Available online at www.sciencedirect.com
ScienceDirect
Journal of Policy Modeling xxx (xxxx) xxx–xxx
Journal of Policy Modeling
www.elsevier.com/locate/jpm

Drafting "better regulation": The economic cost of regulatory complexity☆

Juan de Lucio *,[1], Juan S. Mora-Sanguinetti

Universidad de Alcalá and Banco de España-Eurosystem, Spain

Received 12 May 2021; received in revised form 5 September 2021; accepted 20 October 2021

Abstract

Different public agencies are seeking to draft "better regulation". Complex or poorly drafted norms are more difficult for economic agents to implement, tending to erode economic efficiency. The literature has so far concentrated on the analysis of regulatory complexity as a phenomenon related to the "quantity" of norms. This article guides the process of adopting new regulations, taking into account that norms can also be complex due to new "qualitative" reasons such as linguistic ambiguity or relational structure (references between legal documents). To perform the analysis, we develop new indicators for legibility and regulatory interconnectedness. Specifically, we construct a new database (RECOS – REgulation COmplexity in Spain) by extracting information from 8171 norms (61 million words) which comprise the regulations of all the Spanish Autonomous regions. Our analysis reveals the relationship between measures of "qualitative" complexity and relevant economic (productivity) and institutional (judicial efficacy) variables. This research

☆ The views expressed are those of the authors and should not be attributed to the Banco de España or the Eurosystem. We would like to thank the editor and the referees of this journal for their valuable suggestions and comments. We are also grateful for the suggestions of the participants in the seminars of the Banco de España – Eurosystem, the OECD, the Bank of England, the 8th International meeting in Law & Economics, Université Paris-Nanterre, the GCD-University of A Coruña, the FIDE Foundation and Universidad Autónoma de Madrid. We also thank the comments and suggestions of an anonymous reviewer who evaluated the working paper version of this article (published as DT 2107 of the Banco de España-Eurosystem).

* Corresponding author.
E-mail address: Juan.deLucio@UAH.es (J. de Lucio).

[1] Juan de Lucio acknowledges the financial support from Comunidad de Madrid & UAH (ref: EPU-INV/2020/006) and Comunidad de Madrid (H2019/HUM-5761).

https://doi.org/10.1016/j.jpolmod.2021.10.003

Please cite this article as: J. de Lucio and J.S Mora-Sanguinetti, Drafting "better regulation": The economic cost of regulatory complexity, *Journal of Policy Modeling*, https://doi.org/10.1016/j.jpolmod.2021.10.003

shows that the new dimensions of regulatory complexity matter, yield significant results and should be taken into account in governments' "better regulation" policies.

JEL classification: O43; K2; R11; O47

Keywords: Regulatory complexity; Productivity; Linguistic complexity; Relational complexity; Legal corpus

1. Introduction

Since the 1990s, different national administrations and international organizations have been pursuing various initiatives aimed at achieving "better regulation". The aim is to draft regulations that are more efficient and less "complex" to use both for private agents (citizens and companies) and for public agencies themselves. "Better regulation" policies save money and time, and therefore have better economic features.

It is worth mentioning, among others, the "plainlanguage.gov" initiative in the USA launched in 1994 and promoted again in 2004, the UK "Good Law" principles of 2013 [see also, Office of the Parliamentary Counsel (Cabinet Office) (2013)], the OECD recommendations of 2012 (OECD, 2012), or different initiatives of the European Union, which have been transmitted to national regulations. In the case of Spain, it is worth going back to the 1997 Government Act and highlighting the recent Royal Decree 931/2017 (Mora-Sanguinetti, 2019a).

Many of these initiatives seek to get legislators to reflect on whether drafting a new regulation is really necessary and encourage policy makers to draft the new norms as clearly as possible. However, in order to make the drafting process really efficient and have sufficient support, there is still a lack of more academic work reflecting on the circumstances that make the regulation have negative impacts. More work is also needed to be able to objectify and quantify the "complexity" of regulation. This research thus rationalizes the efforts of public administrations to achieve "better regulation".

As an introduction, it is first necessary to reflect on the nature and effects of regulation in general. In this sense, it can be stated that regulation, as part of the institutional framework, is fundamental to economic development[2] . If it is well designed, it can help to mitigate market failures, such as imperfect information or externalities, and can generally reduce transaction costs[3] .

However, regulation is often perceived in a negative light, because if norms are poorly drafted, they could lead to increased transaction costs, which impact negatively on economic efficiency. Hence the "better regulation" initiatives are aimed at reducing the negative impacts of regulation. The analysis and quantification of these drafting or design problems are precisely the objective

[2] Nedic et al. (2020) show that excessive and complicated regulation does not favor the quality of institutions and find a significant and positive impact from Regulatory Quality on economic growth. In another vein, Costa (2010) claims that greater regulation is essential to solve crime

[3] The study of classical microeconomic theory, specifically the welfare theorems, leads to the same conclusion, as market failures violate the first theorem (see, among others, Mora-Sanguinetti and Salvador-Mora, 2016; Mora-Sanguinetti, 2019a).

of this article. In general, Laffont and Tirole (1993) argue that market failures would be a necessary but not sufficient condition for resorting to regulation since its effects are conditioned by transactional, administrative-political or informational reasons (see also Mora-Sanguinetti and Pérez-Valls, 2021). Specifically, as long as the regulation is not adequately designed or drafted, it could involve both direct and indirect economic costs. The former refers to resources devoted to the compliance of regulations. Indirect costs are associated with changes in agents' behavior (firms, consumers or government). Both direct and indirect costs, may derive in a misallocation of resources. In fact, the economic literature, and more specifically that which analyses the impacts of institutional quality (Acemoglu and Robinson, 2012), makes several references to the costs derived from administrative burdens (red tape) (Hampton, 2005), which are the main motivation for "better regulation" policies (Radaelli, 2007)[4] .

Several studies have also unveiled specific examples of some of the costs involved in regulation: Djankov et al. (2006) shows that improving business regulations, from the worst quartile to the best, increases the annual growth rate by 2.3 percentage points. Coffey et al. (2020) identify sectors affected by regulations and conclude that U.S. economic growth has been dampened by 0.8% per annum due to federal regulations. Bailey and Thomas (2017) show that more intensively regulated industries experience lower enterprise birth rates and slower employment growth. Chambers et al. (2019b) conclude that, in the case of the U.S., a 10% increase in the effective federal regulatory burden would increase the poverty rate by 2.5%. Furthermore, some specific areas of regulation have been intensively analyzed in the literature. One of them is retail trade regulation, which can exert negative impacts on sectoral employment [Bertrand and Kramarz (2002) or Viviano (2008)].

In addition to the potential economic costs, other relevant costs for the public administration also exist, such as judicial costs. The OECD (Palumbo et al., 2013) found that a low regulatory quality, as measured by the World Bank's regulatory quality indicators, was related to more litigation in the countries analyzed. It should be noted that the increase in litigation has a relevant impact on the functioning of the judicial system, implying longer trial durations. All this is important for the activity of lawyers and legal services professionals (Mora-Sanguinetti, 2019b) and has, indirectly, important economic impacts.

1.1. The three dimensions of regulatory "complexity"

One of the reasons why regulation may prove inefficient and therefore a source of concern for "better regulation" initiatives, is due to its "complexity". The concept of complexity refers to problems regarding the "form" rather than the specific topics covered by regulation. A corpus of regulations can in a first instance, be complex because it is too broad (the "quantity" approach), i.e., there is an excessive volume of regulations or the latter proceed from a variety of different sources, making it difficult for economic agents to manage them and verify their validity (Bardhan 2002; Di Vita, 2018). Secondly, norms may be "qualitatively complex" because they are ambiguously or poorly drafted (the "linguistic" approach), complicating the understanding and compliance of consumers and businesses. Additionally, complexity can derive from how rules are connected to each other (the "relational" approach). Norms drafted with more references to other norms demand more resources to be understood and enforced. Fig. 1 refers to these three sources of complexity.

[4] As Gai et al. (2019) emphasize some level of regulation complexity may be necessary.

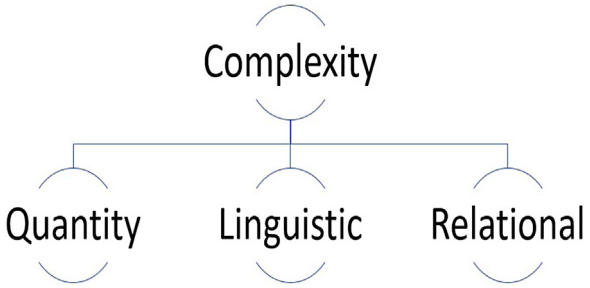

Fig. 1. The three dimensions of regulatory complexity.
Source: Own elaboration.

The first approach to complexity, the volume of norms, has already been studied by some articles in the literature, relying on indicators which measure the number of legal texts. Using this type of indicator, Di Vita (2018) shows that regulatory complexity hinders regional growth in Italy. Rios and Gianmoena (2020) find that government "quality", seems to be one of the most robust drivers of resilience at the regional level. The quality of regulation is a fundamental part of such indicator of government quality. Consequently, the regulatory burden seems to be a relevant proxy for the distortions to which well-functioning economic structures are subjected. Kirchner (2012) demonstrates that at a national level (Australia), the growth in legislation, as measured by the number of pages is, in the short run, negatively related to growth in real income per capita. Mora-Sanguinetti and Pérez-Valls (2021) show that regulation is negatively related to the number of firms in Spain and may have effects on the size distribution of Spanish firms, contributing towards an excessive share of SMEs. This article continues the latter line of research by accounting for words and sentences.

The study of the last two approaches (complexity derived from linguistic or relational problems) has been scarce in the economic literature to date and, therefore, it has not been possible to construct comprehensive recommendations for economic policy based on these measures. As such, the research aimed at constructing measures of complexity in these new dimensions. For this task, we use natural language processing (NLP) techniques in order to measure said dimensions. Moreover, these tools are seeing extensive application in the literature. They have been shown to be useful in measuring economic uncertainty (see, Azqueta-Gavaldón et al., 2020; Ghirelli et al., 2019; Baker et al. 2016). Specifically, authors in the field of financial markets have applied them: Tobback et al. (2017) use NLP techniques to analyze the media's perception of the tone of the ECB's monetary policy discourse; Ehrmann and Talmi (2020) focus on the semantic similarity of subsequent central bank statements in order to prove that volatility rises when changes occur after sequences of similar statements; Hansen et al. (2017) use them to detect issues in Federal Open Market Committee statement[5] ; Hassan et al. (2019) analyze political risk using earnings conference calls, and Calomiris et al. (2020) derived a measure of firm-level regulatory costs from corporate earnings conference calls.

[5] There is an increasing interest in analyzing the drafting and language complexity of central bank statements. Hernández-Murillo and Shell (2014) and Coenen et al. (2017), show that statements have become more complex since the financial crisis. Jansen (2011) shows that more complex statements are associated with higher volatility. Haldane and McMahon (2018) focus on the simplification of central bank communications.

To the best of our knowledge, RegData is the only database that measures regulations quantitively using NLP techniques (see Al-Ubaydli and McLaughlin, 2017). The authors compiled industry-specific federal regulations for the US during the period 1997–2012 (see McLaughlin and Sherouse, 2019, for an update of the database). Davis (2017) uses it to relate regulatory complexity with policy uncertainty, whilst Chambers et al. (2019a) explore the link between regulation and prices. In turn, Coffey et al. (2020) analyzes the relationship between regulation and economic growth and Bailey and Thomas (2017) study industry regulation and how it affects enterprise birth and employment growth. Version 3.2 of the database was released in March 2020, covering the period 1970–2019 and includes, for the first time, measures relating to the complexity of regulatory texts applying a similar methodology to the one developed in this paper, in the form of the Shannon Entropy legibility indicator.

Finally, our analysis is also related to Hurka and Haag (2019) which was the first article to use linguistic and structural-relational complexity measures based on NLP. They show that the length of the decision-making process in the EU is influenced by different types of policy complexity.

Our article develops and presents a new database (RECOS, REgulatory COmplexity in Spain) covering the laws enacted by the regional governments over the whole Spanish democratic period (from 1978 until 2019). As far as we know this is the first effort to build a database using NLP techniques on regional regulations outside the USA. RECOS incorporates new legibility and relational indicators as well as the traditional volume indicators. As explained in Sections 2 and 4, some of these measures of legal complexity are new in the literature. With our new indicators (on regulation readability and interdependency) in hand, we explore the effects that regulatory complexity has on economic activity (productivity) and judicial efficacy (as representative of a relevant institutional dimension) using panel data analysis. All of this serves as the basis for the regulatory policy recommendations developed below.

The study of the Spanish case at the regional level is of particular interest for several reasons: First, it allows us to develop comparative indicators (by region) of regulatory complexity (and analyze its economic impact) without having to compare different languages. This latter problem presents a risk to other potential studies that wish to compare complexity in countries with different languages. Although some regions have their own regional language (in addition to Spanish), we have obtained the official Spanish version of norms. This mitigates comparison problems when using text analysis techniques. Second, Spanish regions also provide an adequate framework for comparison given that they all share the same basic institutional background: on the one hand, the Spanish Constitution establishes a common framework of competences (without denying specificities)[6] . This implies that the set of topics and problems under regulation are more comparable between Autonomous regions than between countries. On the other hand, the international norms affecting Spain, similarly affect all the regions. Third, with regards to the analysis of economic impacts, specifically judicial data, the study of Spain is unique at an international level because it maintains richer judicial databases than many other jurisdictions. Fourth, regarding the construction of robustness checks, Spanish is a highly regulated language, with an official dictionary (developed by the Real Academia Española - RAE), which facilitates the development of objective measures for infrequent words. Finally, beyond the studies mentioned, the case of Spain, has not yet been widely analyzed even though it is one the largest economies of the Euro area.

[6] See, for instance, articles 148 and 149 of the Spanish Constitution.

We will show that the effects of the different complexity measurements are robust to different specifications and seem to be independent between each other. These results suggest that the linguistic and relational complexity dimensions introduced in this document offer important perspectives regarding regulatory complexity, which affect different aspects of economic performance. As such, they prove even more relevant than the traditional measures focused on the quantity of regulation. Public administration policies aimed at achieving "better regulation" must take into account all these dimensions.

The remainder of the article proceeds as follows. Section 2 presents the database and the new indicators used in the analysis. Section 3 relates our indicators of regulatory complexity with measurements of productivity growth and judicial efficacy at the regional level using panel data analysis. Section 4 presents several robustness checks. Section 5 focuses on the main messages of policy advice. The final section presents the main conclusions and future research avenues. The article is completed by an appendix which includes further detailed information.

2. Measuring regulatory complexity

To support our analysis and policy advice, our research analyzes all the regional Laws and Decree-Laws published in the "Official State Bulletin" (*Boletín Oficial del Estado*, BOE) over the democratic period (from 1978 to 2019)[7] . In order to do so, we performed web data extraction (web scrapping) from *BOE.es*. We thus collected a set of regional regulations with the force of Law[8] summing up 8171 norms and 61 million words[9] . We analyze each regulation within the corpus using NLP techniques.

2.1. The quantitative approach: volume of regulations

Fig. 2 shows the total number of norms (second column), the total number of words (third column) and the total number of sentences of said norms (fourth column) for each region. It also graphs the evolution of those measurements during the period of analysis. The markers (small rhombuses) of each series in Fig. 2 highlight the minimum and the maximum level. Usually, the maximum takes place in the most recent years; there is a generalized upward trend in the volume of regulations incorporated each year. We observe that some regions have adopted 3 times more norms than others. For instance, Navarra (NAV) has adopted more than nine hundred regulations whilst La Rioja (RIO) has less than three hundred. Andalusia (AND) norms contain almost 7

[7] With the 1978 Constitution, Spain adopted a decentralized model of territorial and political organization. The regions (Autonomous Communities), have the power to adopt regulations with the force of law, and the Constitution allows them to have a very broad competence ceiling. In strict legal terms, the system is not called "federal" because, among other reasons, access to autonomy is not compulsory, even though all the regions have it [see, among others, López Guerra (1994) or López Guerra et al. (2018)]. Overall, the system has largely converged with the "classical" federal model and is, according to various international indicators, as decentralized as the federal ones (in the Regional Authority Index, RAI, Spain's score would be comparable to that of Canada or the USA and is higher than that of the UK; see Hooghe et al. 2008).

[8] Regulations with the force of law are part of the regulations adopted by the regional administrations, which can also pass lower-level regulations. Taking Aranzadi as a general source [see Mora-Sanguinetti (2019a)], only in 2019 the regions adopted 7970 norms in total, proving to be the level of Spanish administration that adopts more rules (the central administration adopted 1785 and the local-provincial administration 578). Furthermore, in 2020 the pattern continued: regional administrations adopted 9645 regulations, compared to 1929 for the central administration and 676 for the local-provincial administrations.

[9] For an average person reading 8 hours per day, it would take more than a year just to read the corpus.

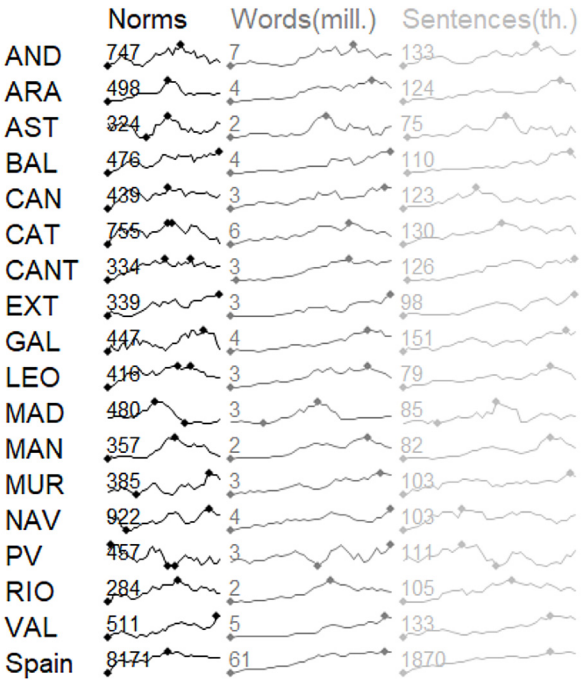

Fig. 2. Basic volume indicators of the regional legal corpus.
Note: AND: Andalusia, ARA: Aragón, AST: Asturias, BAL: Balearic Islands; CAN: Canary Islands, CAT: Catalonia. CANT: Cantabria, EXT: Extremadura, GAL: Galicia, LEO: Castile and Leon, MAD: Community of Madrid, MAN: Castile- La Mancha, MUR: Murcia; NAV: Navarra, PV: Basque Country, RIO: La Rioja, VAL: Valencia. "Spain" is the sum of all the regional norms (thus not including the central administration norms).
Source: Own elaboration.

million words, compared to Asturias (AST) with only 1.8 million. For some regions, the average number of words per sentence is 23 (Rioja), whilst in others such as Andalusia 52 words figure per sentence.

The measurements of the number of norms, sentences, and words are part of the "quantity approach" to complexity. However, these three indicators are independent: the length of the corpus is not necessarily related to the length of the norms within it. Indeed, the correlation between the number of norms and the number of words is 6% and with the number of sentences is 0%.[10] In this paper we will work with the measurement of the number of norms as this is the usual approach to the volume of regulation used in the literature.

2.2. *The relational approach: network analysis*

An additional element that contributes to the complexity of regulation is the external interdependence between norms. Citations and references to other norms increase the cost of knowledge acquisition. That is, a legal text that obliges the reader to consult other legal texts to reach a complete understanding of its content will demand an increased effort in terms of usage and

[10] The correlation between words and sentences amounts to 90%. The correlation between words and paragraphs amounts to 79%. The correlation between sentences and paragraphs amounts to 87%.

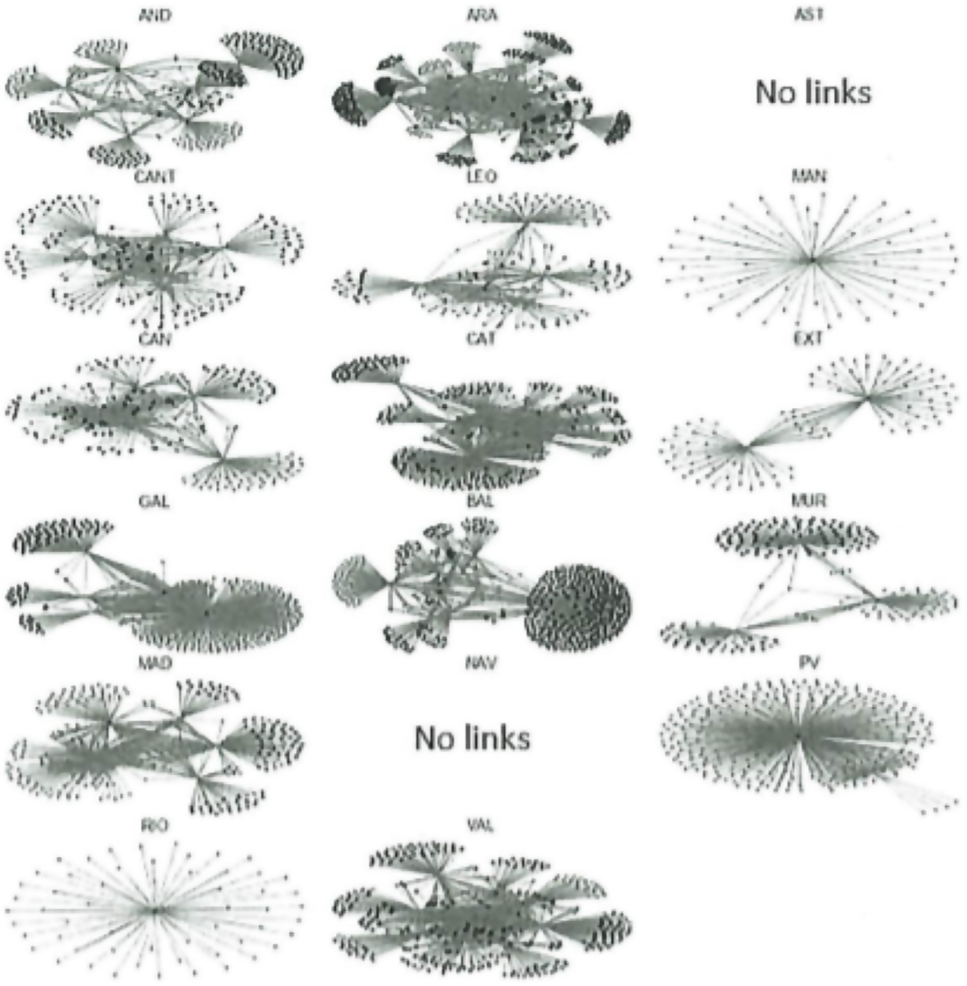

Fig. 3. Regional network of norms (only for norms with more than 50 inward-outward links).
Note: AND: Andalusia, ARA: Aragón, AST: Asturias, BAL: Balearic Islands; CAN: Canary Islands, CAT: Catalonia. CANT: Cantabria, EXT: Extremadura, GAL: Galicia, LEO: Castile and Leon, MAD: Community of Madrid, MAN: Castile- La Mancha, MUR: Murcia; NAV: Navarra, PV: Basque Country, RIO: La Rioja, VAL: Valencia.
Source: Own elaboration.

enforcement. We have built an indicator which computes the average number of norms that are referenced in the norms adopted per each region in a specific year.

Fig. 3 graphs the links between norms per each region during the period analyzed. Each network corresponds to a region. For the sake of simplicity, it only includes those texts with more than 50 links. Fig. 3 illustrates that there are important differences between the networks of different regions. Some regions, like Aragon (ARA) and Valencia (VAL), have much more intricate structures. In contrast, Asturias (AST) or Navarra (NAV) do not have any norms with more than 50 links. We also have regions such as Extremadura (EXT), Castile- La Mancha (MAN) or La Rioja (RIO) with low connectivity within their legal corpus. A more interconnected graph points to a more complex structure of the legal corpus.

The network representation as graphed in Fig. 3 provides a static proxy of the stock of regulations and their relational structure. In order to construct panel data estimations, we need to explore its time variance as well, i.e., the evolution of these network structures over time. Therefore, we also compute the average number of new links generated by the regulation adopted during each year according to Eq. (1).

$$\#Links_{reg,t} = \frac{\sum_{n\,\varepsilon\,(reg,t)} Links_n}{\sum_{n\,\varepsilon\,(reg,t)} n} \tag{1}$$

where $\#Links_{reg,t}$ corresponds to the average number of links that incorporates the norms adopted by a region in the year, t. The numerator is the sum of the links incorporated in the norms, n, adopted during a year. Denominator is the number of norms adopted by the region, *reg*, in year, t.

The last column of Fig. 4 presents our relational indicator. We observe that for all the regions the maximum number of links always occurs in more recent years than the minimum. This reflects the fact that the average number of external links of legal text has increased over the democratic period.

On average, new laws incorporate references to 11 different legal text. Some regions like Aragon (ARA) incorporate more external references, about 18, while other regions such as Navarra (NAV) only incorporate 3 references. In Fig. 3 we observe that Navarra does not possess regulations with more than 50 references, compared to Aragon which maintains a very interconnected structure. The proposed indicator can capture some network characteristics observable in the graphs (size, density and connectivity) as well as their evolution over the period of analysis.

2.3. The linguistic approach to complexity: indicators of legibility

The linguistic approach to complexity analyzes the lexical structure of the texts (paragraphs, sentences and words). Regulations may be "complex" because they are ambiguously or poorly drafted, causing difficulties to consumers and businesses in terms of understanding and compliance. The data presented in the middle column of Fig. 4 follows the μ indicator proposed by Muñoz and Muñoz (2006). This indicator has been used recently in other works analyzing the Spanish language [Brelsford et al. (2018) and Campillo et al. (2020)].[11] It is computed following Eq. (2).

$$\mu_n = \left(\frac{Wo_n}{Wo_n - 1}\right)\left(\frac{\overline{\text{Le}}_n}{\sigma^2_{Len}}\right) \times 100 \tag{2}$$

where, n, refers to norms; Wo, corresponds to the number of words; $\overline{\text{Le}}$, is the average number of letters per word and, σ^2_{Le}, refers to the variance in the number of letters per word[12] . In practice, μ usually takes values between 0 and 100, although it can reach higher values. Greater values of the

[11] Alternatives to our indicator are those derived from the work of Flesch (1948). Fernández (1959) and the Flesch-Szigriszt indicator (1993) adapted the indicator to the Spanish language but used an ad-hoc parameterization which is not without criticism.

[12] For texts with a volume of words similar to those considered in this document, the first parenthesis tends to one and therefore the formula is the inverse of Pearson's coefficient of variation which indicates the relationship between the standard deviation of a sample and its mean, all divided by the standard variation.

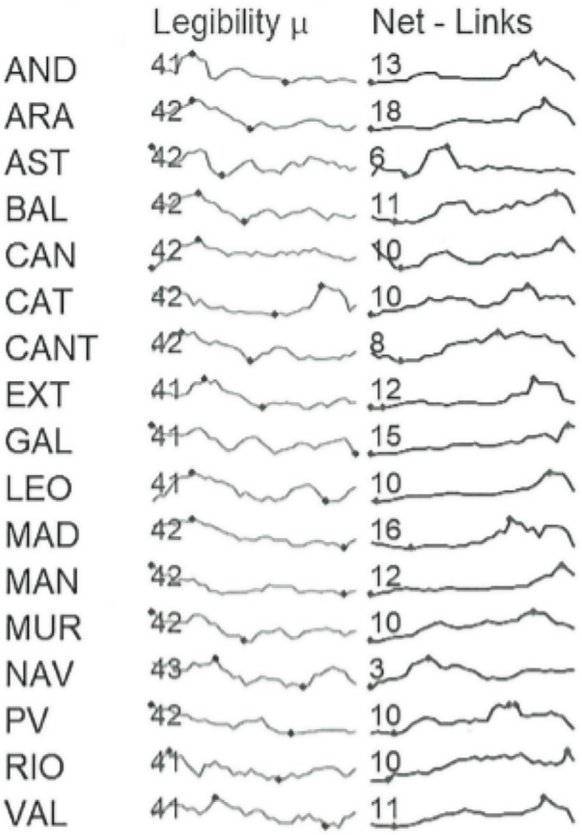

Fig. 4. Basic regulatory complexity indicators at the regional level.
Note: AND: Andalusia, ARA: Aragón, AST: Asturias, BAL: Balearic Islands; CAN: Canary Islands, CAT: Catalonia. CANT: Cantabria, EXT: Extremadura, GAL: Galicia, LEO: Castile and Leon, MAD: Community of Madrid, MAN: Castile- La Mancha, MUR: Murcia; NAV: Navarra, PV: Basque Country, RIO: La Rioja, VAL: Valencia. "Spain" is the sum of all the regional norms (thus not including the central administration norms).
Source: Own elaboration.

μ indicator are associated with better readability. Lower values correspond with more complex regulations, lower legibility.

The final value of the legibility indicator is the mean value of the legibility indicator of the norms approved in the region during the year, $Legib_{reg,t} = \overline{\mu}_{n \in (reg,\, t)}$.

For most regions, we observe a downward trend in legibility, Fig. 4. The small graphs also show the minimum and maximum values of the series. The ease with which a reader can decode norms seems to be lower during the latter period of the series. Please note that the maximum level of legibility usually is located at the beginning of the graph.

The usual values for the legibility indicator variables are between 40 and 44, corresponding to difficult texts (the mean is 41,7, the median is 41,6 and the standard deviation is 1,4). All the data but two are within the range 38–48 which, in accordance with Muñoz and Muñoz (2006) correspond to difficult texts (very difficult texts are below 31 and easy texts have an μ indicator value over 70). Navarra (NAV) has legal norms that are more legible than Andalusia (AND),

Extremadura (EXT) or Valencia (VAL). The variable considered for estimation is the average value of the legibility indicator for the norms adopted during each year.

Lower legibility and higher interconnected laws result in more complex regulations. Their effects are complementary but independent to the measurement of the quantity of regulation, as will be shown in the next sections. The correlation matrix between variables shows that the correlations between the number of links and the number of norms and legibility is only −2%, and −10%, respectively. The number of links and legibility correlates at a level of −33%.

3. Examples of the impact of regulatory complexity on economic efficiency

From a theoretical point of view, more complex regulations are less effective in reducing transaction costs. In the worst case, complexity may entail costs for businesses and citizens and may lead to resource misallocation. In this section, we make use of the variables developed in Section 2 to preliminarily explore correlation with economic and judicial variables. More specifically, we empirically analyze the relationship between regulatory complexity in Spain, focusing on the new indicators (linguistic and relational complexity) and two sets of structural variables at the regional level: (labor) productivity and judicial efficacy. It is part of the future research agenda to analyze in depth each of the different channels of these relationships.

3.1. Comprehensive macroeconomic impacts: labor productivity

We present the relation of regulatory complexity with a comprehensive macroeconomic measure: labor productivity, defined as value added (VA) per hour. We prefer this measure (VA per hour) to a more general one such as the GDP per capita for two reasons: it is more robust to underlying regional demographic and labor market trends (for instance, unemployment or the participation rate) and it is closer to structural economic growth (Harvie et al., 2009). The source of regional macroeconomic data is De la Fuente (2019). Regional information on real VA per hour is available for the period 1977−2017. Those regions with more complex regulations are expected to experiment lower productivity. Within each region, periods with more complex regulations are expected to be negatively related to productivity.

This exploration is based upon the fact that regulatory complexity is supposed to be related to the structural component of productivity. More specifically, the complexity of regulation has a negative impact on productivity through "total factor productivity" (TFP). TFP growth captures the effects of several mixed factors, including the impact of the quality of the institutional environment, such as the regulation of product and labor markets and the capacity of the economy to innovate on the productive use of labor, capital and other inputs. TFP growth is often defined as "technological progress" [see Scarpetta et al. (2002), Mora-Sanguinetti and Fuentes (2012) and Mora-Sanguinetti (2021)].

We estimate Eq. (3) using panel data covering the 17 Spanish regions. Our dependent variable is productivity. As already mentioned, our measures of complexity are: first, the number of norms as a proxy of the volume of regional regulation (we expect a negative relationship); second, average μ legibility as a proxy of linguistic complexity (we expect a positive relationship); and third, our relational variable, measured with the average number of links present in the new norms (we expect a negative relationship). Our estimates include regional and time fixed effects, all variables entering the estimation in logs. As in the rest of estimates, errors have been clustered at the regional level.

Table 1
Labor productivity per hour. Panel data Fixed effects.

	(1)	(2)	(3)	(4)	(5)
# Norms $_t$	0.00232				0.00181
	(0.00158)				(0.00150)
Legibility $_t$		0.0769***		0.0642**	0.0649**
		(0.0261)		(0.0226)	(0.0225)
# Links $_t$			−0.00348*	−0.00257	−0.00221
			(0.00196)	(0.00186)	(0.00221)
Constant	0.551***	0.259	0.544***	0.309	0.315
	(0.146)	(0.196)	(0.140)	(0.177)	(0.183)
Labor productivity $_{t-1}$	0.947***	0.948***	0.949***	0.948***	0.947***
	(0.0148)	(0.0144)	(0.0144)	(0.0146)	(0.0152)
Fixed effects					
Time	Yes	Yes	Yes	Yes	Yes
Regional	Yes	Yes	Yes	Yes	Yes
Observations	583	583	583	583	583
R-squared	0.985	0.985	0.985	0.985	0.985

Source: Own elaboration.
Robust (clustered) standard errors in brackets. ***p < 0.01, **p < 0.05, *p < 0.1.

$$Prod_{reg,t} = Cte + \beta_1 \underbrace{\#Norms_{reg,t-i}}_{\text{Quantity}} + \beta_2 \underbrace{Legib_{reg,t-i}}_{\text{Linguistic}} + \beta_2 \underbrace{\#Links_{reg,t-i}}_{\text{Relational}} + \varepsilon_{reg,t}$$

Dimensions of complexity:

Estimates are presented in Table 1. Columns (1) to (3) present the results when the complexity variables are included independently. The estimates take the form of a traditional convergence regression, which controls for the initial productivity value. All measurements of complexity have the expected sign. The new complexity indicators have a (statistically) significant effect. Column (4) includes both the lexical complexity and the relational complexity dimensions. Both maintain their sign, but the relational complexity dimension loses its statistically significant level when the quantity indicator is also incorporated. Column (5) includes all the complexity indicators simultaneously; the estimated parameters maintain the sign, magnitude and significance levels.

Table 1 is in line with the findings of Di Vita (2018) which approximates the complexity of regulation through the quantity approach (number or volume of norms) and shows that an increase in the number of norms correlates with lower productivity. Our results also show that other dimensions of regulation affect productivity, this being consistent with Coffey et al. (2020) and Djankov et al. (2006). First, legibility is positively related to productivity. Finally, the analysis of our "relational" variable also shows that a greater number of links to other norms in the new regulations correlates negatively with productivity. These results[13] are new in the literature and confirm that when analyzing regulatory complexity, it is useful to examine further factors in addition to the volume of regulations. Interestingly though, these other factors do not seem to be

[13] The effects presented in Table 1 are contemporary. However, there are numerous judicial rulings (from the Constitutional Court) in Spain that suggest that regulation is predictable (and therefore could have anticipated the effects on economic efficiency). Thus, Mora-Sanguinetti and Pérez-Valls (2021) analyzed the advanced effects of the quantitative complexity of regulation (in addition to the contemporary ones) on business demographics. It is debatable whether agents can foresee linguistic or relational complexity beyond quantitative complexity.

related with the volume of regulation (number of norms) which has been traditionally used as a proxy for complexity.

3.2. Effects of complexity on legal costs (litigation and judicial efficacy)

Regulatory complexity may also imply other relevant costs for public management (in addition to having an impact on private agents) such as judicial costs. That is, those related to the effectiveness of the judicial system and the enforcement mechanisms available in the economy. As noted in the introduction, the OECD (Palumbo et al., 2013), found that a low quality of regulation was related to more litigation in the countries analyzed. An increase in litigation is also related to lower judicial efficacy. In other words, the judicial system would show higher rates of congestion or longer trial durations. Judicial inefficacy has important implications for economic efficiency in Spain, reducing credit (Mora-Sanguinetti et al., 2017), investment at the enterprise level (Dejuán & Mora-Sanguinetti, 2021) or the proportion of rented housing (Mora-Sanguinetti, 2012).

Our dependent variables in this case will be a measure of trial duration in the civil jurisdiction following the CEPEJ (2016) approach, Eq. (4). Trial duration is a measure that approximates the congestion of the judicial system. In other words, its ineffectiveness in swiftly resolving cases brought before the courts. The raw data for its calculation comes from the General Council of the Judiciary.

$$Trial\ duration_{reg,t}\,(CEPEJ) = \frac{Pending\ cases_{reg,t}}{Cases\ resolved_{reg,t}} * 365 \tag{4}$$

We analyze this effect by means of Eq. (5). As before, we estimated the equation with panel data covering the 17 Spanish regions. All variables enter the estimation in logs. Following the empirical strategy explained above, we introduce each variable of regulatory complexity independently. All estimates include fixed effects at the regional level and time dummies. As control variables we include productivity and the number of lawyers [see Mora-Sanguinetti and Garoupa (2015) and Carmignani and Giacomelli (2010)]. Errors have been clustered at the regional level.

$$Jud.\,Cong_{reg,t} = Cte + \beta_1\, Complexity_{reg,t-1} + \sum_{k=1}^{K} \delta_k Control^{k}_{reg,t-1} + Jud.\,Cong_{reg,t-1} + \varepsilon_{reg,t} \tag{5}$$

Table 2 shows the results. As indicated previously, the analysis is carried out for civil jurisdiction, which disciplines cases involving private contracts (between citizens and between companies). Specifically, we build the measurements of trial duration until execution. As a robustness check, in a set of estimates (columns 1–3), we exclude family law conflicts and in columns 4–6 we include the whole set of civil cases. The signs are as expected, and the impacts are significant for the relational complexity variables (number of links).

The results suggest that greater relational complexity (that is, the need to consult or understand a greater number of interconnected norms in order to make use of the law), makes the functioning of the courts (which basically have to apply the body of regulation to resolve a specific conflict) more difficult or slower. The work of the courts does not seem to be affected either by the number of rules *per se* or by the linguistic complexity of those rules, possibly suggesting that legal professionals are trained precisely to master these two dimensions.

Table 2
Judicial congestion (trial duration). Panel data Fixed effects.

	(1)	(2)	(3)	(4)	(5)	(6)
	Civil (without family)			Total civil		
# Norms $_{t-1}$	−0.00619			−0.00833		
	(0.0161)			(0.0146)		
Legibility $_{t-1}$		−0.126			−0.0696	
		(0.246)			(0.229)	
# Links $_{t.1}$			0.0232*			0.0212**
			(0.0110)			(0.00996)
Judicial Cong. $_{t-1}$	0.690***	0.689***	0.684***	0.699***	0.699***	0.694***
	(0.0556)	(0.0550)	(0.0563)	(0.0564)	(0.0557)	(0.0569)
Constant	6.820	7.327	7.076	7.003	7.282	7.246
	(6.490)	(6.536)	(6.344)	(6.007)	(6.115)	(5.877)
Fixed effects						
Time	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Regional	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Controls (lawyers, productivity)	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Observations	*272*	*272*	*255*	255	255	255
R-squared	*0.807*	*0.807*	*0.810*	0.783	0.783	0.786

Source: Own elaboration.
Robust (clustered) standard errors in brackets. ***p < 0.01, **p < 0.05, *p < 0.1.

4. Robustness checks

We perform two alternative robustness checks: first we consider two new measures of linguistic complexity: entropy and the percentage of infrequent distinct words (type) in the adopted norms; secondly, we modified the original database to only include rules with a minimum number of words, thus eliminating very short rules.

4.1. Additional measures of linguistic complexity

In this section we introduce two additional measures of linguistic complexity: entropy and the share of infrequent words. These variables replace the μ linguistic complexity index.

4.1.1. Entropy

An additional measure of linguistic complexity is the "entropy" indicator proposed by Katz and Bommarito (2014) and Shannon (1951).[14] To build the entropy indicator we sum up the probability of occurrence of each distinct word, p_w, within the set of total words, W_n, of a norm, n, multiplied by their logarithm in base 2 of this probability (see Eq. 6). Entropy characterizes the uncertainty or variance in a system[15], in this case the normative system. A minimum level of

[14] Another work following this line of research (i.e., quantitative analysis of the law) is Friedrich et al. (2020), who computed entropy indicators based on the written text of opinions published by the U.S. Supreme Court and the German Bundesgerichtshof. As indicated, one advantage of our analysis is that we are not forced to compare different languages since all our sources are officially in Spanish.

[15] For texts with a number of words such as those considered in this document, the first parenthesis tends to one and, therefore, the formula is the inverse of Pearson's coefficient of variation, which indicates the relationship between the standard deviation of a sample and its mean, all divided by the standard variation.

entropy is necessary in order for information to exist within a text. However, a higher degree of entropy is characteristic of more complex legal texts which include a larger variety of concepts.

$$Entropy_n = -\sum_{w \in W_n} p_w \log_2(p_w) \tag{6}$$

The final value for the entropy indicator, specific for each region and year, is the simple mean of the entropy indicator for each norm adopted in the region during the year, $Entropy_{reg,t} = \overline{Entropy_{n \in (reg,\, t)}}$.

4.1.2. Share of infrequent words

We also construct a new index of similarity between the vocabulary used in legal texts and that most frequently used in the Spanish language according to the Royal Academy of the Spanish Language (RAE)[16] . We obtain the percentage of words used in the legal texts which do not appear among the 10.000 more frequently used words in the Spanish language[17] . In Eq. (7), the numerator counts the different words in a norm, Wd_n, that do not appear among the 10.000 most frequent words in Spanish, according to the RAE. The denominator counts the total number of different words in that text. On average, between 55% and 60% of the words in the regional legal corpus do not belong to the list of the 10.000 most frequently used words. The greater the share of infrequent words the lower the legibility. Hence, this is an inversed index of the simplicity of the vocabulary used in the regulatory framework.

$$\text{Freq}_n = 100 \times \frac{\sum_{\text{Wd}_n \; i/n \, \text{RAE}} \text{Wd}_n}{\sum \text{Wd}_n} \tag{7}$$

The entropy indicator is negatively correlated with the legibility indicator μ discussed above (-0.65). The share of *infrequent* words in the Spanish language has a low negative correlation with the entropy indicator (-0.47) and a positive correlation with legibility (0.82).

We again analyzed the impacts of regulatory complexity on productivity and legal costs using these new approaches to complexity. Table 3 presents the estimations: Columns 1 and 2 analyze correlation with productivity. Columns 3 and 4 discuss the implications for legal costs. As carried out so far, the different dimensions of complexity are incorporated independently. In the case of productivity growth, the entropy indicator is significant and with the expected negative sign. In the case of judicial congestion (trial duration), the share of infrequent words is significant with the expected positive value.

4.2. Different data sets

Finally, as an additional robustness check, we have worked with a modified database: we have only taken into account the norms of a certain length (excluding therefore the shorter ones). In the first place, we included only the regulations with more than 200 words and, secondly, the regulations with more than 1000 words.

[16] Information obtained from Current Spanish Reference Corpus, CREA, Real Academia Española - RAE.es

[17] We also used the percentage of (unique) frequent words over the total number of distinct words used in the text. Results do not vary. Results are available on request.

Table 3
Robustness alternative indicators of legibility. Panel data Fixed effects.

	(1)	(2)		(3)	(4)
Dep var: Productivity			Dep var: judicial congestion (civil, without family)		
Entropy $_{t}$	−0.0508*		Entropy $_{t-1}$	0.405*	
	(0.0263)			(0.198)	
Share unique $_{t}$		−0.0319	Share unique $_{t-1}$		0.292**
		(0.0245)			(0.125)
Productivity $_{t-1}$	0.949***	0.949***	Jud Cong $_{t-1}$	0.688***	0.683***
	(0.0149)	(0.0147)		(0.0565)	(0.0566)
Constant	0.871***	0.663***	Constant	4.185	6.138
	(0.214)	(0.138)		(6.445)	(6.490)
Fixed effects			Fixed effects		
Time	Yes	Yes	Time	Yes	Yes
Regional	Yes	Yes	Regional	Yes	Yes
			Lawyers + Prod t-1	Yes	Yes
Observations	583	583	Observations	255	255
R-squared	0.985	0.985	R-squared	0.810	0.810

Source: own elaboration.
Robust (clustered) standard errors in brackets. ***p < 0.01, **p < 0.05, *p < 0.1.

Table 4
Robustness Dep. var.: Growth VA per hour.

	Productivity growth (1) Norms > 200 words	Productivity growth (2) Norms > 1000 words	Judicial cong. (3) Norms > 200 words	Judicial cong. (4) Norms > 1000 words
# Norms	0.00226	0.00139	−0.00811	0.00432
	(0.00159)	(0.00157)	(0.0165)	(0.0196)
Legibility	0.0795**	0.0710***	−0.176	0.0692
	(0.0313)	(0.0210)	(0.255)	(0.469)
# Links	−0.00348*	−0.00336	0.0232*	0.0254**
	(0.00196)	(0.00199)	(0.0110)	(0.0110)
Fixed effects				
Time	Yes	Yes	Yes	Yes
Regional	Yes	Yes	Yes	Yes
Observations	583	583	255	255

Source: Authors' calculations.
Robust (clustered) standard errors in brackets. ***p < 0.01, **p < 0.05, *p < 0.1.

The reason for making these changes is that some regulations with low economic or legal content have the force of law, such as those that just modify a ministerial or departmental structure. Those norms are usually short. When we select longer regulations, we are removing 197 and 2029 norms respectively form the original data set. Table 4 presents the results in terms of the relationship between complexity and productivity (or legal costs) using Eq. (3) for productivity growth (see columns 1 and 2). Columns 3 and 4 present the estimated parameters for judicial congestion (trial duration) following Eq. (5).

Our variables of regulatory complexity always have the expected sign. Legibility seems to have a significant and positive correlation with productivity, while the relational complexity measure

seems to have a positive relationship with legal costs (judicial congestion). Overall, these results show that it may be necessary to select only the regulations with the greatest economic and legal content in order to truly understand the impacts of complexity on productivity and legal costs.

5. Messages for policy modeling

Our analyzes have several policy implications. First, any society needs to build a shared understanding of the importance and meaning of "better regulation". Specifically, regulators and policy makers need to be aware that the way in which they regulate may have economic and social costs. In this sense, in order to improve the process of adoption of new regulations, it is necessary to improve consultations and pre-legislative scrutiny and enhance public awareness and participation.

Second, simultaneously, it is necessary to develop a comprehensive policy which can be expressed as "drafting instructions" and ensure that it is implemented. This have implications in the regulation design stage (checking if new legislation is necessary, desirable or acceptable for adoption), rewriting existing legislation (revising the norms to eliminate redundant, contradictory, unclear regulations as this would serve to enhance the quality of the regulatory framework) and, finally, ensuring that legislation is accessible to society.

From our analysis of the three dimensions of regulatory complexity and their consequences, several recommendations for drafting and rewriting regulations (and thus for achieving "better regulation") can be drawn: first, regulators should use plain language and write clear, concise, well-organized norms with a reader-centered organization; second, regulators should avoid ambiguous or contradictory provisions and use "user friendly" language and writing style; third, regulators should reduce the unnecessary links to other norms.

All of this would be facilitated by the development of checklists, templates and writing guidelines. It seems relevant to recall that the most recent norms, such as the above-mentioned Royal Decree 931/2017, in the case of Spain, provide for the creation of a "Methodological Guide for the preparation of the regulatory impact analysis report". In the specific case of Spain, the "Regulatory Impact Analysis Reports" (designed under the Royal Decree 931/2017 and the Methodological Guide) should be expanded with a series of metrics based on the analysis of the text of the regulation to be adopted, as suggested in this research. Specifically, it is worth thinking of a set of "structural metrics", containing information on the length of the text (and the number of words) of the norm and a reference to the weight of the new norm within the regulatory framework - based on the number of norms it repeals or does not repeal [Article 1. b) of Royal Decree 931/2017]-. This second calculation would require tracking the amount of regulation by subject area (see also Mora-Sanguinetti, 2019a). Furthermore, there must be a clear accounting and an analysis of the number of links to other norms.

More generally, the use and understanding of the regulation would be facilitated by improving its accessibility. In particular, the following considerations may be made: first, law is data which could be made available to society. For instance, if it is possible, writing summaries and compendia of existing legislation and generating non-legal binding documents explaining the law for dissemination among the general public. Second, law is composed mainly by words but, in some cases, its message could be better transmitted with new elements of communication like figures, diagrams, charts, graphs, etc. exploiting the new approaches offered by information technologies. Third, public repositories of legal texts should maintain consolidated norms in one single text as well as provide links to other connected regulations.

These measures, aimed to reduce regulatory complexity, should also lead to an improvement in the functioning of the judicial system; the regulations would be more easily reached and enforced.

6. Concluding remarks

Regulation, as part of the institutional framework, matters for economic efficiency. While the objective of regulation is to reduce transaction costs and mitigate other market failures, its effects may be less beneficial or even counterproductive if the legal corpus is poorly designed. One of the reasons that can make regulation inefficient is its "complexity". The potential negative impacts of regulation have prompted numerous government initiatives aimed at achieving "better regulation".

"Complexity" has been measured by some recent works in terms of "volumes" or quantities of regulations (for instance, the number of norms or pages). In this article we argue that complexity has additional dimensions such as legibility, that is, the ease of reading the norms, and the relational structure of the norm, being the average number of external regulations to which a certain legislative text refers. Natural language processing (NLP) techniques allow us to develop objective indicators on those dimensions and to analyze their effects.

The first contribution to the literature of this article is the construction of a new database (RECOS, REgulatory COmplexity in Spain) which includes information on the laws (and other norms with the force of Law) adopted by the Autonomous Regions since the beginning of the democratic period in Spain. We also build a set of indicators covering the new dimensions of complexity: legibility and network structure. The second contribution of this research is a novel and initial exploration of the correlation of the new dimensions of complexity on various structural variables; namely, labor productivity growth and judicial efficacy.

The results of our estimations show that the new dimensions of regulatory complexity are negatively related to productivity and judicial efficacy. That is, judicial efficacy and labor productivity growth seem to decrease when the norms are less legible and when the legal texts are more difficult to use because they require access to a greater number of regulations (due to their network structure). To the best of our knowledge, this is the first study showing the potential impacts of these new dimensions of regulatory "complexity" on different measures of economic efficiency.

Our research thus rationalizes the efforts of public administrations to achieve "better regulation". Indeed, some countries provide guidelines (or even create public bodies) that try to improve the quality of norms. For instance, the Government of Spain has initiatives in place to improve regulatory quality and to assess the need for new regulation (see Royal Decree 931/2017 and Mora-Sanguinetti, 2019a). Specifically, the quality of regional regulation should never be neglected. At present, nearly 70% of all Spanish regulations are issued at a regional level (see Mora-Sanguinetti and Pérez-Valls, 2021). As already highlighted, based on a regional analysis of the different dimensions of complexity, our research reveals substantial differences in regulatory complexity between the Spanish regions. This provides the opportunity to identify areas for improvement in the quality of regulation.

In terms of economic analysis, this article leaves open the normative question of what the optimal level of regulatory complexity is, taking into account the optimal level of economic growth. Futures avenues of research should also distinguish between pure economic legislation (labor markets, trade and retail, sectors regulation, etc.) and legislation on topics which *prima facie* may have only indirect effects on economic performance (prisons, health, etc.). Addressing the analysis of regulation by theme will also allow to clarify, among other things, if the increase in the regulatory power (increase in competences) of the regions drives the increase in the number of regulations. In this first analysis we decided to follow a more general approach covering all norms,

this strategy precludes the selection bias criticism. Our paper opens the road towards using NLP methodology for the purpose of constructing additional measures of complexity such us in-depth sentence structure (conditional, subordinate, etc.), vocabulary accuracy (use of vague terms) or analysis of the internal structure of regulation.

References

Acemoglu, D., & Robinson, J. A. (2012). *Why nations fail: The origins of power, prosperity, and poverty*. Crown Books.

Al-Ubaydli, O., & McLaughlin, P. A. (2017). Regdata: A numerical database on industry-specific regulations for all United States industries and federal regulations, 1997–2012. *Regulation and Governance*, *11*(1), 109–123.

Azqueta-Gavaldón, A., Hirschbühl, D., Onorante, L., & Saiz, L. (2020). *Economic policy uncertainty in the Euro Area: An unsupervised machine learning approach* ECB Working Papers, 2359.

Bailey, J. B., & Thomas, D. W. (2017). Regulating away competition: The effect of regulation on entrepreneurship and employment. *Journal of Regulatory Economics*, *52*(3), 237–254.

Baker, S. R., Bloom, N., & Davis, S. J. (2016). Measuring economic policy uncertainty. *The Quarterly Journal of Economics*, *131*(4), 1593–1636.

Bardhan, P. (2002). Decentralization of governance and development. *Journal of Economic Perspectives*, *16*(4), 185–205.

Bertrand, M., & Kramarz, F. (2002). Does entry regulation hinder job creation? Evidence from the French retail industry. *The Quarterly Journal of Economics*, *117*(4), 1369–1413.

Brelsford, K. M., Ruiz, E., & Beskow, L. (2018). Developing informed consent materials for non-English-speaking participants: An analysis of four professional firm translations from English to Spanish. *Clinical Trials*, *15*(6), 557–566.

Calomiris, C. W., Mamaysky, H., & Yang, R. (2020). *Measuring the cost of regulation: A text-based approach* NBER-Working Papers, 26856.

Campillo, J., Morato, A., Maqueda, I., & Sánchez-Cuadrado, S. (2020). Readability of Spanish e-government information. In *15th Iberian Conference on Information Systems and Technologies (CISTI)* (pp. 1–4). https://doi.org/10.23919/CISTI49556.2020.9141000

Carmignani, A., & Giacomelli, S. (2010). *Too many lawyers? Litigation in Italian civil courts, Temi di discussione*. pp. 745. Banca d'Italia.

CEPEJ [European Commission for the Efficiency of Justice]. (2016). *European judicial systems efficiency and quality of justice* CEPEJ Studies, No. 23. http://www.coe.int/t/dghl/cooperation/cepej/evaluation/default_en.asp

Chambers, D., Collins, C. A., & Krause, A. (2019). How do federal regulations affect consumer prices? An analysis of the regressive effects of regulation. *Public Choice*, *180*(1–2), 57–90.

Chambers, D., McLaughlin, P. A., & Stanley, L. (2019). Regulation and poverty: An empirical examination of the relationship between the incidence of federal regulation and the occurrence of poverty across the US states. *Public Choice*, *180*(1–2), 131–144.

Coenen, G., Ehrmann, M., Gaballo, G., Hoffmann, P., Nakov, A., Nardelli, S., Persson, E., & Strasser, G. (2017). *Communication of monetary policy in unconventional times* ECB Working Papers, 2080, June.

Coffey, B., McLaughlin, P. A., & Peretto, P. (2020). The cumulative cost of regulations. *Review of Economic Dynamics*, *38*, 1–21.

Costa, A. M. (2010). The economics of crime: A discipline to be invented and a nobel prize to be awarded. *Journal of Policy Modeling*, *32*(5), 648–661.

Davis, S. J. (2017). *Regulatory complexity and policy uncertainty: Headwinds of our own making*. Becker Friedman Institute for Research in Economics Working Papers. No. 2723980.

De la Fuente, A. (2019). Series largas de algunos agregados económicos y demográficos regionales. In *Estudios sobre la economía española (2019/19)*. Fedea.

Dejuán, D., & Mora-Sanguinetti, J. S. (2021). Which legal procedure affects business investment most, and which companies are most sensitive? Evidence from microdata. *Economic Modelling*, *94*, 201–220.

Di Vita, G. (2018). Institutional quality and the growth rates of the Italian regions: The costs of regulatory complexity. *Papers in Regional Science*, *97*(4), 1057–1081.

Djankov, S., McLiesh, C., & Ramalho, R. M. (2006). Regulation and growth. *Economics Letters*, *92*(3), 395–401.

Ehrmann, M., & Talmi, J. (2020). Starting from a blank page? Semantic similarity in central bank communication and market volatility. *Journal of Monetary Economics*, *111*, 48–62.

Fernández, J. (1959). Medidas sencillas de lecturabilidad. *Consigna*, *214*, 29–32.

Flesch, R. (1948). A new readability yardstick. *Journal of Applied Psychology*, *32*, 221–233.

Friedrich, R., Luzzatto, M., & Ash, E. (2020). Entropy in legal language. In *Proceedings of the 2020 Natural Legal Language Processing (NLLP) Workshop* , 6 pp.

Gai, P., Kemp, M. H., Sánchez Serrano, A., & Schnabel, I. (2019). Regulatory complexity and the quest for robust regulation (No. 8). In *Reports of the Advisory Scientific Committee*. European Systemic Risk Board (ESRB), European System of Financial Supervision.

Ghirelli, C., Gil, M., Pérez, J. J., & Urtasun, A. (2019). Measuring economic and economic policy uncertainty and their macroeconomic effects: The case of Spain. *Empirical Economics*, 1–24. https://doi.org/10.1007/s00181-019-01772-8

Haldane, A., & McMahon, M. (2018). Central bank communications and the general public. *AEA Papers and Proceedings*, *108*, 578–583.

Hampton, P. (2005). *Reducing administrative burdens*. London [Hampton Report].

Hansen, S., McMahon, M., & Prat, A. (2017). Transparency and deliberation within the FOMC: A computational linguistics approach. *The Quarterly Journal of Economics*, *133*(2), 801–870.

Harvie, D., Slater, G., Philp, B., & Wheatley, D. (2009). Economic well-being and British regions: The problem with GDP per capita. *Review of Social Economy*, *67*(4), 483–505.

Hassan, T., Hollander, S., van Lent, L., & Tahoun, A. (2019). Firm level political risk: Measurement and effects. *The Quarterly Journal of Economics*, *134*(4), 2135–2202.

Hernández-Murillo, R., & Shell, H. (2014). Rising complexity of the FOMC statement. *Economic Synopses*, *23*(2). Available at. https://research.stlouisfed.org/publications/economic-synopses/2014/11/05/the-rising-complexity-of-the-fomc-statement/

Hooghe, L., Schakel, A. H., & Marks, G. (2008). Appendix B: Country and regional scores. *Regional & Federal Studies*, *18*(2), 259–274.

Hurka, S., & Haag, M. (2019). Policy complexity and legislative duration in the European Union. *European Union Politics*, *21*(1), 87–108.

Jansen, D. J. (2011). Does the clarity of central bank communication affect volatility in financial markets? Evidence from Humphrey–Hawkins testimonies. *Contemporary Economic Policy*, *29*(4), 494–509.

Katz, D. M., & Bommarito, M. J. (2014). Measuring the complexity of the law: The United States code. *Artificial Intelligence and Law*, *22*(4), 337–374.

Kirchner, S. (2012). Federal legislative activism in Australia: A new approach to testing Wagner's law. *Public Choice*, *153*(3-4), 375–392.

Laffont, J.-J., & Tirole, J. (1993). *A theory of incentives in procurement and regulation*. Cambridge, Massachusetts, USA: The MIT Press.

López Guerra, L. (1994). *Introducción al Derecho Constitucional*. Valencia: Tirant Lo Blanch.

López Guerra, L., Espín, E., García Morillo, J., Pérez Tremps, P., & Satrústegui, M. (2018). *Derecho Constitucional Volumen II* (11th edition). Valencia: Tirant Lo Blanch.

McLaughlin, P. A., & Sherouse, O. (2019). RegData 2.2: A panel dataset on US federal regulations. *Public Choice*, *180*(1-2), 43–55.

Mora-Sanguinetti, J. S. (2012). Is judicial inefficacy increasing the weight of the house property market in Spain? Evidence at the local level. SERIEs. *Journal of the Spanish Economic Association*, *3*(3), 339–365.

Mora-Sanguinetti, J. S. (2021). Justicia y economía: la eficiencia del sistema judicial en España y sus impactos económicos. *Papeles de Economía Española*,. N.º 168.

Mora-Sanguinetti, J. S. (2019a). La 'complejidad' de la regulación española ¿Cómo medirla? ¿Qué impacto económico tiene? *Revista ICE*, *907*, 147–162.

Mora-Sanguinetti, J. S. (2019b). *Producción normativa, mercado de la abogacía y litigación en España*. Dossier, Diciembre/2019: Thomson Reuters.

Mora-Sanguinetti, J. S., & Fuentes, A. (2012). *An analysis of productivity performance in Spain before and during the crisis: Exploring the role of institutions* OECD Economics Department Working Papers, No. 973.

Mora-Sanguinetti, J. S., & Garoupa, N. (2015). Do lawyers induce litigation? Evidence from Spain, 2001-2010. *International Review of Law and Economics*, *44*, 29–41.

Mora-Sanguinetti, J. S., & Pérez-Valls, R. (2021). How does regulatory complexity affect business demography? Evidence from Spain. *European Journal of Law & Economics*, *51*, 203–242.

Mora-Sanguinetti, J. S., & Salvador-Mora, S. (2016). El impacto económico de la actividad normativa. Los riesgos de la 'gran convergencia'. *Revista de Economía Industrial*, *398*, 117–128.

Mora-Sanguinetti, J. S., Martínez-Matute, M., & García-Posada, M. (2017). Credit, crisis and contract enforcement: Evidence from the Spanish loan market. *European Journal of Law & Economics*, *44*(2), 361–383.

Muñoz, M., & Muñoz, J. (2006). *Legibilidad Mμ*. Chile: Viña del Mar.

Nedic, V., Despotovic, D., Cvetanovic, S., Djuki'c, T., & Petrovic, D. (2020). Institutional reforms for economic growth in the western balkan countries. *Journal of Policy Modeling*, *42*(5), 933–952.

OECD. (2012). *Recommendation of the council on regulatory policy and governance*. Paris: OECD.

Office of the Parliamentary Counsel (Cabinet Office). (2013). *When laws become too complex. A review into the causes of complex legislation*. London: Cabinet Office.

Palumbo, G., Giupponi, G., Nunziata, L., & Mora-Sanguinetti, J. S. (2013). *The economics of civil justice: New cross-country data and empirics*. OECD Economics Department Working Papers. No. 1060.

Radaelli, C. M. (2007). Whither better regulation for the Lisbon agenda? *Journal of European Public Policy*, *14*(2), 190–207.

Rios, V., & Gianmoena, L. (2020). The link between quality of government and regional resilience in europe. *Journal of Policy Modeling*, *42*(5), 1064–1084.

Scarpetta, S., Hemmings, P., Tressel, T., & Woo, J. (2002). *The role of policy and institutions for productivity and firm dynamics: Evidence from micro and industry data* OECD Economics Department Working Papers, No. 329.

Shannon, C. E. (1951). "Prediction and entropy of printed English". *Bell System Technical Journal*, *30*(1), 50–64.

Tobback, E., Nardelli, S., & Martens, D. (2017). *Between hawks and doves: Measuring central bank communication* ECB Working Papers, No. 2085.

Viviano, E. (2008). Entry regulations and labour market outcomes: Evidence from the Italian retail trade sector. *Labour Economics*, *15*(6), 1200–1222.