1= %name + %var1

if  ((@min({%IND1}) <= 0) or (@min({%DEP1})) <= 0) then
!j = !j +1
    eg_none(!j,1) = %reg
    eg_none(!j,2) = %sec
    eg_none(!j,3) = "n/a"
    eg_none(!j,4) = "n/a"
    eg_none(!j,5) = "n/a"
    eg_none(!j,6) = "n/a"
    eg_none(!j,7) = "n/a"

   eg_const(!j,1) = %reg
   eg_const(!j,2) = %sec
   eg_const(!j,3) = "n/a"
   eg_const(!j,4) = "n/a"
   eg_const(!j,5) = "n/a"
   eg_const(!j,6) = "n/a"
   eg_const(!j,7) = "n/a"

   eg_trend(!j,1) = %reg
   eg_trend(!j,2) = %sec
   eg_trend(!j,3) = "n/a"
   eg_trend(!j,4) = "n/a"
   eg_trend(!j,5) = "n/a"
   eg_trend(!j,6) = "n/a"
   eg_trend(!j,7) = "n/a"

eg_estim_none(!j,1) = %reg
eg_estim_none(!j,2) = %sec
eg_estim_none(!j,3) = "n/a"
eg_estim_none(!j,4) = "n/a"
eg_estim_none(!j,5) = "n/a"
eg_estim_none(!j,6) = "n/a"
eg_estim_none(!j,7) = "n/a"
eg_estim_none(!j,8) = "n/a"
eg_estim_none(!j,9) = "n/a"
```

```
eg_estim_const(!j,1) = %reg
eg_estim_const(!j,2) = %sec
eg_estim_const(!j,3) = "n/a"
eg_estim_const(!j,4) = "n/a"
eg_estim_const(!j,5) = "n/a"
eg_estim_const(!j,6) = "n/a"
eg_estim_const(!j,7) = "n/a"
eg_estim_const(!j,8) = "n/a"
eg_estim_const(!j,9) = "n/a"

eg_estim_trend(!j,1) = %reg
eg_estim_trend(!j,2) = %sec
eg_estim_trend(!j,3) = "n/a"
eg_estim_trend(!j,4) = "n/a"
eg_estim_trend(!j,5) = "n/a"
eg_estim_trend(!j,6) = "n/a"
eg_estim_trend(!j,7) = "n/a"
eg_estim_trend(!j,8) = "n/a"
eg_estim_trend(!j,9) = "n/a"

else

!j=!j+1

call granger(log({%DEP1}),log({%IND1}),1,3)
call granger(log({%DEP1}),log({%IND1}),2,3)
call granger(log({%DEP1}),log({%IND1}),3,3)

endif
next
next
' ----------------------------------------------------------------------------------------------------
' Arguments
'-----------------------------------------------------------------------------------------------------
'series Y              ' dependent variable
'Series G              '  independent variable
'scalar Model          ' 1 = None, 2 = Consant, 3 = Consant and Trend
'scalar Maxlag         ' Maximum number of lags for unit root testing
' ----------------------------------------------------------------------------------------------------

subroutine granger(series Y, series G, scalar Model,scalar Maxlag)
smpl @all

!maxlag=Maxlag

equation gr

'MODEL 1 : none
if Model=1 then

'McKinnon formula (5% level of significance)
mc_binf = -1.941
mc_b1 = - 0.269
mc_b2 = -3.365
mc_b3 = 31.223

    gr.ls Y  G
    gr.makeresid res
    uroot(adf, none, maxlag=!maxlag, save=level) res
   eg_none(!j,1) = %reg
   eg_none(!j,2) = %sec
   eg_none(!j,3) = level(3,1)
   eg_none(!j,4) = level(2,1)
   eg_none(!j,5) =  level(4,1)
   eg_none(!j,6) =  level(1,1)
   eg_none(!j,7) = mc_binf + mc_b1/eg_none(!j,6) + mc_b2/(eg_none(!j,6)^2) + mc_b3/(eg_none(!j,6)^3)

equation eg.ls d(y) c res(-1) d(g) d(y(-1)) d(g(-1))

eg_estim_none(!j,1) = %reg
eg_estim_none(!j,2) = %sec
eg_estim_none(!j,3) = eg.c(2)
eg_estim_none(!j,4) = eg.@tstats(2)
eg_estim_none(!j,5) = eg.@pvals(2)
eg_estim_none(!j,6) = eg.c(3)
```

```
eg_estim_none(!j,7) = eg.@pvals(3)
eg_estim_none(!j,8) = gr.c(1)
eg_estim_none(!j,9) = gr.@pvals(1)

'MODEL 2 : constant
else if Model=2 then

'McKinnon formula (5% level of significance)
mc_binf = -2.862
mc_b1 = -2.890
mc_b2 = -4.234
mc_b3 = -40.040

    gr.ls Y c  G
    gr.makeresid res
    uroot(adf, none, maxlag=!maxlag, save=level) res
   eg_const(!j,1) = %reg
   eg_const(!j,2) = %sec
   eg_const(!j,3) = level(3,1)
   eg_const(!j,4) = level(2,1)
   eg_const(!j,5) = level(4,1)
   eg_const(!j,6) =  level(1,1)
   eg_const(!j,7) = mc_binf + mc_b1/eg_const(!j,6) + mc_b2/(eg_const(!j,6)^2) + mc_b3/(eg_const(!j,6)^3)

equation eg.ls d(y) c res(-1) d(g) d(y(-1)) d(g(-1))

eg_estim_const(!j,1) = %reg
eg_estim_const(!j,2) = %sec
eg_estim_const(!j,3) = eg.c(2)
eg_estim_const(!j,4) = eg.@tstats(2)
eg_estim_const(!j,5) = eg.@pvals(2)
eg_estim_const(!j,6) = eg.c(3)
eg_estim_const(!j,7) = eg.@pvals(3)
eg_estim_const(!j,8) = gr.c(2)
eg_estim_const(!j,9) = gr.@pvals(2)

'MODEL 3 : with linear trend
else if Model = 3 then

'McKinnon formula (5% level of significance)
mc_binf = -3.410
mc_b1 = -4.3904
mc_b2 = -9.036
mc_b3 = -45.37

    gr.ls Y c  G @trend
    gr.makeresid res
    uroot(adf, none, maxlag=!maxlag, save=level) res
    eg_trend(!j,1) = %reg
    eg_trend(!j,2) = %sec
    eg_trend(!j,3) = level(3,1)
    eg_trend(!j,4) = level(2,1)
    eg_trend(!j,5) = level(4,1)
    eg_trend(!j,6) =  level(1,1)
    eg_trend(!j,7) = mc_binf + mc_b1/eg_trend(!j,6) + mc_b2/(eg_trend(!j,6)^2) + mc_b3/(eg_trend(!j,6)^3)

equation eg.ls d(y) c res(-1) d(g) d(y(-1)) d(g(-1))

eg_estim_trend(!j,1) = %reg
eg_estim_trend(!j,2) = %sec
eg_estim_trend(!j,3) = eg.c(2)
eg_estim_trend(!j,4) = eg.@tstats(2)
eg_estim_trend(!j,5) = eg.@pvals(2)
eg_estim_trend(!j,6) = eg.c(3)
eg_estim_trend(!j,7) = eg.@pvals(3)
eg_estim_trend(!j,8) = gr.c(2)
eg_estim_trend(!j,9) = gr.@pvals(2)
endif
endif
endif

delete level gr eg
endsub
```

ECM estimation of the elasticity of substitution with possible structural break

```
'Gregory-Hansen Cointegration Test
'Reference: Gregory, A. W. and Hansen, B. E. (1996). "Residual-Based Tests for Cointegration in Models with Regime Shifts", Journal of Econometrics, Vol. 70, pp. 99-126.
table(2000,30) greg_hans_test_a
!df=0
for %reg  AUT_ BEL_ DEU_ ESP_ FRA_ GBR_ GRC_ ITA_ NLD_ POL_ PRT_ ROU_ AUS_ BGR_ BRA_ CAN_ CHN_ CYP_ CZE_ DNK_ EST_ FIN_ HUN_ IDN_ IND_ IRL_ JPN_ KOR_ LTU_ LUX_ LVA_ MEX_ MLT_ RUS_ SVK_ SVN_ SWE_ TUR_ USA_

for  %sec AtB_  C_  15t16_  17t18_  19_  20_  21t22_  23_  24_  25_  26_  27t28_  29_  30t33_  34t35_   36t37_   E_  F_   50_  51_ 52_ H_  60_ 61_ 62_  63_  64_  J_ 70_  71t74_  L_  M_ N_ O_

%name = %reg + %sec
%var0 = "q"
%var1 = "p"

%DEP1= %name + %var0
%IND1= %name + %var1

if  ((@min({%IND1}) <= 0) or (@min({%DEP1}) <= 0)) then
!df = !df +1
greg_hans_test_a(!df,1) = %reg
greg_hans_test_a(!df,2) = %sec
greg_hans_test_a(!df,3) = "n/a"
greg_hans_test_a(!df,4) = "n/a"
greg_hans_test_a(!df,5) = "n/a"
greg_hans_test_a(!df,6) = "n/a"
greg_hans_test_a(!df,7) = "n/a"
greg_hans_test_a(!df,8) = "n/a"
greg_hans_test_a(!df,9) = "n/a"
greg_hans_test_a(!df,10) = "n/a"
greg_hans_test_a(!df,11) = "n/a"
greg_hans_test_a(!df,12) = "n/a"
greg_hans_test_a(!df,13) = "n/a"
greg_hans_test_a(!df,14) = "n/a"
greg_hans_test_a(!df,15) = "n/a"
greg_hans_test_a(!df,16) = "n/a"
greg_hans_test_a(!df,17) = "n/a"
greg_hans_test_a(!df,18) = "n/a"
greg_hans_test_a(!df,19) = "n/a"
else
group aa log({%IND1})
call greghansen(log({%DEP1}),aa,2,"sic",3)
delete aa

!df = !df +1
greg_hans_test_a(!df,1) = %reg
greg_hans_test_a(!df,2) = %sec
greg_hans_test_a(!df,3) = GHZ(7,2)
greg_hans_test_a(!df,4) = GHZ(8,2)
greg_hans_test_a(!df,5) = GHZ(9,2)
greg_hans_test_a(!df,6) = GHZ(17,2)
greg_hans_test_a(!df,7) = GHZ(18,2)
greg_hans_test_a(!df,8) = GHZ(19,2)
greg_hans_test_a(!df,9) = GHZ(20,2)
greg_hans_test_a(!df,10) = GHZ(21,2)
greg_hans_test_a(!df,11) = GHZ(22,2)
greg_hans_test_a(!df,12) = GHZ(23,2)
greg_hans_test_a(!df,13) = GHZ(24,2)
greg_hans_test_a(!df,14) = GHZ(25,2)
greg_hans_test_a(!df,15) = GHZ(26,2)
greg_hans_test_a(!df,16) = GHZ(27,2)
greg_hans_test_a(!df,17) = GHZ(28,2)
greg_hans_test_a(!df,18) = GHZ(29,2)
greg_hans_test_a(!df,19) = GHZ(30,2)
endif
next
next
' ----------------------------------------------------------------------------------------------------
' Arguments
'-----------------------------------------------------------------------------------------------------
'series Y              ' dependent variable
'group G               ' group of independent variable(s) (including single series)
```

```
'scalar Model        ' 2 = Level Shift, 3 = Level Shift with Trend, 4 = Regime Shift
'scalar Maxlag       ' Maximum number of lags for unit root testing
'string %Criterion  ' Selection criteria for unit root testing (i.e. aic / sic / hqc)
' -----------------------------------------------------------------------------------------------------
subroutine greghansen(series Y, group G, scalar Model, string %Criterion, scalar Maxlag)
smpl @all
!trim = 0.15
!maxlag = Maxlag

!n = @obs(y)
!nindep = G.@count

!lower = @round(@obs(Y)*!trim)
!upper = @round(@obs(Y)*(1-!trim))
matrix(!upper-!lower+1,4) GHtest

equation ghc

Table GHZ

for !i = !lower to !upper

if Model=2 then
'MODEL 2 - C: LEVEL SHIFT MODEL
    ghc.ls Y c G (@trend>!i)
    ghc.makeresid res
    uroot(adf, none, info={%criterion}, maxlag=!maxlag, save=level) res
    GHtest(!i-!lower+1,1) = level(3,1)
    GHtest(!i-!lower+1,2) = level(2,1)
else if Model=3 then
'MODEL 3 - C/T: LEVEL SHIFT WITH TREND MODEL
    ghc.ls Y c @trend G (@trend>!i)
    ghc.makeresid res
    uroot(adf, none, info={%criterion}, maxlag=!maxlag, save=level) res
    GHtest(!i-!lower+1,1) = level(3,1)
    GHtest(!i-!lower+1,2) = level(2,1)
else if Model = 4 then
'MODEL 4 - C/S: REGIME SHIFT MODEL
    for !g = 1 to !nindep
    G.add (@trend>!i)*G(!g)
    next
    ghc.ls Y c (@trend>!i) G
    ghc.makeresid res
    uroot(adf, none, info={%criterion}, maxlag=!maxlag, save=level) res
    GHtest(!i-!lower+1,1) = level(3,1)
    GHtest(!i-!lower+1,2) = level(2,1)
    for !g = G.@count to !nindep+1 step -1
    %name = G.@seriesname(!g)
    G.drop {%name}
    next
endif
endif
endif
next
    vector min_t_lag = @cmin(GHtest)
    vector break = @cimin(GHtest)

    GHZ(7,2) = min_t_lag(1)
    GHZ(8,2) = GHtest(break(1),2)
   if @datestr(@now,"F") = "?" then
    GHZ(9,2) = break(1) + !lower
   else
   GHZ(9,2) = @otod(break(1) + !lower )
   endif
equation estim.ls Y c G (@trend> break(1) + !lower - 1)
estim.makeresid resgreg
series ind = G(1)
equation ecm.ls D(Y) c resgreg(-1) D(ind)

  GHZ(17,2) = ecm.c(2)
  GHZ(18,2) = ecm.@pvals(2)
  GHZ(19,2) = ecm.c(3)
  GHZ(20,2) = ecm.@pvals(3)

equation ecm1.ls D(Y) c resgreg(-1) D(ind) D(Y(-1)) D(ind(-1))
```

```
 GHZ(21,2) = ecm1.c(2)
 GHZ(22,2) = ecm1.@pvals(2)
 GHZ(23,2) = ecm1.c(3)
 GHZ(24,2) = ecm1.@pvals(3)
 GHZ(25,2) = ecm1.c(4)
 GHZ(26,2) = ecm1.@pvals(4)
 GHZ(27,2) = ecm1.c(5)
 GHZ(28,2) = ecm1.@pvals(5)
 GHZ(29,2) = estim.c(2)
 GHZ(30,2) = estim.@pvals(2)

delete res level GHtest break min_t_lag estim ecm resgreg ind ecm1
endsub
```

ARDL estimation of the elasticity of substitution

```
!j=0
table ardl_11_const
table ardl_22_const
for %reg  AUT_ BEL_ DEU_ ESP_ FRA_ GBR_ GRC_ ITA_ NLD_ POL_ PRT_ ROU_ AUS_ BGR_ BRA_ CAN_ CHN_ CYP_
CZE_ DNK_ EST_ FIN_ HUN_ IDN_ IND_ IRL_ JPN_ KOR_ LTU_ LUX_ LVA_ MEX_ MLT_ RUS_ SVK_ SVN_ SWE_ TUR_
USA_

for  %sec AtB_  C_  15t16_  17t18_  19_  20_  21t22_  23_  24_  25_  26_  27t28_  29_  30t33_  34t35_   36t37_   E_  F_   50_  51_ 52_
H_  60_ 61_ 62_  63_  64_  J_ 70_  71t74_  L_  M_ N_ O_

%name = %reg + %sec
%var0 = "q"
%var1 = "p"

%DEP1= %name + %var0
%IND1= %name + %var1

if  ((@min({%IND1}) <= 0) or (@min({%DEP1}) <= 0)) then
!j = !j +1
ardl_11_const(!j,1)=%reg
ardl_11_const(!j,2)=%sec
ardl_11_const(!j,3)="n/a"
ardl_11_const(!j,4)="n/a"
ardl_11_const(!j,5)= "n/a"
ardl_11_const(!j,6)= "n/a"
ardl_11_const(!j,7)= "n/a"
ardl_11_const(!j,8)= "n/a"
ardl_11_const(!j,9)= "n/a"

ardl_22_const(!j,1)=%reg
ardl_22_const(!j,2)=%sec
ardl_22_const(!j,3)="n/a"
ardl_22_const(!j,4)="n/a"
ardl_22_const(!j,5)= "n/a"
ardl_22_const(!j,6)= "n/a"
ardl_22_const(!j,7)= "n/a"
ardl_22_const(!j,8)= "n/a"
ardl_22_const(!j,9)= "n/a"
else

!j=!j+1

equation ardl11.ls d(log({%dep1})) c log({%dep1}(-1)) log({%ind1}(-1))  d(log({%ind1}))
equation ardl22.ls d(log({%dep1})) c log({%dep1}(-1)) log({%ind1}(-1))  d(log({%ind1}))   d(log({%dep1}(-1)))  d(log({%ind1}(-1)))
freeze(a) ardl11.wald c(2)=0 , c(3) =0
freeze(aa) ardl22.wald c(2)=0 , c(3) =0

ardl_11_const(!j,1)=%reg
ardl_11_const(!j,2)=%sec
ardl_11_const(!j,3)=a(6,2)
ardl_11_const(!j,4)= ardl11.c(2)
ardl_11_const(!j,5)= ardl11.@tstats(2)
ardl_11_const(!j,6)= ardl11.c(3)
ardl_11_const(!j,7)= ardl11.@tstats(3)
ardl_11_const(!j,8)= ardl11.c(4)
ardl_11_const(!j,9)= ardl11.@pvals(4)
ardl_11_const(!j,10)= ardl11.@schwarz
```

```
ardl_22_const(!j,1)=%reg
ardl_22_const(!j,2)=%sec
ardl_22_const(!j,3)=aa(6,2)
ardl_22_const(!j,4)=ardl22.c(2)
ardl_22_const(!j,5)= ardl22.@tstats(2)
ardl_22_const(!j,6)= ardl22.c(3)
ardl_22_const(!j,7)= ardl22.@tstats(3)
ardl_22_const(!j,8)= ardl22.c(4)
ardl_22_const(!j,9)= ardl22.@pvals(4)
ardl_22_const(!j,10)= ardl22.@schwarz

delete ardl11 a aa ardl22
endif
next
next
```

NARDL estimation of the elasticity of substitution

```
table(3000,16) nardl_OLS_1lag

!i=0
for %reg  AUT_ BEL_ DEU_ ESP_ FRA_ GBR_ GRC_ ITA_ NLD_ POL_ PRT_ ROU_ AUS_ BGR_ BRA_ CAN_ CHN_ CYP_
CZE_ DNK_ EST_ FIN_ HUN_ IDN_ IND_ IRL_ JPN_ KOR_ LTU_ LUX_ LVA_ MEX_ MLT_ RUS_ SVK_ SVN_ SWE_ TUR_
USA_

for  %sec AtB_  C_  15t16_  17t18_  19_  20_  21t22_  23_  24_  25_  26_  27t28_  29_  30t33_  34t35_   36t37_   E_  F_   50_  51_ 52_
H_  60_ 61_ 62_  63_  64_  J_ 70_  71t74_  L_  M_ N_ O_

%name = %reg + %sec
%var0 = "q"
%var1 = "p"

%DEP1= %name + %var0
%IND1= %name + %var1

if  ((@min({%IND1}) <= 0) or (@min({%DEP1}) <= 0))  then
!i = !i +1
nardl_OLS_1lag(!i,1) = %reg
nardl_OLS_1lag(!i,2) =  %sec
nardl_OLS_1lag(!i,3) =  "n/a"
nardl_OLS_1lag(!i,4) =  "n/a"
nardl_OLS_1lag(!i,5) = "n/a"
nardl_OLS_1lag(!i,6) = "n/a"
nardl_OLS_1lag(!i,7) = "n/a"
nardl_OLS_1lag(!i,8) = "n/a"
nardl_OLS_1lag(!i,9) = "n/a"
nardl_OLS_1lag(!i,10) = "n/a"
nardl_OLS_1lag(!i,11) = "n/a"
nardl_OLS_1lag(!i,12) = "n/a"
nardl_OLS_1lag(!i,13) = "n/a"
nardl_OLS_1lag(!i,14) = "n/a"
nardl_OLS_1lag(!i,15) = "n/a"
else

if ({%DEP1} = kor_64_q) or ({%DEP1} = swe_30t33_q) or ({%DEP1} = tur_m_q) or ({%DEP1} = tur_n_q) then
!i = !i +1
nardl_OLS_1lag(!i,1) = %reg
nardl_OLS_1lag(!i,2) =  %sec
nardl_OLS_1lag(!i,3) =  "n/a"
nardl_OLS_1lag(!i,4) =  "n/a"
nardl_OLS_1lag(!i,5) = "n/a"
nardl_OLS_1lag(!i,6) = "n/a"
nardl_OLS_1lag(!i,7) = "n/a"
nardl_OLS_1lag(!i,8) = "n/a"
nardl_OLS_1lag(!i,9) = "n/a"
nardl_OLS_1lag(!i,10) = "n/a"
nardl_OLS_1lag(!i,11) = "n/a"
nardl_OLS_1lag(!i,12) = "n/a"
nardl_OLS_1lag(!i,13) = "n/a"
nardl_OLS_1lag(!i,14) = "n/a"
nardl_OLS_1lag(!i,15) = "n/a"
else
```

```
genr  d_{%DEP1} = log({%DEP1}) - log({%DEP1}(-1))
genr d_{%IND1} = log({%IND1}) - log({%IND1}(-1))
genr pos = d_{%IND1} >=0
genr d_{%IND1}_p = pos*(d_{%IND1})
genr d_{%IND1}_n = (1-pos)*(d_{%IND1})
genr {%IND1}_p = @cumsum(d_{%IND1}_p)
genr {%IND1}_n = @cumsum(d_{%IND1}_n)

equation eq{%name}.ls d_{%DEP1} c log({%DEP1}(-1)) {%IND1}_p(-1) {%IND1}_n(-1)  d_{%IND1}_p d_{%IND1}_n

 freeze(fr) eq{%name}.wald c(2)=0  , c(3)=0 , c(4) =0
 freeze(assym) eq{%name}.wald -c(3)/c(2) = -c(4)/c(2)
!i = !i +1
nardl_OLS_1lag(!i,1) = %reg
nardl_OLS_1lag(!i,2) =  %sec
nardl_OLS_1lag(!i,3) =  fr(6,2)
nardl_OLS_1lag(!i,4) =  fr(6,4)
nardl_OLS_1lag(!i,5) = c(2)
nardl_OLS_1lag(!i,6) =  eq{%name}.@pvals(2)
nardl_OLS_1lag(!i,7) = c(3)
nardl_OLS_1lag(!i,8) = eq{%name}.@pvals(3)
nardl_OLS_1lag(!i,9) = c(4)
nardl_OLS_1lag(!i,10) = eq{%name}.@pvals(4)
nardl_OLS_1lag(!i,11) = assym(7,4)
nardl_OLS_1lag(!i,12) = c(5)
nardl_OLS_1lag(!i,13) = eq{%name}.@pvals(5)
nardl_OLS_1lag(!i,14) = c(6)
nardl_OLS_1lag(!i,15) = eq{%name}.@pvals(6)

delete fr assym  d_{%DEP1} d_{%IND1} pos d_{%IND1}_p d_{%IND1}_n {%IND1}_p {%IND1}_n eq{%name}
endif
endif
next
next
```

Panel OLS estimation with fixed or random effects

```
!j=0
Table Estimates_fe_re

for %sec  AtB  C  15t16  17t18  19  20  21t22  23  24  25  26  27t28  29  30t33  34t35   36t37   E  F   50  51 52 H  60 61 62  63  64  J 70
71t74  L  M N O

!j = !j +1
%var0 = "q_"
%var1 = "p_"

%dep = %var0 + %sec
%ind = %var1 + %sec

if ((@min({%dep}) <=0) or (@min({%ind}) <=0)) then
Estimates_fe_re(!j,1) = %sec
Estimates_fe_re(!j,2) = "n/a"
Estimates_fe_re(!j,3) = "n/a"
Estimates_fe_re(!j,4) = "n/a"
Estimates_fe_re(!j,5) = "n/a"
Estimates_fe_re(!j,6) = "n/a"
Estimates_fe_re(!j,7) = "n/a"
Estimates_fe_re(!j,8) = "n/a"
Estimates_fe_re(!j,9) = "n/a"
Estimates_fe_re(!j,10) = "n/a"
else
equation ols.ls(cov=cxsur) log({%dep}) c log({%ind})
freeze(randomtest) ols.rcomptest

Estimates_fe_re(!j,1) = %sec
Estimates_fe_re(!j,2) = ols.C(2)
Estimates_fe_re(!j,3) = ols.@pvals(2)
Estimates_fe_re(!j,4) = randomtest(10,2)

equation fe.ls(cx=f , cov=cxsur) log({%dep}) c log({%ind})
```

```
freeze(fixed) fe.fixedtest

Estimates_fe_re(!j,5) = fe.c(2)
Estimates_fe_re(!j,6) = fe.@pvals(2)
Estimates_fe_re(!j,7) = fixed(7,5)

equation re.ls(cx=r, cov=cxsur) log({%dep}) c log({%ind})
freeze(hausman) re.ranhaus

Estimates_fe_re(!j,8) =   re.c(2)
Estimates_fe_re(!j,9) =   re.@pvals(2)
Estimates_fe_re(!j,10) = hausman(7,5)

delete ols randomtest fixed hausman fe re
endif
next
```

GETTING IN TOUCH WITH THE EU

In person

All over the European Union there are hundreds of Europe Direct information centres. You can find the address of the centre nearest you at: https://europa.eu/european-union/contact_en

On the phone or by email

Europe Direct is a service that answers your questions about the European Union. You can contact this service:

– by freephone: 00 800 6 7 8 9 10 11 (certain operators may charge for these calls),

– at the following standard number: +32 22999696, or

– by email via: https://europa.eu/european-union/contact_en

FINDING INFORMATION ABOUT THE EU

Online

Information about the European Union in all the official languages of the EU is available on the Europa website at: https://europa.eu/european-union/index_en

EU publications

You can download or order free and priced EU publications from: https://op.europa.eu/en/publications. Multiple copies of free publications may be obtained by contacting Europe Direct or your local information centre (see https://europa.eu/european-union/contact_en).

EU law and related documents

For access to legal information from the EU, including all EU law since 1952 in all the official language versions, go to EUR-Lex at: http://eur-lex.europa.eu

Open data from the EU

The EU Open Data Portal (http://data.europa.eu/euodp/en) provides access to datasets from the EU. Data can be downloaded and reused for free, for both commercial and non-commercial purposes.

MJ-07-22-269-EN-N

Publications Office
of the European Union

ISBN 978-92-76-52356-7