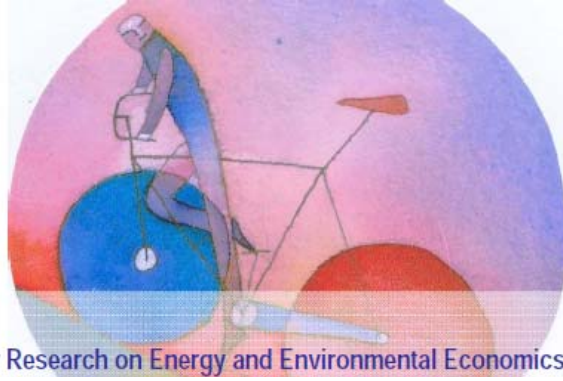


**Bocconi**

**IEFE**

Centre for Research on Energy and Environmental Economics and Policy



Working Paper Series - ISSN 1973-0381

**Environmental Policy Performance and its  
Determinants: Application of a three-level random  
intercept model**

*Marzio Galeotti, Yana Rubashkina, Silvia Salini  
and Elena Verdolini*

**Working Paper n. 71**  
August 2014

***IEFE - The Center for Research on Energy and Environmental  
Economics and Policy at Bocconi University  
via Guglielmo Röntgen, 1 - 20136 Milano  
tel. 02.5836.3820 - fax 02.5836.3890  
www.iefе.unibocconi.it – iefe@unibocconi.it***

This paper can be downloaded at [www.iefе.unibocconi.it](http://www.iefе.unibocconi.it)  
The opinions expressed herein do not necessarily reflect the position of IEFE-Bocconi.



# Environmental Policy Performance and its Determinants: Application of a three-level random intercept model <sup>☆</sup>

Marzio Galeotti<sup>a,b</sup>, Yana Rubashkina<sup>c</sup>, Silvia Salini<sup>a</sup>, Elena Verdolini<sup>d,\*</sup>

<sup>a</sup>*DEMM, Università degli Studi di Milano*

<sup>b</sup>*IEFE, Università Bocconi Milano*

<sup>c</sup>*DEFAP, Università Cattolica del Sacro Cuore, Milano*

<sup>d</sup>*Fondazione Eni Enrico Mattei and CMCC*

---

## Abstract

We propose the use of a two level random intercept model to measure the degree of environmental policy performance of different countries and to study its determinants. Inspired by the literature on multilevel latent models and Item Response Theory (IRT), this framework treats policy commitment as a latent variable which is estimated conditional on the difficulty of the policy portfolio implemented by each country. We contribute to the study and scoring of environmental and energy policies in three main ways. First, the model results in a ranking of countries which is conditional on the complexity of their chosen policy portfolio. Second, we provide a unified framework in which to construct a policy indicator and to study its determinants through a latent regression approach. The resulting country ranking can thus be cleaned from the effect of economic and institutional observables which affect policy design and implementation. Third, the model estimates parameters which can be used to describe and compare policy portfolios across countries. We apply this methodology to the case of energy efficiency policies in the industrial sectors of 29 EU countries between 2004 and 2011. We conclude by highlighting the future possible applications of this approach, which are not confined to the realm of environmental and energy policy.

*Keywords:* Energy policy, environmental policy, ranking, policy portfolios

**JEL codes:** Q58, O57, C33

---

## 1. Introduction

Addressing pressing climate concerns by limiting greenhouse gas emissions, supporting renewable energy sources and increasing energy efficiency is one of the five objectives of Europe 2020, the sustainable growth strategy that EU member states launched in 2010 as a response to

---

<sup>☆</sup>Silvia Salini, Elena Verdolini and Marzio Galeotti gratefully acknowledge funding from, respectively, the MIUR PRIN project MISURA, the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement no. 308481 (ENTRACTE), and the MIUR PRIN 2010-11 Project “Climate changes in the Mediterranean area: evolutionary scenarios, mitigation policies and technological innovation”.

\*Corresponding author

*Email addresses:* marzio.galeotti@unimi.it (Marzio Galeotti), yrubashkina@gmail.com (Yana Rubashkina), silvia.salini@unimi.it (Silvia Salini), elena.verdolini@feem.it (Elena Verdolini)

the recent global economic crisis.<sup>1</sup> To support these goals, which have been declined in national targets for each member state, for 2020 a number of environmental and energy policies have been put in place so far and will be implemented in the years to come. Moreover, a new proposal concerning targets for emissions, renewable energy, and possibly energy efficiency for 2030 is under approval by EU institutions by the end of this year.

A major challenge researchers and policy makers alike have been struggling with is the assessment of such policies and how countries are performing in this respect. This question is important for both policy evaluation and for the research community. Appropriately describing and understanding the past performance of countries with respect to energy and environmental policies, and their ability to implement a more or less complex portfolio of policy instruments, is a crucial step in ensuring that future interventions are drafted in a sound and cost-effective way. An in-depth analysis in this respect is however currently missing, partly due to lack of appropriate data, but also to more complex conceptual problems linked with creating indicators of policy implementation and stringency.

As pointed out in Brunel and Levinson (2013), assessments of environmental and energy policy are often characterized by a series of important shortcomings. First, to address climate and energy concerns countries can choose from a wide array of policy instruments, each of which is characterized by a different level of effectiveness, dynamic efficiency and political acceptability (Fisher and Newell, 2008). This “multidimensionality” translates into the challenge of building a policy indicator able to capture the different aspects of a country’s policy portfolio. Second, countries with worse pollution challenges might impose more stringent options. Not accounting for “simultaneity” would hence provide a biased indicator of environmental policy stringency. Third, the ability of countries to implement certain (lower cost) options might depend crucially on some “initial condition” or on some time varying characteristics, such as the level of energy efficiency of the economy, the age of capital, the industrial composition or the political considerations of the acceptability of stringent policy.

This paper contributes to the literature in several ways. First, it provides a novel approach to assessing and comparing countries’ environmental policy portfolios and performance which address the issue of multidimensionality. In our framework, inspired by the literature on multilevel latent models and Item Response Theory (IRT), each country is assigned a score which accounts for the full set of policy instruments it implements. The resulting country ranking hence accounts for the complexity of the chosen policy portfolio.

Second, our approach accounts for the comparative advantage determined by “initial conditions” or time varying characteristics, such as the level of energy efficiency of the economy. This is done through the use of a latent regression model which cleans each country’s score from the effect of specific observables likely to affect policy instrument choice and implementation.

Finally, given this set-up our analysis also addresses the issue of “simultaneity”. The proposed three-level random intercept model with a latent regression provides a unified framework to both create an indicator of policy stringency and to study its determinants. It is therefore of

---

<sup>1</sup>The Europe 2020 strategy includes five main objectives: ensuring 75 % employment of 20/64-year-old; Getting 3 % of the EU’s GDP invested into research and development; limiting greenhouse gas emissions by 20 % or even 30 % compared to 1990 levels, creating 20 % of EU energy needs from renewables and increasing energy efficiency by 20 %; reducing school dropout rates to below 10 %, with at least 40 % of 30/34-year-old completing tertiary education; ensuring 20 million fewer people are at risk of poverty or social exclusion. The environment and energy objective summarizes the so-called 20-20 Climate and Energy Package approved in 2007 by the EC and subsequently translated into a set of five directives approved in 2009.

potential relevance also for future applications on a variety of research questions, where a key requirement is the creation of a policy indicator cleaned from reverse causality and from the effect of covariates.

The rest of this paper is organized as follows. Section 2 provides a review of the available literature and highlights the contributions of this paper. Section 3 presents the proposed statistical model. Its empirical application, which focuses on energy efficiency policies in Europe, is presented in Section 4. We describe therein the data and report the empirical results, which include country rankings which account for (a) the complexity of the policy mix put into place and (b) the effect of economic and institutional observables. Section 5 concludes with a summary of main results, policy implications and a list of future research avenues.

## 2. Literature Review

Assessing the economic impact of policy decisions is of central interest to Economics. As environmental and energy policy has become increasingly active, especially in Europe, in the last decade, several papers have appeared aiming to ascertain the consequences of decisions concerning energy efficiency, renewable energy sources, emission reductions, and the like, for innovation activity, economic growth or overall economic performance. A critical issue is of course the definition of an appropriate indicator of policy intervention and stringency. This is a topic that has received recently increasing attention. Brunel and Levinson (2013) provide a comprehensive review of the literature in this respect.

Popular proxies for regulatory stringency are data on private sector abatement expenditures (Pollution Abatement Costs, or PACs). Such data inform on the level of financial effort a given firm/sector has to face to comply with given standards (Lanjouw and Mody, 1996; Jaffe and Palmer, 1997; Berman and Bui, 2001; Hamamoto, 2006; Rubashkina et al., 2014). The use of this indicator is based on the assumption that profit maximizing firms typically face marginal abatement costs that are increasing in pollution abatement. Other popular indicators of choice include reductions in emissions or pollutants or indicators based on energy use (Cole and Elliot, 2003; Gollop and Roberts, 1983). Changes in regulation-based measures have also been used to judge the level of policy stringency (Popp, 2003, 2010). An extensive literature also resorts to building general composite indexes through the use of aggregation techniques. The data used to this end include information on the presence or absence of a given policy (0-1 indicators) or scores from surveys of government officials or business leaders (Tobey, 1990; Kellenberg, 2009). Finally, many have resorted to *ad hoc* data sets which are tailored to answering a specific research question (Jeppsen and Folmer, 2001).

Brunel and Levinson (2013) nicely describe the main conceptual issues that plague almost all previous efforts to create an index of energy and environmental policy stringency. First, creating a reliable indicator is challenging due to the issue of “multidimensionality”. Governments regulate various aspects of energy production and environmental protection, namely air, water, toxic chemicals, but also energy efficiency and renewable energy production. Moreover, policy instruments can be aimed at regulating pollution directly, through either a command-and-control or a market-based approach. In addition, environmental and energy policies *per se* can be combined with policies aimed at addressing the knowledge market failure externality, and can stimulate the creation and diffusion of less polluting technologies.<sup>2</sup> Such heterogeneity in policy responses and

---

<sup>2</sup>Environmental (and energy) policy directly targets the environmental externality by regulating pollutants or emis-

in the sectors targeted makes it hard to build an indicator that is at the same time comprehensive and detailed enough to capture changes in all different aspects of a country's policy portfolio.

Second, while policy makers and researchers ideally would want to measure the effect of policy on other important outcome variables such as industry location, trade patterns, economic growth or knowledge transfer, the variables measuring the stringency of environmental regulation are plagued by simultaneity and endogeneity. One must therefore bear in mind that policies are often jointly determined with other outcome variables and that they themselves are not exogenous, but are the result of forces within the economic system.

Finally, some countries/sectors might have a "comparative" advantage with respect to others in implementing strict environmental policy. This might be due to their industrial composition, but also to the vintage of capital or to the fact that they are more polluting to begin with. This gives them the option to implement low-cost high-reduction (or high-efficiency improvement) policies.

Hence, PACs are plagued with reverse causality issues and in the presence of market or behavioral failures they no longer successfully measure the level of regulatory pressure (Berman and Bui, 2001).<sup>3</sup> Emission or energy-use based indicators are also likely to mirror changes other than regulatory stringency, such as for example factor prices. Moreover, when used at the disaggregated level, it is often hard to build indicators that can be used in cross sectoral or cross country analyses due to the heterogeneity of the regulated pollutants. Proxies based on normative prescriptions do not account for the level of actual enforcement of a given policy and might also be subject to issues of reverse causality (Brunnermeier and Cohen, 2003; Shimshack and Ward, 2005). Finally, most composite indicators are built using simple approaches such as the sum of policy instruments.

Vona and Nicolli (2012) propose two different aggregate indicators to measure the level of renewable environmental policies in European countries. First, an average-based indicator which uses information on the timing of adoption of a given policy instrument (namely, an average of dummy variables indicators equal to zero before the instrument is put into place and equal to one afterwards). Second, a more complex indicator is built using principal component analysis (PCA) relying both on dummy variables and on intensity of specific policy instruments such as Renewable Energy Certificates of Feed-in Tariffs. This approach is more sophisticated than previous efforts in this sense, since the factor loadings resulting from the PCA in Vona and Nicolli (2012) can be interpreted as importance weights which vary by item/policy. In their approach, however, PCA is built using binary indicators of the presence/absence of a given policy instrument and no consideration is given to the number of policies in place at any given time.<sup>4</sup>

In this paper, we propose an approach inspired by IRT to describe and study the level of environmental policy within a set of countries. We extend the basic IRT model to a multilevel framework and use it to build indicators of environmental policy performance. While the basic approach we propose has been applied in the statistical literature of scoring, its application to

---

sions. On the one hand, command-and-control policy instruments include mandates and standard, which set a minimum requirement for firms to comply with. On the other hand, market-based approaches such as taxes and permits allow firms to respond more flexibly to comply with the regulation. Conversely, technology policy targets the knowledge market failure and supports R&D in greener and more efficient technologies with, among other options, research subsidies and investments.

<sup>3</sup>Moreover, if the data are used at the aggregate level, such as sectors or countries, changes in PACs might result from changes due to unobserved heterogeneity rather than from changes in regulatory stringency.

<sup>4</sup>Ferrari and Salini (2011) also holds that the presence of dummy variables should require Categorical Principal Components Analysis (CATPCA) which is not based on the assumption of linear correlation.

longitudinal data and to the context of environmental and energy policy is new. Moreover, the use of a latent regression allows cleaning of the indicator from the effect of observables which are likely to affect both policy instrument choice and implementation.

There are three main outcomes of interest from our proposed model. First, we estimate different weights (or difficulty levels) for each different policy instrument implemented over the sample period. Second, we rank countries (a) conditional on the specific complexity of each country's policy portfolio and (b) accounting for advantages deriving from "initial conditions" or specific country observables. Third, the resulting country score represents an indicator of policy implementation which is cleaned from issues of reverse causality and simultaneity.

In the next section we present our extended multi-level model. The empirical application which focuses on policies supporting energy efficiency in Europe follows.

### 3. Empirical Strategy

The starting point of our analysis is the Rasch model (RM) (Rasch, 1980), a statistical model for dichotomous data used as a psychometric tool in social sciences to compare the performance of various subjects in questionnaires/tests. The RM assumes that the performance of each subject is the outcome of two different aspects: the "difficulty" of each of the items the subject is faced with, and the "ability" of each subject, which is modeled as a latent variable. In this framework, the probability of observing a positive/correct response for each item for any given individual is a decreasing function of item difficulty/complexity and an increasing function of the subject latent trait, or "ability".

Unlike other aggregation techniques, this model scores each subject and each item under consideration along a continuum. Extensions of the RM have been subsequently developed to allow handling ordinal observations, as in our case, and are generally classified within the framework of IRT models (Mair and Hatzinger, 2007). The application of these models has recently extended beyond psychometrics to psychology, education, and medicine, among others. For instance, Bacci (2012) and Bacci and Bartolucci (2012) apply IRT to the scoring of quality of life and Bacci and Caviezel (2011) use it to score teaching evaluation. Recently, these models found application also in organizational and management studies in particular for financial issues (Soutar and Cronish-Ward, 1997), marketing and consumer behavior (Fischer et al., 2006; Salzberg and Sinkovics, 2006) and tourism management (Oreja-Rodriguez and Yanes-Estevez, 2007). Ferrari et al. (2005) explore the validity and constraints of this approach as a tool to quantify the degree of vulnerability of historical-architectonical buildings in Northern Italy. Finally, Murray and Mills (2012) applies a similar methodology to the scoring of energy insecurity in the United States.

There are two main features of interest in this approach. First, these models allow the treatment of many interrelated variables with the aim of summarizing data and highlighting possible latent factors. Second, the data requirements are relatively limited. Specifically, the model requires knowledge of how countries "perform" with respect to different "items".

To apply the IRT approach to the realm of environmental policy performance, we define each item as a specific instrument to address environmental policy, and we score countries on the number of policies active at a given time in each given "item". Hence, in our framework countries are the observed subjects and the different policy instruments that they can implement represent the different items. Based on these models, each country can be characterized by a score on an ordinal scale for each of the different policy instruments considered. However, we modify

the traditional approach of IRT models to study the performance of countries in a longitudinal framework.

Specifically, we propose a three-level random intercept ordinal logistic model for adjacent item scores where each policy  $i$  (one of the different environmental policy instruments a country can choose to implement) is the first level unit, measurement  $t$  (the time period in which responses to each policy are observed) is the second level unit and country  $j$  is the third level unit. This model can be written as a multilevel Partial Credit Model (PCM) as follows: Bacci and Caviezel (2011) :

$$P(Y_{itj} = m | \theta_{0tj}, \theta_{00j}) = \frac{\exp\left[\sum_{k=0}^m (\theta_{0tj} + \theta_{00j} - (\beta_i + \tau_{ik}))\right]}{1 + \sum_{l=1}^{M-1} \exp\left[\sum_{k=0}^l (\theta_{0tj} + \theta_{00j} - (\beta_i + \tau_{ik}))\right]} \quad (1)$$

where  $P(\cdot)$  is the probability that  $Y_{itj}$ , namely the score on policy  $i$  ( $i = 1, \dots, I$ ) in time  $t$  ( $t = 1, \dots, T$ ) for country  $j$  ( $j = 1, \dots, n$ ), assumes a given value  $m$ . The parameter  $\beta_i$ , which in the traditional model indicates the average difficulty of the  $i$ th item, is now associated with the average difficulty of the  $i$ th policy instrument. The parameter  $\tau_{ik}$  indicates the different threshold in each item score. A threshold is intended as the point in which two adjacent item scores have the same probability to be chosen. In a PCM the distance between thresholds for each item score can differ; in addition, thresholds can differ freely from one item to another (Masters, 1982). Simpler Rasch polytomous models exist in which threshold values for all the items are assumed to be equal even if the distance between thresholds can differ (Andrich, 1978). In our context, however, a PCM is more coherent, since difficulty levels can be assumed to differ between policies and the country should then receive a partial credit (score for each policy) equivalent to the relative level of difficulty of the performance achieved.

The main difference between the present formulation and the “traditional” PCM lies in the fact that we account for two levels of clustering of the data, namely time and country, rather than only one. Hence, the presence of two new random effects  $\theta_{0tj}$  and  $\theta_{00j}$ , instead of the traditional subject (country) random effect  $\theta_j$  indicating the level of the latent trait for the  $j$ th country. In our framework, the second level residuals  $\theta_{0tj}$  indicate the deviation of the latent variable  $\theta$  (e.g. the composite indicators) for year  $t$  and country  $j$  from the average value of country  $j$ : accordingly, they allow for an analysis of time within each country. On the other hand, third-level residuals  $\theta_{00j}$  indicate the deviation of the latent variable for county  $j$  from the average value of the population: they thus allow for a ranking of countries in terms of the mean level of the latent variable.  $\theta_{0tj}$  and  $\theta_{00j}$  are independent and normally distributed with mean zero and constant variances.

The third-level residuals  $\theta_{00j}$  represent the source of information of greatest use in our case as they allow the ranking of countries with respect to their observed composite indicator of environmental and energy policies over the sample period. Moreover, the second level residuals,  $\theta_{0tj}$ , allow tracking the performance of each country throughout the sample period with respect to its average. They are both obtained as the expected a posteriori (empirical) Bayes predictions, namely the posterior distributions of the country parameters given the policy responses. In addition, the threshold parameters can be used to characterize the level of complexity/difficulty of the policy portfolio implemented in each year.

Our statistical framework therefore provides (1) threshold parameters measuring the intrinsic difficulty/probability of observing a given categorical response for each item/policy instrument; (2) time-country specific intercepts for each country over time and (3) country-specific parameters which allow for an overall country ranking in the period under consideration. The elements



(2) and (3) are derived conditional on the policy instrument difficulty levels.

A problem with the country indicators (2) and (3) mentioned above is that they are derived from observing the performance of each country, but they do not account for the fact that such performance can be affected by some “initial condition” or by the presence of enabling factors that make the choice of a given policy portfolio more or less likely. The ability of each country to implement, say, energy efficiency policies can indeed be thought of as the combination of two different effects. On the one hand, some institutional and economic characteristics are likely to affect each country’s ability to pursue efficiency improvements. For example, richer countries might have more room to phase out old capital equipment. On the other hand, conditional on these observable characteristics some countries might be truly more committed to environmental policy than others.

The ranking emerging from a descriptive model like the one discussed so far does not discriminate between these two different effects. A country, for instance, could rank highest because it has better starting conditions or due to a true commitment to environmental policy with respect to other countries, or both. To disentangle these effects, we combine the multilevel PCM with a latent regression model which explains the variation in the time and country posterior Bayes estimates using a set of covariates. De Boeck and Wilson (2004) refer to this case as the *latent regression Rasch model*, namely a *subject explanatory model*, which includes subject properties as explanatory variables for the differences in between-subject scores.

Given the three level structure of our model, we include a vector of covariates  $\mathbf{X}_j$  for the third-level residuals  $\theta_{00j}$  for each country  $j$ :

$$\theta_{00j} = \mathbf{X}_j \boldsymbol{\zeta} + \varepsilon_{00j} \quad (2)$$

and a vector of covariates  $\mathbf{Z}_{tj}$  for the second level residuals  $\theta_{0tj}$  for each year  $t$  and for each country  $j$

$$\theta_{0tj} = \mathbf{Z}_{tj} \boldsymbol{\gamma} + \varepsilon_{0tj} \quad (3)$$

The new statistical model becomes:

$$P(Y_{itj} = m | \theta_{0tj}, \theta_{00j}) = \frac{\exp \left[ \sum_{k=0}^m (\mathbf{X}_j \boldsymbol{\zeta} + \mathbf{Z}_{tj} \boldsymbol{\gamma} + \varepsilon_{0tj} + \varepsilon_{00j} - (\beta_i + \tau_{ik})) \right]}{1 + \sum_{l=1}^{M-1} \exp \left[ \sum_{k=0}^l (\mathbf{X}_j \boldsymbol{\zeta} + \mathbf{Z}_{tj} \boldsymbol{\gamma} + \varepsilon_{0tj} + \varepsilon_{00j} - (\beta_i + \tau_{ik})) \right]} \quad (4)$$

The random effects for the second and third level that are obtained as the expected a posteriori (empirical) Bayes predictions are now the residual components  $\varepsilon_{0tj}$  and  $\varepsilon_{00j}$ . These residuals are cleaned from the “comparative advantage” effect or any causality between the policy indicator and the covariates. They can therefore be interpreted as measures of commitment to environmental policy.

In the next section, the three level random intercept model is applied to a specific data set for EU countries, highlighting the potential of this approach to score and study policy implementation.

## 4. Empirical Application

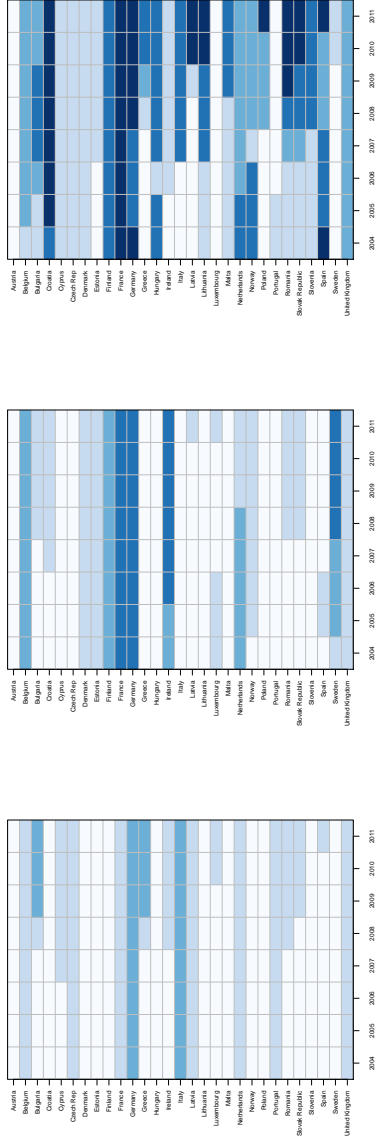
### 4.1. Data Description

The data we use in this application is given by a set of policy indicators for a sample of 29 countries over the years 2004-2011. The data are extracted from the *Mesures d'Utilisation Rationnelle de l'Energie* (MURE) database, which collects information on the adoption time of selected energy efficiency policy measures for the rational use of energy and for the promotion of end-use renewables in the manufacturing sector of EU Member States. The MURE database includes the national policies that have macro-economic impact, imposing a quality threshold which eliminates low-impact policies.

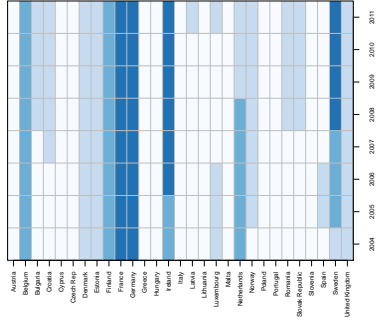
We are interested in scoring the performance of different EU countries which address the environmental externality by implementing policies which improving the efficiency of energy use. We thus use information on five different environmental policy instruments: Regulatory policies, Voluntary measures, Financial instruments, Fiscal/tax reductions and Information/Education. Regulatory policies include norms and standards, such as energy efficiency levels for various kinds of equipment and production processes or products, which often are based on the phase out of old technologies. Voluntary agreements include the creation of industry/government co-operation, as well as various industry initiatives aimed at promoting higher levels of energy efficiency. Financial instruments include investment subsidies and low interest loans, as well as incentives and subsidies for energy audits. Fiscal measures include tax credits and exemptions which are put in place to target higher levels of efficiency within industrial sectors. Information and education policies and measure are aimed at increasing the awareness of technology users and their knowledge about opportunities for efficiency improvements.

We construct an ordered categorical variable for each policy instrument by counting the specific policies which are active in any given year. The number of categories varies for each instrument. For Regulatory policies we have: 0, 1, 2 or more; Voluntary measures: 0, 1, 2, 3 or more; for Financial instruments: 0, 1, 2, 3, 4 or more; for Fiscal/tax reductions: 1, 2, 3 or more; for Information/Education: 0, 1, 2, 3 or more. Figure 1 provides an overview of the data by type of instrument and country.

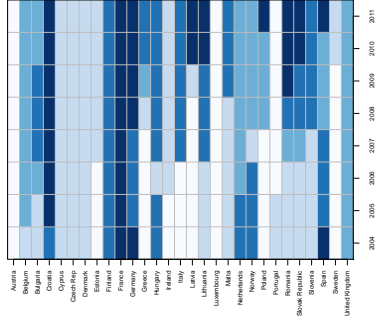
As apparent from Figure 1, some countries such as Germany, France and Finland, score higher than other countries. However, they also exhibit lower variation throughout the sample period. Most Eastern European countries, such as for example Romania, increase their commitment to energy efficiency towards the middle of the sample period. Regulatory measures are among the least implemented across the sample, while Financial instruments are widely used. Voluntary policy and Education are relatively unexploited in most countries.



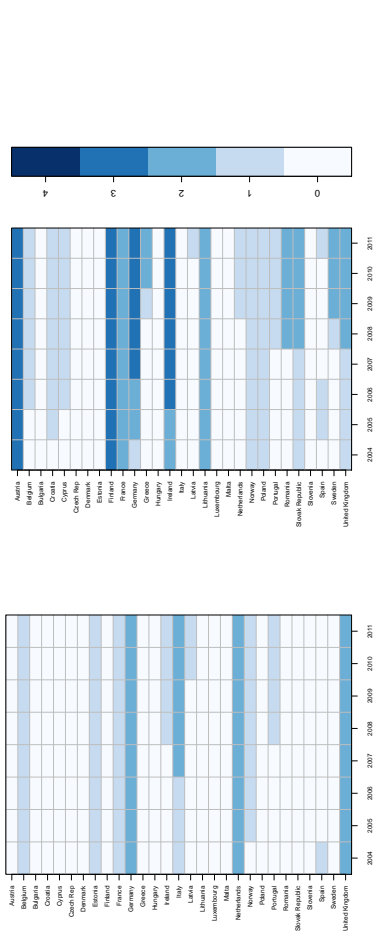
(a)



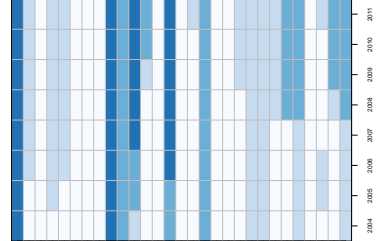
(b)



(c)



(d)



(e)

Figure 1: Policy variables distributions per country and year. (a) Regulatory Policy, (b) Voluntary agreements, (c) Financial instruments, (d) Fiscal/tax reduction, (e) Information/Education

In theory, we could have extended the number of “items” (policy instruments) to also include those policies which promote energy efficiency by addressing the knowledge externality through, for example, R&D investment or support to innovation. We chose not to do so for two reasons. First, the timing of the impact of these two policy instruments is very different. Policies aimed at supporting innovation work on a longer time frame than policies directly regulating the environmental externality. Their impact is likely to occur with a lag because the invention of new technologies requires time for development. Second, given this difference in timing, the availability of novel improved technologies is in fact an enabling factor that would make the implementation of stringent environmental policies more likely. In fact, setting stringent performance standards or costly taxes on energy use is more politically and socially feasible if highly efficient equipment is already available to firms and consumers to purchase. Hence, measures related to increasing the supply of efficiency technologies are considered as entering the latent regression that explains policy implementation in our framework.<sup>5</sup>

The multilevel PCM specifications with and without latent regression on the subject parameters are estimated using the GLLAMM routine in STATA (Rabe-Hesketh et al., 2004; Bacci and Caviezel, 2011).

#### 4.2. Multilevel Partial Credit Model

Estimates for each of the parameter thresholds which are reported in Table 1.

<b>Difficulty Thresholds</b>	<b>Coeff.</b>	<b>SE</b>	<b>p-value</b>
Regulatory Policy - Threshold 1	0.47	0.22	0.03
Regulatory Policy - Threshold 2	1.87	0.31	0.00
Voluntary Measures - Threshold 1	0.51	0.23	0.03
Voluntary Measures - Threshold 2	1.01	0.29	0.00
Voluntary Measures - Threshold 3	0.87	0.35	0.01
Financial Instruments -Threshold 1	-0.73	0.26	0.00
Financial Instruments -Threshold 2	0.35	0.26	0.19
Financial Instruments -Threshold 3	-0.17	0.28	0.53
Financial Instruments -Threshold 4	1.14	0.30	0.00
Fiscal/Tax Reductions - Threshold 1	1.29	0.23	0.00
Fiscal/Tax Reductions - Threshold 2	2.43	0.41	0.00
Information/Education - Threshold 1	0.43	0.23	0.07
Information/Education - Threshold 2	0.57	0.27	0.03
Information/Education - Threshold 3	1.10	0.33	0.00
Variances of random effects			
Second Level Variance (Time)	9.46e-21	1.953e-11	0.00
Second Level Variance (Country)	0.73	0.22	0.00

Table 1: Item/policy thresholds

<sup>5</sup>This notwithstanding, the results including items indicating support to technology development (and excluding these variables from the latent regression) are available from the authors upon request, and do not significantly differ from the ones presented here.

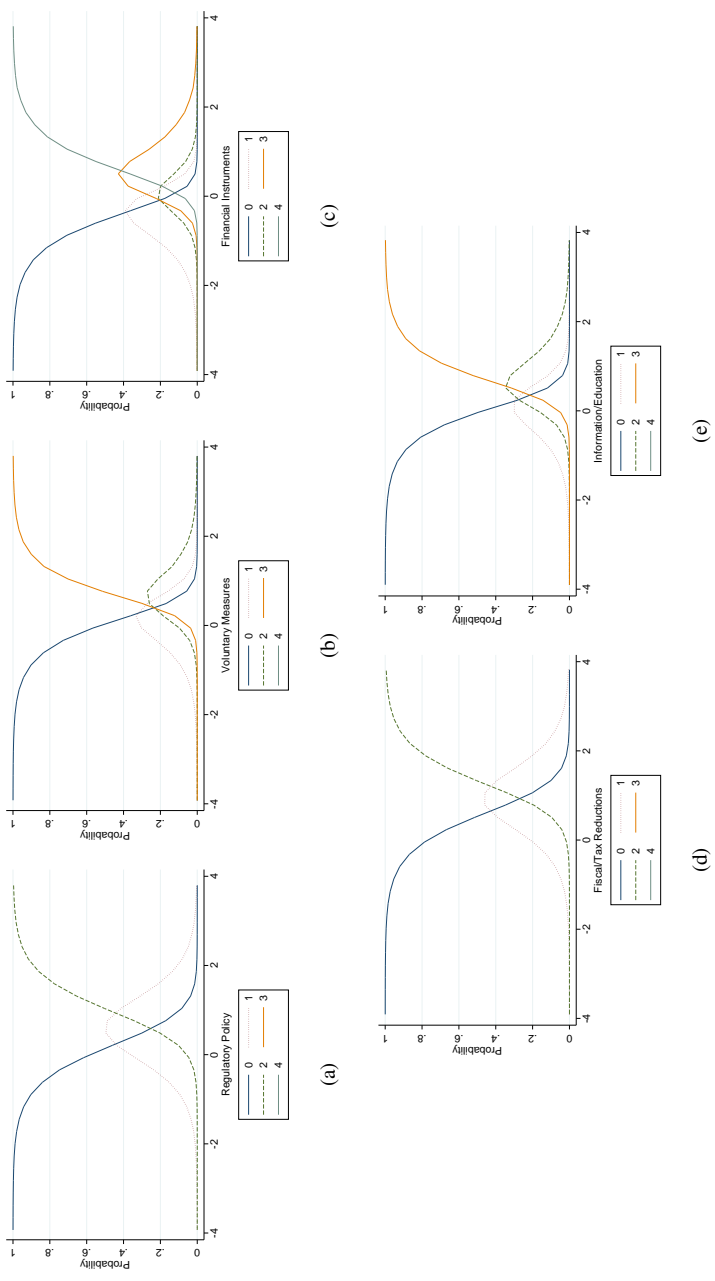


Figure 2: Category probability curves. (a) Regulatory Policy, (b) Voluntary agreements, (c) Financial instruments, (d) Fiscal/tax reduction, (e) New Market Instruments, (F) Information/Education

Conditional on the estimated item thresholds, the variance of the second level (time) and third level (country) residuals (and associated standard errors SE) are estimated at  $9.46e-21$  (SE:  $1.953e-11$ ) and  $0.730$  (SE:  $0.22$ ), respectively. Both effects are significant. Note however that the second level variance, which indicates deviations over time of each country's aggregate score from its own mean, is extremely low in absolute value, while the third level variance, indicating the variation of country scores from the overall mean, is higher in absolute value. The between country variation over our sample period is thus much higher than within country variation over time.

Given the estimated variance of the second and third level residuals, we can obtain the latent traits as the posterior Bayes estimates from the model. We focus first on the country latent trait, and present a ranking of countries in Figure 3 which also displays confidence intervals for the estimates. The ranking emerging from the multilevel PCM indicates that, conditional on the difficulty of the chosen policy portfolio, Luxembourg and Portugal are among the worst performing countries in Europe in terms of addressing energy efficiency concerns. Note that confidence intervals for the different countries in our sample greatly overlap, indicating that the performance of different EU countries is not strikingly different. Exceptions are Germany and France, which appear at the top of the ranking and whose confidence intervals do not overlap those of other countries.

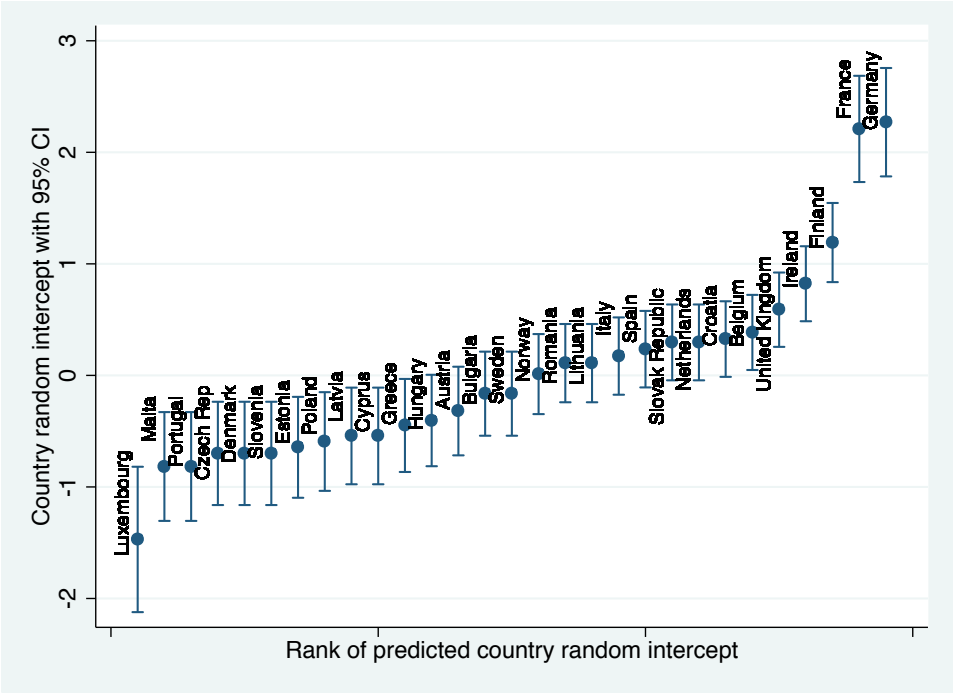


Figure 3: Rank of predicted country random intercept for multilevel PCM

Finally, we plot the posterior Bayes estimates for each year and country in Figure 4. The figure displays the evolution of the time profile (second level residual) which differs between countries.

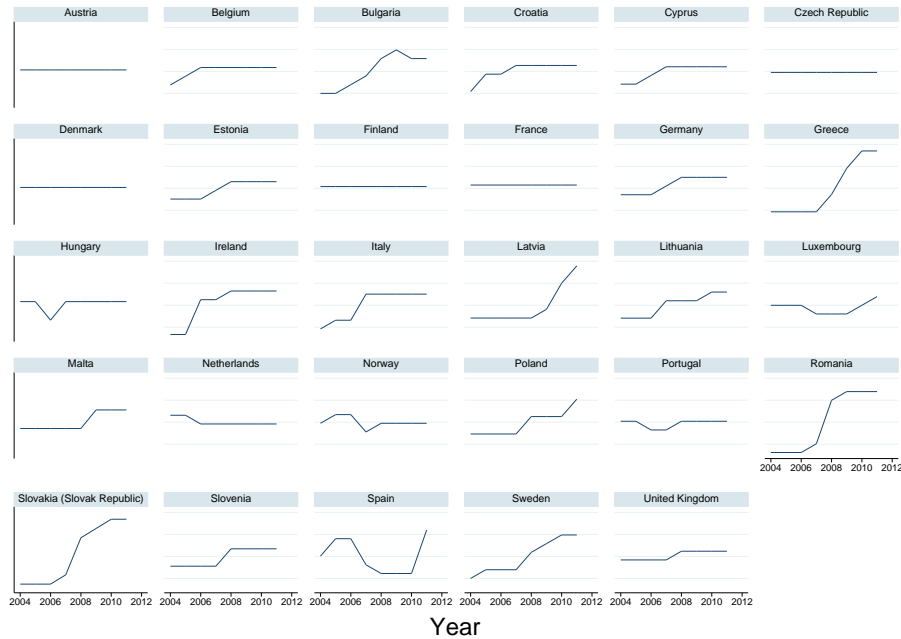


Figure 4: Predicted random effect for each year, by country

Note that among the New Member countries Bulgaria, Latvia, Lithuania, Poland, Romania and the Slovak Republic considerably increase their policy effort over time. Conversely, the improvements during the early years in Croatia and Cyprus are not sustained over time. As far as the Old Member countries are concerned, Greece, Ireland and Sweden show highest increase of their policy effort over time. Italy also increases its commitment, but only in the early years of the sample. Several countries that score low on policy performance (Figure 3) do not show improvements over time. This is the case for Luxembourg, Poland, Malta, the Czech Republic and Slovenia.

#### 4.3. Explanatory Multilevel Partial Credit Model

In the previous section, we presented results on the latent trait and time profile of countries' performances based on a PCM. However, these latent traits (and hence the ranking emerging from Figure 3) are likely affected by certain observable country characteristics. The fact that Germany and France score best in the sample might be the result of a true higher commitment and effort to tackle energy efficiency in these countries, or simply the result of the fact that these countries are among the richest in the sample.

A number of observable characteristics are thus likely to affect each country's rank in Figure 3. First of all, the likelihood of targeting efficiency in the industrial sector will depend on the weight of this sector within each economy. Our expectation in this respect is that the higher the share of GDP accounted for by manufacturing, the higher the commitment to fostering energy efficiency. On the one hand, the potential savings from increased energy efficiency would be

high. On the other hand, industry lobbies are likely to put pressure on governments to establish measures such as fiscal incentives aimed at lowering the costs of energy inputs. Second, countries with higher levels of efficiency might have already reaped the low hanging fruits, and either not feel the need to address the issue of energy efficiency through regulation or find it harder to implement any additional policy. Third, the level of dependence of a given country on energy inputs is also a factor likely to affect its propensity to tackle energy efficiency as a way to achieve energy security. Finally, the availability of energy efficient technologies within any given economy will increase the likelihood that pro-efficiency regulation is passed given that a certain degree of improvement should in theory be easily reached.

As explained in the previous section, we thus extend the model to condition each of the estimated latent traits on these observables. Specifically, we select the share of manufacturing in GDP (WorldBank, 2013), the level of energy efficiency (WorldBank, 2013), the share of energy imports (WorldBank, 2013) and the number of energy efficiency patent applications to the EPO by applicants in a given country (OECD, 2013). Descriptive statistics of these variables by country and for the overall sample are displayed in Table 2. All these variables vary greatly between countries, with the exception of the share of manufacturing in GDP, which exhibits low within country variation. For this reason, we include lagged values of energy efficiency, share of energy imports and efficiency patents as time-varying country-level regressors. The share of manufacturing in GDP is instead included as a country-varying regressor to explain third level (between countries) variation.



Country	Energy Efficient Patent Applications to EPO	SD	Energy Intensity	SD	Fuel Imports % of Merchandise Imports	SD	Manufacturing in VA % of GDP	SD	Law and Order (mean)
All countries	32.48	82.16	249.1	159.3	13.06	5.187	16.86	4.826	5.030
Austria	26.97	12.31	133.9	5.360	11.34	1.502	19.54	0.625	6
Belgium	19.99	6.795	190.5	8.498	12.80	2.207	16.10	1.193	5.475
Bulgaria	0.750	0.664	792.6	97.28	14.90	8.148	15.86	0.756	4.261
Croatia	0.125	0.332	241.1	13.80	16.40	17.92	16.63	0.911	4.917
Cyprus	2.250	1.003	188.0	9.323	18.20	3.527	8.198	0.700	4.315
Czech Republic	2.875	2.663	411.4	41.31	8.520	1.391	24.44	1.103	5.204
Denmark	85.92	45.97	96.79	4.428	6.666	1.342	13.79	0.769	6
Estonia	0.875	0.930	500.7	45.59	14.67	3.872	16.49	1.017	4
Finland	20.50	11.83	228.4	14.92	15.74	2.943	22.09	2.372	6
France	89.81	50.77	155.8	5.994	14.20	1.802	12.47	0.951	5.148
Germany	385.6	195.4	149.3	7.270	11.47	1.428	22.24	1.517	5.407
Greece	3.563	2.636	156.2	6.859	18.72	4.526	9.38	.	3.969
Hungary	9.625	7.883	93.29	4.149	9.505	2.937	22.97	1.312	4.787
Ireland	54.75	24.03	126.8	3.847	14.35	3.596	17.96	0.926	5.404
Italy	0.375	0.698	344.1	32.11	14.00	2.101	11.71	1.026	4.865
Latvia	0	0	403.8	58.77	25.20	5.571	19.15	1.622	4
Lithuania	5.063	4.608	146.6	10.41	9.402	1.347	7.980	1.394	6
Luxembourg	0.313	0.557	185.6	11.57	15.25	9.377	15.21	1.213	4.567
Malta	55.13	25.60	156.9	5.831	14.34	2.635	13.73	0.602	6
Netherlands	19.19	10.10	115.9	29.22	10.65	0.878	9.650	0.418	6
Norway	3.563	3.127	362.5	6.776	14.05	1.060	18.70	0.221	4.506
Poland	3.650	2.694	165.3	7.661	11.80	2.483	14.36	0.729	5.090
Portugal	0.563	0.585	461.3	59.95	12.71	1.569	29.90	0	3.682
Romania	1.125	0.930	439.4	69.28	11.42	1.061	22.79	1.729	4.593
Slovak Republic	0.750	0.664	241.5	13.56	16.15	1.993	22.63	1.655	4.673
Slovenia	49.74	29.86	149.9	9.161	12.34	3.082	14.97	1.136	4.694
Spain	27.73	13.54	165.5	11.68	9.925	1.567	18.36	1.788	6
Sweden	69.05	24.47	119.8	8.061	9.925	2.497	12.11	0.900	5.465
United Kingdom									

Table 2: Descriptive statistics of covariates of the latent regression

One major concern with our approach lies in the nature of our data, which are descriptive in nature and do not account for the level of implementation or monitoring of a given efficiency policy. Our assessment should therefore be thought of as an analysis of the potential effects of these policies in each country, or simply of a country's "policy effort". The best way to address such an issue would be to use a scoring system for our dependent variable weighted by the level of policy implementation or stringency in each country. Since we cannot proceed in this way, we instead include a proxy measuring the level of respect for Law and Order (ICRG, 2011) as an explanatory variable of between-country variance. This index is compiled by the International Country Risk Guide and measures on a scale between 0 and 6 the degree with which a given country successfully implements the law and respects it. We use the average value of the Law and Order variable over the sample period (also reported in Table 2) to account for the fact that the policy variables we observe do not convey information about stringency or implementation.

The results on covariates coefficients emerging from the explanatory model are presented in Table 3. Estimated coefficients are significant, with the exception of the Law and Order variable, and are in line with expectations. The availability of energy efficient technologies in a given country, measured by the number of energy efficient patent applications at the EPO, has a positive impact on the second level random effect. The same can be said for the percentage of fuel imports: countries which are more dependent from abroad for their fuel supply tend to commit to energy efficiency more, on average and *ceteris paribus*. Energy intensive countries, on the other hand, are less likely to commit to energy efficiency policy. With respect to the covariates at the country level, those countries in which manufacturing has a high share in value added are more likely to pursue energy efficiency.

<b>Variable</b>	<b>Coeff.</b>	<b>SE</b>	<b>t</b>
<i>Covariates of Second Level Random Effect (Time)</i>			
Energy Efficient Patent Applications to EPO, (t-1)	0.003	0.001	3.00
Energy Intensity, (t-1)	-0.004	0.001	4.00
Fuel Imports, % of Merchandise Imports (t-1)	0.046	0.012	3.83
<i>Covariates of Third level random effects (Country)</i>			
Manufacturing in Value Added, % of GDP (mean)	0.085	0.033	2.57
Law and Order (mean)	-0.219	0.270	0.81

Table 3: Results on Second and Third Level Covariates

Conditioning the country-level random effect on these observed covariates, the ranking of countries is shown to change in Figure 5. Specifically, controlling for the observable covariates improves the ranking position of those countries which previously scored low because of some inhibiting "initial conditions", namely (1) the manufacturing sector accounts for a smaller share of their economy, (2) they are more energy intensive than others (3) their availability of new efficient technologies is lower and (4) their dependence from foreign energy imports is lower. Norway, for example, improves its ranking: it has very low energy imports compared to other countries, and manufacturing as a share of value added is low. Eastern European countries also improve their ranking, due mostly to the fact that they have a lower number of patents in new efficient technologies. On the contrary, Ireland, Italy and Spain perform considerably worse when their "initial conditions" are taken into account.

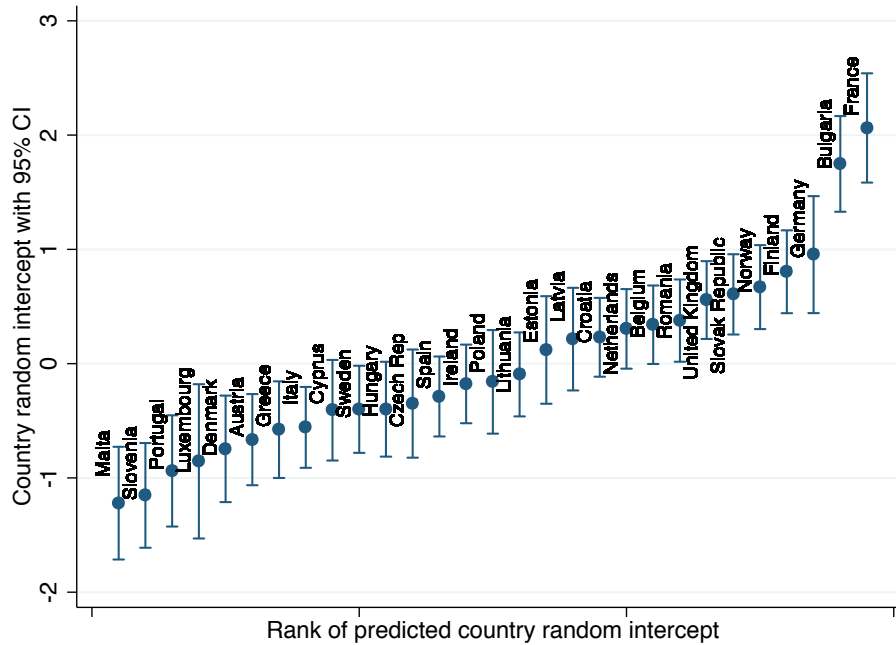


Figure 5: Rank of predicted country random intercept for multilevel PCM explanatory model

## 5. Conclusion

This paper presented a novel approach to assessing and comparing countries' energy efficiency or environmental policy portfolios and performance. The method, inspired by the literature on multilevel latent models and Item Response Theory (IRT), results in a ranking of European countries which accounts for the inherent difficulty of a given policy instrument mix. In addition, the model is extended to deal with the longitudinal nature of our data and to adjust the ranking as a result of country-specific economic and institutional observables which are likely to affect energy efficiency regulation.

We believe this approach is a promising one in assessing countries' commitment to environmental and energy policy since it is able to overcome a number of shortcomings of the previous literature on policy indicators. First, while the basic approach we propose has been applied in the statistical literature of scoring, its application to the context of environmental policy is novel. Second, we extend the basic IRT model to a multilevel framework which is consistent with the nature of the policy data available. Third, this approach allows attributing different weights or difficulty levels to the different policy instruments included in the policy portfolio. Thus, our assessment is conditional on the specific complexity of each country's policy portfolio. Fourth, we combine the multilevel IRT model with a latent regression, thereby allowing each country's scores to be conditioned on the country's observed characteristics.

We apply this methodology to data on efficiency policy targeting industrial sectors in 29 EU countries in the years 2004-2011. Our results show that accounting for economic and institutional characteristics changes the ranking of countries with respect to energy efficient policy. Specifi-

cally, the position of those countries with worse “initial conditions” but which choose to regulate energy efficiency nonetheless, demonstrating a higher than average commitment, improves.

Given the nature of our data, namely count of policies which are implemented in each country in any given year, our ranking informs only on the general commitment of a given country, and does not shed any light on the actual level of stringency of the given policy. While we try to control for this shortcoming in our empirical setting by adding covariates capturing the level of respect for law and order, our results should be interpreted accordingly.

We believe our methodology is a step in the right direction of creating an index of policy commitment, and possibly stringency. Moreover, its application to other fields of study where similar data on policy implementation is available is also very promising. This might include, among other, labour or monetary policy. The model presented in this paper could also be fruitfully extended to account for the presence of random slopes (as opposed to only random intercepts) and to better study the effect of time. To this end, focusing on a wider sample of countries would be beneficial, since variation in policy responses of the EU member states is necessarily limited given the common framework under which these policies are developed. Our current research is moving in this direction.

## References

- Andrich, D., 1978. Relationships between the thurstone and rasch approaches to item scaling. *Applied Psychological Measurement* 2 (3), 451–462.
- Bacci, S., 2012. Longitudinal data: different approaches in the context of item-response theory models. *Journal of Applied Statistics* 39 (9), 2047–2065.
- Bacci, S., Bartolucci, F., 2012. A multidimensional latent class rasch model for the assessment of the health-related quality of life. arXiv preprint arXiv:1211.2635.
- Bacci, S., Caviezel, V., 2011. Multilevel IRT models for the university teaching evaluation. *Journal of Applied Statistics* 38 (12), 2775–2791.
- Berman, E., Bui, L. T. M., 2001. Environmental Regulation and Productivity: Evidence from Oil Refineries. *Review of Economics and Statistics* 83 (3), 498–510.
- Brunel, C., Levinson, A., 2013. Measuring environmental regulatory stringency. *Oecd working paper 2013/05*, OECD. URL DOI : 10 . 1787/18166881
- Brunnermeier, S. B., Cohen, M. A., 2003. Determinants of Environmental Innovation in US Manufacturing Industries. *Journal of Environmental Economics and Management* 45, 278293.
- Cole, M., Elliot, R., 2003. Do environmental regulations influence trade patterns? Testing old and new trade theories. *The World Economy* 26 (8), 1163–1186.
- De Boeck, P., Wilson, M., 2004. *Explanatory Item Response Models. A Generalize Linear and Nonlinear Approach*. Springer, New York.
- Ferrari, P. A., Annoni, P., Salini, S., 2005. A comparison between alternative models for environmental ordinal data: Nonlinear pca vs rasch analysis. In: *Statistical Solutions to Modern Problems: Proceedings of the 20th international Workshop on Statistical Modelling*, eds. AR Francis, KM Matawie, A. Oshlack and GK Smyth, Sydney: University of Western Sydney. pp. 173–177.
- Ferrari, P. A., Salini, S., 2011. Complementary use of rasch models and nonlinear principal components analysis in the assessment of the opinion of europeans about utilities. *Journal of classification* 28 (1), 53–69.
- Fischer, G. H., Frewer, L., Nauta, M., 2006. Toward Improving Food Safety in the Domestic Environment: a multi-item Rasc scale for the Measurement of the Safety Efficacy of Food-Handling Domestic Practices. *Risk Analysis* 26 (5), 1323–1338.
- Fisher, C., Newell, R., 2008. Environmental and technology policies for climate mitigation. *Journal of Environmental Economics and Management* 55 (), 142–162.
- Gollop, F., Roberts, J., 1983. Environmental regulations and productivity growth: The case of fossil-fueled electric power generation. *Journal of Political Economy* 91 (4), 654–674.
- Hamamoto, M., 2006. Environmental Regulation and the Productivity of Japanese Manufacturing Industries. *Resource and Energy Economics* 28, 299–312.
- ICRG, 2011. International country risk guide database.
- Jaffe, A., Palmer, K., 1997. Environmental Regulation and Innovation: A Panel Data Study. *Review of Economics and Statistics* 79 (4), 610–619.
- Jeppsen, T., Folmer, H., 2001. The Confusing Relationship between Environmental Policy and Location Behavior of Firms: A Methodological Review of Selected Case Studies. *Annals of Regional Science* 35, 523–546.
- Kellenberg, D., 2009. An empirical investigation of the pollution haven effect with strategic environment and trade policy. *Journal of International Economics* 78 (2), 242–255.
- Lanjouw, J., Mody, A., 1996. Innovation and the International Diffusion of Environmentally Responsive Technology. *Research Policy* 25 (4), 549–571.
- Mair, P., Hatzinger, R., 2007. Cml based estimation of extended rasch models with the erm package in r. *Psychology Science* 49 (1), 26.
- Masters, G. N., 1982. A rasch model for partial credit scoring. *Psychometrika* 47 (2), 149–174.
- Murray, A. G., Mills, B. F., 2012. An application of dichotomous and polytomous rasch models for scoring energy insecurity. *Energy Policy* 51 (C), 946–956.
- OECD, 2013. Patents by main technology and by international patent classification (ipc). URL /content/data/data-00508-en
- Oreja-Rodriguez, J., Yanes-Estevez, V., 2007. Perceived Environmental Uncertainty in Tourism: a New Approach Using the Rasch Model. *Tourism Management* 28, 1450–1463.
- Popp, D., 2003. Pollution Control Innovations and the Clean Air Act of 1990. *Journal of Policy Analysis and Management* 22 (4), 641660.
- Popp, D., 2010. Exploring the Link between Innovation and Diffusion: Adoption of NOx Control Technologies and U.S. Coal fired Power Plants. *Environmental and Resource Economics* 45, 319–352.
- Rabe-Hesketh, S., Skrondal, A., Pickles, A., 2004. *GLLAMM manual*. Working Paper Series, Paper No. 160., University

- of California, Berkeley, Division of Biostatistics.  
URL <http://www.bepress.com/ucbbiostat/paper160/>
- Rasch, G., 1980. Probabilistic Models for Some Intelligence and Attainment Tests. University of Chicago Press, Chicago, IL.
- Rubashkina, Y., Galeotti, M., Verdolini, E., 2014. Environmental regulation and competitiveness: Empirical evidence on the porter hypothesis from european manufacturing sectors. Tech. Rep. IEFE Bocconi Working Paper n.69.
- Salzberg, T., Sinkovics, R., 2006. Conceptualization and Measuring the organizational environment: A multidimensional Approach. *Journal of Management* 17 (4), 681–700.
- Shimshack, J. P., Ward, M. B., 2005. Regulator Reputation, Enforcement and Environmental Compliance. *Journal of Environmental Economics and Management* 50, 519–540.
- Soutar, G. N., Cronish-Ward, S., 1997. Ownership Patterns for Durable Good and Financial Assets: a Rasch Analysis. *Applied Economics* 29 (7), 903–911.
- Tobey, J., 1990. The effects of domestic environmental policies on patterns of world trade: An empirical test. *KYKLOS* 43 (2), 191–209.
- Vona, F., Nicolli, F., 2012. The Evolution of Renewable Energy Policy in OECD Countries: Aggregate Indicators and Determinants. FEEM Nota di Lavoro 2012/051, FEEM.
- WorldBank, 2013. The world bank, world development indicators.