Comparative analysis of regulatory governance regimes in the OECD

Eric Brousseau
Carlos Gonzalez Regalado
Governance and Regulation Chair
University Paris Dauphine PSL

Very first and incomplete draft

June 25, 2021

1 Introduction

From the 1990s, most countries have experienced liberalization of their economy, which translated into the introduction of more competition in utilities and the privatization of key operators in the related industries. To back up this movement, governments were incited to establish independent regulators in charge of promoting fair competition between new entrants and incumbents and guaranteeing public interest. The model of 'independent' regulators was for a long time US-specific (Balleisen, 2015). In most other countries in the world, utilities had been operated by organizations, public or private, under direct monitoring of ministries, either though command and control mode of governance or through contractual relationships. From the 1990s, the principle of "good" governance has been promoted by various inter-governmental organizations, among which the World Bank, the OECD, or the EU, to promote a model of agency immune from the undue influence of the business (in response to Stigler's analysis of the capture of regulators by businesses) and from the undue influence of governments (especially because their sensitivity to electoral cycles may prevent them from making decisions aligned with the long term collective interest). The independence of regulators has also been understood as a significant driver of the transnationalization of the related service industries, allowing to attract foreign investments in infrastructures, and supporting the consolidation of specialized operators benefiting from economies of scale and economies of experience. That said, most countries "imported" the independent regulator model their own way.

Indeed, as pointed out by Balleisen (2015) the US model of an independent agency is both the result of a specific path of historical evolution and a response to the US institutional specificity. In particular, in the US, no federal ministries of telecommunication, energy or transportation pre-existed to the development of sectoral agencies. In Europe, to this opposite, such ministries with their bureaucratic lines were in place. Beyond these differences in terms of public administration tradition, the countries that adopted the independent regulator model are also characterized by contrasting size, federal vs. unitary logic, governmental involvement in the industry, legal tradition, etc. Thus, in practice, independent regulators have been settled on different bases. For instance, in some countries, the regulators are focused on a given industry, while in others, the regulatory agency is in charge of all network industries. Also, in some countries, agencies are specialized by issues (e.g., safety, market performance, externalities), while in others, each sectoral regulator manages the interlinking among those issues. Of course, the actual mandate, the legal status, and the means available to regulators are also specific to each jurisdiction, being dependent on its institutional specificities and also on the political compromise behind the implementation of this model. As with any institutional arrangement, the equilibrium behind the establishment and the operation of a regulator might evolve. Because regulatory agencies are in practice different, and because their characteristics might evolve, it is essential to develop methodologies to compare the organizational specificities of alternative institutional arrangements both across countries and industries, and along time. Such methodologies are necessary to identify the drivers of alternative institutional design, understand the outcomes of these alternatives, and compare their performances. More generally, a better understanding of the determinant and outcomes of regulatory agencies characteristics must be understood in the context of heterogeneous social preferences, national institutional patterns, and intrinsic sector characteristics (i.e., the specificity of market structures and technological choices, and the stakeholders' political salience).

Several comparative governance studies have already quantitatively assessed regulatory regime heterogeneity and its association with institutional determinants and outcomes (Trillas and Montoya, 2013a). However, many of the methodologies relied upon to measure 'governance regimes' tend to be idiosyncratic, and their results are hardly comparable across studies (and therefore across time, jurisdictions and industries). Two standard limits are at play. First, in many studies, the relevant dimensions to characterize the institutional arrangements are assumed ex-ante and measured independently from each other, without checking whether or not these dimensions are actually the most relevant one to contrasts governance patterns, and independent of each other. Second, these dimensions are usually assessed through a set of measured proxies, which are then aggregated into an index. The methodology relied upon to aggregate the "measures" generally assume that each of them contributes equally and linearly to the predefined indexes, while the relationship among each 'proxy' and the dimension measured is not discussed and analyzed. Such methodological

biases might compromise the results' validity and conclusions.

One of the central issues when comparing alternative institutional or organizational characteristics is that they might differ along with multiple characteristics and that reducing these characteristics to the most contrasting ones is a challenge. Automated textual analysis is precisely aimed at identifying the contrasts and similarities among texts/documents by identifying the way words/expressions are articulated among each other in a given text covering a set of topics and how different texts are similar or different along with the different topics (which "weigh" and "value" vary across documents). The contributions of words to topics and topics to document are computed by measuring frequency and co-occurrence of words in documents and do no result from any a priori hypothesis on how they should be articulated among each other. Moreover, the identification of topics in the whole corpus is based on the will to identify the more contrasting ones and to consider expressions/descriptors that are correlated as contributing to a common topic, hence reducing the number of topics to the most significant vectors of differentiation/characterization of the considered set of documents. The analysis allows therefore characterizing every single document/description by the way it contributes to the various topics identified in the whole corpus of all documents/descriptions. The documents/descriptions can be then compared among themselves thanks to a common metric built without any a priori about the most relevant dimension and without biases in aggregating the primary information (i.e., the descriptors in the document) into indexes (i.e., the topics). For instance, if several descriptors proxy very similar characteristics, they will not increase this characteristic's contribution to the measure of the overall variance within the corpus and the measure of the contrasts among documents. In other words, measurement biases will not impact the results, and a priori about the "weight" of single descriptors will not matter.

We apply these principles to textual descriptions of regulatory agencies' status in different countries and sectors, together with the characterization of their relationships with various stakeholders (such as the executive and the legislative), and of their duties and means. We rely on surveys managed by the OECD among its member states, which has the advantage of providing us with observation based on common descriptors. This allows us to characterize what we qualify as "regulatory governance regimes," describing the de facto status and the operations performed by "sectoral" regulators in a set of countries. Four dimensions explain most of the variance and can allow us to compare these regulators among them, but also across time. We can relate these descriptions of "regulatory regimes" and regulatory regimes evolutions with various descriptors of industry performances (such as the volume of investments, the level of activity, retail and wholesale prices, safety index, etc.) to explore potential causal links. The "regulatory governance regimes" can also be linked to other institutional, sectoral, or economic structural patterns to identify regularities and how the industry performances mentioned above might be explained by a combination of these structural patterns, including the "regulatory governance regime."

This exploratory approach aims at comparing regulatory governance across industries and countries to identify potential contrast and similarities. The objective is also to identify potential stylized facts that deserve greater scrutiny to investigate causal relationships between industry performance and status or means of the regulatory agency and institutional and economic determinants or regulatory agencies' features.

2 Methodology

2.1 Previous literature and contribution

Several studies have attempted to quantitatively compare regulatory governance regimes — i.e., the institutional and organizational characteristics of sectoral regulators — and to link them to institutional/political determinants and industrial/market outcomes. The methodological heterogeneity of these studies and their potential flaws impact the validity of their conclusions and the comparability of their results.

First, scores or indexes (mostly measuring regulatory independence) diverge in the weights assigned to each governance descriptor (usually captured by survey data). Many studies rely on ad-hoc weighting assumptions, which makes the results very sensitive to the observation lenses. Some studies even consider only one institutional characteristic; as regulatory independence, by taking into account the existence of an independent regulatory agency only, e.g., (Bortolotti et al., 2011). Many studies relying on multi-dimensional descriptors tend to assign equal weight to all elements and sum the presence or absence of this or that characteristic (Gilardi, 2002, 2005; OECD, 2016). This tends to ignore potential redundancies among the observed characteristics and make the results very sensitive to the observation tool and the number of descriptors chosen ex-ante. Characterization of regulators is then challenging to compare across studies.

Another group of methods relaxes the weighting assumptions and models the information variation as a function of a latent governance trait, i.e., unobserved significant governance trait characterizing a regulator and its behavior (Hanretty and Koop, 2012; Koop and Hanretty, 2018). For instance, the existence of a "cooling period" preventing a commissioner or top executives from a regulatory agency to work for the industry immediately after he or she leaves his/her office is contributing to the strengthening of the independence of the regulatory agency vis-à-vis regulated operators. The methodology might also lead to determining whether this "descriptor" is aligned with all the other sources of potential undue influence (e.g., executive power, members of the parliament, judiciary, activist groups, etc.), or whether independence from the industry is a very specific trait as compared to independence from political or administrative

influences. Such a methodology reveals the relevant latent traits and their relationships and might also point out the significant characteristics of a regulatory regime.

Partly because of the biases in characterizing regulatory governance regimes, but also because it is often assumed that regulation is different in practice, from one industry to another, most existing studies rely on a sectoral approach and compare regulatory governance regimes across countries, and do not try to compare regulators across industries. This limits the capability to disentangle the governance regulatory regime influence from more generic country-level institutional characteristics when studying the impact of public governance on outcome variables such as firm leverage (Cambini and Rondi, 2012), investment to capital ratio (Sutherland et al., 2011), or sectoral added value (Égert, 2009). Sectoral approaches do not allow identifying potentially similar or divergent regulatory governance patterns by industries and their potential impact on supply-side performance. For instance, Trillas and Montoya (2013b) rely on an instrumental variables approach to highlight a strong association between regulatory independence and telecommunication penetration rates in Latin America. Similarly, Edwards and Waverman (2006) suggest that regulatory independence is correlated to lower interconnection rates across the EU telecommunication industry. However, it is not clear whether regulatory independence is related to another institutional pattern in both cases.

Our study proposes using text analysis algorithms to circumvent the measurement/characterization shortcomings discussed above, following Blei et al. (2003). We exploit the "Indicators on the Governance of Sectoral Regulators (PMR-RM)" survey performed by the OECD to identify co-occurrence patterns in the pooled data, without making any assumptions about the weight of each descriptor and the number of relevant descriptive dimensions of governance. Our results identify four dimensions: the independence from the executive; the scope of the discretion of the agency (i.e., the degree to which its powers are formally framed); the scope of the instruments relied upon to monitor market coordination; and the scope of instruments aimed at guaranteeing transparency and compliance. To a certain extent, the first two dimensions hence describe the relationship between the regulatory agency and the governmental and societal actors, while the last two characterize the levers in the hands of regulators to weigh on market players' behaviors.

Our 'measurement' methodology allows us to highlight structural patterns and their evolution for each sectoral regulator, allowing comparisons across countries or industries. They allow exploring potential causal relationships between governance and performance in the studied industries, namely energy, e-communication, air, and rail transportation. The identified relevant governance traits seem to determine industrial and market outputs, although their impact varies across industries. Of course, a detailed and fine-grained analysis would be needed to demonstrate any causal inferences. The suggested relation-

ships are, however, enlightening and suggest appropriate research directions.

2.2 Data and preprocessing

We use the "Indicators on the Governance of Sector Regulators" (PMR-RM) database (Casullo et al., 2018) to describe governance patterns. It draws from a survey to document regulatory agencies' (RA) institutional characteristics, management practices, and formal relationships with governmental and market stakeholders. Since 2013 it has been managed every five years (and therefore in 2013 and 2018) to gather information about economic regulators in transportation (air, rail), utilities (energy, water), e-communication, infrastructures (roads, and ports) in 45 OECD and non-OECD countries.

Our study focuses on four network industries: energy, e-communication, rail, and airports for 23 European OECD countries, for which we benefit from consistent data from the survey in 2013 and 2018 (since additional industries were considered in the second wave) and on industry performances. The covered countries include Austria, Belgium, Czech Republic, Denmark, Estonia, Finland, France, Germany, Great Britain, Greece, Hungary, Ireland, Italy, Luxembourg, Netherlands, Norway, Poland, Portugal, Slovakia, Slovenia, Spain, Switzerland, and Sweden. This subset covers 184 regulator-level observations.

We evaluate survey comparability between periods, sectors, and countries. The PMR-RM survey experienced considerable variation between 2013 and 2018. The number of questions increased from 52 to 76, and the range of possible answers to some of them evolved. Besides, data is not available for all countries and sectors.

The first issue requires identifying the between-period non-overlapping questions and evaluating whether their exclusion affects our analysis. Most non-overlapping questions either describe an informal provision (not part of our analysis) or describe a regulator's second rank type of action, e.g., beyond publishing its decision, the agency makes it available online. The remaining questions describe budgetary agency practices and are excluded from the analysis.[1] Second, since some questions have a richer subset of potential answering options in 2018 compared to 2013, we adopt a conservative posture and keep the 2013 menu of possible replies. For instance, if in 2013 the choice for performing a given activity was between either the agency or the government, while in 2018 an additional option "cooperation between the two" was added, we keep the reply provided in 2013 to avoid identifying an evolution that did not occur. As a result, our dataset covers 38 questions/descriptors covering the *de jure* aspects

[1]The way budgetary practices are dealt with is different between 2013 and 2018. We, however, checked ex-post that it did not affect too much our characterization of regulatory regimes since budgetary information is aligned with the "independence" dimension.

of regulatory governance.

Before applying text analysis methods, we convert the survey database into textual data, since the OECD is made of scores. Every question has its unique descriptor. For dichotomous inquiries (yes / no), positive (negative) responses are assigned to a positive (negative) version of a descriptor, while categorical questions use their unique descriptor plus an additional term that characterize the answer. We converted survey data into a collection of 184 documents (92 for each period). The number of documents' terms is homogeneous, as shown graphically in fig. 5. Documents with term counts beyond two standard deviations from the descriptor mean (52.8) are removed from the corpus.

The next step is to identify non-useful (non-discriminant) descriptors because they are either excessively frequent or too sporadic. We remove four descriptors with a frequency below 10 in the corpus, and the term *regulator*, which was extremely common among documents. As a result, we use 74 unique descriptors set.

2.3 Methodology

2.3.1 Topic modelling

Topic modeling (probabilistic models) describes textual documents in a corpus (collection of documents) as mixtures over latent topics in the corpus. Topics are defined and labeled according to word (term) co-occurrence. Topics summarize highly dimensional feature spaces (words), facilitating document classification. The scholars' preferences for this unsupervised method have been significantly increasing in economics, political science, and management for the past years. In economics, Bandiera et al. (2017) uses topic modeling to detect latent traits between managers based on their time usage.

We consider that topic modeling has advantages relative to other survey aggregation methods. First, the Latent Dirichlet Allocation (LDA) helps to model a regime rather than standalone dimensions (e.g., independence, accountability). Second, dimensions and descriptors' weights are identified by document (regime) heterogeneity. Third, dimension scores are fully comparable across dimensions, sectors, countries, and periods. [2]

[2]Different techniques are also available to treat survey information. The "aggregation" methods assign or estimate scores for individual questions and aggregate them according to a defined hierarchical structure. The aggregation methods range from the equally weighted component approach as in Casullo et al. (2018) to factor analysis techniques as in Hanretty and Koop (2012); Jordana et al. (2018).

2.3.2 Latent Dirichilet Allocation

We use LDA to model regulatory regimes as mixtures of latent dimensions. The probability of observing particular arrangements (descriptors) depends on each regime's dominant dimension(s).

First, we introduce the notation used by Blei et al. (2003):

α: Dirichilet prior on the distribution of topics over documents.

β: Dirichilet prior on the distribution of descriptors over topics.

θ: topic distribution vector

z_n: *n-th* topic in a document

w_n is the specific descriptor in a document

N: number of descriptors in a given document

Next, a governance provision or feature in the network industries may be described as follows:

1. Draw $\theta \sim \text{Dir}(\alpha)$
2. For each provision w_n:
   (a) Draw a topic (dimension) $z_n \sim \text{Multinomial}(\theta)$
   (b) Draw a word (descriptors) w_n from $p(w_n|z_n, \beta)$, a multinomial probability conditioned on topic z_n

We will use the terms *topics* or *dimensions* indistinctly, the same treatment holds for *words* and *descriptors*. The previous generative process explains the way the n−th descriptors appear in our dataset. At the corpus level, we draw once the α parameter, which determines the dimension weights θ and the β parameter that specifies the weight of words within a topic. Given θ, we draw a topic z_n for every word w_n in the document. Finally, given the topic z_n, we draw a word conditional on the given topic z_n and the β.

A Dirichlet n-dimensional random variable θ can take values that lie in simplex (k-1), which means that $\sum_{n=1}^{k} \theta_n = 1$. In our context, we interpret θ as the influence that a particular dimension (topic) imposes on the regulatory regime—different weights of θ capture the observed differences in regulatory management across sectors and countries.

The LDA has its disadvantages. The model assumes that the number of topics n is fixed and known. To estimate the "best" number of topics, we follow the Probabilistic Coherence score developed by Jones (2019). This score calculates a measure of pairwise top-term topic correlation, correcting for meaningless word correlations. The fig. 7 plots the coherence score between 1 and 20 topics. The visualization shows that the highest score is achieved by selecting four topics.

2.3.3 Model inference

We are interested in the estimation posterior distribution of our latent topics given a document θ as shown in eq. (1). However, this distribution is intractable for exact inference (Blei et al., 2003). Following Griffiths et al. (2004), we applied Gibbs sampling to approximate the latent posterior distribution. The algorithm assigns a topic randomly to every word in a document. Next, it draws a topic for one word, holding the previous topic/term distribution fixed. The process is repeated for every word in the corpus until convergence is reached.

$$p(\theta, z \mid w, \alpha, \beta) = \frac{p(\theta, z, w \mid \alpha, \beta)}{p(w \mid \alpha, \beta)} \tag{1}$$

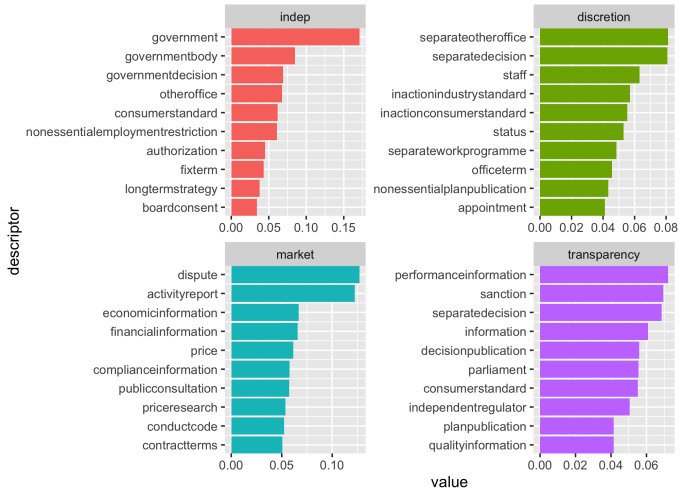
The estimation requires to specify the hyperparameters α and β. Their choice depends on the number of topics and vocabulary size. For α we set a value of 10, which is close to Griffiths et al. (2004) value ($50/T = 4$). Higher values of α smooth the topic distribution over documents. In our case, we expect that latent dimensions are balanced in a governance regime, i.e., governance regimes are not defined by only one (a few) dimension(s). For β, we opt for a value of 0.05, which lies in the midpoint of the literature's values. Lower values of β assign a specific word to only one specific topic, i.e., certain words appear only in one topic. The estimation uncovers four latent dimensions, as shown in fig. 1. More detailed results are provided in the appendix. They area interpreted and commented in the next section.

3 Results

3.1 Four Dimensions Characterizing Regulatory Governance Regimes

The fig. 1 highlights the most frequent words/expressions contributing to each dimension. Such a list helps the analyst interpreting the main institutional

Figure 1: Descriptor distribution by dimension



concepts behind each cluster (e.g., "market monitoring"). We complemented our analysis/interpretation of each dimension by considering semantic networks and correlation with an MCA analysis based on the scores computed by the OECD.

The four latent dimensions characterizing network industries' governance regimes might be defined as follows:

- Independence from the government (*independence*) estimates the extent to which the regulatory agency is protected against the executive's undue influence. The dimension's descriptors portray legal provisions limiting the executive power to dismiss the agency head/board and review/overturn agency decisions.
- Discretion (*Discretion*) is inverse to the number of legal provisions specifying the agency's obligations and proportional to its freedom in managing its resources. It is an inverse measure of the degree of the formalization of the agency decision making process and established delegation of authority. In that sense it could also be considered as a measure of "informality".
- Scope of the market monitoring (*market*), which measures the variety of levers that a regulator can handle to monitor operators' activity on markets (license, tariffs, conflict settlement, obligations imposed to market operators, supervision of their economic and financial performances). This dimension reflects the RA's ability to monitor the competitive process, oversight players' behaviors, and manage economic incentives.

- Transparency (*transparency*) aggregates measures of the obligations or (public) reporting imposed to the RA and obligations of reporting imposed by the RA to market player. It reflects the role of regulators in reducing information asymmetries, in particular, to ensure compliance and limit behavioral drifts. Here, the regulator seems to be considered an 'intermediary' between operators in the industry, public authorities and the users, whose expertise contributes to disclose unbiased information.

While independent from each other, the two first dimensions characterize the degree of independence and autonomy of the regulatory agency, while the last two grasp the channels relied upon by the agency to fulfill its mandate and reveal, therefore, the latter.

The previous definitions are supported by the correlations observed between our scores and the OECD Indicators on the Governance of sector Regulators ($OECD - RSRI$), which use the same underlying dataset. [3] The correlations shown in table 12 keep the expected signs and support our dimensions definitions. We arranged the variables to link high performance in a dimension with a high numerical score. Regarding governance, the linear association between $OECD$ independence and our autonomy scores ranges between 0.55 to 0.58, depending on the period. The *discretion* measure is negatively associated with OECD Scope (-0.47,-0.58) and Accountability scores (-0.25,-0.47). The negative coefficients suggest that our "discretion" dimension captures a degree of informality or the delegation of power to the RA, which is inversely proportional to its obligation in terms of formal accountability. Therefore, we can expect that regulators benefiting from a high score in terms of discretion would derive their actual authority from their relationship with the other participants in the power system, and with the industry stakeholders since it is not granted with a broad set of formal levers of power. It might echoes in-depth institutional characteristics. For instance, in Great Britain there is a long tradition of "trustees" benefitting from significant de facto authority, while its scope is not established in details "de jure". To the opposite, in France, detailed legal provisions establish the jurisdiction of each decision maker in the public system. That said the level of discretion can also result from a policy choice, either to weaken the authority of the RA, or to the opposite to allow it to choose the most appropriate levers to rule. The *market* coefficients are strongly linked to the OECD scores in a matter of scope (and therefore of accountability because of the correlation between the latter; see above), which confirms that it captures the RA's ability to influence industry players. Finally, transparency shows significant positive correlations with the three OECD Indicators on the Governance of Sector Regulators. We already noticed that Accountability and

[3] The OECD Indicators on the Governance of sector Regulators ($OECD - RSRI$) measure (at the sector level) by equally aggregating survey information in three governance dimensions: The regulator degree of insulation from undue political and market influence (Independence), the accountability of the regulator vis-a-vis other stakeholders (accountability), the range of activities that the regulator performs (Scope of action) (OECD, 2016)

Scope are correlated among themselves. Besides, here we notice the positive correlation with Independence. Our "transparency" indicator seems to grasp the idea that an independent and accountable RA combine capability of gathering information from industry players and obligations of transparency toward other stakeholders involved in the regulatory game.

Overall, our methodology seems useful to identify two dimensions related to the "status" of the RA, instead of the executive's sole distancing: independence and discretion, the latter being both an asset and a weakness in terms of the level of formal authority. Also, compared to the OECD scope score, our methodology disentangles two channels through which regulators operate; to make it short: market design and transparency (to ensure compliance). Interestingly, even if not surprising, there is a correlation between the RA's mission/status and the tools relied upon. Thus the OECD measures on the latter are correlated with measures of the former. Our approach allows us to take into consideration these relationships, and therefore to identify more significant vectors of contrasts among regulatory governance regimes. This illustrates the aggregation issue pointed out in section 2.1.

Beyond the list of frequent terms per topic, the model in section 2.3.3 estimates the topic distribution in every document θ_k. The individual topic contribution ranges between 0 (no contribution) to one (full contribution), and their sum is equal to one ($\sum_{k=1}^{K} \theta_k = 1$). We interpret these distributions as the influence of the considered dimension on the characterization of the governance regime. [4] The table 9 presents a summary of the dimension scaled scores θ for 2013 and 2018.

Next, we test whether our dimensions are relevant to characterize alternative governance regimes. Distributional similarities between dimensions' scores would suggest that their number is not accurate and should be revised. In fig. 6, each dimension displays different distributional characteristics. The *independence* score distribution is skewed to the left (skewness = -0.46) and a lower standard deviation (1.05) compared to the *discretion* distribution (skewness = 0.45, sd = 1.21). Moreover, the *independence* influence seems to impact regimes more evenly. Almost 65% of independence scores lie within one standard deviation from the mean in contrast to the 40% of the *discretion* score. The *market* score distribution is heavily skewed to the left (skewness = -0.78) with a standard deviation of 0.74, while the *transparency* distribution seems more symmetrical (skewness = 0.05) and more dispersed (sd = 0.91). Despite the distribution asymmetries, the scope capabilities are evenly shared by a large number of regimes. Close to 80% of the scope dimensions lie within one standard deviation from their means. Overall, this highlights that independence and the market design tools are indeed insufficient to characterize governance regimes.

[4] Note that the descriptors in the *indep* panel (top left) in fig. 1 suggest the existence of a "government proximity" dimension. We defined the "Independence from the government" as government distance to the regulator, ($\theta_{indep} = -\theta_{govdep}$).

Both discretion and transparency are important dimensions of contrasts among RAs.

3.2 Cross-Industry and Cross-Country Comparisons

At the sector level, as shown in fig. 8, energy and e-communication show relatively low variation across dimensions, suggesting that regulatory regimes in those sectors are relatively standardized across countries. Moreover, each dimension's mean score is similar (while *discretion* is significantly lower in e-communication). This suggests that both network industries tend to be governed by similar institutional arrangements. By contrast, regulatory regimes for transportation industries seem to be characterized by much more heterogeneous arrangements across countries and industries. Overall, *discretion* tends to be high and divergent across regulators in these industries, and they are granted fewer regulatory tools than their counterparts in network industries. The contrasts between railways regulators and airport regulators are essentially based on their independence and the reliance on market design tools (both higher for rail regulators).

3.2.1 The Comparison Among Regulatory Governance Regimes

We use score time averages (2013-2018) *mean* and growth rates *growth* to compare regimes across sectors and countries. Time averages assess a regulatory governance regime's structural characteristics, while growth rates attempt to capture institutional evolution. These metrics will be used in later sections to explore the correlation between regulatory governance and industry performances.

Regarding structural characteristics, our scores expose industry differences in regime configurations. We used the ANOVA test to find statistically significant differences between sector *mean* scores. The results are presented graphically in fig. 2. [5])

The *governance* panel highlights significant contrast among the governance arrangements across industries—the discretion dimension matters in building these contrasts. When considering network industries, an inverse correlation

[5]We tested the normality and homogeneous variance conditions using the Kruskal-Wallis (column *pnormm* in table 15) and Levene tests (see table 13), respectively. Except for the discretion dimension for e-communication, we could not reject the null hypothesis that the sector distributions are normal. Regarding the homogeneous variance condition, the Levene test applied at the sector level data did not reject the null hypothesis that distributions hold the same variance. Besides, we tested the mean score significant differences using Tukey's test (see table 14.

seems to hold between independence and discretion. More independent regulators benefit from less discretion (e-communication vs. energy), which is consistent with what would predict a rational theory of institutional design. This inverse correlation seems to hold for airport regulation: characterized by non-independent RA with weak status. However, it does not work for rail regulators, characterized by both a high degree of independence and discretion. This configuration reflects the resistance to railways liberalization in Europe, which has been managed at a slower pace than reforms in telecommunication and energy. It seems that, to comply with the European Union successive "Railway packages" [6], members state created independent RA but failed to grant them with formal authority. Overall *independence* (red) characterize regulatory governance in e-communication and energy regimes, while *discretion* (green) is the mark of transportation infrastructure regulators.

The *scope* panel exhibits an association between *market* (cyan) and *transparency* (purple) scores. There is also a clear ranking from the e-communication to the airport sectors; telecommunication operators are granted a wide scope of regulatory means to design markets and ensure transparency. Interestingly, the gap between market and transparency scores is higher for energy and railways, suggesting that while regulators in these industries may establish tariffs and organize markets, they have fewer capabilities to organize transparency, which becomes a significant factor of discrimination sectoral regulatory regimes.

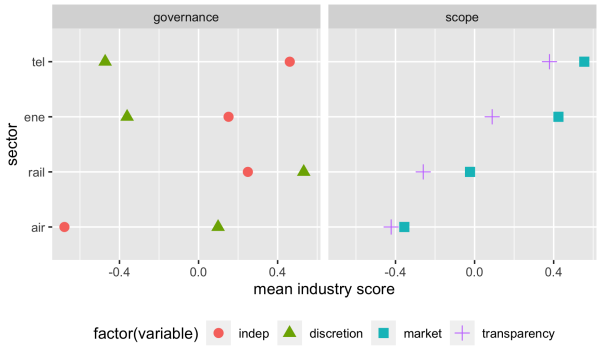


Figure 2: ANOVA predicted mean scores by sector

When considering transnational comparisons (country sectors' scores average), the fig. 9 presents national average scores sorted by the distance between

[6] Between 2001 and 2016, four legislative packages were adopted with the aim of gradually opening up rail transport service markets for competition, making national railway systems interoperable, and defining appropriate framework conditions for the development of a single European railway area. These include charging, and capacity allocation rules, common provisions on licensing of railway undertakings and train driver certification, safety requirements, the creation of the European Agency for railways and rail regulatory bodies in each Member State as well as rail passenger rights.

independence and *discretion* scores. It highlights a potential high influence of national specific institutional characteristics/political equilibria on RA's status. The figure highlights a strong heterogeneity not only in terms of score level but also in terms of hierarchy and gap size between *independence* and *discretion* scores, while practices of RA seem to be more "parallel" (the country-level scores are relatively close) and overall *market* (cyan) are higher than *transparency* (purple) scores. Interestingly, the scores in regulatory levers seem to be correlated to the gap between independence and discretion, rather than to the degree of independence or discretion. Putting it another way, powerful agencies (proxied by the scope of regulatory tools they operate), tend to be either independent from the government (with precise delegation of authority), or granted with a lot of discretion (but very closely linked with the executive). Also, it can be pointed out that the ranking among countries characterized by a predominance of independence over discretion and those by the inverse pattern does not reflect the usual (e.g., World Bank) ranking of countries in terms of public administration performance or market-friendly governance. For instance, we find the United Kingdom on one extremum of the spectrum and Switzerland and Denmark on the other end. Country size could be one factor playing a role in explaining the stringency of discretion/informality. We already mention the role of political/legal culture. Path dependency in institutional evolutions as well as contrasted preferences in terms of "social contract" are probably the key underlying factor: establishing an "independent and efficient RA", and the will to do so, are highly dependent on the pre-exisiting institutional structures, which are slow to evolve and difficult to reform, and of the socio-political equilibria. The lower variability of the "scope" scores compared to the governance ones suggests that (market and transparency) practices have diffused more evenly across European governance regimes than the institutional model of the "independent" regulator. [7].

3.2.2 Comparing Regulatory Governance Evolutions

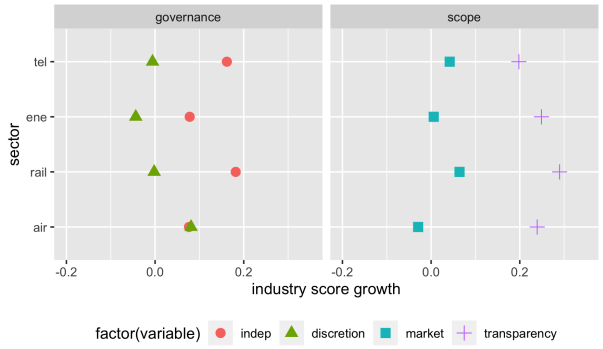
When considering evolution [8] between the two periods — cf. fig. 3 —, the average evolution at the sector level highlight a relative stability of two indicators (*discretion* and *market*) for all sectors, while progress are made in matter of *independence* (but for air-transportation RA) and transparency (also in all sectors). The relative stability of *discretion* is in line with the idea that this characteristic is intrinsically linked to the general institutional framework/political equilibrium in most countries and is difficult to transform quickly, whereas increasing the RA's formal independence from the executive is easier to imple-

[7] We use a Mann-Whitney U test to check whether the observed differences between groups significant

[8] The regime's evolution is measures in percent change relative to the initial situation by comparing scores between the two periods. We define percent change as $PCH = \frac{rawscore_t - rawscore_{t-1}}{rawscore_{t-1}}$

ment. The increasing transparency (and the stability of market design capability) might be linked to the fact that relative performances on the matter lag for most agencies; and the ability to develop the related practices incrementally and therefore at relatively low political and organizational costs. [9]. The fig. 4 highlights, however, that the modest evolution of the mean evolution of the *discretion* and of the *market* scores does hinder an evolution of the distribution of the scores on the matter, which is also contrasted across sectors, with an overall tendency to a homogenization around a common norm of lower discretion and higher transparency for e-communication, energy, and rail regulatory governance. This pattern of evolution does not hold for airport regulators that are only becoming more transparent (but which discretion does not evolve)

Figure 3: ANOVA evolution scores and residuals by sector

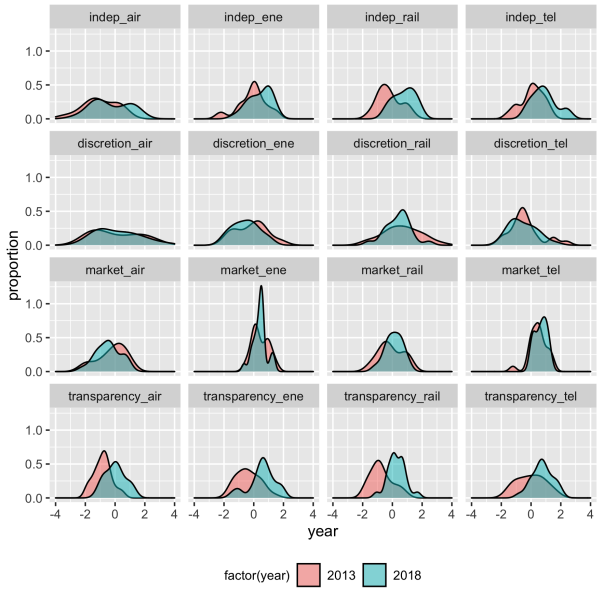


Cross-national comparisons confirm that with few exceptions, *independence* and *transparency* scores tend to increase for all countries as shown in fig. 10. In the case of *independence*, 19 out of 23 countries display reinforcement, and three countries register a non-significant negative evolution (the United Kingdom, Austria, and Denmark). Only Sweden exhibits a regression in matter of *independence*. *Discretion* shows a divergent pattern. Only 8 out of 23 countries exhibit a growth on the matter. The changes of Norway, Estonia,

[9]The columns *meandif* and *pval* in table 15 show the average period differences (2013-2018) and its statistical significance (t-test). The *independence* scores increased between 8% and 18%, while transparency rates range between 20% and 29%, depending on the sector. In contrast, the *market* and *discretion* influence changes are modest and statistically insignificant. The *market* score variations are positive for rail (6%), e-communication (4%) and energy (0.06%) and negative for air (3%). The negative changes in *discretion* averages suggest a reduction in the agency discretionary space to perform regulatory activities. These variations can be assessed graphically in fig. 4. The *evolution* scores do not differ systematically across sectors at the dimension level, as shown graphically in fig. 3. We formally check the significance of the differences using Tukey's test, which does not reject the null hypothesis that sector averages are the same (see columns *evoldif* and *evolpval* of table 14). Note that the rail sector experienced higher regime shift in *independence* (18%), *transparency* (29%), and *market*(6%), i.e., the rail regulatory agencies have benefited from legal provisions or change their practices favoring transparency, higher autonomy from the government, and superior ability to monitor market behaviors.

Germany, and Denmark are not statistically different from zero. In line with the recommendations pushed forward by the OECD and the European Union, structural reforms have targeted independence, and discretion seems more to be a structural characteristics of each country. *Transparency* was reinforced in 22 out of 23 countries in our sample[10] (with the only exception of Finland), while reinforcements in matter of market design levers are modest (and only 12 out of 23 countries show positive increments).[11]. These figures confirm that, over the period, most countries adopted a relatively parallel evolution of their regulatory governance regime. The focus was on guaranteeing more independence to RAs, and pushing them to promote transparency (soft-law based/sunshine regulation).

Figure 4: Time/sector kernel estimate distributions (Gaussian)



3.3 Regulatory Governance Regime and sector Performance

We now explore the potential relationships between our measured regulatory regimes and a set of industry performance measures. At this exploratory stage, we do not claim to demonstrate any causal relationships but rather to identify co-variations that might explain either the potential impact of a regulatory governance regime over industry performance or vice versa, the potential con-

[10] All of the changes are statistically different from zero.

[11] Negative changes are, however, small, and four of them are not statistically different from zero.

straints sectoral organizations impose on the design of regulatory governance. To put it another way, our goal is to exhibit stylized facts about the relationships between regulatory governance regimes and industry performance in terms of capacity, quality and price. Our aim is to explore the revealed impact of regulatory governance regime, which should be a combination of the mandate of the RA (decided by the legislator and the executive) and the policy of the RA (which should be linked to its mandate and to the institutional and political constraints he faces. Our scores highlight significant associations to the regulated industries' capacity, prices, quality of service (with specific variations on this notion accross industries: from ecologic performance to safety). We test this associations using a linear regression model :

$$y_{ic} = \alpha + \beta x_{ic,m} + \gamma controls_{ic} + \epsilon_{ic} \tag{2}$$

In eq. (2), y_{ic} measures the sector c performance (capacity, price, quality, coverage) for country i. The variable $x_{ic,m}$ stands for the dimension structure score for country i in industry c. β stands for the effect of dimension structure or change on sector performance. Besides, we include a list of general and sector-specific controls ($controls_{ic}$) to account for differences in income, country size, institutional quality, and geographical position. The full list of dependent variables and sector controls are shown in table 32 for energy, table 33 for e-communication, table 34 for rail, andtable 35 for air sector. γ is a vector that captures the effect of each control variable on sector performance and ϵ_{ic} measures unobserved sector-specific heterogeneity. We use 5-year sector averages to capture long-run sector characteristics and reduce data variability in one particular period.

European mandates seem to be an essential driver of governance evolution between periods, even if national and sectoral constraints could influence the local rule adoption. For instance, the Directive 2009/72/EC defines the governance principles for electricity regulators, including transparency in rule adoption and publication, public consultation, and accountability. Thus, a regime variation measure is less sensitive to national and industry long-run effects. In this regard, we inspect whether institutional evolution (measured by *growth* dimension scores) drives significant changes in outcomes variables, as shown in eq. (3).

$$\Delta y_{ic,t} = \theta + \kappa x_{ic,g} + \mu controls_{ic} + \nu_{ic} \tag{3}$$

The eq. (3) searches for robust correlations by controlling for unobserved time-invariant heterogeneity. The expression tests whether sector performance time variation $\Delta y_{ic,t} = y_{ic,t} - y_{ic,t-1}$ is linearly correlated to time changes in our dimension scores $x_{ic,g}$. This setup allows the inclusion of time invariant

controls, in contrast to other approaches such as panel data methods, and the identification of the effects (possibly different) of long-run regimes (β) and its changes (κ). The vector $controls_{ic}$ measures the income (long-run), institutional quality, and other sector-specific performance determinants (variation), while the vector μ captures their effects on performance. Finally, the term ν_{ic} captures unobserved time-variant heterogeneity. Summary tables report significant coefficients at 5% level. The regression tables in the following subsections present relevant governance effects on sector performance. Each row represents one regression equation.

We manage the discussion sector by sector, relying on a sector-specific index of performance. In each case, we consider, first, the impact of RA governance measures (i.e. *independence* and *discretion*), and second the impact of RA levers measures (i.e. *market* (cyan) and *transparency*).

3.3.1 Energy

As shown in table 1, *independence* from the government is associated with larger electricity generation capacity (model 1), including renewable generation and for exportation, except for generation capacities based on gas. Increasing independence, seems, however, favoring the securitization of domestic electricity supply, to the cost of imports (model 4) or of increasing CO2 emitting capacities (models 5,6)[12], even if it favors system efficiency (particularly wind conversion rates, as shown in table 17). Overall, this might suggest that more independent regulators favor investments by operators but are less sensitive than governments to decarbonization objectives since their main mandate is to guarantee both security of supply (and of investments) and low prices for the users. The level of *discretion* is negatively associated with more 'traditional' electricity production capacities (i.e., gas and hydro; table 16) and positively with renewable capabilities (model 3). However, as *discretion* expands, CO2 emitting production capabilities increase (model 8), and a renewable generation contraction occurs (model 17). Thus discretion does not seems to be favorable to decarbonization. Also, *discretion* shows significant associations with higher consumer prices and price increments (models 2,7). Discretion does not play in favor of consumers.

As shown in table 2, RAs benefiting from more levers to monitor *market* and industry operators' remuneration seem to confirm our insight that RAs' mandate is not oriented toward CO2 emissions, but rather consumers/users protection. Operators are then led to invest in domestic generation capabilities, even if they are emitting CO2 (models 1,2,3). A reinforcement of the regulator's arsenal aimed at providing economic market incentives to the operators is associated with increasing CO2 emitting capabilities (model 4) and renewable

[12] A 1% increment in *independence* score is correlated with almost 1% increment on the new CO2 capacity per capita.

Table 1: Energy regression OLS estimates for governance dimensions

	class	dimen	category	coef	pval	N
1	mean	indep	elecprod	922.632	0.025	22
2	mean	discretion	price_ind	21.962	0.044	22
3	mean	discretion	renewprod	409.169	0.028	22
4	growth	indep	imports	8.019	0.010	22
5	growth	indep	intcomb	0.269	0.006	19
6	growth	indep	solcap	-0.001	0.040	20
7	growth	discretion	price_hou	3.718	0.002	22
8	growth	discretion	totcomb	6.884	0.029	20

The column *class* shows whether the regression uses structural *mean* or variation *growth* dependent variable, as explained in section 3.3. The column *dimen* shows the type of institutional variable, *capacity* describes total system variables, *quality* whether energy is produced by CO2 or renewable technologies, or efficiently (relative measure) and *price* consumer retail prices per GWh. **System variables:** *elecprod* total production in GWh, *imports* electricity imports in GWh, $price_hou$ and $price_ind$ electricity consumer prices in USD per GWh. **Renewable:** *renewprod* production of electricity based in renewables and biofuels in GWh, *solcap* solar electricity capacity in MWe. **CO2:** *totomb* electricity capacity based in all combustion technolies in MWe, *intcomb* electricity capacity based in internal combustion machines in MWe. **Controls:** absolute country latitude, rule of law (WB), sector market regulation (OECD), GDPpc growth rate, tax revenue percentage (GDP), change in system capacity in MW. Complete regression tables are found in table 16 and table 17. Heteroskedasticity − robust standard errors.

capabilities (model 5) simultaneously, which seems to confirm that decarbonization does not rank high in the hierarchy of objectives of the regulators. Low energy price seems the main driver of their policy. A positive *transparency* variation is correlated to lower electricity prices for industrial consumers (model 6). In particular, an average *transparency* growth rate (23%) is associated with a 4% decrease in the average industrial consumer price.

Table 2: Energy regression OLS estimates for scope dimensions

	class	dimen	category	coef	pval	N
1	mean	market	CO2sh	1.020	0.029	20
2	mean	market	elecprod	1,054	0.011	22
3	mean	market	exportelec	-414.150	0.010	22
4	growth	market	othcomb	11.930	0.024	18
5	growth	market	solarpv	38.390	0.037	21
6	growth	transparency	price_ind	-3	0.027	22

The column *class* shows whether the regression uses structural *mean* or variation *growth* dependent variable, as explained in section 3.3. The column *dimen* shows the type of institutional variable, *capacity* describes total system variables, *quality* whether energy is produced by CO2 or renewable technologies or more efficiently (relative measure) and *price* stands for consumer retail prices per GWh. **System dependent variables:** *elecprod* total production in GWh, *exportelec* electricity exports in GWh, $price_ind$ electricity consumer prices in USD per GWh. **Renewable sources:** *solarpv* solar photovoltaic electricity capacity, *CO2sh* CO2 electricity production in GWh, *othcomb* electricity capacity based in internal combustion machines in MWe. **Controls:** absolute value of country latitude, rule of law (WB), market regulation score (OECD), GDP per capita growth rate, tax revenue percentage (GDP), change in system capacity in MW. Complete regression tables are found in table 18 and table 19. Heteroskedasticity − robust standard errors.

3.3.2 E-communication

The relationship between regulatory governance and performance in e-communication is challenging given the available performance indicators on this market characterized by the marketing of a diversity of services that are partly substitutable

and partly complementary (think for instance to voice and digital communication, with the development of voice on IP or videoconferencing on the Internet; or to fix and mobile telecommunication). During the period under consideration a central challenge has been the development of broadband access to the Internet. In the same time, to optimize their operations, telecommunication operators have been bundling their services, providing in particular bundled fix and mobile access to digital network, as well as join subscription to Internet, Cable TV and telephony. One of the issue is that statistical systems did not adapt at the pace of evolution of the technology and marketing of e-communication services and continue to differentiate these services. Overall, it seems that RAs concentrated their efforts in promoting the development of a broadband infrastructure, enhancing the overall quality of e-communication services. Pricing of services was not their main targets, at least according to the correlations we observe.

At first sight, discretion seems to have a higher impact than the independence from the government. As shown in table 3, countries with higher *independence* scores exhibits higher development of broadband based service (VoIP, model 2) to the cost of higher fix-broadband prices (model 1). The reinforcement of the RA independence seems however to have a positive impact on the high-speed infrastructure (models 6), on the (decreasing) fix-broadband prices (model 7), and on the use of digital services (model 8). Higher scores in the matter of *discretion* are associated with higher fix-broadband coverage (model 4) and lower mobile coverage (model 3), while prices are positively correlated with discretion, in particular for mobile services (model 5). As a matter of evolution, increasing *discretion* has a positive impact on broadband adoption, primarily via mobile (models 9), and triggers a decrease in prices (model 10) and a higher coverage.

As pointed out by table 4, scores in terms of reliance on *market* levers do not show a strong association with performance variables, e.g., service coverage or prices. Although, the score is negatively associated mobile termination rates (model 1). Besides, reinforced *market* practices do not seem correlated to any industry outcome change between 2013 and 2018. Since the e-communication sector has been liberalized for a while and since the European Union has been promoting convergence of regulatory policies on the matter, very similar regulatory policies seem to be at play when it is the question to organize markets and competition among operators. However, policies in mater of *transparency*, seem to have a significant impact on prices (models 4,5) and total number of subscriptions (model 2). Transparency scores are higher in countries with lower fixed penetration rates (model 3), but when transparency increases, it positively impacts subscriptions and penetration rates. Reinforced transparency highlights a positive correlation with all types of broadband technology subscriptions (model 6). These associations are accompanied by a price increase for fix-broadband internet service (model 7), where a 1% percent increase in *transparency* relates to a 5% increase in the average price change. Transparency might allow users

Table 3: E-communication regression OLS estimates for governance dimensions

	class	varia	dimen	category	coef	pval	N
1	mean	price	indep	prixfixbb	0.040	0.020	20
2	mean	capacity	indep	telVoIP	1,227	0.044	19
3	mean	capacity	discretion	accpath	-835.9	0.022	21
4	mean	quality	discretion	fixprate	0.770	0.006	21
5	mean	price	discretion	prixmob	0.730	0.039	20
6	growth	quality	indep	cabsubs	0.030	0.044	19
7	growth	price	indep	ppprice	-0.150	0.049	19
8	growth	capacity	indep	voipsubs	41.860	0.032	17
9	growth	quality	discretion	mobpenrate	0.330	0.032	20
10	growth	price	discretion	ppprice	-0.450	0.005	19

The column *class* shows whether the regression uses structural *mean* or variation *growth* dependent variable, as explained in section 3.3. The column *dimen* shows the type of institutional variable, *capacity* describes absolute system variables (e.g., total numbers of service subscribers), *quality* describes relative measures (e.g., number of subscribers by 100 habitants) and *price* stands for consumer retail prices in USD for telecom services. **Capacity dependent variables:** $telVoIP$ and *voipsubs* number of subscribers of VoIP services, *accpath* number of access paths for telecom services in thousands of access. **Quality:** $fixprate$ penetration rate for fix broadband internet services, *mobpenrate* penetration rate for mobile services. **Prices:** $prixfixbb$ and *pprice* price for 5gb fix internet bundle in PPP USD, $prixfix$ price for 5gb fix internet bundle in USD. **Controls:** Rule of law index (WB), market regulation score (OECD). GDP per capita growth rate average (2013-2018), population density, percentage of population living in urban areas, number of hotel nights spent by tourists in a given country. Complete regression tables are found in table 22 and table 23. Heteroskedasticity − robust standard errors.

to identify the variety of potential access to the Internet, even without a price effect (since the market is already significantly competitive).

Overall, increasing transparency being the more significant regulatory governance evolution in the e-communication industry (cf. fig. 4), it is not surprising that the evolution of broadband access and use is linked to that governance dimension; which has however not a negative impact on prices. More generally, regulatory governance in e-communication seems to be facing a trade-off between the development of high-speed Internet coverage and the seek for lower prices.

Table 4: E-communication regression OLS estimates for scope dimensions

	class	varia	dimen	category	coef	pval	N
1	mean	price	market	mobterm	-0.220	0.036	20
2	mean	capacity	transparency	accpathtot	1,302	0.016	21
3	mean	quality	transparency	fixtotal	-1.320	0	21
4	mean	price	transparency	prixfix	-1.720	0.018	20
5	mean	price	transparency	prixmob	-2.020	0.016	20
6	growth	quality	transparency	accpath	0.220	0.050	20
7	growth	price	transparency	pcprice	0.010	0.023	19

The column *class* shows whether the regression uses structural *mean* or variation *growth* dependent variable, as explained in section 3.3. The column *dimen* shows the type of institutional variable, *capacity* describes absolute system variables (e.g., total numbers of service subscribers), *quality* describes relative measures (e.g., number of subscribers by 100 habitants) and *price* stands for consumer retail prices in USD for telecom services. **Capacity dependent variables:** *accpathtot* number of access paths for telecom services in thousands, *accpath* number of access paths for telecom services per 100 habitants. **Quality:** *fixtotal* number of subscribers to fix broadband internet service per 100 habitants. **Prices:** *pcprice* price for 5gb fix internet bundle as percentage of per capita income, *prixmob* price for 5gb mobile internet bundle in PPP USD, $prixfix$ price for 5gb fix internet bundle in PPP USD, *prixmob* price for 5gb mobile internet bundle in PPP USD, *mobterm* price of termination rates in another network in USD. **Controls:** Rule of law index (WB), market regulation score (OECD). GDP per capita growth rate average (2013-2018), population density, percentage of population living in urban areas, number of hotel nights spent by tourists in a given country. Complete regression tables are found in table 21 and table 20. Heteroskedasticity − robust standard errors.

3.3.3 Railways

In the case of transportation industries,the multi-product nature of the activity (e.g. local vs long-distance transportation, passengers vs. freight, etc.) and the marketing methods (e.g. Yield management, subscription) makes it difficult to benefit from comparable aggregated statistics to grasp economic performance. What we found, in any case, is that regulatory regimes seem to impact quality, especially safety, rather than retail prices. We observe also impact on volumes of traffic. It is also important to keep in mind that in matter of transportation, competition is largelly inter-modal and not only intra-modal; and that users tend to consider mobility in terms of inter-modality, since no single mode of transportation is able to respond to demand for trips.

When considering railways' regulatory governance, the correlations with structural variables seem inconsistent with those of evolution. As shown in table 5, higher *independence* relates to high deployed safety capacity (model 1) [13]. An increment of autonomy is associated with a modest increment in derailments (model 5), which account for 4% of total incidents (ERA, 2017), while also with decreased personal incidents rate (model 3) and a positive variation in safety capacity deployment (model 4). These patterns suggest that more independent regulators tend to pressure operators essentially on safety. [14]

Discretion seems also supports safer rail operations. A high *discretion* score negatively correlates with total incidents in rail operations per km (model 2). An increment only impact (negatively) on freight traffic (model 6) and on safety capacity deployment (reported in table 24).

The scope dimensions show two contrasting patterns; on the one hand, *market* coordination levers seem to favor superior industry output at the expense of safety (see table 6). Reliance on *market* regulatory tools, hence competitive pressure, is associated with superior freight traffic (model 2), and higher train delays (model 1), and safety incidents (reported in table 26). Growth in the *market* score is only associated with superior freight traffic (model 3) and investments in infrastructure (model 4). Transparency does not show a significant correlation at the structural level, but its increments are linked to incident reduction (models 5,6). For instance, a 1% shift in *transparency* score accounts for a 3% decrease in the average value of accidents at track crossings. This result seems associated with rail safety assessment as part of the Common Safety Method (Commission Decision 2009/460/EC). Transparency and reporting might allow member states to identify main safety concerns and take further

[13] High independence is correlated with incidents that involve passengers (table 24). However, these events account only for 3% of total incidents that involve persons in 2015

[14] In their safety overview, ERA (2017) highlights that unauthorized person accidents and level-crossing incidents account for almost 85% of all incidents (91% of incidents with casualties) Most of the level-crossing incidents are caused by user misuse (25% of all incidents), and only 53% of level-crossings use automatic mechanisms.

Table 5: Rail Transport regression OLS estimates for governance dimensions

	class	varia	dimen	category	coef	pval	N
1	mean	quality	indep	activecross	0.010	0.013	17
2	mean	quality	discretion	allaccidkm	-0.050	0.045	17
3	growth	quality	indep	accunpeop	−0.001	0.013	19
4	growth	quality	indep	activecrosskm	0.001	0.006	18
5	growth	quality	indep	derail	0.030	0.034	19
6	growth	capacity	discretion	goodskm	-0.060	0.031	19

The column *class* shows whether the regression uses structural *mean* or variation *growth* dependent variable, as explained in section 3.3. The column *varia* shows the type of variable: *capacity* describes absolute system variables (e.g., length of rail tracks), *quality* describes safety measures (absolute or relative) (e.g., number of rail incidents by km of track). **Quality dependent variables:** *activecross* number of automatic devices to handle crossings in tracks, *allaccidkm* total number of incidents in the rail system per track km, *goodskm* volume of goods transported by rail system in metric tons per km, *accunpeop* number of incidents that ended in the injure of unathorized persons in rails, *derail* number of yearly derailments countrywide, *activecrosskm* number of automatic devices to handle crossings in tracks per track km.**Controls:** market regulation score (OECD), GDP in PPP USD, total system rail track length, GDP per capita growth rate, rule of law score (WB). Complete regression tables are found in table 24 and table 25. Heteroskedasticity − robust standard errors.

action. This information flow requires active stakeholder engagement and coordination as highlighted by ERA (2018). All in all, what seems significant here is that there is a tension between market pressure, which seems to lead to higher volume provided by the operators, and quality both in terms of punctuality and safety. It seems that transparency levers are relied upon to deal with this necessity to pressure operators on the quality of they deliver.

Table 6: Rail Transport regression OLS estimates for scope dimensions

	class	varia	dimen	category	coef	pval	N
1	mean	quality	market	delaymin	5,205.950	0.020	16
2	mean	capacity	market	freight	2,330.970	0.045	17
3	growth	capacity	market	goodskm	0.070	0.007	19
4	growth	capacity	market	tracklen	55.280	0.017	19
5	growth	quality	transparency	accicross	-0.100	0.046	19
6	growth	quality	transparency	accidempl	-0.080	0.007	19

The column *class* shows whether the regression uses structural *mean* or variation *growth* dependent variable, as explained in section 3.3. The column *varia* shows the type of variable: *capacity* describes absolute system variables (e.g., length of rail tracks), *quality* describes safety measures (absolute or relative) (e.g., number of rail incidents by km of track). **Capacity dependent variables:** *freight* volume of goods transported by rail system in metric tons, *tracklen* rail system track length in km. **Quality:** *delaymin* minutes of delay product of rail incidents, *goodskm* volume of goods transported by rail system in metric tons per km, *accicross* number of incidents in road crossings, *accidempl* number of incidents that end in injure of rail employees. **Controls:** market regulation score (OECD), GDP in PPP USD, total system rail track length, GDP per capita growth rate, rule of law score (WB). Complete regression tables are found in table 26 and table 27. Heteroskedasticity − robust standard errors.

As pointed out above, safety seems to be central in railways regulatory governance, and a regime on the matter is influenced by this issue, which differs significantly from e-communication and energy regulatory governance. Both *independence* and *discretion* variations are mostly associated with relevant operational incident reductions. The agency scope shows contrasting performance. While *market* supports high industrial output, increments in *transparency* are associated with safer operation.

3.3.4 Airports

In the case pf airports, governance dimensions show similarities with what was obseved in the case of railways in their correlations with industry performance. As shown in table 7, *independence* is negatively associated with the number of operating airports (model 1), and a lower number of operational incidents (model 2). A reinforcement in autonomy seems related to an additional decrease in operating airports (model 4). Thus a more independent regulator seems to select airports in favor of safer — , possibly larger, ones.

Our correlations seem aligned with the air transport sector's perspectives. The sector faces a growing unaccommodated demand and low incentives for new capacity deployment (EUROCONTROL, 2013). [15] Therefore, operators and regulators attempt to keep stable management costs and efficiency under the pressure of additional traffic.

Discretion is associated with inferior airport operation incidents (model 3). However, reinforced *discretion* is correlated to an increase in total system incidents, including gate-to-gate operations (model 5), and a growth in the number of operating airports (non-significant coefficient and not reported). This seems to confirm the fact that the when the regulatory choices lead to operate more airports, smaller, less reliable airports are operated; explaining the raise of gate-to-gate incidents (EUROSTAT, 2020).

Table 7: Air regression OLS estimates for governance dimensions

	class	dimen	category	coef	pval	N
1	mean	indep	airptot	-1.700	0.043	15
2	mean	indep	operincid	-1.160	0.021	15
3	mean	discretion	operincid	-0.930	0.004	15
4	growth	indep	numairpo	-0.170	0	16
5	growth	discretion	totdisrup	3,908.340	0.001	16

The column *class* shows whether the regression uses structural *mean* or variation *growth* dependent variable, as explained in section 3.3. The column *varia* shows the type of variable: *capacity* describes system variables (e.g., number of operating airports, number of arrivals), *quality* describes safety measures (absolute or relative) (e.g., number of air incidents). **Capacity dependent variables:** *airptot* and *numairpo* number of total airports operating in a country. **Quality:** *operincid* number of incidents related to airline operations, *totdisrup* total number of disruptions of any kinds in air operations. **Controls:** product market regulation in air sector, average number of aircrafts in a country airspace (measure of traffic), GDP per capita growth rate, number of hotel nights spent by tourists in a given country, country GDP in constant USD, air sector national accounts added value in constant USD. Complete regression tables are found in table 28 and table 29. Heteroskedasticity − robust standard errors.

As in the case of railways, reliance on *market* instruments is positively correlated to the volume of activity and capacity; in our case proxied by the aircraft fleet capacity operating in the country (models 1,2) in table 8. A 1%

[15] Slow economic recovery (2008 crisis), the European market maturity, and the perspective of industry growth outside the European Union have raised capital costs. EUROCONTROL (2013) highlights that capacity deployment projections have been re-estimated from a 38% (2008 benchmark) increase in 2030 to a 17% increase in 2035.

score increase accounts for an 8% increment in large capacity aircraft (150-250 passengers). Increasing transparency, on its side, is related to increasing traffic (as proxied by the number of flight arrivals (model 3), and the number of operating aircraft (model 5)). While traffic development goes with progress in terms of system disruptions (model reported in table 31), increased transparency triggers reduced airport operational incidents (model 4). A 1% increment in the *transparency* score is associated with a 23% decrease in average incident change. The combination of market incentives and transparency requirements seems, as in the case of railways, to be relied upon by RAs to push operators to accomodate growing demand, while avoiding to downgrade safety. EUROCONTROL (2013) highlights that airport operators and airlines need to optimize aircraft fleets, local available runways usage, and improve flight scheduling to deal with the growing unaccommodated demand.

Table 8: Air regression OLS estimates for scope dimensions

	class	dimen	category	coef	pval	N
1	mean	market	air150	4.990	0.002	15
2	mean	market	air250	4.480	0.001	15
3	growth	transparency	arrivals	664.390	0.008	14
4	growth	transparency	capacinci	-283.830	0.042	14
5	growth	transparency	numplanes	0.910	0.004	14

The column *class* shows whether the regression uses structural *mean* or variation *growth* dependent variable, as explained in section 3.3. The column *varia* shows the type of variable: *capacity* describes system variables (e.g., number of operating airports, number of arrivals), *quality* describes safety measures (absolute or relative) (e.g., number of air incidents). **Capacity dependent variables:** *numplanes* number of operating aircrafts registered in a country, *arrivals* number of year arrivals in all airports countrywide, *air*150 number of aircrafts with capacity below or equal to 150 passengers, *air*250 number of aircrafts with capacity from 150 to 250 passengers.**Quality:** *capacinci* number of incidents related to airport land operations. **Controls:** product market regulation in air sector, average number of aircrafts in a country airspace (measure of traffic), GDP per capita growth rate, number of hotel nights spent by tourists in a given country, country GDP in constant USD, air sector national accounts added value in constant USD. Complete regression tables are found in table 30 and table 31. Heteroskedasticity − robust standard errors.

In sum, more independent RAs seems to be able to favor the development of larger airports, with a positive impact on safety. Reliance on market levers has a positive impact on traffic, while the latter's development might translate into a higher number of incidents. To try to control for that later effect, RAs favor increased transparency, which seems to be a successful strategy. As said, more in-depth investigation would be needed to actually test causality.

4 Discussion and conclusion

Our methodology suggests four independent dimensions to characterize governance regimes:

- the independence from the government
- the regulatory agency discretion's level (which might also be interpreted in terms of the informality of its powers)

- the scope of market monitoring capabilities of the regulator
- the RA capability and obligation to organize transparency between the supply side and the other stakeholders of the regulatory game; in particular public authorities and the users.

These dimensions seem to capture long-run sectoral and national aspects of regulatory governance. At the sector level, energy and e-communication exhibit similar regulatory governance logic driven by the RA's high independence and low discretion. The latter benefits of strong capabilities to frame operators' behavior on the market and have been increasing their role of guarantor of transparency over the past years. The regulatory governance regime differs in the transportation industries. While rail RAs tend to be as independent as their counterpart in the e-communication and electricity industries, they benefit of less formal status and less clear delegation of authority. This is accentuated for the RAs responsible for airports, which remain in the governmental sphere. Across countries, institutional arrangements governing RA's status tend to differ a lot, and differences on the matter seems to persist. At the same time, there is a common trend toward developing the role of RAs as agents of transparency, which lies behind their ability to monitor markets and the economics of operators.

Moreover, our governance and scope dimensions (both the structures and evolutions) exhibit significant correlations with industry performance. However, strong contrasts exist across industries, suggesting that the actual role of regulators differs from one industry to the other. This might reflect partly differences in terms of "maturity" since, in Europe, the implementation of independent sectoral regulatory agencies started with the liberalization of the telecommunication markets in the 1990s, followed by the Energy industry ten years later, and by railways and airports mostly from the 2010s. Younger RAs seem characterized by more informality and access to a lower set of regulatory tools. However, it is not for sure that sectoral regulators converge toward a common model since they operate in industries with highly contrasted economies. In the e-communication sector, regulatory governance seems critical to the performance in terms of quality (broadband) of service. In contrast, in the electricity industry, the RA's main driver seems to be the energy price, even at the cost of electricity's environmental quality. In the transportation industries, the focus is on the volume/development of traffic and superior safety.

While our results are partly in line with previous studies that have attempted to establish a link between regulatory governance and performance, they differ on two main grounds. First, while independence from the government has already been identified as a significant dimension, we point out that the degree of discretion/formalization of RA's powers matters as well. We also highlight that RAs have clearly another role than designing markets and setting tariffs: promoting transparency. Moreover, in the past years in Europe,

the most significant evolutions in the matter of regulatory frameworks have concerned these two overlooked dimension of regulatory governance regimes: discretion/formalism and transparency. Our exploratory study clearly calls for more in-depth analyses of these dimensions and consequences for sectoral performances.

The drivers of the potential causal relationships identified in this study request further investigations. We have not accounted for other features that impact performance, such as national political priorities, unobserved sector characteristics that also influence operators' economic incentives and decisions. Further research is needed to understand better the potential loops between sector characteristics and governance regimes and between governance evolutions and industrial performances. Such detailed analyses are necessary to provide policymakers with relevant knowledge to design superior regulatory institutions.

5 Appendices

5.1 Topic modelling

Figure 5: Term distribution per document

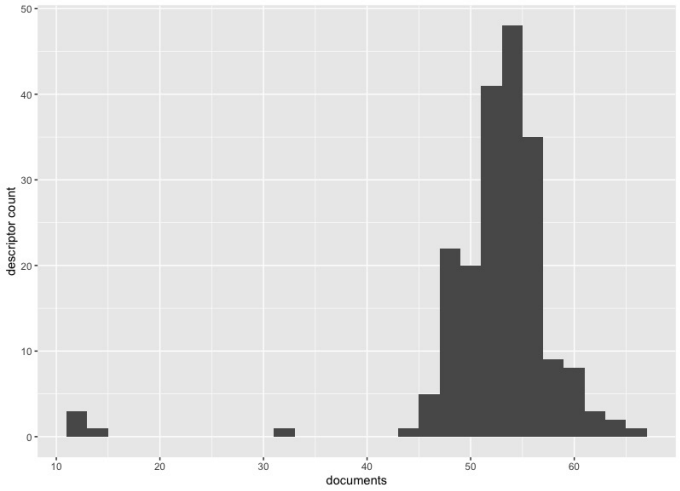


Figure 6: Histogram of the agency mean dimension score

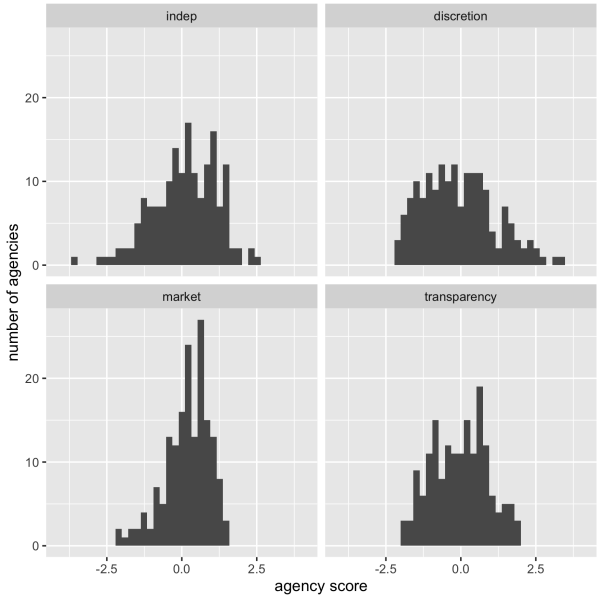
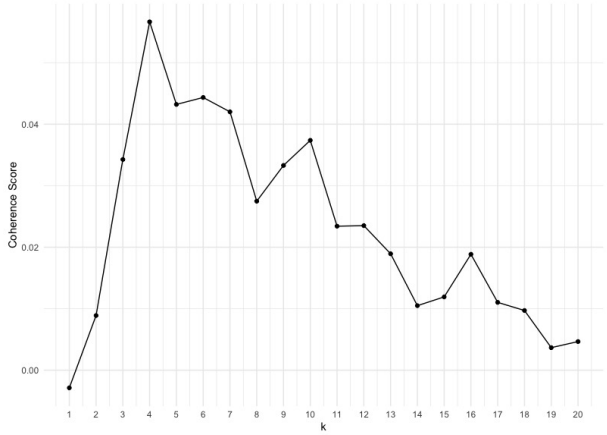


Figure 7: Coherence score - optimal number of topics



5.2 Summary statistics

5.2.1 Summary statistics and correlation tables

Table 9: Summary Statistics for LDA dimension weight estimates (scaled)

sector	indicator	year	mean	sd	min	max	n
air	indep	2,013	-1.022	1.161	-3.509	0.690	18
air	indep	2,018	-0.316	1.262	-2.601	1.421	18
air	discretion	2,013	0.047	1.608	-1.983	2.760	18
air	discretion	2,018	0.152	1.476	-2.008	3.353	18
air	market	2,013	-0.231	0.960	-2.135	0.944	18
air	market	2,018	-0.476	0.830	-2.166	0.853	18
air	transparency	2,013	-0.838	0.599	-1.840	0.460	18
air	transparency	2,018	0.008	0.677	-1.098	1.150	18
ene	indep	2,013	-0.107	0.959	-2.364	1.571	23
ene	indep	2,018	0.410	0.774	-1.327	1.499	23
ene	discretion	2,013	-0.169	1.049	-1.747	1.762	23
ene	discretion	2,018	-0.553	0.900	-2.040	1.098	23
ene	market	2,013	0.438	0.555	-0.512	1.460	23
ene	market	2,018	0.410	0.459	-0.659	1.318	23
ene	transparency	2,013	-0.375	0.794	-1.610	1.380	23
ene	transparency	2,018	0.553	0.875	-1.460	1.910	23
rail	indep	2,013	-0.214	0.798	-1.610	1.380	22
rail	indep	2,018	0.781	0.730	-0.500	1.910	21
rail	discretion	2,013	0.640	1.274	-1.791	3.221	22
rail	discretion	2,018	0.397	0.870	-1.574	2.444	21
rail	market	2,013	-0.107	0.826	-1.610	1.181	22
rail	market	2,018	0.105	0.584	-1.194	1.080	21
rail	transparency	2,013	-0.747	0.744	-1.825	0.878	22
rail	transparency	2,018	0.279	0.595	-1.055	1.688	21
tel	indep	2,013	0.077	0.789	-1.610	1.380	22
tel	indep	2,018	0.863	0.826	-0.337	2.530	22
tel	discretion	2,013	-0.376	1.081	-2.103	2.389	22
tel	discretion	2,018	-0.536	0.932	-1.943	1.610	22
tel	market	2,013	0.462	0.599	-1.221	1.460	22
tel	market	2,018	0.653	0.411	0	1.307	22
tel	transparency	2,013	-0.009	0.940	-1.499	1.688	22
tel	transparency	2,018	0.746	0.702	-0.786	1.989	22

Table 10: Pearson correlations for dimension mean-scores

var1	var2	agency	country	air	rail	ene	tel
indep	transparency	0.410	0.314	0.111	0.267	0.413	0.459
indep	market	0.331	0.163	0.214	0.207	-0.016	0.029
transparency	market	0.388	0.598	0.234	0.220	-0.043	0.234
indep	discretion	0.384	0.527	0.629	0.478	0.611	0.399
transparency	discretion	-0.512	-0.558	-0.462	-0.479	-0.336	-0.522
market	discretion	-0.581	-0.654	-0.516	-0.599	-0.442	-0.630

Table 11: Pearson correlations for dimension evolution-scores

var1	var2	agency	country	air	rail	ene	tel
indep	transparency	0.419	0.314	0.111	0.267	0.413	0.459
indep	market	0.234	0.163	0.214	0.207	-0.016	0.029
transparency	market	0.093	0.598	0.234	0.220	-0.043	0.234
indep	discretion	0.284	0.527	0.629	0.478	0.611	0.399
transparency	discretion	-0.453	-0.558	-0.462	-0.479	-0.336	-0.522
market	discretion	-0.467	-0.654	-0.516	-0.599	-0.442	-0.630

Table 12: Significant Pearson correlation coefficients TA and OECD scores

avg_var	indic	year_indic	correlcoef
indep	ACC	2018	0.479
indep	IND	2013	0.558
indep	IND	2018	0.580
discretion	ACC	2013	-0.473
discretion	SCO	2013	-0.584
discretion	SCO	2018	-0.470
market	ACC	2013	0.534
market	ACC	2018	0.542
market	IND	2018	0.405
market	SCO	2013	0.771
market	SCO	2018	0.655
transparency	ACC	2013	0.546
transparency	ACC	2018	0.557
transparency	IND	2018	0.543
transparency	SCO	2013	0.558
transparency	SCO	2018	0.442

Table 13: Levene test for equality of sector distribution variance

dimension	meantest	evoltest
indep	0.122	0.289
discretion	0.075	0.505
market	0.005	0.46
transparency	0.375	0.612

Table 14: Tukey test between sector distributional differences

sector1	sector2	dimension	meandif	meanpval	evoldif	evolpval
ene	air	indep	0.821	0.001	0.002	1
ene	rail	indep	-0.120	0.939	-0.104	0.317
rail	air	indep	0.941	0	0.105	0.388
tel	air	indep	1.139	0	0.085	0.570
tel	ene	indep	0.318	0.420	0.084	0.507
tel	rail	indep	0.198	0.785	-0.020	0.988
ene	air	discretion	-0.461	0.275	-0.125	0.556
ene	rail	discretion	-0.883	0.002	-0.042	0.965
rail	air	discretion	0.422	0.367	-0.083	0.825
tel	air	discretion	-0.556	0.141	-0.087	0.805
tel	ene	discretion	-0.095	0.980	0.038	0.973
tel	rail	discretion	-0.978	0.001	-0.004	1
ene	air	market	0.778	0	0.035	0.929
ene	rail	market	0.428	0.014	-0.058	0.699
rail	air	market	0.350	0.094	0.093	0.394
tel	air	market	0.911	0	0.071	0.624
tel	ene	market	0.134	0.775	0.036	0.907
tel	rail	market	0.561	0.001	-0.022	0.978
ene	air	transparency	0.504	0.051	0.010	1
ene	rail	transparency	0.335	0.275	-0.041	0.959
rail	air	transparency	0.169	0.826	0.050	0.943
tel	air	transparency	0.784	0.001	-0.041	0.968
tel	ene	transparency	0.279	0.431	-0.051	0.925
tel	rail	transparency	0.614	0.007	-0.092	0.693

Table 15: Time differences and normality tests statistics

variable	sector	avgmean	sdmean	avgevol	sdevol	meandif	pval	pnormm	pnorme
indep	air	-0.678	1.047	0.076	0.265	-0.582	0.936	0.788	0.093
indep	rail	0.249	0.677	0.182	0.153	-0.971	1	0.4	0.696
indep	ene	0.152	0.676	0.078	0.209	-0.517	0.983	0.118	0.626
indep	tel	0.461	0.677	0.162	0.169	-0.808	1	0.697	0.205
discretion	air	0.099	1.306	0.081	0.44	0.008	0.492	0.446	0.006
discretion	rail	0.532	0.897	-0.002	0.228	0.226	0.209	0.647	0.513
discretion	ene	-0.361	0.748	-0.044	0.277	0.384	0.079	0.173	0.131
discretion	tel	-0.472	0.888	-0.006	0.212	0.142	0.246	0.026	0.228
market	air	-0.355	0.772	-0.029	0.23	0.264	0.123	0.461	0.006
market	rail	-0.023	0.571	0.064	0.209	-0.175	0.805	0.447	0.219
market	ene	0.424	0.343	0.006	0.142	0.028	0.43	0.691	0.83
market	tel	0.555	0.421	0.042	0.127	-0.164	0.884	0.42	0.16
transparency	air	-0.422	0.538	0.239	0.254	-0.854	1	0.203	0.096
transparency	rail	-0.26	0.485	0.29	0.269	-1.021	1	0.527	0.423
transparency	ene	0.089	0.545	0.249	0.311	-0.929	0.999	0.39	1
transparency	tel	0.379	0.675	0.198	0.236	-0.786	1	0.761	0.367

5.2.2 Summary statistics graphs

Figure 8: TA Score box plots by sector - scaled scores

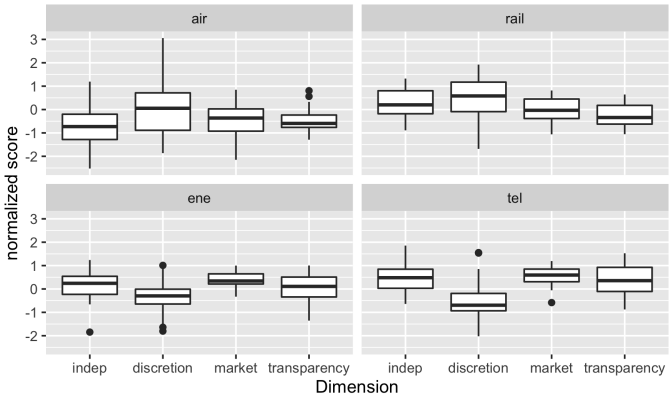


Figure 9: Dimension national averages

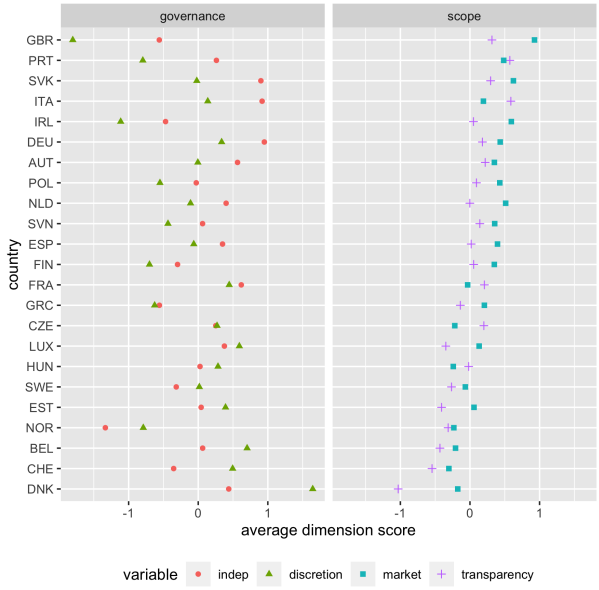
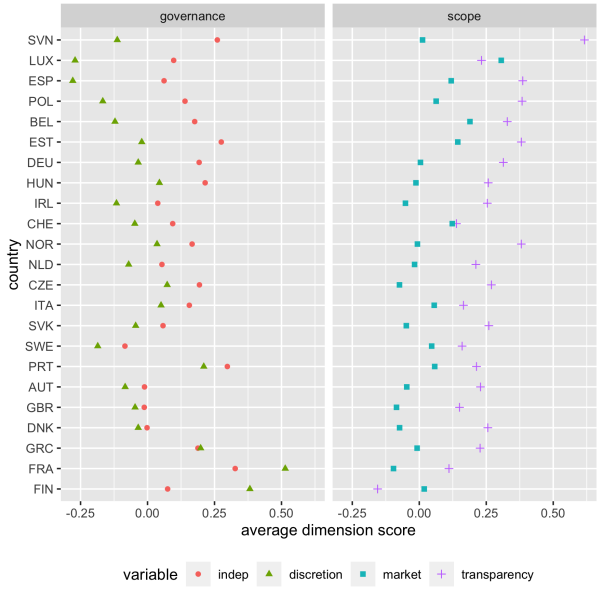


Figure 10: Dimension National Evolution Scores



5.3 Regression tables

5.3.1 Energy

Table 16: Energy regression OLS estimates for Governance *mean* variables

var	gas	solarsur	renewprod	price_ind	hydroprod	exportelec	renewprod.1	elecprod	gas.1
indep_mean						191.5131	620.4069	922.632	-680.373
						(69.41025)	(181.8705)	(364.3199)	(217.7151)
discretion_mean	-528.4476	412.9042	409.1687	21.96157	-2928.95				
	(166.0969)	(174.7952)	(164.9765)	(9.899704)	(1289.467)				
captotcap	-0.07661	0.052396	0.082084	-0.004596	1.230435	0.032437	0.101413	0.056891	-0.065625
	(0.111708)	(0.053397)	(0.060419)	(0.002313)	(0.258757)	(0.018991)	(0.058522)	(0.054929)	(0.115315)
dis_equ	337.4918	-64.32178	223.1411	-26.91209	-7.379974	-45.46125	242.5701	-293.1488	304.5241
	(170.1391)	(113.0668)	(134.4615)	(8.198546)	(896.663)	(49.40928)	(132.5203)	(126.3413)	(163.9183)
ele_ind	18.11337	-1.246861	-2.733485		-4.599501	-3.005614	-4.147501	-17.23339	20.11745
	(5.734556)	(4.547674)	(1.978511)		(20.61011)	(1.700268)	(2.023774)	(5.185602)	(6.537703)
pmr_elec	2051.933	330.002	777.8526	100.4971	-5067.06	-853.6547	283.6566	-492.3604	2923.057
	(2033.979)	(1710.792)	(1106.97)	(62.44803)	(7031.376)	(426.3)	(1134.199)	(972.2684)	(2098.675)
nygdppcapkdzg	-1441.544	321.6297	361.8704	22.63531	-9073.06	-49.68165	72.37978	228.2639	-1071.611
	(705.3101)	(338.2226)	(242.2195)	(34.05805)	(2292.533)	(130.784)	(240.5352)	(379.5197)	(803.2321)
gctaxtotlgdzs	144.576	79.05196	682.0099	-49.95163	-11246.93	-293.5387	443.4676	-1138.004	286.8014
	(890.8261)	(1113.842)	(480.7065)	(39.6725)	(4021.883)	(189.6416)	(510.2935)	(683.8995)	(868.9075)
rgdpe_penn	0.006549	0.000672	0.002696	0.000359	-0.052893	0.000127	0.001417	0.012422	0.007416
	(0.003863)	(0.002096)	(0.002728)	(0.000102)	(0.011869)	(0.000925)	(0.002527)	(0.003035)	(0.003789)
N	18	18	19	22	22	22	22	22	N
r2	0.9011	0.9074	0.72	0.9142	0.7258	0.7622	0.7409	0.9672	r2

System variables: *elecprod* total electricity production in GWh, *exportelec* electricity exports in GWh.**Renewable:** *renewprod* production of electricity based in renewable sources in GWh (EUROSTAT classification), **CO2:** *gas* production of electricity based on gas in kilogram of oil equivalent (ktoe, 11.63 MWh). **Controls:** *captotcap* total system generation capacity in MWe, *dis_equ* absolute value of country latitude, *ele_ind* industrial consumer electricity price, *nygdppcapkdzg* GDP per capita growth rate, *gctaxtotlgdzs* tax revenue percentage in national budget, *pmr_elec* market regulation score (OECD, higher scores represent lower barriers to competition), *rgdpe_penn* GDP in constant USD. Summary statistics of each variables are found in table 32 Heteroskedasticity − robust standard errors in parenthesis.

Table 17: Energy regression OLS estimates for Governance *growth* variables

var	price_hou	consPC	perenew	totcomb	pcgen	imports	windconv	solcap	intcomb
indep_growth					7.130443	8.018553	0.000831	-0.00067	0.269436
					(2.411024)	(2.673139)	(0.00026)	(0.000288)	(0.077964)
discretion_growth	3.718279	1.520725	-0.008709	6.883739					
	(0.976294)	(0.647922)	(0.002809)	(2.746245)					
capdif					0.004248	-0.001411	-1e-06	-2e-06	0.000371
					(0.003621)	(0.002409)	(1e-06)	(1e-06)	(0.001185)
captotcap	-0.000286	0.00152	-7e-06	0.005687					
	(0.005624)	(0.001769)	(8e-06)	(0.007032)					
dis_equ	-4.618439	5.383244	0.087133	-33.82603	24.58809	35.19547	0.001568	0.001152	0.323042
	(9.413106)	(6.079257)	(0.019567)	(13.29462)	(16.93258)	(13.9286)	(0.000943)	(0.001595)	(0.570591)
ele_ind	0.148011	0.13747	0.000895	-0.97201	0.583248	0.477248	-1.6e-05	-2.5e-05	0.018495
	(0.39515)	(0.185045)	(0.000451)	(0.337636)	(0.287746)	(0.209607)	(2.8e-05)	(4.7e-05)	(0.010382)
gdp	-0.306202	-12.45458	0.20402	-51.07284	-17.54792	-3.785293	-0.000146	0.002128	0.328989
	(25.64099)	(9.929145)	(0.04409)	(30.86446)	(6.121536)	(3.767324)	(0.000933)	(0.001024)	(0.469727)
pmr_elec	-34.76091	114.5131	0.569337	-168.3827	185.8199	109.1629	0.012576	-0.000223	-4.446918
	(100.7404)	(41.27312)	(0.121069)	(101.7721)	(85.28114)	(75.40224)	(0.008258)	(0.013385)	(6.003392)
nygdppcapkdzg	-55.17571	104.4289	0.01078	75.46355	144.4187	9.466684	-0.008432	-0.001622	0.408187
	(38.27071)	(23.39653)	(0.06734)	(48.77213)	(37.87256)	(13.57049)	(0.002804)	(0.00552)	(1.105121)
gctaxtotlgdzs	-37.12663	138.9555	0.295768	-4.769603	226.8207	81.88775	-0.011745	-0.012221	-0.666864
	(46.20404)	(35.90542)	(0.122555)	(59.46887)	(64.38649)	(43.54897)	(0.00439)	(0.010754)	(1.646644)
N	22	20	20	20	20	22	20	20	19
r2	0.2987	0.7951	0.9786	0.748	0.7442	0.6089	0.6718	0.5397	0.416

System variables: *pcgen* per capita electricity production in GWh, *imports* electricity imports in GWh, *price_ind* electricity price for industrial consumers in USD per GWh. **Renewable:** *solcap* solar electricity capacity in MWe, *windconv* electricity produced using wind sources divided by the total wind capacity, *hydro* electricity capacity based in hydropower. **CO2:** *intcomb* and *othcomb* electricity capacity based in internal combustion machines in MWe. **Controls:** *capdif* period difference generation capacity in MWe, *dis_equ* absolute value of country latitude, *ele_ind* industrial consumer electricity price, *nygdppcapkdzg* GDP per capita growth rate, *gctaxtotlgdzs* tax revenue percentage in national budget, *pmr_elec* market regulation score (OECD, higher scores represent lower barriers to competition), *gdp* GDP in constant USD. Summary statistics of each variables are found in table 32 Heteroskedasticity − robust standard errors in parenthesis.

Table 18: Energy regression OLS estimates for Scope *mean* variables

var	elecprod	natgassh	CO2sh	windgrate	windcap	newcapy	exportelec	totcap
market_mean	1054.805	7.694201	1.017747	0.029856	-130.0688	-272.5763	-414.1456	-453.1903
	(358.0604)	(3.104294)	(0.40576)	(0.012762)	(48.01594)	(108.2541)	(136.8174)	(185.148)
captotcap	0.196928	0.000142	0.000191	2e-06	-0.011746	-0.057327	0.01311	0.768542
	(0.043353)	(0.000348)	(5.4e-05)	(1e-06)	(0.006999)	(0.015137)	(0.01884)	(0.026038)
dis_equ	-97.05054	-2.356329	-0.097498	0.004045	19.59289	49.68168	-38.89753	100.0352
	(92.28043)	(0.925007)	(0.083413)	(0.003117)	(12.41234)	(31.32502)	(47.28245)	(49.81055)
ele_ind	2.904208	-0.003732	0.003467	-7.8e-05	-0.034723	-0.701017	-1.97456	-0.30613
	(1.755173)	(0.02555)	(0.001965)	(7.2e-05)	(0.621884)	(1.063686)	(1.110284)	(2.535911)
pmr_elec	-307.0361	2.121805	-0.415076	-0.012503	37.03241	489.2404	-556.5202	1160.493
	(628.1997)	(4.996507)	(0.714418)	(0.027066)	(143.9944)	(420.9482)	(355.4948)	(477.9121)
nygdppcapkdzg	-102.0134	-0.874849	-0.0918	-0.026374	17.10778	40.00167	43.9396	-187.3175
	(187.8727)	(3.503073)	(0.24485)	(0.008565)	(58.70098)	(129.8939)	(122.4144)	(236.9213)
gctaxtotlgdzs	262.8523	-11.38329	0.369352	-0.041627	27.57048	227.7059	-31.16829	453.5245
	(275.8318)	(5.540935)	(0.43415)	(0.017378)	(109.3227)	(258.1851)	(166.0583)	(298.4733)
rgdpe_penn	-0.007233	-1e-06	-7e-06	0	0.001229	0.004366	0.000964	0.014516
	(0.001967)	(1.9e-05)	(2e-06)	(0)	(0.000377)	(0.000794)	(0.000969)	(0.001435)
N	22	20	20	20	20	20	22	20
r2	0.8782	0.6368	0.88	0.686	0.8554	0.8856	0.7992	0.9995

System dependent variables: *totcap* total system capacity in MWe, *elecprod* total production in GWh, *newcapy* yearly growth rate of new deployed system capacity in percentage, *exportelec* electricity exports in GWh, *price_ind* electricity price for industrial consumers in USD per GWh. **Renewable sources:** *renewprod* production of electricity based in renewables and biofuels in GWh, *solarsur* solar panels surface in m^2, *hydroprod* electricity production based in hydroelectric plants in GWh, *windgrate* yearly growth rate of new system deployed wind capacity in percentage, *windcap* wind electricity capacity in MWe. **CO2 sources:** *gas* gas-based electricity produced in kilogram of oil equivalent (ktoe, 11.63 MWh), *natgassh* share of natural gas in electricity output in percentage, *CO2sh* CO2 technology electricity production in GWh. **Controls:** *captotcap* total system generation capacity in MWe, *dis_equ* absolute value of country latitude, *ele_ind* industrial consumer electricity price, *nygdppcapkdzg* GDP per capita growth rate, *gctaxtotlgdzs* tax revenue percentage in national budget, *pmr_elec* market regulation score (OECD, higher scores represent lower barriers to competition), *rgdpe_penn* GDP in constant USD. Summary statistics of each variables are found in table 32. Heteroskedasticity − robust standard errors in parenthesis.

Table 19: Energy regression OLS estimates for Scope *growth* variables

var	solarpv	othcomb	price_ind
market_growth	38.3938	11.93457	
	(16.33782)	(4.418271)	
transparency_growth			-2.996146
			(1.212834)
capdif			-0.00015
			(0.00472)
captotcap	-0.10611	-0.030174	
	(0.015541)	(0.004067)	
dis_equ	80.07472	26.31493	0.469309
	(36.71269)	(8.102074)	(5.944399)
ele_ind	0.302847	-0.012361	
	(1.157639)	(0.253171)	
gdp	663.9564	165.9719	9.780258
	(84.96778)	(17.36037)	(3.931562)
pmr_elec	338.5195	225.4796	-63.47445
	(182.0332)	(74.1332)	(61.41701)
nygdppcapkdzg	-92.01582	-27.57465	-2.92636
	(98.06669)	(33.23933)	(14.62252)
gctaxtotlgdzs	163.1922	21.35686	13.29657
	(123.4109)	(37.58155)	(28.62713)
N	21	18	18
r2	0.9404	0.8404	0.9361

System dependent variables: $consPC$ electricity consumption per capita in GWh, $price_hou$ electricity price for household consumers in USD per GWh. **Renewable sources:** *solarpv* solar photovoltaic electricity capacity, *perrenew* yearly growth rate of new system deployed renewable capacity in percentage. **CO2 sources:** *totomb* and *othcomb* electricity capacity based in all internal combustion technologies in MWe. **Controls:** *captotcap* total system generation capacity in MWe, *dis_equ* absolute value of country latitude, *ele_ind* industrial consumer electricity price, *nygdppcapkdzg* GDP per capita growth rate, *gctaxtotlgdzs* tax revenue percentage in national budget, *pmr_elec* market regulation score (OECD, higher scores represent lower barriers to competition), *rgdpe_penn* GDP in constant USD. Summary statistics of each variables are found in table 32 Heteroskedasticity − robust standard errors in parenthesis.

5.3.2 E-communication

Table 20: E-communication regression OLS estimates for scope *growth* dimensions

var	pcprice	penrate	accpath	cabsubs	totsubs
transparency_growth	0.008476	0.175436	0.221049	0.044571	0.287369
	(0.003154)	(0.074681)	(0.100345)	(0.017051)	(0.08944)
fix_broad_100	-0.024578				
	(0.019311)				
gdp	-0.022185	0.53558	-0.220101	0.127651	0.289701
	(0.035079)	(0.3565)	(0.477189)	(0.037534)	(0.465064)
pmr	-0.004105	-11.80108	0.464027	0.24355	-11.14732
	(0.218928)	(6.079729)	(2.768864)	(0.346336)	(5.249404)
npopdnst	-0.000433	0.000546	-0.014754	-0.002242	-0.016589
	(0.000643)	(0.024869)	(0.015852)	(0.002902)	(0.022703)
fix_bprice			-0.554055	-0.091416	-0.699121
			(0.384395)	(0.047985)	(0.378433)
mob_bprice		-0.136276			
		(0.12969)			
nights	0.002688	-0.058973	0.035974	-0.008181	-0.038704
	(0.002528)	(0.020507)	(0.03552)	(0.002629)	(0.035068)
nygdppcapkdzg		-0.188892			
		(1.226448)			
spurbtotlinzs	-0.047828	1.12518	4.16893	0.172853	1.290843
	(0.104202)	(2.153122)	(2.009066)	(0.353986)	(2.375703)
rule_wb	0.372966	-9.539293	-0.946497	0.055453	-6.355056
	(0.266377)	(4.661781)	(3.208414)	(0.592442)	(3.741083)
N	19	20	20	20	19
r2	0.3934	0.5381	0.5776	0.6241	0.6876

Capacity dependent variables: *mobsubs* number of subscribers to mobile broadband services. **Quality:** *bbsubs*100 number of subscribers to fix broadband internet service per 100 habitants, *mobpenrate* penetration rate for mobile services. **Prices:** *pprice* price for 5gb fix internet bundle in PPP USD. **Controls:** *rule_wb* rule of law index from the World Bank, *pmr* telecom market regulation score from OECD (higher values show higher support for market competition). *rgdpe_penn* GDP in PPP USD, *npopdnst* population density, *spurbtotlinzs* percentage of population living in urban areas, *nights* number of hotel nights spent by tourists in a given country, fix_bprice average price for fix and mobile broadband services in PPP USD. Summary statistics of each variables are found in table 33. Heteroskedasticity − robust standard errors in parenthesis.

Table 21: E-communication regression OLS estimates for scope *mean* dimensions

var	fix100low	fixtot	mobterm	accpathtot	fixprate	fixtotal	fix100	prixfix	prixmob
market_mean	0.845317	-56658.72	-0.217578						
	(0.331209)	(24382.35)	(0.090024)						
transparency_mean				1302041	-1.293476	-1.320625	-1.337509	-1.721405	-2.015619
				(464440.6)	(0.255003)	(0.243481)	(0.263749)	(0.621008)	(0.710632)
fix_broad_100								-1.07599	-0.670534
								(0.366758)	(0.592341)
pmr	2.328794	-231727.4	0.345476	-3504518	2.992274	3.00351	3.124162	4.806679	-0.456085
	(1.213525)	(138776.3)	(0.216179)	(2610092)	(2.065594)	(2.002638)	(2.592419)	(1.703813)	(2.351073)
npopdnst	-0.004195	-412.0885	0.002592	12217.94	0.000876	-0.000996	-0.001267	-0.00414	-0.021189
	(0.005332)	(444.5797)	(0.000746)	(10821.24)	(0.006642)	(0.006559)	(0.007267)	(0.008493)	(0.017177)
fix_bprice	-0.181236	4704.125		64550.64	-0.20164	-0.211048	-0.172223		
	(0.132058)	(9589.399)		(264638.7)	(0.15141)	(0.146502)	(0.223292)		
mob_bb_penrate			0.004687						
			(0.006397)						
nights	0.015546	-621.0986	-0.00148	29286.57	0.010496	0.010965	0.008962	-0.040403	-0.109382
	(0.018145)	(980.7844)	(0.002406)	(33339.54)	(0.014924)	(0.014754)	(0.019564)	(0.028334)	(0.033819)
nygdppcapkdzg			0.136858						
			(0.063171)						
spurbtotlinzs	-1.475392	-111005	-0.089569	-241128.4	2.491411	2.802557	3.24595	6.171125	2.153008
	(0.805647)	(79866.89)	(0.118724)	(1265760)	(0.780003)	(0.720634)	(1.229441)	(1.367849)	(3.680011)
rgdpe_penn	-2e-06	0.112326	0	43.91722	1e-06	1e-06	1e-06	9e-06	1.9e-05
	(3e-06)	(0.11486)	(0)	(5.024587)	(2e-06)	(2e-06)	(3e-06)	(4e-06)	(5e-06)
rule_wb	1.667939	-314249.9	-0.082948	-4965209	3.321087	2.976541	1.784598	-1.108354	-4.995374
	(1.373395)	(132993.5)	(0.28375)	(3078189)	(1.495009)	(1.505223)	(1.682501)	(3.111963)	(2.468287)
N	20	19	20	21	21	21	20	20	20
r2	0.6017	0.5494	0.7187	0.9912	0.8477	0.8487	0.8125	0.6059	0.6518

Capacity dependent variables: $fixtot$ number of subscribers to fix broadband internet service, *server* number of internet servers available in a country, *accpath* number of access paths for telecom services. **Quality:** $fixprate$ penetration rate for fix broadband internet services, $fix100$ number of subscribers to fix broadband internet service per 100 habitants, $fix100low$ number of subscribers to low-speed fix broadband internet service per 100 habitants. **Prices:** $prixfix$ price for 5gb fix internet bundle in PPP USD, $prixmob$ price for 5gb mobile internet bundle in PPP USD, $mobterm$ price of termination rates in another network in USD. **Controls:** $rule\_wb$ rule of law index from the World Bank, *pmr* telecom market regulation score from OECD (higher values show higher support for market competition). $rgdpe\_penn$ GDP in PPP USD, $npopdnst$ population density, $spurbtotlinzs$ percentage of population living in urban areas, $nights$ number of hotel nights spent by tourists in a given country, fix_bprice average price for fix and mobile broadband services in PPP USD. Summary statistics of each variables are found in table 33. Heteroskedasticity − robust standard errors in parenthesis.

Table 22: E-communication regression OLS estimates for governance *mean* dimensions

var	server	fixprate	prixmob	prixfix	fix100	accpath	telVoIP	prixfixbb	landli
indep_mean							1227790	0.042146	-851839.6
							(532485.3)	(0.015492)	(358780.9)
discretion_mean	1032.236	0.770641	0.73183	0.024248	-0.663838	-835946.7			
	(412.3172)	(0.230808)	(0.311497)	(0.008635)	(0.297508)	(319045)			
fix_broad_100			0.023956	-0.023855				-0.008269	
			(0.359059)	(0.013966)				(0.009952)	
pmr	-7893.629	0.287964	-0.996348	-0.023251	-1.764443	-637660.6	-620053.9	-0.077502	-664689.1
	(4083.536)	(1.960541)	(1.903697)	(0.087635)	(2.659266)	(2852977)	(1276125)	(0.09696)	(2254060)
npopdnst	29.19601	0.004672	0.011056	-0.000403	-0.019495	9014.28	-4290.117	-0.000348	7121.56
	(21.78903)	(0.009484)	(0.009799)	(0.000381)	(0.008001)	(10215.62)	(6122.6)	(0.000321)	(8359.55)
fix_bprice	404.7259	0.161626			-0.711761	-319615.6	278877.6		25179.19
	(224.4883)	(0.147108)			(0.234432)	(274799.6)	(190955.4)		(152114.1)
nights	-59.03848	-0.015859	-0.041692	-0.003865	0.054205	57612.44	-33484.38	-0.004104	40210.66
	(24.14568)	(0.015628)	(0.031449)	(0.001093)	(0.022912)	(36831.39)	(14725.75)	(0.001048)	(19012.57)
spurbtotlinzs	458.6984	2.029297	0.961974	0.223054	-3.864636	205611.5	729767.3	0.1613	142210.9
	(2291.74)	(1.211452)	(1.600968)	(0.059884)	(0.93349)	(1632289)	(559750.7)	(0.047281)	(873204.9)
rgdpe_penn	0.00524	3e-06	5e-06	1e-06	-9e-06	41.05072	10.61219	0	1.167534
	(0.004012)	(2e-06)	(4e-06)	(0)	(3e-06)	(5.704789)	(1.779096)	(0)	(3.934059)
rule_wb	8410.973	5.702149	-3.545942	-0.430276	9.907512	-7276818	3914841	-0.369936	-2754472
	(2540.235)	(1.99965)	(3.829511)	(0.125517)	(2.300331)	(2738688)	(2030179)	(0.119118)	(1365662)
N	21	21	20	20	21	21	19	20	21
r2	0.707	0.7932	0.5359	0.8148	0.8282	0.9909	0.9317	0.8435	0.8299

Capacity dependent variables: $telVoIP$ number of subscribers of VoIP services, *landli* number of landline subscribers, *accpathtot* number of access paths for telecom services. **Quality:** $fixprate$ penetration rate for fix broadband internet services, $fix100$ and $fixitu100$ number of subscribers to fix broadband internet service per 100 habitants. **Prices:** $prixfix$ and $prixfixbb$ price for 5gb fix internet bundle in PPP USD, $prixmob$ price for 5gb mobile internet bundle in PPP USD. **Controls:** $rule\_wb$ rule of law index from the World Bank, *pmr* telecom market regulation score from OECD (higher values show higher support for market competition). $rgdpe\_penn$ GDP in PPP USD, $npopdnst$ population density, $spurbtotlinzs$ percentage of population living in urban areas, *nights* number of hotel nights spent by tourists in a given country, fix_bprice average price for fix and mobile broadband services in PPP USD. Summary statistics of each variables are found in table 33. Heteroskedasticity − robust standard errors in parenthesis.

Table 23: E-communication regression OLS estimates for governance *growth* dimensions

var	ppprice	mobpenrate	mobsubs	bbsubs100	ppprice.1	voipsubs	cabsubs
indep_growth					-0.149343	41.86174	0.02741
					(0.06657)	(16.18913)	(0.011902)
discretion_growth	-0.44641	0.331822	81.58199	0.408568			
	(0.124691)	(0.133698)	(31.20843)	(0.138559)			
fix_broad_100	-0.472402				-0.428646		
	(0.448526)				(0.663264)		
gdp	0.433275	-1.347462	344.5446	-1.692517	-0.122876	499.8525	0.060193
	(0.597726)	(0.610543)	(249.8055)	(0.77083)	(0.951962)	(75.09492)	(0.046835)
pmr	-2.73734	-4.309135	4326.147	-5.638853	2.160783	446.1506	0.03486
	(5.318986)	(6.278435)	(1856.861)	(7.497756)	(6.39717)	(422.0932)	(0.669656)
npopdnst	-0.015612	0.027893	5.221454	0.045838	0.022452	4.266321	-0.001932
	(0.019556)	(0.024842)	(6.901828)	(0.025592)	(0.014592)	(2.309611)	(0.002571)
fix_bprice						15.68655	-0.120851
						(46.3468)	(0.050404)
mob_bprice		-0.013203	93.64867	0.121931			
		(0.147861)	(42.76748)	(0.165334)			
nights	0.005145	0.0751	15.40989	0.096938	0.022099	-18.2502	-0.001617
	(0.035421)	(0.03262)	(13.16374)	(0.040716)	(0.067622)	(5.394723)	(0.004393)
nygdppcapkdzg		2.089168	731.2595	2.072755			
		(0.995649)	(194.1195)	(1.457811)			
spurbtotlinzs	2.282722	-9.502063	187.6668	-8.134859	-1.143994	-132.9306	0.187391
	(3.305546)	(2.890715)	(1174.543)	(3.456588)	(3.215978)	(315.4793)	(0.275598)
rule_wb	5.509252	15.45004	5500.934	21.99715	8.042235	-211.549	0.111145
	(3.395972)	(7.742872)	(1093.597)	(8.312667)	(5.791605)	(512.1559)	(0.684223)
N	19	20	20	20	17	19	N
r2	0.5137	0.4689	0.7927	0.5064	0.957	0.4915	r2

Capacity dependent variables: *voipsubs* number of subscribers of VoIP services, *landli* number of landline subscribers, *accpath* number of access paths for telecom services. **Quality:** *totsubs* number of subscribers to fix broadband internet service per 100 habitants, *cabsubs* and *cabsubs*.1 number of cable internet subscribers per 100 habitants, *mobpenrate* penetration rate for mobile services, *accpath* telecom access paths per 100 habitants. **Prices:** *pprice* price for 5gb fix internet bundle in PPP USD, *pcprice* price for 5gb fix internet bundle as percentage of per capita income. **Controls:** *rule_wb* rule of law index from the World Bank, *pmr* telecom market regulation score from OECD (higher values show higher support for market competition). *rgdpe_penn* GDP in PPP USD, *npopdnst* population density, *spurbtotlinzs* percentage of population living in urban areas, *nights* number of hotel nights spent by tourists in a given country, fix_bprice average price for fix and mobile broadband services in PPP USD. Summary statistics of each variables are found in table 33. Heteroskedasticity − robust standard errors in parenthesis.

5.3.3 Rail Transport

Table 24: Rail Transport regression OLS estimates for governance *mean* dimension

var	fallback	passinjurtot	allaccidkm	activecross	injurpasskm
indep_mean				0.011793	1.08445
				(0.003824)	(0.216034)
discretion_mean	48825.09	0.592615	-0.050583		
	(19431.51)	(0.242838)	(0.021707)		
pmr	-66672	-1.309245	-0.018754	-0.030305	-2.940149
	(148101.9)	(1.517681)	(0.154142)	(0.027476)	(1.084991)
nygdppcapkdzg	-539814.5	-0.935285	0.162279	-0.054874	-1.687914
	(174325.6)	(1.628982)	(0.257176)	(0.025586)	(0.915558)
mean_rai_35	5.589703	0.000167	2e-05	6e-06	0.000359
	(30.46726)	(0.000353)	(3.9e-05)	(7e-06)	(0.000221)
rgdpe_penn	-0.112622	2e-06	0	0	-2e-06
	(0.36135)	(4e-06)	(0)	(0)	(2e-06)
rule_wb	908176.5	-2.96204	-0.525559	-0.039008	-3.21056
	(161267.6)	(1.181995)	(0.225068)	(0.04323)	(1.401852)
wage_d302a9_st	21.54706	0.000274	-1.3e-05	1.3e-05	0.000474
	(23.58796)	(0.000122)	(3.7e-05)	(2e-06)	(0.000103)
N	17	17	17	17	17
r2	0.858	0.7202	0.6552	0.6736	0.8337

The column *class* shows whether the regression uses structural *mean* or variation *growth* dependent variable, as explained in section 3.3. The column *varia* shows the type of variable: *capacity* describes absolute system variables (e.g., length of rail tracks), *quality* describes safety measures (absolute or relative) (e.g., number of rail incidents by km of track). **Quality dependent variables:** *activecross* number of automatic devices to handle crossings in tracks, *injurpasskm* number of incidents that ended in passenger injury per track km. **Controls:** *pmr* market regulation score (OECD), *rgdpe_penn* GDP in PPP USD, *mean_rai_35* total system rail track length. Summary statistics of each variables are found in table 34. Heteroskedasticity − robust standard errors in parentheses.

Table 25: Rail Transport regression OLS estimates for governance *growth* dimension

var	activeleverkm	goodskm	accunpeop	derail	activecrosskm
indep_growth			-0.001022	0.028111	0.000887
			(0.000348)	(0.011596)	(0.000256)
discretion_growth	-0.000447	-0.064901			
	(0.000193)	(0.026216)			
dis_equ	0.001417	-0.059316			
	(0.000697)	(0.072476)			
gdp	0.002868	-0.063107	0.003378	0.142767	-0.003483
	(0.00319)	(0.106653)	(0.002896)	(0.07142)	(0.001876)
pmr	9.6e-05	-0.393697	0.003464	0.041635	-0.007797
	(0.008545)	(0.514458)	(0.008556)	(0.271844)	(0.005586)
nygdppcapkdzg			-0.025894	-0.490054	-0.013854
			(0.02175)	(0.402229)	(0.008646)
mean_rai_35	-4e-06	-1.2e-05	-4e-06	-0.000309	4e-06
	(4e-06)	(9e-05)	(3e-06)	(6.6e-05)	(2e-06)
ravar_mean_i14	-0.017249	0.632937			
	(0.003849)	(0.359531)			
rule_wb			-0.022118	-0.849578	-0.018733
			(0.028049)	(0.311254)	(0.012834)
valk_d49t53_st	0	0	0	4e-06	0
	(0)	(1e-06)	(0)	(1e-06)	(0)
N	19	19	19	19	18
r2	0.5966	0.576	0.6099	0.8942	0.6974

The column *class* shows whether the regression uses structural *mean* or variation *growth* dependent variable, as explained in section 3.3. The column *varia* shows the type of variable: *capacity* describes absolute system variables (e.g., length of rail tracks), *quality* describes safety measures (absolute or relative) (e.g., number of rail incidents by km of track). **Quality dependent variables:** *accunpeop* number of incidents that ended in the injure of unathorized persons in rails, *activecrosskm* number of automatic devices to handle crossings in tracks per track km, *accicross* number of incidents in road crossings, *accidempl* number of incidents that end in injure of rail employees. **Controls:** *pmr* market regulation score (OECD), *rgdpe_penn* GDP in PPP USD, *mean_rai_35* total system rail track length. Summary statistics of each variables are found in table 34. Heteroskedasticity − robust standard errors in parentheses.

Table 26: Rail Transport regression OLS estimates for scope *mean* dimensions

var	freight	delaymin	injurpasskm	econimpact
market_mean	2330.967	5205.946	0.008424	6789514
	(1002.304)	(1792.402)	(0.003042)	(2608213)
pmr	-5104.667	-10136.6	-0.014432	-5910992
	(5540.682)	(8446.741)	(0.014796)	(1.7e+07)
nygdppcapkdzg	-4355.181	-10239.59	-0.005882	-18800000
	(4537.028)	(9159.014)	(0.013645)	(15100000)
mean_rai_35	4.276011	6.41523	0	10574.54
	(1.298366)	(2.221519)	(4e-06)	(3765.725)
rgdpe_penn	-0.021797	-0.036666	0	-35.95551
	(0.011593)	(0.024325)	(0)	(39.01572)
rule_wb	1483.263	-5186.012	-0.037033	-440987.1
	(6675.819)	(9648.567)	(0.01718)	(16400000)
wage_d302a9_st	0.269584	-0.141679	2e-06	4008.124
	(0.502989)	(0.824602)	(2e-06)	(1386.85)
N	17	16	17	17
r2	0.8217	0.881	0.5584	0.8767

The column *class* shows whether the regression uses structural *mean* or variation *growth* dependent variable, as explained in section 3.3. The column *varia* shows the type of variable: *capacity* describes absolute system variables (e.g., length of rail tracks), *quality* describes safety measures (absolute or relative) (e.g., number of rail incidents by km of track). **Capacity dependent variables:** $freight$ volume of goods transported by rail system in metric tons. **Quality:** $delaymin$ minutes of delay product of rail incidents, $injurpasskm$ number of incidents that end in passenger injury per track km, $econimpact$ economic impact in USD of rail incidents, $fallback$ signals made in case of ATP communication failure, $passinjurtot$ number of incidents that end in passenger injury, $allaccidkm$ total number of incidents in the rail system per track km. **Controls:** pmr market regulation score (OECD), $rgdpe\_penn$ GDP in PPP USD, $mean\_rai\_35$ total system rail track length. Summary statistics of each variables are found in table 34. Heteroskedasticity − robust standard errors in parentheses.

Table 27: Rail Transport regression OLS estimates for scope *growth* dimensions

var	goodskm	tracklen	accicross	accidempl
market_growth	0.066559	55.28378		
	(0.02001)	(19.74639)		
transparency_growth			-0.102378	-0.084804
			(0.045412)	(0.025749)
dis_equ	-0.139743	-123.63		
	(0.067132)	(45.8665)		
gdp	-0.042071	-364.6775	0.41409	0.284582
	(0.091444)	(117.7816)	(0.311852)	(0.164491)
pmr	-0.726052	-1354.248	-0.259799	0.014267
	(0.526271)	(354.4423)	(0.873714)	(0.506825)
nygdppcapkdzg	0.289878	674.0361	0.185388	2.061711
	(0.38828)	(497.7444)	(1.236131)	(0.755293)
mean_rai_35	-7.2e-05	0.346434	-0.000469	-0.000404
	(0.000103)	(0.115206)	(0.000373)	(0.000183)
ravar_mean_c25			1.2e-05	8e-06
			(1.6e-05)	(8e-06)
rule_wb				
valk_d49t53_st	0	-0.004157	-2e-06	-1e-06
	(1e-06)	(0.000939)	(2e-06)	(1e-06)
N	19	19	19	19
r2	0.6546	0.791	0.544	0.7586

The column *class* shows whether the regression uses structural *mean* or variation *growth* dependent variable, as explained in section 3.3. The column *varia* shows the type of variable: *capacity* describes absolute system variables (e.g., length of rail tracks), *quality* describes safety measures (absolute or relative) (e.g., number of rail incidents by km of track). **Capacity dependent variables:** *goodskm* and *goodskm*.1 volume of goods transported by rail system in metric tons per km, *tracklen* rail system track length in km. **Quality:** *activeleverkm* number of automatic devices to handle crossings in tracks per km. **Controls:** *pmr* market regulation score (OECD), *rgdpe_penn* GDP in PPP USD, *mean_rai_35* total system rail track length. Summary statistics of each variables are found in table 34. Heteroskedasticity − robust standard errors in parentheses.

5.3.4 Air Transport

Table 28: Air regression OLS estimates for governance *mean* dimension

var	operincid	operincid.1	aipmain	airptot
indep_mean		-1.160263	-0.719815	-1.703685
		(0.39057)	(0.290233)	(0.691937)
discretion_mean	-0.927337			
	(0.224943)			
pmr	-12.54804	-8.012944	3.898383	9.385553
	(3.71466)	(2.774684)	(2.188433)	(5.327671)
airflights	1.4e-05	2e-06	-3e-06	-7e-06
	(6e-06)	(6e-06)	(6e-06)	(1.5e-05)
nygdppcapkdzg	12.98137	9.727678	-6.687607	-14.96657
	(4.072082)	(3.779362)	(2.53163)	(6.386715)
nights	-0.149415	-0.119082	0.08473	0.122043
	(0.040911)	(0.035958)	(0.018542)	(0.051814)
rgdpe_penn	1.8e-05	2.6e-05	1e-06	0
	(5e-06)	(6e-06)	(6e-06)	(1.3e-05)
valk_d51_st	-0.002007	-0.001673	0.000573	0.001374
	(0.000472)	(0.000466)	(0.000308)	(0.000762)
N	15	15	15	N
r2	0.8777	0.9297	0.844	r2

The column *class* shows whether the regression uses structural *mean* or variation *growth* dependent variable, as explained in section 3.3. The column *varia* shows the type of variable: *capacity* describes system variables (e.g., number of operating airports, number of arrivals), *quality* describes safety measures (absolute or relative) (e.g., number of air incidents). **Capacity dependent variables:** *aipmain* number of large commercial airports operating in a country, *airptot* number of total airports operating in a country. **Quality:** *operincid* number of incidents related to airline operations.**Controls:** *pmr*: product market regulation in air sector, $airflights$ average number of aircrafts in a country airspace (measure of traffic), *nygdppcapkdzg* GDP per capita growth rate, *nights* number of hotel nights spent by tourists in a given country, $rgdpe_penn$ country GDP in constant USD, $valk_d51_st$ air sector national accounts added value in constant USD. For additional control information refer to table 35. Heteroskedasticity − robust standard errors in parentheses.

Table 29: Air regression OLS estimates for governance *growth* dimension

var	totdisrup	numairpo
indep_growth		-0.169532
		(0.026562)
discretion_growth	3908.336	
	(821.3871)	
gdp	19650.03	0.138654
	(9723.635)	(0.136115)
pmr	2426.411	0.074127
	(24733.81)	(0.813815)
airflights	0.08967	-1e-05
	(0.098674)	(2e-06)
nights	-1630.702	0.004899
	(532.9435)	(0.008166)
rule_wb	59430.88	-3.672242
	(56711.66)	(1.02675)
N	16	16
r2	0.855	0.9138

The column *class* shows whether the regression uses structural *mean* or variation *growth* dependent variable, as explained in section 3.3. The column *varia* shows the type of variable: *capacity* describes system variables (e.g., number of operating airports, number of arrivals), *quality* describes safety measures (absolute or relative) (e.g., number of air incidents). **Capacity dependent variables:** *numairpo* number of total airports operating in a country, *arrivals* number of year arrivals in all airports countrywide, *numplanes* number of operating aircrafts registered in a country. **Quality:** *capacinci* number of incidents related to airport land operations. **Controls:** *pmr*: product market regulation in air sector, $airflights$ average number of aircrafts in a country airspace (measure of traffic), *nygdppcapkdzg* GDP per capita growth rate, *nights* number of hotel nights spent by tourists in a given country, $rgdpe_penn$ country GDP in constant USD, $valk_d51_st$ air sector national accounts added value in constant USD. For additional control information refer to table 35. Heteroskedasticity − robust standard errors parentheses.

Table 30: Air regression OLS estimates for scope *mean* dimension

var	air250	air150
market_mean	4.476154	4.994706
	(0.842899)	(1.052329)
pmr	-21.06812	-19.36173
	(4.852095)	(6.71555)
airflights	4e-05	4.2e-05
	(1.1e-05)	(1.9e-05)
nygdppcapkdzg	24.54264	25.64561
	(6.248572)	(10.19166)
nights	-0.152139	-0.289778
	(0.067372)	(0.083071)
rgdpe_penn	3.4e-05	4.4e-05
	(1e-05)	(1.3e-05)
valk_d51_st	-0.002519	-0.00191
	(0.000905)	(0.001087)
N	15	15
r2	0.9793	0.9387

The column *class* shows whether the regression uses structural *mean* or variation *growth* dependent variable, as explained in section 3.3. The column *varia* shows the type of variable: *capacity* describes system variables (e.g., number of operating airports, number of arrivals), *quality* describes safety measures (absolute or relative) (e.g., number of air incidents). **Capacity dependent variables:** $air150$ number of aircrafts with capacity below or equal to 150 passengers, $air250$ number of aircrafts with capacity from 150 to 250 passengers.**Quality:** *operincind* number of incidents related to airline operations. **Controls:** *pmr*: product market regulation in air sector, $airflights$ average number of aircrafts in a country airspace (measure of traffic), $nygdppcapkdzg$ GDP per capita growth rate, $nights$ number of hotel nights spent by tourists in a given country, $rgdpe_penn$ country GDP in constant USD, $valk_d51_st$ air sector national accounts added value in constant USD. For additional control information refer to table 35. Heteroskedasticity − robust standard errors parentheses.

Table 31: Air regression OLS estimates for scope *growth* dimension

var	totdisrup	capacinci	arrivals	numplanes
market_growth				
transparency_growth	151.2698	-283.8258	664.3851	0.913646
	(15.96126)	(110.0897)	(168.7554)	(0.198186)
gdp	-815.3353	558.1283	-1180.638	-0.526386
	(338.9403)	(2153.295)	(1920.907)	(3.496685)
pmr	-6434.93	17829.38	-4591.767	-14.11368
	(2170.754)	(9382.646)	(8052.042)	(12.20801)
airflights	0.019647	0.017727	0.010579	3e-06
	(0.003277)	(0.0193)	(0.017361)	(2.7e-05)
nygdppcapkdzg	10296.56	-21295.77	1999.503	16.28361
	(2373.485)	(10852.65)	(9398.181)	(13.75266)
nights	17.59089	-111.5151	100.6455	-0.068853
	(19.93378)	(120.4289)	(120.4423)	(0.176202)
rule_wb				
valk_d51_st	-1.16274	2.037845	-0.666831	-0.00038
	(0.277474)	(1.195069)	(1.094078)	(0.001525)
N	14	14	14	14
r2	0.9587	0.8023	0.7896	0.8197

The column *class* shows whether the regression uses structural *mean* or variation *growth* dependent variable, as explained in section 3.3. The column *varia* shows the type of variable: *capacity* describes system variables (e.g., number of operating airports, number of arrivals), *quality* describes safety measures (absolute or relative) (e.g., number of air incidents). **Quality:** *totdisrup* total number of disruptions of any kind in air operations. **Controls:** *pmr*: product market regulation in air sector, $airflights$ average number of aircrafts in a country airspace (measure of traffic), *nygdppcapkdzg* GDP per capita growth rate, *nights* number of hotel nights spent by tourists in a given country, $rgdpe_penn$ country GDP in constant USD, $valk_d51_st$ air sector national accounts added value in constant USD. For additional control information refer to table 35. Heteroskedasticity − robust standard errors parentheses.

5.3.5 Summary of main variables

Table 32: Energy sector main dependent and independent variables

	count	mean	sd	min	max
energy_transparency_growth	24	23.3908	31.30449	-38.95349	96.22093
energy_indep_growth	24	12.8737	24.65712	-18.26211	97.66924
energy_market_growth	24	1.07955	14.07552	-27.45455	27.27273
energy_bureau_growth	24	-3.189624	27.72122	-43.71765	62.7907
lev_enenew13	21	224.68	449.4059	-299	1574.83
capOTHCOMB_TOTAL	20	230.0686	583.2736	-56.08525	2532.65
pro_TOTPRO_Impor_b	24	1273.205	2448.952	-3136.214	8580.164
mean_enenew47	21	-.0476191	.8576823	-2	1.6
ELE_IND	23	-224.644	182.8657	-582.0122	73.83592
capINTCOMB_TOTAL	20	3.419913	12.6303	-17.99721	34.33411
lev_enenew26	21	.0657143	.04556	0	.17
mean_enenew30	21	.0285714	.0293744	-.01	.11
pro_TOTPRO_Impor_b	23	46.07296	230.8647	-784.3334	501.5268
lev_enenew38	21	54.26525	389.5657	-871.71	575.46
lev_enenew44	18	-6.893334	10.41312	-43.92	2.260002
mean_enenew48	20	.32	1.070956	-1.6	3.2
lev_capOTHCOMB_TOTAL	20	174.7744	515.8631	-124	2313
lev_capSOLARPV_TOTAL	23	1156.692	2319.613	-5	9839
lev_capCOMBINED_TOTAL	22	-117.5203	583.0404	-1985.49	846
mean_enenew9	21	1.795238	2.194315	.0200001	7.860001
lev_enenew49	21	-52.31001	230.2	-585.8101	317.76
mean_enenew52	21	41.28333	80.47345	-2.199997	292.4
lev_ELE_HOU	23	-218.1411	284.4144	-674.1957	562.2061
gdp	23	8.336106	10.37013	.1553978	38.27657
lag_pmr_elec	23	-2.082174	.6799128	-3.23	-.87
dis_equ	23	50.75279	7.476148	39.16258	67.46999
capTOTCAP_MAINTOT	24	38635.76	49004.22	1682.015	195559.5
ELE_IND	23	1297.618	359.5961	572.0584	2233.161
m_nygdppcapkdzg	23	2.335652	1.926905	.82	9.62
rest_m_gctaxtotlgdzs	23	.538261	1.537731	-3.719999	3.59

mean: Dependent variables are the (5-year) average change between 2009-2013 to 2014-2018. Variables with no *mean* specification stand for mean values corrected by country population size. *lev*: Dependent variable change between 2013 and 2018. $capOTHCOMB\_TOTAL$: CO2 electricity generation capacity based on technologies other than internal combustion in MW, $pro\_TOTPRO\_Impor\_b$: yearly energy imports in MW from third-countries, *enenew*47: number of electricity producers that cover more than 5% of the industry supply, ELE_IND: industrial consumer retail prices in USD, *enenew*26: wind and solar electricity generation capacity in MW, *enenew*30: Wind generation efficiency conversion in MW, *enenew*38:total electricity generated per capita in MW, *enenew*44 market share of largest electricity producer in percentage, $capSOLARPV\_TOTAL$: solar photovoltaic electricity generation capacity in MW, $capCOMBINED\_TOTAL$: electricity generation capacity based on gas and steam (fuel efficient) in MW, *enenew*9: percentage of renewable electricity generation production, *enenew*49: electricity consumption per capita in MW. *enenew*52: number of electricity retailers, ELE_HOU: household consumer prices in USD, *gdp*: Gross domestic product PPP base (), dis_equ: absolute latitude value, lag_rail_pmr: rail product market regulation index (OECD) for 2013, $capTOTCAP\_MAINTOT$: total electricity generation capacity in MW, $gctaxtotlgdzs$: tax revenue as a percentage of GDP.

Table 33: E-Communication sector main dependent and independent variables

	count	mean	sd	min	max
telecom_transparency_growth	22	18.52249	23.81705	-16.30252	74.37908
telecom_indep_growth	22	22.36652	29.36359	-14.8087	100.6742
telecom_market_growth	22	3.847192	12.48907	-11.74056	35.36173
telecom_discretion_growth	22	-.5082586	20.65252	-31.57895	41.17647
lev_bbwp100tot	23	33.74087	12.45864	7.240005	58.14
mean_bbp100cab	24	1.827083	1.298472	0	4.329999
mean_telaccpath100	24	1.934584	8.260808	-10.35001	15.45
level_mob_bb_penrate	23	34.35	11.91868	7.65	57.75
lev_fixbbas5gbtelpripc	23	.2030435	.3896669	-.8599999	.8900001
lev_bbp100cab	23	1.344348	1.273206	-.3999996	3.87
voip	20	1642.716	3239.391	-160.8322	13876
lev_fixbbas5gbtelpripppp	23	12.58043	11.25792	-7	38.45
mean_bbwp100stan	24	40.02333	13.87384	19.93	88.89
totmob	24	2143.423	5951.085	-8687.292	21733
mean_mob_bb_penrate	23	40.20696	14.55367	24.72	82.6
lev_fixbbas5gbtelpripppp	23	12.58043	11.25792	-7	38.45
gdp	23	8.336106	10.37013	.1553978	38.27657
rule_wb	24	1.363833	.5874164	.194	2.052
lag_telecom_pmr	23	.9534783	.619337	.27	2.66
fix_broad_100_Total	22	35.29864	6.799405	20.2	46.78
le_enpopdnst	23	151.0174	119.2623	14.55	511.48
rest_m_spurbtotlinzs	23	1.025217	.8991659	-.6599998	2.920002
nights	23	129	158.5514	2.9	471.2

mean: Dependent variables are the (5-year) average change between 2009-2013 to 2014-2018. Variables with no *mean* specification stand for mean values corrected by country population size. *lev*: Dependent variable change between 2013 and 2018. $bbwp100tot$: total number of subscriptions of fix internet broadband connection by 100 people, $bbp100cab$: total number of subscriptions of cable internet connection by 100 people, $telaccpath100$: total number of access paths by 100 people, $mob\_bb\_penrate$: mobile communication services penetration rates in percentage, $fixbbas5gbtelpripc$: estimated price of a fix internet connection (5gb) per capita in USD, *voip*: total number of VoIP subscriptions, $fixbbas5gbtelprippp$: estimated price of a fix internet connection (5gb) USD corrected for purchase parity, $bbwp100stan$: total number of mobile broadband subscribers per 100 habitants, *totmob*: change in total number of mobile services subscriptions in thousands, $fix\_broad\_100\_Total$: total number of internet subscriptions per 100 habitants, $le\_enpopdnst$: population density per square km, $rest\_m\_spurbtotlinzs$: percentage of pupulation living in urban areas. *gdp*: Gross domestic product PPP base (), dis_equ: absolute latitude value, lag_rail_pmr: rail product market regulation index (OECD) for 2013.

Table 34: Rail sector main dependent and independent variables

	count	mean	sd	min	max
rail_transparency_growth	21	28.97117	26.93066	-23.61546	73.44538
rail_indep_growth	21	25.00061	21.41624	-8.766803	60.57835
rail_market_growth	21	6.409455	20.89313	-27.16279	48.23529
rail_discretion_growth	21	-.2110982	22.82539	-34.95798	41.62791
mean_ss00	22	-.8363636	2.483957	-8	2.6
lev_ps24	22	-.0186364	.0362919	-.15	.01
mean_n03	22	-3.418182	3.939984	-17.2	.4000015
mean_tgoodsrltot	22	1229.325	2983.439	-3013.2	12971.59
mean_c25	22	-33717.45	76280.69	-315841.8	0
lev_t16	20	-.0095	.0308605	-.1	.04
mean_n02	22	-.9818181	2.902156	-9.4	4.2
mean_n21	22	.0045455	.0147122	-.03	.04
mean_us10	22	-.0181818	.05679	-.17	.12
lev_n07	22	-8.727273	24.85612	-102	34
lev_n10	22	.0036364	.2954504	-.4399999	.9300001
lev_r03	22	528.5621	2220.278	-2170	9013.299
mean_i14	22	.2718182	1.20712	-.6999999	5.5
mean_r06	22	-.4872721	2.096269	-5.790001	3.43
mean_t29	22	.0104545	.0450901	-.1	.16
lev_n10	22	.0036364	.2954504	-.4399999	.9300001
lev_tk10	22	-.0281818	.2085468	-.45	.7399999
mean_i14	22	.2718182	1.20712	-.6999999	5.5
mean_r06	22	-.4872721	2.096269	-5.790001	3.43
mean_t29	22	.0104545	.0450901	-.1	.16
mean_rai_35	22	-34.42509	580.8192	-1813.9	1582.5
gdp	46	7.949646	9.77064	.1322581	38.27657
dis_equ	46	50.75279	7.392613	39.16258	67.46999
VALK_D49T53_ST	48	114329.5	281322.3	1136.157	1477314
lag_rail_pmr	23	3.273478	1.081122	.25	5.41

mean: Dependent variables are the (5-year) average change between 2009-2013 to 2014-2018. *lev*: Dependent variable change between 2013 and 2018. $ss00$: total operational accidents that involved rail employees, $ps24$: Total accidents that involve passengers, $n03$: total accidents at lever crossings, $tgoodsrltot$: total freight traffic, $c25$: estimated costs of delays due to operational incidents, $t16$: active lever crossings per line km., $n02$: train derail incidents, $n21$: total number of accidents that involve the transport of dangerous goods, $us10$: total number of accidents that involve unauthorized personnel, $n07$: total number of suicides in railways, $n10$: total number of incidents per line km, $r03$: rail line lenght in km, $i14$: accidents precursors before incidents per line km, $r06$: freight transport in tons per line km, $t29$: total lever crossings per line km, $tk10$: total number of faltal victims involved in rail incidents, rai_35: total line length in km. *gdp*: Gross domestic product PPP base (), *dis_equ*: absolute latitude value, lag_rail_pmr: rail product market regulation index (OECD) for 2013, $VALKD49T53ST$ transport industry added value.

Table 35: Air sector main dependent and independent variables

	count	mean	sd	min	max
air_transparency_growth	16	23.92396	25.39048	-12.79904	94.35216
air_indep_growth	16	13.12934	26.51194	-26.35372	57.81898
air_market_growth	16	-2.926813	22.99693	-37.2093	53.84615
air_discretion_growth	16	8.115754	43.9825	-43.97727	128.9157
mean_air_5	23	-26.7913	35.52737	-98.4	28.6
mean_airifrarrivalsnm	22	18586.86	21216.43	-1489.5	80932
mean_airercapacityatc	21	11968.62	205896.8	-620901.6	618735.1
mean_airerdisruptions	21	3551.462	11105.15	-10175.3	42600.8
mean_air_19	23	-4.41087	7.300762	-28.6	2
mean_air_14	23	-6.763043	8.828649	-30	3.5
lev_airsmindic	22	11.40909	7.048533	0	33
mean_air_20	23	.0608696	1.141864	-2.85	2.8
mean_airercapacityatc	21	11968.62	205896.8	-620901.6	618735.1
gdp	46	7.949646	9.77064	.1322581	38.27657
airflights	44	1007015	834804	62017	3257894
dis_equ	46	50.75279	7.392613	39.16258	67.46999
lag_air_pmr	23	.9291304	1.226169	0	3.55
nights	46	129	156.7798	2.9	471.2

The dependent variables are the (5-year) average change between 2009-2013 to 2014-2018. $mean\_air\_5$: total number of aircrafts, $mean\_airifrarrivalsnm$: total number of arrivals, $mean\_airercapacityatc$: total number of operational disruptions per year, $mean\_airerdisruptions$ total number of disruptions per year, $mean\_air19$: total number of main airports (more than 25000 passengers), $mean\_air14$: number of small size aircrafts, $levairsmindic$: safety performance index, $mean\_air20$: total number of airports, gdp: Gross domestic product PPP base (), $airflights$ total number of flights, dis_equ: absolute latitude value, lag_air_pmr: air product market regulation (OECD) for 2013, $nights$ number of nights a year tourist spend in a hotel on a given country.

References

Balleisen, E. J. (2015). The dialectics of modern regulatory governance. In *Business Regulation*. Edward Elgar Publishing Limited.

Bandiera, O., Hansen, S., Prat, A., Sadun, R., Bennedsen, M., Burgess, R., Dessein, W., Gibbons, B., Henderson, R., Hermalin, B., Ingram, P., Khandelwal, A., Limodio, N., Mcmahon, M., Schoar, A., Scur, D., and Tadelis, S. (2017). Nber Working Paper Series Ceo Behavior and Firm Performance.

Blei, D. M., Ng, A. Y., and Jordan, M. I. (2003). Latent Dirichlet allocation. *Journal of Machine Learning Research*, 3(4-5):993–1022.

Bortolotti, B., Cambini, C., Rondi, L., and Spiegel, Y. (2011). Capital Structure and Regulation: Do Ownership and Regulatory Independence Matter? *Journal of Economics and Management Strategy*, 20(2):517–564.

Cambini, C. and Rondi, L. (2012). Capital structure and investment in regulated network utilities: Evidence from EU telecoms. *Industrial and Corporate Change*, 21(1):31–71.

Casullo, L., Durand, A., and Cavassini, F. (2018). The 2018 Indicators on the Governance of Sector Regulators - Part of the Product Market Regulation (PMR) Survey.

Edwards, G. and Waverman, L. (2006). The effects of public ownership and regulatory independence on regulatory outcomes : AA study of interconnect rates in EU telecommunications. *Journal of Regulatory Economics*, 29(1):23–67.

Égert, B. (2009). Infrastructure Investment in Network Industries : The Role of Incentive Regulation and Regulatory Independence. *CESifo Working Paper*, (956):27.

ERA (2017). Railway Safety in the European Union: Safety overview 2017. Technical report, European Union Agency for Railways, Luxembourg.

ERA (2018). 2018 assessment of achievements of safety targets. Technical report, European Union Agency for Railways.

EUROCONTROL (2013). Challenges of Growth 2013 Summary Report Network Manager nominated by the European Commission. Technical report, EUROCONTROL.

EUROSTAT (2020). Air transport statistics - Statistics Explained. Technical Report June, Eurostat.

Gilardi, F. (2002). Policy credibility and delegation to independent regulatory agencies: A comparative empirical analysis. *Journal of European Public Policy*, 9(6):873–893.

Gilardi, F. (2005). The formal independence of regulators: A comparison of 17 countries and 7 sectors. *Swiss Political Science Review*, 11(4):139–167.

Griffiths, T. L., Steyvers, M., Blei, B., and Blei, J. (2004). Finding scientific topics A first step in identifying the content of a document is determining which topics that document addresses. We describe a generative model for documents, introduced. *Pnas*, 101(SUPPL.1):5228–5235.

Hanretty, C. and Koop, C. (2012). Measuring the formal independence of regulatory agencies. *Journal of European Public Policy*, 19(2):198–216.

Jones, T. (2019). Topic modeling.

Jordana, J., Fernández-i Marín, X., and Bianculli, A. C. (2018). Agency proliferation and the globalization of the regulatory state: Introducing a data set on the institutional features of regulatory agencies. *Regulation and Governance*, 12(4):524–540.

Koop, C. and Hanretty, C. (2018). Political Independence, Accountability, and the Quality of Regulatory Decision-Making. *Comparative Political Studies*, 51(1):38–75.

OECD (2016). The governance of regulators: being an independent regulator. Technical report, Organisation for Economic Co-operation and Development (OECD).

Sutherland, D., Araújo, S., Égert, B., Koźluk, T., Elmeskov, J., Schneider, J.-L., Nicoletti, G., Schiantarelli, F., Berlingieri, G., Spiezia, V., Garcilazo, J. E., and Sinha, I. (2011). Public Policies and Investment in Network Infrastructure.

Trillas, F. and Montoya, M. A. (2013a). Independent regulators: theory, evidence and reform proposals. *info*, 15(3):39–53.

Trillas, F. and Montoya, M. A. (2013b). Independent regulators: theory, evidence and reform proposals. *Info*.