Panel regression analysis of electricity prices and renewable energy in the European Union

Cech, Marek and Janda, Karel

16 October 2016

Panel regression analysis of electricity prices and renewable energy in the European Union[#]

Marel Cech[*] – Karel Janda[**]

Abstract. This paper provides an econometric panel data model with data collected from 13 member states of the European Union over the period between 2010 and 2013 analysing two energy and climate relationships. First, it investigates the impact of the share of renewable energy sources in the final electricity production on the European consumer electricity prices. Second, it analyzes whether the replacement of fossil fuels by renewable energy causes a significant decrease in the greenhouse gases (specifically carbon dioxide) emissions. The results of our model analysis suggest that household electricity prices in the studied countries increase with the deployment of renewable electricity production. On the contrary, a negative effect of the renewables used in energy consumption on the CO_2 emissions produced was found by the model regression.

Key words: electricity price, renewable energy sources, energy policy, European Union

JEL classification: Q20, Q40, Q47, Q48, Q54

[#] This project has received funding from the European Union's Horizon 2020 Research and Innovation Staff Exchange programme under the Marie Sklodowska-Curie grant agreement No 681228.We also acknowledge support from the Czech Science Foundation (grants 15-00036S and 16-00027S) and from University of Economic, Prague (institutional support IP100040). Karel Janda acknowledges research support provided during his long-term visits at Australian National University and McGill University. The views expressed in the paper are those of the authors and not necessarily those of our institutions.

[*] Marek Cech; Institute of Economic Studies, Charles University, Opletalova 26,11000 Praha 1, Czech Republic, E-mail: marekcechx@seznam.cz

[**] Karel Janda; Department of Banking and Insurance, University of Economics, Prague, W. Churchilla 4, 13067 Praha 3, Czech Republic and Institute of Economic Studies, Charles University, Opletalova 26,11000 Praha 1, Czech Republic, E-mail: Karel-Janda@seznam.cz

Introduction

In this paper, the influence of using renewable energy sources (instead of fossil and nuclear resources) in the EU electricity production on the EU end-user electricity prices is estimated by employing an econometric panel data analysis. Moreover, the impact of renewables in the EU energy production on the amount of CO2 emissions produced by each region is estimated by the model as well. In the following sections, we provide a review of past researches done on the same or closely related topics, data set and methodology characterisation and theoretical background along with the practical application of the model itself.

The relationship between the modern energy policies, regarding the significant increase in renewable energy (electricity) production, and the changes in energy (electricity) prices have been analysed by many research papers over the last decade. The empirical and theoretical studies using different methodologies and data sets have shown ambiguous results; in some cases they were even contradictory. Mostly, a positive response of the electricity prices to the increased proportion of renewables in RES-E production was found. However, some studies came to the opposite conclusion using arguments specific for the analytical methodology used.

Paraschiv, Erni & Pietsch (2014) analysed the impact of renewable energy promotion (wind and PV) in Germany on the changes in electricity prices. Their analysis revealed that the deployment of RES-E technologies enhance extreme price changes. While the results of their dynamic fundamental model implied that renewable energy caused a decrease in market spot prices, the prices for final consumers (which we are interested in for our analysis) increased overall due to the feed-in tariff costs added to the spot prices. Fernández, Ortiz & Bernat (2013) used their study to analyse the RES-E deployment in Spain and Germany, the EU members with very similar electricity systems both having significant role in the EU energy production. According to the study, public funding, set by the EU to promote investment in renewable energy generation facilities, means an additional cost to electricity pricing systems and can but does not have to lead to an increase in the electricity price for final consumers (depending on aspects specific for each country).

Moreno & López (2011) proposed to use panel data model with the aim of explaining the household electricity prices as a function of several economic variables related to renewable energy sources and electricity market regulation. Their results, using panel data set provided by Eurostat and covering 27 EU countries from 1998 to 2009, suggested that electricity prices increased with the deployment of RES-E, mainly due to high initial generation, distribution and transmission costs. González, de Miera & Vizcaíno (2008) in their study agreed with the general opinion that the private costs of RES-E generation were in most cases above those of conventional electricity but they stressed the fact that it was important to consider the social benefits provided by RES-E production, including the environmental aspects, which some studies had overlooked. On the case of Spanish RES-E generation, they showed that a reduction in the wholesale price of electricity (caused by lower costs of the energy component of the price, see Section 3.2.1) could be greater than the increased costs for the consumers arising from the RES-E support schemes (usually feed-in systems in the EU). Therefore, the net effect of RES-E on retail prices can be to reduce, not raise. A similar analysis was provided by Würzburg, Labandeira & Linares (2013) regarding the Austrian and German region. Their study also showed that the net effect of RES-E production can be positive to final consumers (i.e. decreasing the retail prices) depending on the region and assessment method chosen.

The other research question to be analysed by the model in this paper is whether the amount of CO_2 emissions produced by the EU countries significantly depends on the share of renewables in the EU energy production. Vast majority of researches based on this topic showed that there is sufficient evidence that the RE participation in the total EU energy production had an important impact on the carbon dioxide emissions produced by the economy. However, the fossil-based energy industry causing the majority of greenhouse gas emissions has not been typical only for the EU. Shafiei & Salim (2014) showed this fact using the data from all OECD countries; Özbugday & Erbas (2015) proved the long-run reduction in CO_2 emissions caused by the replacement of fossil fuels by RE sources in the energy production processes in thirty six different countries; Moore, Lewis & Cepela (2010) came to the same conclusion while studying the United States energy production.

For our econometric panel data analysis we have chosen to study the effect of the EU RES-E production on the electricity prices. According to Moreno (2011), Paraschiv (2014), and the observed increasing trend in both the EU electricity prices and RES-E share in electricity production, we expect our model to show a positive impact of RES-E on the prices. On the contrary, regarding the analysis of the impact of RE promotion on the EU CO_2 emissions, we expect it to be negative.

Data and Methodology

Data Set Summary

The data set encompasses 4 subsets of data for each of the 14 selected European regions reflecting a 4-year time period (from 2010 to 2013). The areas include thirteen European countries, namely Belgium (BE), the Czech Republic (CZ), Germany (DE), Spain (ES), Finland (FI), France (FR), the United Kingdom (GB), Italy (IT), Netherlands (NL), Poland (PL), Portugal (PT), Romania (RO), and Sweden (SE) along with a compound region called EU27. The data for EU27 are used only for comparisons with the individual member states and are excluded from the econometric analysis. They were computed either as an average or as an aggregation (specified for each data subset) of the data collected from the 27 EU member states which had entered the EU before the enlargement in June 2013.

The countries are selected according to their energy production share in the total EU energy production (regarding the data collected by Eurostat in 2013). The countries with the highest shares are included in the analysis excepting Denmark (2.4%) for which a sufficient amount of data needed for further analysis was not provided by the data sources. In addition, Portugal (with only 0.6% share in the total EU energy generation) is involved in the data set as it is a country with the highest share of renewable energy sources used for the electricity production. Altogether, the collected data describe 89.2% of the EU energy production (see Figure 1).

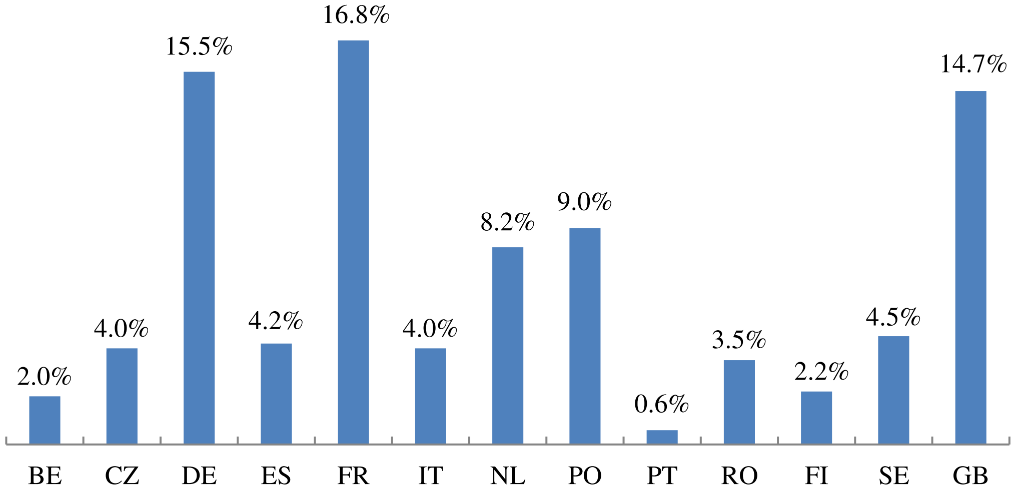


Figure 1: EU Member States' Shares in the Total EU Energy Production

Source: Eurostat: Energy Production 2013

The 4 mentioned subsets incorporate the information about each region's:

(i) electricity prices for domestic households (EUR/kWh)

(ii) the percentage share of electricity generated by using renewable energy sources in the total electricity production

(iii) the percentage share of renewable energy in gross final energy consumption

(iv) the amount of CO_2 emissions (Mt) produced by the region in total, per capita and per unit of energy production

Data adjusted to *per capita* or *per unit of production* values are incorporated in the analysis since they enable us to clearly compare the data from different regions regardless of either the area's population or the level of production, respectively. The base currency used in the data set is EUR. The unit of measurement of each variable is mentioned in each specific case of the model application and interpretation.

Data Sources

The examined data have been acquired from several resources. The electricity prices for households have been provided by Eurostat using the new methodology of data collection (from 2007 onwards) and excluding all taxes and levies. The proportions of electricity generated by using renewable energy sources in total electricity production for each of the 14 regions were obtained from Global Energy Statistical Yearbook 2014 published by Enerdata. The percentage shares of renewable energy in gross final energy consumption have been found in the Eurostat database as well as the electricity prices mentioned above. The data are submitted on the basis of an Annual Joint Questionnaire (Eurostat/IEA/United Nations Economic Commission for Europe) employing an internationally agreed methodology.

The accuracy of the basic data depends on the quality of the national statistical systems. However, Eurostat verifies to the highest possible extent whether the reported data respect the prescribed methodology. Hence the data are considered to be highly comparable and accurate. The last subset of the econometric model data set is the amount of CO_2 emissions (in Mt) produced by fuel combustion by each region in total, per capita and per unit of energy production. The source of these data was again the already mentioned Global Energy Statistical Yearbook from 2014 which can be found on the Enerdata website.

Variables

Country Each of the examined European regions is assigned a natural number from 1 to 14 as follows: 1 = EU27, 2 = BE, 3 = CZ, 4 = DE, 5 = ES, 6 = FR, 7 = IT, 8 = NL, 9 = PO, 10 = PT, 11 = RO, 12 = FI, 13 = SE, 14 = GB. The numbers altogether form an id dimension for the panel data. Each id variable is constant for all time periods and has only data ordering function in the panel data analysis.

Year Our data set consists of 4 time periods (2010 to 2013, yearly) which are the same for each of the researched countries and serve as time variables of the panel data model. The year 2010 was chosen as a starting point since it has been the first year in which the Renewable Energy Directive 2009 (see Section 4.1) was already in force. All sufficient data for the year 2014 were not found at the time of our research. Hence the data set ends with 2013 data.

Electricity Prices (EUR/kWh) For each country in the data set, the variable *elprice* reflects the average electricity price for households comprised of electricity basic price, transmission, system services, distribution and other services, and excluding taxes and levies. For the variable EU27 as a country aggregation, the values are calculated by weighting the twenty seven EU member states' national prices with the latest available national consumption for the households.

Electricity from Renewable Energy (%) The values of the variable *elfromRE* are computed as the ratio between the electricity production from selected renewable energies (hydro, wind, geothermal and solar) and the total electricity supply for end-users for each id and time variable of the panel data set.

Renewable Energy in Energy Consumption (%) The variable *REcons* serves as an indicator measuring how intensive is the use of renewable energy and, by implication, the degree to which renewable fuels have submitted fossil and/or nuclear fuels.

CO_2 Emissions (Mt) The total amount of CO_2 emissions produced by each region each year is represented by the variable *CO2*. The units of measurement are metric tons. The variables *CO2percap* and *CO2perprod* correspond to the level of carbon dioxide emissions adjusted to the region's population and the total energy production, respectively. These variables serve for an initial data set analysis and comparison of the examined countries. However, in the econometric model, only the variable *CO2* is included since we study the impact of RE sources on the total amount of carbon dioxide emitted.

In Table 1, the summary of the researched data set is presented by using the Stata statistical software. The number of observations reflects the fact that the data from 13 regions over the 4 mentioned time periods are included in the computation. The data for EU27 have been excluded from the summary as they could distort the results. They represent either averages or summations of the values from the countries already included in the statistics. According to Table 1, the electricity price (represented by the variable *elprice*) paid in the selected European regions by households is estimated to be 12.665 EUR cents per kWh on average. While the lowest average price, 7.95 EUR cents per kWh, was paid by consumers in Romania in 2012, the highest average electricity price in the data set, 17.72 EUR cents per kWh, applied to Spanish households in 2013.

Table 1: Summary of the Variables

Variable	Number of observations	Mean	Standard deviation	Min	Max
elprice	52	.12665	.0239	.0795	.1772
elfromRE	52	25.625	15.4716	7.4	62.5
REcons	52	16.8	12.6922	3.3	51.9
CO2	52	234.2673	200.3998	38.3	756.8
CO2percap	52	7.1735	2.428	3.3716	12.71
CO2perprod	52	4.8833	2.9153	1.0943	12.9667

Source: Authors data set and Stata computation

Regarding the variable *elfromRE*, the minimum proportion of electricity generated by using renewable energy sources in the total electricity production was recorded in Poland in 2010 at the level of 7.4% while the maximum share of 62.5% was monitored in Portugal in 2013. The overall mean percentage value of renewable energy participation in the total European electricity production was 25.625% over the examined 4-year time period for our data set, while the average share for the EU27 countries was about 2% higher, specifically 27.8%. In seven out of the thirteen countries in the data set, the overall average proportion was below the 25.625% level, namely in Poland (9.7%), the Czech Republic (9.8%), the United Kingdom (11.8%), the Netherlands (12.7%), Belgium (13.4%), France (15.2%) and Germany (22.6%). The above average participations of renewable energy in electricity generation were seen in Sweden (54.8%), Portugal (52.1%), Finland (35.1%), Spain (34%), Italy

(31.9%) and Romania (30.1%).[1] Concerning the values of the *REcons* variable, we can see that the percentage share of renewable energy in the gross final energy consumption measured in the countries included in the data set ranges from 3.3% to 51.9% having the mean at 16.8% level. The values substantially vary due to the differences in the aims of energy policies and approaches to production and consumption of renewable energy in the examined European countries albeit there are some targets set by the EU.

The least intensive use of renewable energy was seen in the United Kingdom in each of the examined time periods whereas, by contrast, Sweden each year showed the highest degree to which renewable sources of energy have substituted fossil and/or nuclear energy sources. Apart from Sweden, also three other countries from the data set exceeded on average the mean value, specifically Finland (33.8%), Portugal (24.5%) and Romania (22.7%). However, the below average values were found in most of the studied regions: in the United Kingdom (4%), the Netherlands (4.4%), Belgium (6.3%), Poland (10.5%), the Czech Republic (10.7%), Germany (12%), Italy (12.8%), France (13%) and Spain (14.2%).[2] See Figure 2 on the next page for a graphical summary of these values along with the average shares of renewable energy in the electricity production.

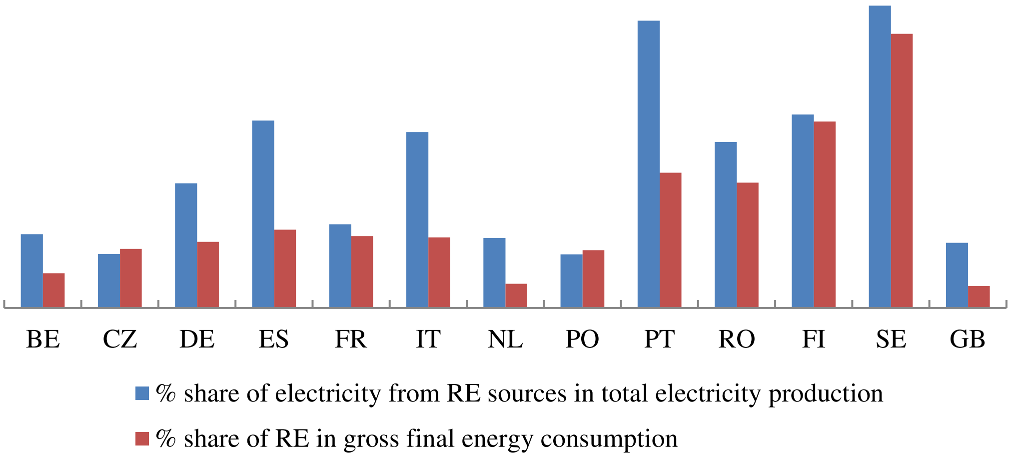
The last three variables from the summary are associated with the amount of carbon dioxide emissions produced by each country in the data set. According to Table 1, the mean level of CO_2 emissions produced by the countries from our sample was approximately 234.27 Mt a year. However, the individual values varied considerably, from the minimum at 38.3 Mt per year observed in Sweden in 2013 to the maximum at 756.8 Mt per year seen in Germany in 2013. Since the Swedish surface area is almost 1.2 times larger than the German one, it is clear that size of the region's surface does not imply larger carbon dioxide emissions produced.[3]

---

[1] The values were computed as an arithmetic average of the percentage shares of electricity generated by using RE sources in the total electricity production found in the data set for each of the countries.

[2] The figures were obtained by averaging the percentage representations of RE in the gross final energy consumption of the selected European regions using the data in the data set.

[3] The surface areas for Germany and Sweden were found at the Eurostat website.



	BE	CZ	DE	ES	FR	IT	NL	PO	PT	RO	FI	SE	GB
■	27.8	9.8	22.6	34.0	15.2	31.9	12.7	9.7	52.1	30.1	35.1	54.8	11.8
■	13.8	10.7	12.0	14.2	13.0	12.8	4.4	10.5	24.5	22.7	33.8	49.7	4.0

Figure 2: RE in the EU Electricity Production and Energy Consumption

Source: Authors computation using the data in the data set.

Nevertheless, some other variables can influence the level of pollution generated by a region, such as the region's population or the level of the energy production. Hence, the data adjusted to *per capita* and *per unit of energy production* values are included in this initial data set analysis. As we can see in Table 1, the average amount of CO_2 emissions per inhabitant was 7.1735 Mt a year. The lowest carbon footprint observed in our sample was left by an average Romanian in 2013, approximately 3.37 Mt a year, whereas the highest amount of carbon dioxide produced per capita was seen in Finland in 2010, 12.71 Mt a year.

Albeit the variance of the mentioned *per capita* values is relatively high, the values *per unit of energy production* vary even more across the data set. The mean amount of carbon dioxide emissions produced per 1 Mtoe of energy was 4.8833 Mt. The least has been emitted by the Swedish energy production, 1.0943 Mt/Mtoe in 2012. The most polluting (in terms of carbon dioxide emissions) energy production has been found in Italy, emitting 12.9667 Mt of CO_2 per Mtoe of energy generated in

2010. However, a decreasing trend of CO_2 emissions in Europe has been seen in majority of the researched countries. The total amount of produced carbon dioxide has been reduced over the 4-year time period in 11 out of the 13 countries. France and Germany represented the only exceptions. In terms of *per capita* values, the figures decreased in all regions apart from Germany and Portugal. Eventually, regarding the quantity of CO_2 emitted per Mtoe of energy production, all the regions excluding Great Britain, Germany and France saw a decline in the emission level. This short summary implies that Germany is the only country which has not been able to cope with cutting down the greenhouse gas emissions by any measure.

Theoretical Framework

In our model, we use panel data with the 13 selected European countries as the cross-sectional units, and years from 2010 to 2013 as the time dimension. The addition of a time component to the static nature of cross-sectional data brings with it a greater leverage on questions of causality. Due to this fact we can more effectively estimate the causal effect of one variable on the other with a panel data set. More specifically, in this paper we are interested in two major research questions, whether a higher share of electricity from RE in total electricity production causes an increase in consumer prices of energy, and whether a higher proportion of RE in gross final energy consumption leads to a considerable decrease in CO_2 emissions produced by the European countries.

Before we formulate our model for the estimation of the mentioned effects, there is another rationale for using more complex panel data analysis instead of simple cross-sectional analysis. If we use cross section from only one period (e.g. 2010) and run a simple regression with one independent variable, we probably obtain results suffering from omitted variable problems. One possible solution is to try to control for more factors, affecting the dependent variable, in a multiple regression analysis. However, many factors can be hard to realize and control for. In this case, we can use panel data to view the unobserved factors affecting the dependent variable as consisting of two types, those that are constant for each cross-sectional unit and those that vary over time, and manipulate with them differently in the analysis.

First Differences Estimation

We can write a panel data model with a single observed explanatory variable, letting *i* denote the cross-sectional unit and *t* the time period, as:

$$y_{it} = \delta_1 + \delta_2 d2011_t + \delta_3 d2012_t + \delta_4 d2013_t + \beta_1 x_{it} + a_i + u_{it} \qquad (1)$$

In the notation, i = 2, 3 ... 14 denotes the countries in the data set according to their assigned id numbers, *t* = 2010, 2011, 2012, 2013 stands for the time period. The variables $d2011_t, d2012_t, d2013_t$ are binary variables equal to one for *t* = 2011, 2012 or 2013, respectively, otherwise they equal to zero. Due to the inclusion of the yearly dummy variables in the model, we allow the intercept to change over time. The variable a_i captures all unobserved, time-constant factors which influence y_{it} and is called *unobserved effect* or *fixed effect* since it is fixed over time. The error u_{it} is referred to as the *idiosyncratic error*. It represents unobserved factors changing over time and affecting y_{it}.

Since we assume that the unobserved effect a_i is uncorrelated with x_{it} in our analyses, we can use the *first-differences* (FD) estimation to obtain the estimate of β_1 and eliminate the unobserved effects from the regression equation (1). By using the differencing method, we acquire the following equation for *t* = 2011, 2012, 2013:

$$\Delta y_{it} = \delta_2 \Delta d2011_t + \delta_3 \Delta d2012_t + \delta_4 \Delta d2013_t + \beta_1 \Delta x_{it} + \Delta u_{it} \qquad (3)$$

If the equation (5.3) satisfies the first four assumptions listed below, the FD estimator (pooled OLS estimator) is unbiased. If all six assumptions are satisfied, usual standard errors and test statistics are valid.

Assumption FD.1. For each *i*, the model is:

$$y_{it} = \beta_1 x_{it1} + \ldots + \beta_k x_{itk} + a_i + u_{it}\,, \qquad t = 1 \ldots T$$

where the parameters β_j are to be estimated and a_i is the unobserved effect.

Assumption FD.2. Each period we observe the same random sample.

Assumption FD.3. Each explanatory variable changes over time (for at least some *i*) and no perfect linear relationships exist among the explanatory variables.

Assumption FD.4. For each *t*, the expected value of the idiosyncratic error given the explanatory variables in all time periods and the effect $\boldsymbol{a_i}$: $\boldsymbol{E(u_{it}|x_{itj}, a_i) = 0}$, or by implication, $\boldsymbol{E(\Delta u_{it}|x_{itj}) = 0}$.

Assumption FD.5. The variance of the differenced errors, conditional on all explanatory variables, is constant: $\boldsymbol{Var(\Delta u_{it}|x_{itj}) = \sigma^2}$ for $t = 2 \dots \mathrm{T}$. Hence the differenced errors are homoskedastic.

Assumption FD.6. The differenced errors are serially uncorrelated. It means that for all $t \neq s$, the differences in the idiosyncratic errors are uncorrelated (conditional on all explanatory variables): $\boldsymbol{Cov(\Delta u_{it}, \Delta u_{is}|x_{itj}) = 0.}$

Fixed Effects Estimation

The other method for estimation of the unobserved effects panel data models, eliminating the *fixed effect* a_i, is the *fixed effects* (FE) transformation (or *within* transformation). Again, we consider an unobserved effects model with a single explanatory variable. For each *i* we then have:

$$y_{it} = \beta_1 x_{it} + a_i + u_{it}\,, \qquad t = 1 \dots T \tag{4}$$

$$\bar{y}_i = \beta_1 \bar{x}_i + a_i + \bar{u}_i \tag{5}$$

where the equation (5) represents the equation (4) averaged over time. To eliminate the factors in a_i, we subtract (5) from (4) and obtain:

$$\ddot{y}_{it} = \beta_1 \ddot{x}_{it} + \ddot{u}_{it}\,, \qquad t = 1 \dots T \tag{6}$$

Since we have disposed of the fixed effects included in a_i, we can use the pooled OLS to estimate β_1, as well as in the FD case. The obtained *fixed effects* or *within* estimator is then unbiased if the first four assumptions, identical to FD.1 through FD.4 listed above, are fulfilled. Under all six assumptions (the fifth and sixth

FE assumptions are mentioned below), the FE estimator of β_1 is the best linear unbiased estimator. Hence, the linear unbiased FD estimator should be worse than the FE estimator under such conditions.

Assumption FE.5. The variance of the errors, conditional on all explanatory variables and the unobserved effect, is constant: $\boldsymbol{Var(u_{it}|x_{itj}, a_i) = Var(u_{it}) = \sigma_u^2}$ for $t = 1 \ldots \mathrm{T}$. Hence the errors are homoskedastic.

Assumption FE.6. The idiosyncratic errors are uncorrelated (conditional on all explanatory variables and a_i): $\boldsymbol{Cov(u_{it}, u_{is}|x_{itj}, a_i) = 0,}$ for all $t \neq s$.

Further information regarding the FD and FE estimation processes along with a comparison of these two methods are included in the theoretical appendix.

Practical Applications of the Theory

In this section, we estimate our panel data model specifications using the theory explained in the previous section and the appendix. Each specific model equation with a single observed explanatory variable allows us to control for a predefined factor that is expected to affect the dependent variable.

Electricity Price and Renewable Energy

In our first model specification, we estimate the following equation:

$$\begin{gathered} \ln(elprice_{it}) = \delta_1 + \delta_2 d2011_t + \delta_3 d2012_t + \delta_4 d2013_t \\ +\beta_1 \ln(elfromRE_{it}) + a_i + u_{it} \end{gathered} \tag{7}$$

where i = 2,3 … 14 denotes the 13 European countries according to their assigned id numbers serving as the control group; t = 2010, 2011, 2012, 2013 stands for the time period over which the data have been collected; *d2011*, *d2012*, *d2013* are year dummy variables; a_i is the unobserved effect; and u_{it} is the idiosyncratic error. Using the Stata software, we estimate the model to discover whether there is a significant relationship between the proportion of RES-E in total electricity production in the EU (the variable *elfromRE*) and the European prices of electricity for households (the variable *elprice*). According to the reviewed literature and the fact that the electricity generation from RE sources is relatively uncompetitive, uncertain and connected with high initial costs; we expect it to have a positive effect on the electricity prices in the EU. As we decided to use a *log-log* model, the estimated coefficient β_1 on the variable *elfromRE* signifies the elasticity of electricity price with respect to the share of renewable energy sources in the total EU energy production.

We use FD and FE estimation methods to obtain the estimate of β_1 since the variable *elfromRE* is expected to be correlated with the unobserved effects in a_i (fixed or roughly constant over the 4 years in each of the countries). Factors assumed to be contained in a_i are e.g. already built infrastructure for power plants using fossil, nuclear or renewable energy sources; the access to fossil and nuclear energy sources; and the natural conditions suitable for development of renewable energy generation in each of the countries (such as the weather, duration of average day and sun light, terrain structure, geographical location etc.).

First Differences

To obtain unbiased and consistent pooled OLS estimator and valid test statistics using the FD estimation method, all six FD assumptions have to be satisfied. We verify these assumptions using Stata, running the FD regression and obtaining the parameters' estimates for the following equation:

$$\begin{gathered}\Delta\ln(elprice_{it}) = \propto_1 d2011_t + \propto_2 d2012_t + \propto_3 d2013_t \\ +\beta_1 \Delta\ln(elfromRE_{it}) + \Delta u_{it}\end{gathered} \tag{8}$$

According to Stata outputs of several tests (see Section B.1 in Appendix B) we consider the assumptions to be fulfilled. The estimate of β_1 is $\hat{\beta}_1 = .16967$ (standard deviation is equal to .04822) with *p-value* = .001 (see Table 2). Hence, the variable ln*elfromRE* is statistically significant at 5% (or even 1%) significance level as .001 < .01. Since we have already estimated the value of the coefficient β_1, we can now interpret the relationship between the dependent and independent variables. For instance, a 10% increase in the share of renewable energy sources in the total EU electricity production is predicted to cause an increase of the electricity price in the examined European countries by approximately 1.67% on average based on our collected data. The coefficients on all three year dummy variables *d2011*, *d2012* and *d2013* are statistically significant at 5% significance level with *p-values* equal to .004, .02 and .036 respectively. These variables serve as different intercepts for each of the years from 2011 to 2013 and account for secular changes (e.g. market trends) influencing the dependent variable that are not being modelled.

The *R-squared* of the model specification is R^2 = .5515. It implies that approximately 55.15% of the variation in the electricity prices in the EU countries is expected to be explained by the variation in the independent variables included in the model. The value of the *R-squared* is not very high albeit the model includes the time dummy variables which often cause a noticeable increase in the *R-squared* since they often account for effects that explain much of the variation in the dependent variable. While separately regressing the variable ln*elprice* solely on ln*elfromRE*, we indeed obtain the *R-squared* with a lower value, specifically R^2 = .2154. Hence, the variation in the share of renewable energy sources in the total electricity production is estimated to explain about 21.5% of the variation in the electricity prices in the studied European regions.

Fixed Effects

As well as in the case of FD estimation, the assumptions needed for acquiring an unbiased and consistent pooled OLS estimator have to be verified before we interpret our regression results. In Section B.1, Appendix B, we describe the justification of each assumption's verification. Once all the six FE assumptions are fulfilled, we can estimate the model equation (7) and interpret the outcome of the regression using FE transformation.

The results of the FE regression run in Stata (see Table 2) show a positive effect of the explanatory variable ln*elfromRE* on the dependent variable ln*elprice*. Specifically, e.g. a 10% increase in the proportion of the RE sources in the total EU electricity production is estimated to cause approximately 1.92% increase in the electricity price for the European households. The variable ln*elfromRE* is statistically significant at 5% significance level as well as all the time dummy variables included in the model. The exact FE (and FD) regression results can be seen in Table 2 on the following page. In addition, an interesting part of the FE regression output is *Rho* denoting the proportion of the total variation of dependent variable which is explained by the fixed effect a_i. In our case, *Rho* = .9805, hence only less than 2% of the total variation in ln*elprice* is caused by the idiosyncratic error.

Fixed Effects versus First Differences

In Table 2, we can see the summary of the FD and FE regression results obtained by using Stata. Both estimation methods indicate a positive effect of the participation of the RE sources in the European electricity production on the prices of electricity. Both estimates of the coefficient on the variable ln*elfromRE* are very statistically significant. However, using the FE transformation, the coefficient (.192486) is estimated to be larger than the FD estimate (.169669) and the expected ln*elfromRE* standard errors in the FE estimation are lower. It implies that the FE estimate is more significant, both statistically and economically.

While noticing the values of the *R-squared*, we have to take into consideration the fact that each of them has a different meaning. The *R-squared* from the FD regression denotes that approximately 55% of the sample variation in the ln*elprice* is explained by the variation in the independent variables included in the model. On the contrary, the value of the *within R-squared* from the FE regression means that about 71% of the ln*elprice* variation within each of the countries in the data set over the 4 years (excluding the fixed effects a_i) is explained by the explanatory variables. Since both the FD and FE assumptions were satisfied before running the regressions, the FE estimator is considered to be the best linear unbiased estimator and thus better than the FD estimator. Moreover, during the FD estimation we lose the first year observations due to which we can miss some important data.

Table 2: Regression Results (ln*elprice* on ln*elfromRE*)

ln*elprice*	FD	FE
ln*elfromRE*	.169669*** (.0482181)	.192486*** (.0464593)
d2011	.037073*** (.0119352)	.035378** (.0134709)
d2012	.044697** (.018255)	.040625** (.0154366)
d2013	.053352** (.0244815)	.046486** (.0191039)
R^2	.5515	.7118
N	39	52

$*p < 0.1$, $**p < 0.05$, $***p < 0.01$

Source: Authors data set and Stata computation

CO_2 Emissions and Renewable Energy

For this model specification, we use the same approach as in the previous case. We base our analysis on the estimation of the following equation:

$$CO2_{it} = \delta_1 + \beta_1 REcons_{it} + a_i + u_{it} \quad (10)$$

where i = 2,3 … 14 denotes the 13 European countries; t = 2010, 2011, 2012, 2013 stands for the time component; a_i is the fixed effect; and u_{it} is the idiosyncratic error. The variables *CO2* and *Recons* are described in Section 5.2.3. The major aim of estimating this model specification is to find the answer to the question whether an increase in the proportion of renewable sources of energy in total energy consumption of the specified EU member states (the variable *REcons*) has a significant effect on the level of CO_2 emitted by these regions (the variable *CO2*). Since RE resources are considered to be the "cleaner" alternative to the fossil-based energy production, we expect the growth of its share in total energy consumption to have a negative effect on the CO_2 emissions generated. Again, we estimate the model using the Stata software. We assume that the variable *REcons* is correlated with the fixed unobserved effects a_i (including e.g. the already built infrastructure for power plants or the natural conditions such as the weather, average day duration etc.) hence we use FD and FE estimation methods to obtain the estimates of β_1 as well as we have done it in the previous section.

First Differences

First we have to verify the six FD assumptions before we interpret our model results. The first three assumptions (FD.1 through FD.3) are verified directly by considering the format of the model equation (10) and the data set. The other three assumptions can be satisfied by using several tests (regarding endogeneity, autocorrelation and heteroskedasticity) and running regression of the following equation:

$$\Delta CO2_{it} = \beta_1 \Delta REcons_{it} + \Delta u_{it} \tag{11}$$

Once all the FD assumptions are considered to be fulfilled (see Section B.2., Appendix B) we can focus on the results of the FD regression. The estimate of β_1 is approximately $\hat{\beta}_1 =$ -3.745 with *p-value* = .017. Hence, the variable *REcons* is statistically significant at 5% significance level (.017 < .05). The minus sign of the value of $\hat{\beta}_1$ indicates that our initial expectations about the variables' relationship were correct. According to the results of the FD regression, the relationship between the variables *REcons* and *CO2* can be interpreted as follows: if the proportion of RE resources in the total energy consumption increases by e.g. 1 percentage points, the amount of CO_2 emissions produced by the examined European regions is estimated to decrease by approximately 3.745 megatons per year on average. In addition, the *R-squared* of the model specification is R^2 = .1571. Hence, approximately 15.71% of the variation in the level of CO_2 emissions caused by the EU countries is estimated to be explained by the variation in the renewable energy sources' participation in total energy consumption in the EU countries.

Fixed Effects

To obtain the estimate of β_1 from the equation (10) and then to be able to compare the results with the FD estimation, we use the FE transformation as well as in the previous section. Since, the assumptions FE.1 through FE.6 are considered to be satisfied (see Section B.2, Appendix B), we can proceed to FE regression results. The regression output indicates a negative effect of the explanatory variable *REcons* on the dependent variable *CO2*. Specifically, an increase in the share of RE sources in the EU energy consumption by e.g. 1 percentage point is estimated to cause a decrease in the yearly amount of CO_2 emitted by the EU countries by approximately 5 megatons on average (see Table 5.3). The only explanatory variable of the model, *REcons*, is statistically significant at 5% significance level. In addition, the *Rho* of the FE regression, denoting the proportion of the total variation of dependent variable explained by the fixed effect a_i, is equal to .99765. It implies that only approximately .00235% of the total variation in *CO2* is caused by the idiosyncratic error.

Fixed Effects versus First Differences

The outputs of both the FD and FE regressions are summarized in Table 3. The FD estimation as well as the FE transformation indicates that the proportion of RE sources in the EU countries' energy consumption has a negative effect on the CO_2 emission level, as we expected. For both estimation methods, the estimates of the coefficient on *REcons* are statistically significant. By using the FE method, we have obtained an estimate with noticeably higher negative effect (-5.0017) than in the case of the FD estimation (-3.74481). The standard errors of the β_1 estimates are lower for the FE estimator (1.097557) than those acquired by the FD regression (1.359718). It implies that the FE estimate is both statistically and economically more significant.

The value of the *R-squared* for the FD regression denotes that approximately 15.71% of the sample variation in *CO2* is explained by the variation in *REcons*. By contrast, the *R-squared* obtained from the FE regression is so called *within R-squared* indicating that about 35.34% of the *CO2* variation within each of the countries in the data set over the 4-year period (excluding the unobserved effects a_i) is explained by the variation in *REcons*. Albeit in both FD and FE estimations we have verified all assumptions necessary to acquire an unbiased consistent estimator, only the FE estimator is considered to be the best linear unbiased estimator under FE.1 through FE.6. Hence we assume that it performs better than the FD estimator.

Table 3: Regression Results (*CO2* on *REcons*)

CO2	FD	FE
REcons	-3.74481** (1.359718)	-5.0017*** (1.097557)
R^2	.1571	.3534
N	39	52

*p < 0.1, **p < 0.05, ***p < 0.01

Source: Authors data set and Stata computation

Discussion of the Model Results

Electricity Price and Renewable Energy

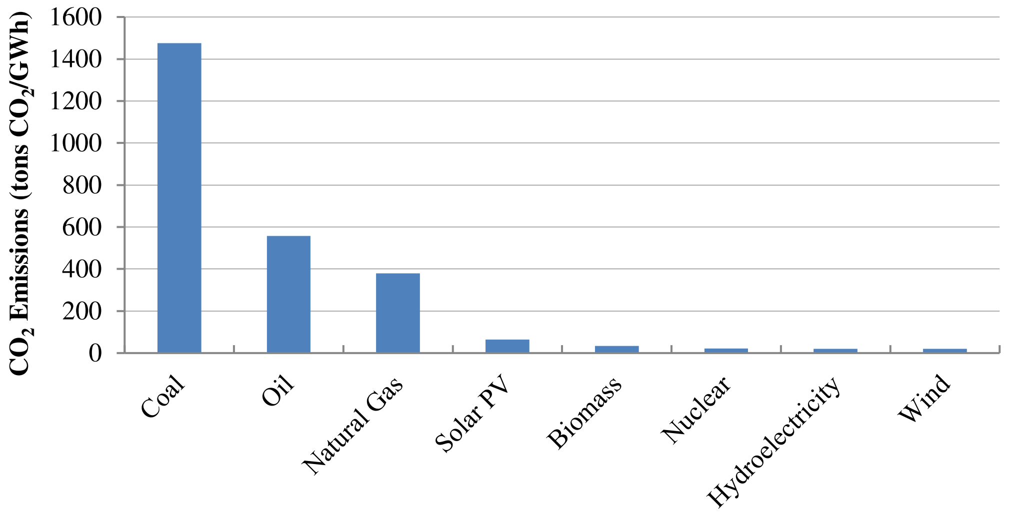
As expected from the literature review our econometric model showed a positive effect of the RES-E share in the total electricity production on the final price of electricity for the EU households. We used electricity prices excluding taxes and levies in our analysis since these financial charges considerably vary across the countries in the data set and are specific to each member state's economic and political regime. Hence, we specifically analysed the impact of the rising support for RES-E production (binding for all EU members) on the energy and network element of the EU electricity prices. Since power stations using the RE sources (mainly wind, hydro and solar power) are connected with high initial construction, transmission and distribution costs creating an additional cost burdens for electricity end-users (including households), it makes sense that the mentioned impact on the EU electricity prices has been showed to be positive and significant.

The high initial investments, regarding the energy and network components of the electricity prices, are needed mainly for building infrastructure, construction of power plants, and transmission and distribution of the power. These investments are very similar for each EU member state (deciding to build a new RES-E network) and are expected to increase the cost of providing renewable electricity, especially during early years. They include for instance: prospecting for publicly acceptable and suitable place with good access to RE resources and transmission lines; developing standards and permitting issues for renewables; marketing costs of communicating the benefits of renewables to consumers who are used to buying electricity from traditional sources; and installation, operation and maintenance costs including power plant constructions but also e.g. worker trainings regarding the proper treatment of the new technologies.

CO_2 Emissions and Renewable Energy

The results of the second model specification indicate that the increase in the use of RE sources in the total EU energy consumption leads to a decline in the amount of CO_2 emitted by the EU. This regression output corresponds not only to the past researches regarding the same topic, but also to a *lifecycle approach* of analysing the level of CO_2 emissions produced by each energy source. Since distinct electricity generation methods (drawing energy from different sources) produce carbon dioxide (and other greenhouse gases) in varying quantities through construction, operation (including fuel supply activities) and decommissioning, the *lifecycle approach* accounts for emissions from all phases of each electricity production project (construction, operation and decommissioning) attempting to calculate the global warming potential of electrical energy sources. Observing the lifecycle emissions from electricity generation allows us to fairly compare the different generation methods on a per kilowatt-hour basis (see Figure 3).

The data in Figure 3 were obtained from the World Nuclear Association (WNA) Report 2011 reviewing over twenty studies assessing the greenhouse gas emissions produced by different forms of electricity generation. It is noticeable that all renewable sources included in the statistic (wind, solar PV, biomass and hydro power) perform substantially better than each of the fossil-based fuels with respect to the level of CO_2 emitted. Hence, according to the WNA Report and the *lifecycle approach*, it is rational to expect that the replacement of fossil fuels in the EU electricity generation by RES-E production results in a decrease in the amount of carbon dioxide produced.



					Renewable Sources			
	Coal	**Oil**	**Natural Gas**	**Nuclear**	**Solar PV**	**Biomass**	**Hydro**	**Wind**
Mean	1476	557	379	22	65	34	20	20

Figure 3: Lifecycle CO_2 Emissions by Source (in t/GWh)

Source: WNA Report 2011

Conclusion

The aim of this paper was to create a sufficient overview of the EU renewable energy and climate policy, its targets towards next few years and the impacts of the increasing share of renewables in the EU energy consumption and production on final consumers and the environment. More precisely, we focus on renewables in the *electricity production* (RES-E) since it plays a decisive role in achieving the EU renewable energy targets and since the changes in electricity prices affect vast majority of the EU inhabitants.

The core of this paper is the econometric model analysing the effects of the renewable energy use on the electricity prices for final consumers and the amount of carbon dioxide emissions produced in the EU a year. We have decided to use panel data analysis as, while using the first differences and fixed effects methods of estimation, it allows for the effects that are unobserved and fixed over time in our model to be correlated with the explanatory variables and eliminated through the regression. Hence we can dispose of the potential omitted variable problem and study the effects of explanatory variables on the dependent variables over a given time period. The results of our model analysis suggest that household electricity prices in the studied EU member states increase with the deployment of RES-E production. Such effect on prices was anticipated, since the majority of renewable energy technologies increase electricity generation, distribution and transmission costs. Moreover, in the EU the largest part of investments for electricity production over the last few years was devoted to new wind power stations and solar photovoltaics which are connected with the highest initial costs when compared to conventional generation methods. On the contrary, a negative effect of the renewables used in the EU energy consumption on the CO_2 emissions produced was found by the model regression, as it had been expected while formulating the model since the lifecycle CO_2 emissions (covering construction, operation and decommissioning of the power stations) were considerably lower for renewable sources in comparison with fossil-based fuels.

This paper serves well as an overview in the field of renewable energy and electricity production, consumption and pricing in the EU. It provides the essential background for this topic along with the detailed analysis of two specific impacts of the deployment of renewable energy technologies on the European level. However, within the scope of this paper, we cannot hope to cover all the possible consequences of the promotion of renewable energy sources in Europe. Nevertheless, this fact makes a space for further research and study. Such work could concern, for instance, the question how the rapid replacement of fossil fuels by renewables in the EU electricity production affects the changes in each particular component comprising the value of the EU electricity prices (energy, network and taxes/levies component separately); or how e.g. the economic development, employment in rural areas and security of energy supply can be affected by this trend. In addition, it would be also interesting to repeat this study in a few years and ascertain whether the high initial costs of renewable energy power stations gradually pay off and allow the EU electricity prices to decrease, taking the advantage of the relatively low operation and maintenance costs of RES-E stations and zero costs of obtaining the energy source (as wind, water and solar energy can be usually used free of charge unlike oil, coal or natural gas).

Bibliography

Aboumahboub T., K. Schaber, U. Wagner & T. Hamacher (2012): "On the CO_2 Emissions of the Global Electricity Supply Sector and the Influence of Renewable Power-Modelling and Optimization." Energy Policy 42: pp. 297 – 314.

Buchan, D. (2014): "Costs, Competitiveness and Climate Policy: Distortions Across Europe." The Oxford Institute for Energy Studies: 3 – 20.

ČECH, Marek. The Impact of Renewable Energy on the EU Electricity Prices and CO_2 Emissions, Prague 2015. 38 p. Bachelor thesis, Charles University in Prague, Faculty of Social Sciences, Institute of Economic Studies.

Crofl D., I. Preston, P. Guertler & J. Carrington (2012): "Impact of Future Energy Policy on Consumer Bills." ACE-CSE: 6 – 12; 106.

Drukker D. M. (2003): "Testing for Serial Correlation in Linear Panel-Data Models." The Stata Journal 3(2): pp. 168 – 177.

Enerdata (2014): "Global Energy Statistical Yearbook 2014." Web, 9 February 2015.

European Commission (2010): "Analysis of Options to Move Beyond 20% Greenhouse Gas Emission Reductions and Assessing the Risk of Carbon Leakage." Communication from the Commission.

European Commission (2011): "Technical Assessment of the Renewable Action Plans." Publications Office of the European Union. ISBN 978-92-79-21049-5.

European Commission (2013): "Guidance on the Use of Renewable Energy Cooperation Mechanism." Communication from the Commission.

European Commission (2013): "The EU Emissions Trading System (EU ETS)." European Union Publications Office. ISBN 978-92-79-32962-3.

European Commission (2014): "A Policy Framework for Climate and Energy in the Period from 2020 to 2030." Communication from the Commission.

European Commission (2014): "Energy Prices and Costs in Europe." Communication from the Commission.

European Parliament and European Council (2009): "Directive 2009/28/EC."

Eurostat (2014): "Energy Price Statistics." Web, 10 February 2015.

Eurostat (2014): "Energy from Renewable Sources." Web, 10 February 2015.

Eurostat (2014): “Share of Renewable Energy in Gross Final Energy Consumption.” Web, 20 February 2015.

Fernández P. F., E. V. Ortiz & J. X. Bernat (2013): “The Deployment of Electricity Generation from Renewable Energies in Germany and Spain: A comparative Analysis Based on a Simle Model.” Energy Policy 57: pp. 552 – 562.

Gerardi W. & P. Nidras (2013): “Estimating the Impact of the RET on Retail Prices.” Sinclair Knight Merz: pp. 1 – 20.

Gerbelová H., F. Amorim, A. Pina, Ch. Ioakimidis, P. Ferrao & M. Melo (2014): “Potencial of CO_2 Taxes as A Policy Measure Towards Low-Carbon Portuguese Electricity Sector by 2050.” Energy Journal 69: pp. 113 – 119.

González P. del R., G. S. de Miera & I. Vizcaíno (2008): “Analysing the Impact of Renewable Electricity Support Schemes on Power Prices: The case of Wind Electricity in Spain.” Energy Policy 36(9): pp. 3345 – 3359.

Götz B., A. Voss, M. Blesl & U. Fahl (2012): “Comparing Different Types of Support Systems for Renewable Electricity.” Institute of Energy Economics and the Rational Use of Energy (IER), University of Stuttgart.

International Energy Agency (2012): “Energy Technology Perspectives 2012: Pathways to a Clean Energy System.” IEA: pp. 2 – 7.

International Energy Agency (2014): “Key World Energy Statistics.” IEA: pp. 6 – 57.

Menanteau P., D. Finon & M.-L. Lamy (2003): “Prices versus Quantities: Choosing Policies for Promoting the Development of Renewable Energy.” Energy Policy 31(8): pp. 799 – 812.

Moore M. R., G. McD. Lewis & D. J. Cepela (2010): “Markets for Renewable Energy and Pollution Emissions: Environmental Claims, Emission-Reduction Accounting, and Product Decoupling.” Energy Policy 38(10): pp. 5956 – 5966.

Moreno B. & A. J. López (2011): “The Impact of Renewable Energies and Electric Market Liberalisation on Electrical Prices in the European Union: An Econometric Panel Data Model.” Department of Applied Economics (University of Oviedo): pp. 27 – 31. ISBN 978-1-61208-138-0.

Nicolosi M. & M. Feursch (2010): “Implications of the European Renewables Directive on RES-E Support Scheme Designs and its Impact on the Conventional Power Markets.” International Association for Energy Economics: pp. 25 – 29.

Özbugday F. C. & B. C. Erbas (2015): “How Effective Are Energy Efficiency and Renewable Energy in Curbing CO_2 Emissions in the Long Run? A Heterogenous Panel Data Analysis.” Energy Journal 82: pp. 734 – 745.

Paraschiv F., D. Erni & R. Pietsch (2014): “The Impact of Renewable Energies on EEX Day-Ahead Electricity Prices.” Energy Policy 73: pp. 196 – 210.

Renewable Energy Policy Network (2014): “Renewables 2014: Global Status Report.” REN21: pp. 4 – 12. ISBN 978-3-9815934-1-9.

RWE Corporate Website. “How the Electricity Price Is Determined.” Web, 27 February 2015.

Shafiei S. & R. A. Salim (2014): “Non-Renewable and Renewable Energy Consumption and CO_2 Emissions in OECD Countries: A Comparative Analysis.” Energy Policy 66: pp. 547 – 556.

Shogren J. F. (2013): “Encyclopaedia of Energy, Natural Resource, and Environmental Economics.” University of Wyoming: pp. 15 – 20; 128. ISBN 978-0-12-375067-9.

United Nations Environment Programme (2012): “The Emissions Gas Report 2012.” UNEP: pp. 12 – 30. ISBN 978-92-807-3303-7.

Wooldridge J. M. (2002): “Econometric Analysis of Cross Section and Panel Data.” Massachusetts Institute of Technology: pp. 280 – 310. ISBN 0-262-23219-7.

Wooldridge J. M. (2009): “Introductory Econometrics: A Modern Approach.” Michigan State University: pp. 443 – 505. ISBN 978-0-324-58162-1.

World Energy Council (2013): “World Energy Resources 2013 Survey.” World Energy Council: pp. 5 – 23. ISBN 978-0-946121-29-8.

World Nuclear Association (2011): “Comparison of Lifecycle Greenhouse Gas Emissions of Various Electricity Generation Sources.” WNA: pp. 2 – 8.

Würzburg K., X. Labandeira & P. Linares (2013): “Renewable Generation and Electricity Prices: Taking Stock and New Evidence for Germany and Austria.” Energy Economics 40(1): pp. S159 – S171.

Appendix: Theoretical Framework of the Panel Data Model

Since our data set used for the econometric analysis consists of both cross-sectional and time series dimensions following the same units over time, we call it *panel data set*. In other words, by panel data we mean data containing repeated measures of the same variable taken from the same set of cross-sectional units over time. In our applications the units are the 13 selected European countries and time periods are years from 2010 to 2013.

A.1 First Differences Estimation

In Section 5.2.1 we use a single observed explanatory variable model, letting *i* denote the cross-sectional unit and *t* the time period, as:

$$y_{it} = \delta_1 + \delta_2 d2011_t + \delta_3 d2012_t + \delta_4 d2013_t + \beta_1 x_{it} + a_i + u_{it} \qquad (5.2)$$

where i = 2,3 … 14 denotes the countries in the data set according to their assigned id numbers (see Section 5.2.3.), t = 2010, 2011, 2012, 2013 stands for the time period and the variables $d2011_t, d2012_t, d2013_t$ are yearly binary variables. The intercept for t = 2010 is δ_1, for t = 2011 it is $\delta_1 + \delta_2$, for t = 2012 it equals to $\delta_1 + \delta_3$, and when t = 2013 we have the intercept of $\delta_1 + \delta_4$. Since 2010 is in our case considered to be the base year, the three dummy variables help us to find the influence of the time when the data were observed (2011, 2012 or 2013) on the value of the dependent variable, holding all factors influencing the dependent variable fixed, and compare this value with the value in 2010. For instance, the coefficient δ_2 on the year dummy variable $d2011_t$ shows us what the difference between the values of y_{it} in 2011 and 2010 is, holding all other factors affecting y_{it} fixed.

The variable a_i captures all unobserved, time-constant factors which influence y_{it} (such as geographical features of a country; different historical factors with an effect on y_{it} or even some not exactly constant factors which are, however, roughly constant over the relatively short time period). Generally, it is called *unobserved effect* or *fixed effect* since it is fixed over time. Due to the variable a_i, the model in (5.1) is also called *fixed effects model*. The error u_{it} is often referred to as the *idiosyncratic* (specific) or *time-varying error*. It represents unobserved factors changing over time and affecting y_{it}. The *idiosyncratic error* along with the *unobserved effect* is called the *composite error* $v_{it} = a_i + u_{it}$.

To estimate the parameter of interest, β_1, we can generally use directly the method of pooled OLS. However, for pooled OLS to produce a consistent estimator of β_1, we have to assume that the unobserved effect a_i is uncorrelated with x_{it}. Since we will assume the opposite in our analyses, the estimator in this case would be biased and inconsistent. If we want to allow the unobserved factors included in a_i affecting y_{it} to be correlated with x_{it}, we can use *differencing* method to obtain the *first-differences* (FD) estimator. The key assumption in this case is that the *idiosyncratic errors* are uncorrelated with the explanatory variable in each time period:

$$Cov(x_{itj}, u_{is}) = 0, \text{ for all } t, s, j \qquad (5.2)$$

It implies that the explanatory variables are *strictly exogenous* after we take out the unobserved effect a_i. If a_i is correlated with x_{itj}, then under (5.2), x_{itj} will be correlated with the *composite error:* $v_{it} = a_i + u_{it}$. To eliminate a_i by using *differencing* method, we (or any statistical software we use) just difference adjacent periods and then run pooled OLS regression. In our 4-period case, we subtract time period one from time period two, time period two from time period three and finally time period three from time period four. We obtain the following equation for t = 2011, 2012 and 2013:

$$\Delta y_{it} = \delta_2 \Delta d2011_t + \delta_3 \Delta d2012_t + \delta_4 \Delta d2013_t + \beta_1 \Delta x_{it} + \Delta u_{it} \qquad (5.3)$$

If the equation (5.3) satisfies the first four assumptions of the listed below, a pooled OLS estimator (the FD estimator in this case) is unbiased. To acquire consistent OLS estimator, Δu_{it} has to be uncorrelated with Δx_{it}. Moreover, we must assume that Δu_{it} are uncorrelated and homoskedastic over time for the usual standard errors and test statistics to be valid. Hence we will further test serial correlation and heteroskedasticity in the first-differenced equation in our model specifications. The important assumptions for the *first differences* estimation are as follows:

Assumption FD.2. For each *i*, the model is:

$$y_{it} = \beta_1 x_{it1} + \ldots + \beta_k x_{itk} + a_i + u_{it}\,, \qquad t = 1 \ldots T$$

where the parameters β_j are to be estimated and a_i is the unobserved effect.

Assumption FD.2. Each period we observe the same random sample.

Assumption FD.3. Each explanatory variable changes over time (for at least some *i*) and no perfect linear relationships exist among the explanatory variables.

Assumption FD.4. For each *t*, the expected value of the idiosyncratic error given the explanatory variables in all time periods and the effect $\boldsymbol{a_i}$: $\boldsymbol{E}(\boldsymbol{u_{it}}|\boldsymbol{x_{itj}}, \boldsymbol{a_i}) = \boldsymbol{0}$, or by implication, $\boldsymbol{E}(\Delta \boldsymbol{u_{it}}|\boldsymbol{x_{itj}}) = \boldsymbol{0}$.

Assumption FD.5. The variance of the differenced errors, conditional on all explanatory variables, is constant: $Var(\Delta \boldsymbol{u_{it}}|\boldsymbol{x_{itj}}) = \sigma^2$ for $t = 2 \ldots \mathrm{T}$. Hence the differenced errors are homoskedastic.

Assumption FD.6. The differenced errors are serially uncorrelated. It means that for all $t \neq s$, the differences in the idiosyncratic errors are uncorrelated (conditional on all explanatory variables): $Cov(\Delta \boldsymbol{u_{it}}, \Delta \boldsymbol{u_{is}}|\boldsymbol{x_{itj}}) = 0$.

A.2 Fixed Effects Estimation

The other method for estimation of the unobserved effects panel data models is the *fixed effects* (FE) transformation which is, as well as the FD estimation, one of the ways to eliminate the *fixed effect* a_i which is expected to be correlated with the explanatory variable(s) in any time period. In our model specifications we will compare the results of the FD and FE estimations and test which of them is more efficient under certain assumptions. For the description of the FE transformation (also called the *within transformation*), we consider an unobserved effects model with a single explanatory variable, for each *i* we then have:

$$y_{it} = \beta_1 x_{it} + a_i + u_{it}\,, \qquad t = 1 \ldots T \tag{5.4}$$

$$\bar{y}_i = \beta_1 \bar{x}_i + a_i + \bar{u}_i \tag{5.5}$$

where the equation (5.5) represents the equation (5.4) averaged over time, with $\bar{y}_i = T^{-1}\sum_{t=1}^{T} y_{it}$ and likewise for $\bar{x}_i$ and $\bar{u}_i$. To eliminate the fixed factors in a_i appearing in both equations we subtract (5.5) from (5.4) and obtain:

$$\ddot{y}_{it} = \beta_1 \ddot{x}_{it} + \ddot{u}_{it}\,, \qquad t = 1 \ldots T \tag{5.6}$$

where $\ddot{y}_{it} = y_{it} - \bar{y}_i$ is the *time-demeaned data* on *y* (and similarly for $\ddot{x}_{it}$ and $\ddot{u}_{it}$).

Now we have disposed of the fixed effects included in a_i and as well as in the FD estimation we can use the pooled OLS to estimate β_1. The pooled OLS estimator based on time-demeaned variables is called the *fixed effects* or *within estimator* since the OLS on (5.6) uses time variation in *y* and *x* within each cross-sectional observation. The assumptions for the *fixed effects* estimation are listed below:

Assumption FE.3. See **Assumption FD.1.**

Assumption FE.2. See **Assumption FD.2.**

Assumption FE.3. See **Assumption FD.3.**

Assumption FE.4. See **Assumption FD.4.**

As we can see, the first four assumptions are identical to the assumptions for the FD estimator. Under them, the FE estimator is unbiased (as well as in the case of first differences). The key assumption is the strict exogeneity assumption (FE.4.).

Assumption FE.5. The variance of the errors, conditional on all explanatory variables and the unobserved effect, is constant: $Var(u_{it}|x_{itj}, a_i) = Var(u_{it}) = \sigma_u^2$ for $t = 1 \dots$ T. Hence the errors are homoskedastic.

Assumption FE.6. The idiosyncratic errors are uncorrelated (conditional on all explanatory variables and a_i): $Cov(\boldsymbol{u_{it}}, \boldsymbol{u_{is}}|\boldsymbol{x_{itj}}, a_i) = 0$, for all $t \neq s$.

Under the all first six assumptions, the FE estimator of β_1 is the best linear unbiased estimator. Hence, the linear unbiased FD estimator should be worse than the FE estimator under such conditions.

A.3 Fixed Effects versus First Differences

While comparing two different estimators we often use unbiasedness and consistency as the criteria. However, since both FE and FD estimators are unbiased under the Assumptions FE.1 through FE.4 as well as asymptotically consistent (with T fixed as $N \to \infty$), the decision on which estimator is better to use then depends on considering some other factors.

Hence we focus on the error structure. If $\boldsymbol{u_{it}}$ is serially uncorrelated, the FE estimator is more efficient and used rather than the FD estimator. On the contrary, when $\boldsymbol{u_{it}}$ follows a random walk (i.e. very substantial positive autocorrelation), then the $\Delta\boldsymbol{u_{it}}$ is serially uncorrelated and the FD estimator is more efficient. We can also test directly whether the differenced errors ($\Delta\boldsymbol{u_{it}}$) are serially uncorrelated. If the null hypothesis of no serial correlation is rejected and there is an evidence of substantial negative autocorrelation in the differenced errors, the FE estimator is considered to perform better.

Appendix B: Practical Applications of the Theoretical Model

Based on the theoretical background regarding the econometric panel data analysis offered in the main body of this paper and in Appendix B we estimate our model with its several specifications using the *first differences* and *fixed effects* estimation methods. Using the Stata software, we test the assumptions that have to be fulfilled for obtaining a reliable slope estimate for the independent variable along with its standard deviation. The slope estimate is necessary for measuring the partial effect of the independent variable on the dependent variable. Moreover, the Stata output includes *p-values* for test statistics (which are helpful while testing hypotheses, recognizing statistical significance etc.) and the value of *R-squared* as well. The *R-squared*, a goodness-of-fit measure, denotes the proportion of the sample variation in the dependent variable explained by the independent variable.

In the *fixed effects* regression, we obtain three distinct values of *R-squared*. Nevertheless, we often do not have to focus on all of them. The first is called the *overall* R^2 and is interpreted as the usual *R-squared* from the regression of the dependent variable on the explanatory variable. The second one is called the *between* R^2 obtained from the regression of time-demeaned data which consists in collapsing the data and removing the time component by taking the means of our variables for each panel unit individually. It implies the *between* R^2 measures the variation between the individual cross-sectional units. However, since we are interested in a good amount of within information (the variation within one individual over time) that can be exploited by the FE estimator, we rather focus on the value of the *within* R^2 offering the goodness-of-fit measure for individual mean de-trended data taking no account of all the between information in the data.

B.1 Electricity Price and Renewable Energy

For the first model specification we estimate the equation:

$$\ln(elprice_{it}) = \delta_1 + \delta_2 d2011_t + \delta_3 d2012_t + \delta_4 d2013_t + \beta_1 \ln(elfromRE_{it}) + a_i + u_{it} \tag{5.7}$$

with i = 2,3 … 14 denoting the 13 European countries; t = 2010, 2011, 2012, 2013 stands for the time period; *d2011*, *d2012*, *d2013* are yearly dummy variables; a_i is the unobserved effect; and u_{it} is the idiosyncratic error. The equation (5.7) is called *log-log* model specification since the natural logarithm transformed values of *y* are being regressed on natural logarithm transformed values of *x*. The output of the *log-log* model regression is interpreted as the percentage change in the value of the dependent variable caused by 1% change in the value of the explanatory variable.

First Differences

While using the *first difference* regression in Stata, the assumptions FD.1 through FD.6 have to be verified and fulfilled for us to obtain unbiased and consistent OLS estimator and valid test statistics (see Section 5.3.1). The first assumption is fulfilled since the log transformation ensures the desired linearity in parameters. The second and third assumptions can be verified as well due to the way we have collected the data set (see Section 5.2.1) and since the value of *elfromRE* changes over time. Moreover, if there is found a perfect collinearity while running the regression, Stata omits the problematic variable and states the fact to inform us. The last three assumptions will be inspected after running the *first difference* regression and obtaining the parameters' estimates for the following equation:

$$\Delta\ln(elprice_{it}) = \propto_1 d2011_t + \propto_2 d2012_t + \propto_3 d2013_t + \beta_1 \Delta\ln(elfromRE_{it}) + \Delta u_{it} \tag{5.8}$$

According to Stata output (using commands *.predict res, r* and *.summ res, d*), the expected value of the idiosyncratic errors from the regression equation (5.8) is $\boldsymbol{E}(\Delta \boldsymbol{u}_{it} | \boldsymbol{x}_{itj})$ = .00001 which is really close to zero. Hence we consider the fourth FD assumption to be verified. Next, we test for heteroskedasticity using Breusch-Pagan Lagrange multiplier test (obtained by Stata command *.bpagan* ln*elfromRE d2011 d2012 d2013*). The Breusch-Pagan Chi-squared statistics yields χ^2 = 4.937

with *p-value* = .1764. Hence there is not enough evidence of heteroskedasticity as we cannot reject the null hypothesis of homoskedasticity at 5% or even 10% significance level (.10 < .1764). Finally, we have to verify the last FD assumption that there is no serial correlation between the differences in the idiosyncratic errors conditional on all explanatory variables in the model. We use the Wooldridge test for autocorrelation in panel data models (Stata command *.xtserial* ln*elprice* ln*elfromRE d2011 d2012 d2013*). The F statistics yields $F = 4.389$ with *p-value* = .0581. Thus we do not reject null hypothesis of no autocorrelation at 5% significance level and there is not enough evidence of serial correlation between $\Delta \boldsymbol{u_{it}}$.

Fixed Effects

The other method of obtaining the estimate of β_1 from the equation (5.7) is the *fixed effects* (or *within*) transformation. Before we estimate the model using the Stata software we again have to verify the assumptions needed for acquiring an unbiased and consistent OLS estimator. The first three assumptions FE.1 through FE.3 (see Section 5.3.2) are fulfilled as well as the FD.1 through FD.3 since we estimate the same model specification using the same data set as in the previous case. However, the strict exogeneity assumption (FE.4) has to be tested in a different way than in the *first difference* estimation. First, we specify the equation (5.7) as:

$$\ln(elprice_{it}) = \delta_1 + \beta_1 \ln(elfromRE_{it}) + \pi_1\, w_{i,t+1} + a_i + u_{it} \tag{5.9}$$

where $w_{i,t+1}$ is a subset of the explanatory variables of the model in the time $(t+1)$, in our case it is the variable $\ln(elfromRE_{i,t+1})$, for t = 2010, 2011, 2012. According to Wooldridge (2002), under strict exogeneity, the parameter $\pi_1 =$ 0. While estimating the equation (5.9) in Stata, we obtained the expected value of $\hat{\pi}_1$ = .0016 with the *p-value* equal to .210, hence the null hypothesis $H_0\colon \hat{\pi}_1 = 0$ cannot be rejected at 5% (or even 20%) significance level and we consider the FE.4 assumption to be verified. Finally, in order to be sure that the FE estimator is unbiased and consistent, the last two assumptions of the *fixed effects* estimation, FE.5 and FE.6, have to be fulfilled as well. We verify them by using the Breusch-Pagan test and

Wooldridge test, respectively, as well as in the case of the FD estimation and neither serial correlation of the idiosyncratic errors nor heteroskedasticity is found in the model.

B.2 CO_2 Emissions and Renewable Energy

In Section 5.4.2, we use the same approach as in Section 5.4.1. Our second model specification is based on the estimation of the following equation:

$$CO2_{it} = \delta_1 + \beta_1 REcons_{it} + a_i + u_{it} \tag{5.10}$$

where, as well as in the model equation (5.7), i = 2,3 … 14 denotes the 13 European countries according to their assigned id numbers; t = 2010, 2011, 2012, 2013 is the time dimension of the panel data set; a_i is the fixed effect; and u_{it} is the idiosyncratic error. For the description of the variables *CO2* and *REcons*, see Section 5.2.3. In comparison to the model equation (5.7), the time dummy variables *d2011*, *d2012*, *d2013* are excluded from (5.10) since they showed to be very statistically insignificant in this model regression and the results fit better without including them.

First Differences

Before we use the *first difference* regression in Stata, we have to verify the six FD assumptions needed for acquiring the unbiased and consistent estimator and valid test statistics (see Section B.1). The first three assumptions, i.e. FD.1 through FD.3, are verified directly by considering the format of the model equation, the way the data set has been collected and the fact that we have a model with a single explanatory variable hence there cannot be any linear relationship among the explanatory variables (FD.3).

The assumption of strict exogeneity in the explanatory variables, FD.4, can be tested the same way as in Section 5.4.1. We run the FD regression and obtain the parameters' estimates for the following equation:

$$\Delta CO2_{it} = \beta_1 \Delta REcons_{it} + \Delta u_{it} \tag{5.11}$$

Then we use the commands *.predict resid, r* and *.summ resid, d* in Stata and look at the expected value of the idiosyncratic errors from the equation (5.11) which is approximately equal to zero ($\boldsymbol{E}(\Delta \boldsymbol{u}_{it} | \boldsymbol{x}_{itj}) = .0001$). Thus, the FD.4 assumption is also considered to be fulfilled. The last two assumptions, FD.5 and FD.6, are tested by the Breusch-Pagan test and Wooldridge test, respectively (see Section B.1 in this appendix for more information). The Breusch-Pagan Chi-squared statistics yields $\chi^2 = 3.637$ with *p-value* = .0565 and the Wooldridge F statistics yields $F = 1.784$ with *p-value* = .2064. Hence there is not enough evidence of either heteroskedasticity or serial correlation between the differences in the idiosyncratic errors as we cannot reject the null hypotheses of homoskedasticity and no autocorrelation, respectively, at 5% significance level.

Fixed Effects

As well as in Section B.1, we also use the *fixed effects* (or *within*) transformation to obtain the estimate of β_1 from the equation (5.10) and then compare the results with the FD estimation. As in the previous cases, the assumptions needed for acquiring an unbiased and consistent OLS estimator have to be verified first. The assumptions FE.1 through FE.3 (see Section B.1) are fulfilled as well as the FD.1 through FD.3 as we estimate the same model equation (5.10) with the same data set in both cases. To verify the assumption FE.4, we specify the equation (5.10) as:

$$CO2_{it} = \delta_1 + \beta_1 REcons_{it} + \pi_1 w_{i,t+1} + a_i + u_{it} \tag{5.12}$$

where $w_{i,t+1}$ is a subset of the $REcons_{i,t+1}$, for t = 2010, 2011, 2012. According to Wooldridge (2002), under strict exogeneity, the parameter π_1 has to be equal to 0. By using Stata, we obtained the expected value $\hat{\pi}_1 = .008$ with the *p-value* equal to .678. Thus, the null hypothesis $H_0: \hat{\pi}_1 = 0$ cannot be rejected at 5% significance level and we consider the FE.4 assumption to be fulfilled. To verify the last two assumptions, FE.5 and FE.6, we once more use the Breusch-Pagan test and Wooldridge test, respectively, as well as in the case of the FD estimation. Since neither serial correlation of the idiosyncratic errors nor heteroskedasticity is found, we can proceed to the regression results.