Legal uncertainty and its consequences: A natural language processing approach

Maximiliano Sosa Andrés[1]

Uppsala University

Abstract

Legal uncertainty is widely recognized as negative for economic growth, and yet empirical evidence is limited due to the difficulty of producing comparable measures across time and regions. This paper develops a new index of legal uncertainty (LUI) based on legal texts, by extracting the information content in all laws approved in a country and year to calculate the unforecastable component in legal content. I present the index construction, its robustness and comparison with institutional measures. When considering a full set of developing and developed economies I show great variability in legal uncertainty across countries and institutional frameworks. I then apply the index to provide evidence on the impact of legal uncertainty over the financial decisions of firms. Results show that, in the event of an increase in legal uncertainty, firms react by slowing down investments and laying off workers.

Keywords: Legal uncertainty; natural language processing; investment growth, unemployment

JEL classification: D80, E22, E66, K00, L51

First Version: June 2022

This Version: September 2022

Preliminary and incomplete

---

[1]Department of Economics, Uppsala University, SE-751 20 Uppsala, Sweden. Email: maximiliano.sosa_andres@nek.uu.se.

I would like to thank Eva Mörk, Luca Repetto, Hannes Mueller, Mikael Carlsson, Jacob Dalgaard Christensen, the participants of the "Applied Econ and Machine Learning" group at the Institut d'Anàlisi Econòmica in Barcelona and seminar participants in Uppsala for valuable suggestions. Additional thanks to Oscar Calvo-González and Germán Reyes for providing their code

"The increasing calculability of the functioning of the legal process constituted one of the most important conditions for the existence of economic enterprise intended to function with stability and, especially, of capitalistic enterprise, which cannot do without legal security."

Max Weber (1921), Economy and Society

"Another obvious advantage of establishing as soon as possible [clear and definite rules]: predictability. Even in simpler times uncertainty has been regarded as incompatible with the Rule of Law. Rudimentary justice requires that those subject to the law must have the means of knowing what it prescribe."

US Supreme Court Justice Antonin Scalia (1989), The Rule of Law as a Law of Rules

1. Introduction

The importance of the Rule of Law in economic outcomes is embedded within the New Institutional Economics framework, which rose from contributions from Coase, North, Ostrom and Williamson. According to a definition by North "institutions are the rules of the game in a society or, more formally, are the humanly devised constraints that shape human interaction" (North (1990), p.3). Still, legal certainty also matters for economic growth. Max Weber went as far as suggesting that legal certainty was required for the development of capitalism (Trubek, 2008). Legal uncertainty undermines the Rule of Law and its power to govern behavior. Davis (2017) defines legal uncertainty as the lack of clarity regarding fact patterns. One reason why there may lack of clarity in the legal framework may be due to the fact that constant changes in the legal environment make it difficult for economic agents to comprehend the state of the world and anticipate where that state will be next. In the words of Justice Scalia, legal certainty requires predictability (Raban, 2009).

Still, measuring legal uncertainty has proven difficult and is contingent on subjectivity. This makes any analysis of the relation between legal uncertainty and economic decision making challenging. This paper contributes to this debate by providing a new and more objective Legal Uncertainty Index (LUI) based on legal texts. The proposed index relies heavily on computational linguistics and natural language processing (NLP),[1] by extracting the information content in all laws approved in a particular country and period of time. By the use of topic modelling I propose a classification of the content of laws in a country, from which it is possible to obtain the share dedicated to each topic in a given period. These topic shares are the main input to calculate how much uncertainty there might be in the legal framework, from one period to the next. In this regard, I consider legal uncertainty as the risk arising from changes in the regulatory framework that cannot be forecasted by economic agents. I then use the resulting LUI to examine how firms react to an increase in legal uncertainty.

Uncertainty has been widely analyzed in the economic sphere. The theoretical foundations trace back to Bernanke (1983). In partial equilibrium settings, uncertainty makes

[1]By natural language processing we refer to the application of computational techniques to the analysis and synthesis of natural language and speech.

optimization hard for agents and can make firms delay hiring or investment, if they are subject to fixed costs, or if financial constraints become more binding due to increases in interest rates. The underlying assumption is the irreversibility of capital outlays (Carlsson, 2007). Consumers are also expected to reduce consumption if they are risk averse in a "precautionary savings" style. Within this literature, policy uncertainty has received particular attention. Rodrik (1991), Hassett and Metcalf (1997), Fernandez-Villaverde et al. (2015) and Pástor and Veronesi (2013) provide the theoretical foundations by studying the negative economic effects of fiscal, regulatory and political uncertainty. Fluctuations in policy uncertainty contribute to slow hiring and investment, as companies are generally reluctant to make essential or costly decisions in unpredictable regulatory environments (Bloom, 2014).

Under this framework, I ask whether legal uncertainty can also make firms freeze investments and reduce hiring or increase lay-offs. By considering panel data regressions at the country level and a VAR analysis in a case study I show that an increase in legal uncertainty is associated with a decrease in investment growth and an increase in the national unemployment rate. These results are in line with previous studies that show that firms become more conservative under high uncertainty, when the cost of borrowing increases (Jens, 2017; Kelly, Pástor and Veronesi, 2016). Under this scenario, firms spend less on capital (Carlsson, 2007; Gulen and Ion, 2015) and engage in fewer mergers and acquisitions (Bonaime, Gulen and Ion, 2018). The results on this paper contribute to the literature on the economic consequences of policy uncertainty and give evidence on how legal uncertainty is a channel through which policy uncertainty impacts firm decision making.

This paper provides an additional component for policy uncertainty by extending the scope to legal uncertainty. The law is a tool over which a government has a monopoly in the ability to legislate and to produce, repeal and amend legislation. Naturally, the degree to which laws or legislation can be used by a national government varies across borders. But despite these differences in institutional frameworks across countries we can consider laws as a comparable tool for the implementation of a wide variety of public policy.

My paper also relates to the literature on policy uncertainty measurement, by offering an innovative way of quantifying uncertainty using NLP. Pioneering in this strand of literature is Baker, Bloom and Davis (2016) and their influential "Economic Policy Uncertainty" index. Their measure considers the frequency of newspapers articles that contain a group of terms for the economy, uncertainty and policy. Their empirical application shows that policy uncertainty is associated with greater stock price volatility, and anticipates investment, output and employment declines in the US. The index has since been widely expanded to other countries, policy areas and more sophisticated uses of natural language processing techniques (Azqueta Gavaldon, 2017; Baker et al., 2019; Husted, Rogers and Sun, 2020), as well as implemented across economic studies.[2] This NLP revolution in policy uncertainty has also considered other measures in addition to newspaper coverage. Hassan et al. (2019) use quarterly earnings conference calls, which allows for firm-level measures. Alternatively,

[2]See Al-Thaqeb and Algharabali (2019) for a review

Hansen and McMahon (2016) consider press releases from the FED, while Cruz Tadle (2021) analyze meeting minutes of the Federal Open Market Committee. The application of NLP techniques for measuring uncertainty has also expanded beyond economic uncertainty. Caldara and Iacoviello (2022) measures geopolitical uncertainty based on newspaper articles covering geopolitical tensions, while Mueller and Rauh (2018) calculates the probability of an outbreak of conflict using a similar approach. Closer to my approach, Calvo-González, Eizmendi and Reyes (2018) consider presidential speeches in Spanish speaking countries to proxy for policy volatility. To the best of my knowledge, my paper is the first to consider laws from different countries to develop a measure of legal uncertainty using NLP.

Finally, this article also contributes to the growing literature using NLP in the legal domain.[3] Many applications refer to text classification (Boella, Di Caro and Humphreys, 2011; Sulea et al., 2017) to determine what a text is about and feature extraction such as the identification of facts and principles within a legal text, which are useful for dealing with large amounts of documents. Alternatively, topic modelling has also been applied. The literature goes back to seminal work by Gerrish and Blei (2011), who use topic modelling from the text of US bills to forecast Congressional votes.[4] I build on this approach to study topic distributions directly from legal texts and across countries.

There have been some attempts at measuring legal uncertainty at the national level. An example is the Index of Legal Certainty from the Fondation Pour le Droit Continental (Deffains and Kessedjian, 2015), which uses survey based measures from legal experts to determine how accessible and stable the law is in a group of 13 countries. Alternatively, Djankov et al. (2003) construct a cross-country index of procedural formalism of dispute resolution by analyzing legal and court procedures to evict a tenant for non-payment of rent and to collect a bounced check. Still, these measures rely on subjective appraisals of the regulatory framework, and are not available on a periodic basis as they are one-time measured.

The use of newspaper data to construct uncertainty measures has also limitations. Newspaper coverage changes with political cycles (Le Moglie and Turati, 2019), and therefore the uncertainty measures derived may be biased by the electoral calendar. Moreover, newspapers are (mostly) for-profit companies who may want to maximize readership by providing coverage based on consumer preferences, thus over-representing easy-to-read news (Ho and Liu, 2015). Moreover, media companies also have their own political agendas, which impacts topic selection and presentation (Gentzkow and Shapiro, 2010; Mullainathan and Shleifer, 2005)

By considering the regulatory framework directly we can bypass these shortcomings and implement a more objective measure. This implies analysing the content of legal texts as given, considering laws as the direct source of policy. Designing a systematic measure of legal uncertainty would also allow for comparisons across time and countries, which is currently

[3] See a summary of the latest applications in Mumcuoglu et al. (2021)

[4] Other recent examples include Osnabrugge, Ash and Morelli (2021), who trained a machine classifier to learn topics from party platforms and use it to classify parliamentary speeches, and Vannoni, Ash and Morelli (2021) who measure discretion in legal texts and apply it to US states.

limited.

To calculate the LUI, I collect two datasets, a cross-country dataset and one for Chile that will serve as case study. For the cross-country data I obtain all legal texts for a group of 44 emerging and developed economies from all corners of the world, for the period of 1990 to 2020. These texts are the basis for the construction of the LUI, as well as additional text analysis measurements that will serve as controls, including length, complexity and sentiment of the wording. Results at the country level show great variability in legal uncertainty, with large swings in countries like Botswana, Estonia, Indonesia, Jamaica, Mexico, Poland, and Vietnam and more stables legal frameworks in Austria and Italy. To investigate these results further I compare the LUI with other institutional measures and find that the best performing countries in terms of institutional quality and Rule of Law, along with democracies and parliamentary systems, are all associated with lower legal uncertainty.

These comparisons are only available at the yearly level, as the cross-country dataset does not allow to identify the exact date of enactment. To go deeper in the understanding of the index I consider an individual country case study. Chile is an ideal choice given its rapid economic growth since the 1990s and the institutional transformation experienced since the end of Augusto Pinochet's dictatorship in 1990. For this I obtain all legal texts enacted in Chile from 1990 to 2021. Moreover, the data allows to identify the date of enactment, which enables the construction of the LUI on, for example, a monthly basis. The Chilean case shows that the topics covered in laws are somewhat stable across time, but the composition can have significant changes, particularly when major new pieces of legislation are enacted or new governments with innovative policy agendas take office. My results show that during the first quarter of each new administration there is a jump in the LUI, particularly during the center-left *"Concertación"* administrations.

The rest of this paper is organized as follows: Section 2 motivates and provides the methodological aspects of the index. Section 3 describes the data needed for the construction of the datasets. Section 4 presents the LUI index for the full dataset and Chile as case study, together with a comparison based on institutional measures. Section 5 offers the main application by analyzing the impact of the LUI on investment growth and unemployment. Finally, section 6 concludes.

2. Measuring Legal Uncertainty

2.1. Methodology motivation

Several alternatives have been proposed within the empirical literature to proxy for uncertainty, including macroeconomic and stock market volatility (Bloom, 2009), economic forecast disagreements (Bachmann, Elstner and Sims, 2013) and pure econometric exercises such as generalized ARCH applications that measure uncertainty on inflation and output by the respective conditional variances (Fountas, Karanasos and Kim, 2006). However, how appropriate these measures are depends on how correlated they are with the stochastic process. For instance, the volatility of stock markets can change even when the uncertainty regarding economic fundamentals is stable, for example under shifts in risk aversion.

Conversely, I consider the approach by Jurado, Ludvigson and Ng (2015) as an objective measure for uncertainty. In their framework, what matters for economic decision making is not whether an indicator has become more or less variable, but rather whether it has become more or less predictable. In this setting uncertainty is not the same as volatility of the variable.

When applying this approach to the legal sphere, I aim to capture uncertainty regarding what the legal framework is compared to what expectations agents had about that framework. Under this setting, economic agents make decisions based on the present and their expectation of the following period. I consider uncertainty as the part of the legal outcome that the agents are unable to forecast.

The challenge resides in measuring what that framework is at any given point in time. For this purpose I assume that laws enacted in a given country and period reflect the legal framework and the goal of policy makers. Therefore, I use the legal texts as the basis for modelling, and attempt at determining what there content is in terms of topics. To do so I apply NLP techniques to determine objectively what those legal texts cover and, more importantly, to quantify the salience of the topics in each period.

Topic models are statistical models that extract the main themes contained in large, unstructured collections of documents known as corpus. This is done by using statistical relationships between the terms in the documents. Topic models assume that there exist a number of K latent topics in a corpus of D documents with a total vocabulary size N. The objective is to find a topic distribution over each document and a term distribution over each topic. The models determine which words are most important for discriminating between topics instead of imposing this feature by choice. Topic modelling requires little knowledge about the texts themselves, as the analysis is purely probabilistic.

The classification of documents by topic allows to capture their content without making any prior assumption. Moreover, the output can be easily interpreted. In the particular case of legal texts, topic modelling allows to characterize the areas on which legislation has been enacted in a quantifiable way. The topic share distributions will be the input to construct my measure of legal uncertainty, by considering the shares as an objective proxy for content

To combine this with the approach by Jurado, Ludvigson and Ng (2015) I assume that economic agents make rational expectations about future legislation based on the content of past and present day laws. Since shares proxy for content, I assume the agents use them as input for forecasting in the following period. The uncertainty measure is then the unforecastable component of the legal content, as proxied by topic shares. In the following sections I describe the methodology in detail.

2.2. Implementation of Topic Analysis

NLP requires some handling of the data to enhance accuracy and reduce the dimensionality problem, for which I use standard Python packages. First I remove punctuation and special characters, common English language stop words such as "a" or "that" as well as numbers. I also set all words to lower case. Then I stem all words in each law using the Snowball algorithm from Porter (1980).[5] This ensures that the algorithm captures the lack of distinct meanings between different verb conjugations or words with common roots. For example, the words "economics", "economy" and "economized" would be stemmed to "economi" as one common token. Furthermore, I also consider groups of either two or three words combined, called bigrams and trigrams, such as "property tax" or "senate budget committee". I do this to ensure I capture the appropriate meanings of those terms, which differ to when they are used as single words. This improves the accuracy of term representation but at the same time increases the dimensionality of the model.

The most widely used topic model is Latent Dirichlet Allocation (LDA), first introduced by Blei, Ng and Jordan (2003). LDA “discovers” the topics in each document and the proportion in which the topics represent the documents. LDA considers documents to be random probability distributions over topics, and topics random probability distributions over words. Word order does not matter for LDA, as the algorithm uses term frequencies across the corpus in a bag-of-words approach. LDA is particularly useful method when we do not have any prior regarding the structure of the data, as the topics, the document topic distribution and the word topic distribution are latent structure.

Appendix A.1 presents the technical details of LDA. The most important choice by the researcher when implementing LDA is the number of topics. Still, there are some tools at our disposal to select this variable. I choose the number of topics that maximizes the Coherence Score suggested by Röder, Both and Hinneburgz (2015), which evaluates how semantically interpretable the given sets of words are as topics. The coherence score shows how well the terms support each other according to their similarity to all other terms within their topic.

Once the topics have been established I obtain the topic shares of each law, and then calculate the average share of each topic by country and period of time. In other words I calculate the mean of each topic share θ for each topic k in year t and country i for all laws D, as follows:

$$\theta_{k,i,t} = \sum_{d=1}^{D_{i,t}} \frac{\hat{\theta}_{d,k,i,t}}{D_{i,t}} \tag{1}$$

where $\hat{\theta}_{d,k,i,t}$ is the estimated share of topic k in document d at country i and year t

2.3. Index Construction

As mentioned, the main measure of legal uncertainty is based on the approach by Jurado, Ludvigson and Ng (2015). In their framework, uncertainty in period t is the unforecastable component of the future value of the series. More formally:

$$U_{i,t} = \sqrt{E_t[(\theta_{k,i,t+1} - E_t[\theta_{k,i,t+1}|I_t])^2]} \tag{2}$$

where the expectation $E(|I_t)$ is taken with respect to information I_t available at time t. The main rationale behind equation 2 is that uncertainty rises when the expectation today of

[5]The Python packages used for this can be found on https://snowballstem.org/

the squared error in forecasting increases, conditional on available information as of today.

The implementation of the uncertainty measure à la Jurado, Ludvigson and Ng requires replacing the conditional expectation in $E[\theta_{i,t+1}|I_t]$ by a forecast, from which I can obtain an estimate of the forecast error. A standard approach is to consider the r previous realizations $\theta_{i,t-r}$, thus estimating an AR process:

$$\theta_{k,i,t} = \beta_0 + \sum_{r=1}^{R} \beta_r \theta_{k,i,t-r} + \epsilon_{i,t} \tag{3}$$

where $\theta_{k,i,t}$ is a vector of size kx1 with the proportions of each topic as observations in year t and country i, and $\epsilon_{i,t}$ is the error term. Then the one period forecast is:

$$\hat{\theta}_{k,i,t+1} = \hat{\beta}_0 + \sum_{r=0}^{R-1} \hat{\beta}_r \theta_{k,i,t-r} \tag{4}$$

where $\hat{\beta}_r$ are the OLS estimates of β_r. The variance of the error term $\epsilon_{i,t}$ provides the variation needed to calculate the uncertainty index. This variance is measured across country and time as the AR process moves each period. Therefore, the variance is not measured across a single observation.

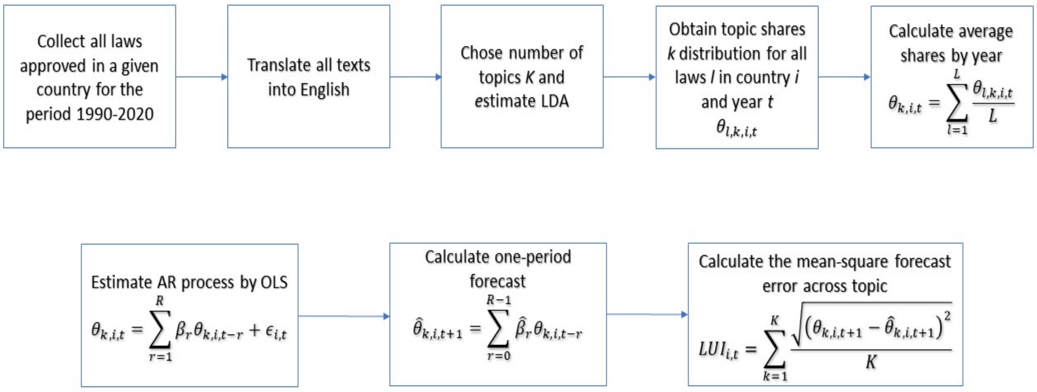
For the yearly-level index I assume an AR structure in the behavior of the $\hat{\theta}_{k,i,t+1}$ with $R = 3$ lags and forecast it considering all information available up to year t. I consider 3 lags under the assumption that a typical term in office in democratic countries is 4 years. All results in section 4 and 5 are unchanged to considering alternative AR processes with $R = 2$ or $R = 4$ lags respectively.

I then calculate the mean-square forecast error across topics to obtain a measure of uncertainty in country i and year t, as follows:

$$LUI_{i,t} = \sum_{k=1}^{K} \frac{\sqrt{(\theta_{k,i,t+1} - \hat{\theta}_{k,i,t+1})^2}}{K} \tag{5}$$

Figure 1 offers a schematic representation of the construction of the LUI step by step. This methodology allows to consider the variation in topic shares as input for forecasting, and links the output of the LDA model with Jurado, Ludvigson and Ng (2015) framework of measuring uncertainty.

Figure 1
LUI construction



Note: Schematic representation. Source: Author

3. Data description

The main data input for the construction of the Legal Uncertainty Index are legal texts, as enacted in chronological order in a given country. For the purpose of this study I construct two datasets, a cross-country dataset and an individual country level that will serve as case study.

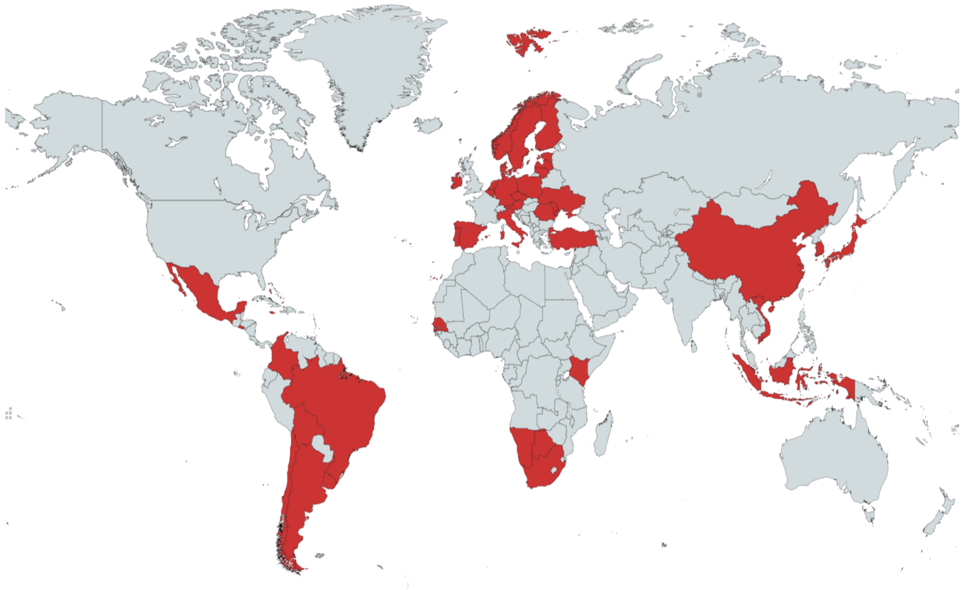
My main data set is the complete set of laws enacted by 44 countries for which I have access to, during the period 1990-2021. I obtain a total of 551,145 laws, available from from Global Regulation database.[6] For each law in country i and year t I have country of origin, year of enactment and title of the law and full text. The texts have been translated into the English language using machine learning, mostly from Microsoft's Translator Text API, although some documents were translated using Google. This will enable the implementation of a single LDA model for all countries.

I use the term Law broadly, as classified by Global Regulation. Admittedly, the definition will vary by country and will include laws, acts, constitutional amendments and code enactments, statutes and other regulations. I exclude case law, agency regulations and technical standards to ease comparison. Nonetheless, I cannot distinguish the laws by how enforceable they are or their hierarchical level within a country's legislation.

I restrict the sample of countries to those for which there are at least 300 laws in the considered period and for which the year of enactment could be identified. Figure 2 shows the

[6]See https://global-regulation.com/

Figure 2
Data availability



Source: Author based on data from Global Regulation

countries for which I obtained the full database of laws from Global Regulation.[7] I have good coverage of both Western and Eastern Europe and Latin America, and some other emerging economies in Africa and Asia-Pacific. Although time availability varies there is an average of 25 years of data per country.

I complement this data with basic annual economic statistics, including GDP, investment rate, total exports and imports, unemployment, real exchange rate, inflation rate and interest rate, all available from CEPII. To compare the Legal Uncertainty Index with other institutional quality measures, I also gather the Worldwide Governance Indicators available from the World Bank, complemented with institutional measures constructed by the Polity project and the Database on Political Institutions by Cruz, Keefer and Scartascini (2021). Finally, I also consider the World Uncertainty index (WUI) from Ahir, Bloom and Furceri (2022). This index is an adaptation of Baker, Bloom and Davis (2016) for a group of 143 countries, using the Economist Intelligence Unit country reports instead of newspaper articles. See a definition of each variable together with time availability and source in Table B.8 in the Appendix. Descriptive statistics are available in Table B.6.

For the individual case study I consider Chile. For this I obtain all laws approved in Chile from January 1990 to December 2021, available from the Library of the National Congress.[8] I gather 2,479 documents in total, for which I obtain date of enactment, title and main text

[7] Due to proprietary reasons Global Regulation limited the number of countries to which I could access the entire data set. In particular, although Australia, Canada, New Zealand, Pakistan, Russia, The Philippines, Sri Lanka, Switzerland, the United Kingdom and the United States meet the eligibility criteria and are available in the Global Regulation website for searching purposes they are not available for the purposes of this study

[8] See https://www.bcn.cl/leychile/Consulta/buscador_avanzada

of the bill in Spanish language.[9] I consider as law all documents classified as "*ley*" by the Library so as to avoid classification discrepancies.

I complement the legal data with monthly/quarterly economic statistics. At the monthly level I obtain the unemployment rate, available from the National Statistics Institute from January 1990. At the quarterly level I obtain the total investment growth rate and the gross domestic product, available from January 1990, all from Chile's Ministry of Finance. Finally, and as the main benchmark to compare the new LUI, I obtain the Economic Policy Uncertainty index à la Baker, Bloom and Davis, available for Chile on a monthly basis since 1993 onward, from Cerda, Silva and Valente (2016). [10]

Descriptive statistics are available in Table 1. There are on average 6.5 laws per month during the selected period, which provide enough monthly variation to construct the LUI. Still, some months do have fewer laws. Nonetheless, the EPU index also shows great variability in the selected period, which suggests policy uncertainty does vary significantly over time.

TABLE 1
DESCRIPTIVE STATISTICS - CHILE

	Mean	Std.Dev	Min	Max	N
Number of Laws	6.52	3.78	1.00	32.00	380
EPU	104.91	52.05	25.97	345.40	323
Investment Growth	2.62	15.30	-25.64	31.97	377
Inflation	0.44	0.60	-1.20	4.90	380
Unemployment	7.90	1.57	5.12	13.09	380
Fiscal Deficit	-0.06	1.36	-3.30	3.00	211
Stock Exchange Variation	0.35	4.82	-15.41	14.05	147
Business Trust Index	1.39	13.66	-41.10	31.80	209

Notes: Descriptive statistics for Chile for period January 1990 - December 2021 on a monthly basis (depending on availability).

4. The Legal Uncertainty Index and its Applications

4.1. Legal Uncertainty Across Countries

In this section I implement the Legal Uncertainty Index for all countries for which I gathered legal texts. The first step implies running the LDA model. Following Calvo-González, Eizmendi and Reyes (2018) I implement the LDA model by pooling all laws from all countries considered, to allow cross-country comparability. If we implement a separate LDA by country we cannot extract the same topics in each implementation. Conversely, by pooling all laws together we ensure that all documents are classified according to the same topics.[11]

[9] The website built-in searcher identifies 2,514 laws instead but there was no text available for 35 of those.

[10] See Table B.7 in the Appendix for detailed descriptions of the source and availability of each variable.

[11] The underlying assumption is that there is one distribution that produces all legislation in the world. Then each law in each country draws from a vector of topic proportions that are the same in all countries.

Having one LDA model is needed when doing the cross country analysis. If the topics are not comparable across country then the variation captured would not be either. Of course, this implies losing country specificities. Nonetheless, these should be captured by the shares themselves. For example, if a certain topic does not matter in a country then it will get a 0% share in that topic.

As discussed, when estimating an LDA model the most important choice of the researcher is the number of topics. The topic modelling presented considers 18 topics, which maximizes the Coherence score.[12] As a first analysis I make a qualitative interpretation of the topics by looking at words comprising each topic. Figure 3 presents the word clouds of each topic, where word size indicates how important that word is for the topic in question. We can make a direct inference of the topics by analyzing the word clouds. For example, topic 2 deals with private sector employment, while topic 5 is related to taxation and state funding. Topic 6 is related with the judicial system while topic 7 is about business regulation. Topic 10 accounts for national investment projects and topic 11 for international agreements. Finally, topic 17 si related to education and other social services.

Figure 3
Topic Words

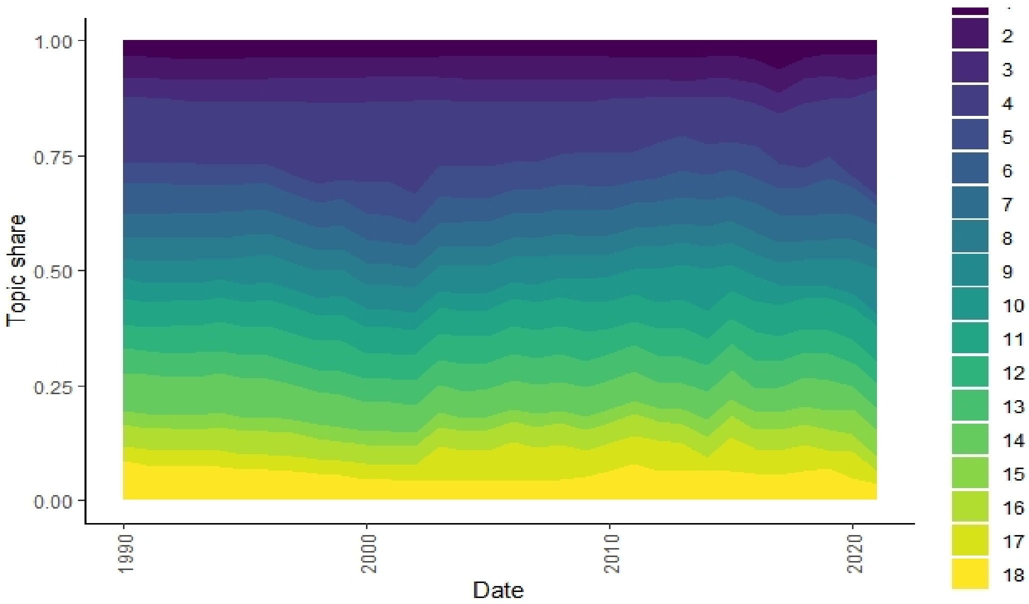
Source: Author based on data from Global Regulation

Next I calculate the average topic shares for country and year. In figure 4 I plot the average share of each topic for all countries combined, for the period 1990 to 2020. Topics

[12] Incidentally, this is also the number of topics used when performing topic modelling in US laws in Ash, Morelli and Vannoni (2022)

are relatively stable in the aggregate. However, at the individual country level there is much stronger variation in topic shares by year.

FIGURE 4
TOPIC SHARES



Source: Author based on data from Global Regulation

I now turn to construct a series of measures of legal uncertainty by exploiting the variation in topic shares by year in each country. The baseline LUI considers the baseline methodology with an AR(3) process for the forecast and 18 topics. Nonetheless, results are unchanged to considering different AR structures. The correlation to the LUI considering AR(2) and AR(4) processes is 0.97 and 0.98 respectively. Furthermore, the correlation to the LUI considering 15 and 21 topics is 0.85 and 0.97 respectively. Therefore, results are not driven by either the number of topics or the AR process considered for the forecast estimation.

As a first representation I show the main descriptive statistics of the LUI in each country. Results are available in Table 2. The table is indicative of the significant variation in legal uncertainty across countries. The worst performers in the sample are Botswana, Estonia, Indonesia, Jamaica, Mexico, Poland, and Vietnam, as measures by the average LUI index across time. Conversely, Austria and Italy show the most predictable legal environments, as shown by the lowest average LUI in the sample.

Table 2
LUI Statistics

	Mean	Standard Deviation	T
Antigua and Barbuda	3.59	2.50	25
Argentina	3.87	2.45	8
Austria	0.33	0.17	15
Belgium	3.03	4.40	22
Bermuda	0.71	0.76	27
Bolivia	0.88	1.24	24
Botswana	8.85	8.17	21
Brazil	1.09	1.16	17
Chile	0.74	0.96	27
China	0.97	1.03	18
Colombia	2.25	3.40	23
Czech Republic	1.45	1.87	25
Denmark	0.62	0.30	10
El Salvador	4.48	4.66	25
Estonia	10.86	14.48	17
Finland	6.11	6.30	29
Germany	1.26	1.02	27
Indonesia	2.43	3.26	25
Ireland	0.73	0.95	24
Italy	0.35	0.38	29
Jamaica	7.82	5.54	19
Japan	6.50	6.11	21
Korea	2.18	4.76	11
Latvia	0.53	0.30	22
Lithuania	1.70	0.94	7
Mexico	8.70	9.46	28
Namibia	1.63	0.93	7
Netherlands	1.42	2.23	29
Norway	2.92	2.60	29
Pakistan	2.91	2.25	22
Poland	9.21	5.98	23
Portugal	4.51	5.48	21
Romania	0.99	1.21	25
Senegal	1.68	1.72	15
South Africa	2.48	4.25	24
Spain	1.96	2.47	29
Sweden	1.68	1.09	29
The Bahamas	5.52	4.40	29
Turkey	1.85	1.51	20
Ukraine	0.57	0.40	28
Uruguay	1.40	3.54	28
Vietnam	7.37	9.03	23
Total	3.11	5.10	927

Notes: Main statistics for the LUI for period 1990 - 2020 on a yearly basis - by country.

To further investigate the LUI results by country I compare the index to other yearly measures of institutional quality. The polity variable from the Polity5 Project is the most popular measure of a country's political regime. The polity score consists of six component measures that record key qualities of executive recruitment, constraints on executive authority and political competition. It also records changes in the institutionalized qualities of governing authority. The score ranges from -10 (hereditary monarchy) to +10 (full democracy). Table 3 shows the mean LUI score by democracy status according to the polity score, as well as the standard error of the mean difference. Countries characterized as full democracies have less legal uncertainty than less democratic societies. To test for statistical significance I consider a t-test in mean difference, and find the difference statistically significant at the 1% confidence level.

Another measure of institutional quality refers to state capacity. For this I consider the measure of "Regulatory Quality" index available from the Worldwide Governance Indicators (WGI) at the World Bank. Regulatory quality captures perceptions of the ability of the government to formulate and implement sound policies and regulations that permit and promote private sector development. Therefore, as opposed to the LUI it is not a measure of regulatory uncertainty but rather of efficiency. I classify my sample of countries according to the Regulatory Quality index and consider the top 25 percentile as the best performers. These countries have less legal uncertainty than those that perform more poorly in Regulatory Quality, and the difference is statistically significant at the 1% level. Therefore, more regulatory efficiency is associated with less legal uncertainty.

Alternatively, the WGI also offer the Rule of law index, which captures perceptions of the extent to which agents have confidence in and abide by the rules of society, and in particular the quality of contract enforcement, property rights, the police, and the courts, as well as the likelihood of crime and violence. Therefore, the Rule of Law index captures how well is the legal framework enforced. The best performers in Rule of Law also have less legal uncertainty than those that perform more poorly, and the difference is statistically significant at the 1% level.

The institutional design of a country can also have an impact on the level of legal uncertainty. Seminal work by Linz (1994) argued that presidentialism was inherently more prone to democratic breakdown than parliamentarism. According to Linz (1994), presidentialisms lack parliamentarism's safety valve - the confidence vote - that allows for the removal of a government from office in case of crisis without disrupting the constitutional order, and also create the incentives and conditions that encourage such crises. Presidentialisms are open to political outsiders, a problem reinforced by the single-person nature of the office of presidency. Moreover, legislators under presidentialism are less inclined than under parliamentarism to support the executive, because lack of confidence does not jeopardize the survival of the government. This generates inter-branch conflict without the constitutional means for resolving them, particularly under divided government where one party holds the office of the presidency and another holds the Legislature. Repetto and Sosa Andrés (2022) show that under divided government there is an increase in interbranch conflict as well as more ideological polarization among politicians. Therefore, it is natural to compare the LUI under parliamentary versus presidential systems. For this I consider the classification of government systems available at the Database of Political Institutions. Table 3 shows that parliamentary systems have less legal uncertainty than presidential systems, and the difference is statistically significant at the 5% confidence level. Therefore, I find evidence in support of Linz (1994) thesis of the instability of presidential systems.

The DPI also includes the year a new government takes office, which I use to create a variable equal 1 if year t includes a change in government, and 0 otherwise. I find no impact of a new government in legal uncertainty. One possible explanation for this result is that the year aggregation of the data could hide the true variability in legislation from a new government that would be observed during the "honeymoon" period.

Alternatively, there could be differential effects coming from the ideological sign of the new government. To test this I consider the ideological classification of governments available at the DPI. I find no difference in LUI on left and right wing governments. However, I do find an impact in new governments when considering their ideological sign. A new government from the right is associated with an increase in legal uncertainty, while a new government from the left has the opposite effect. Therefore, when a right wing government takes office they implement legislative changes that are more significant compared to left wing governments. This results are somewhat surprising, as right wing governments are associated with stability compared to left wing governments (Bechtel, 2009). Still, recent work by Nguyen, Castro and Wood (2020) argues that, given the fact that right wing governments are more prone to promote economic freedom and freedom of trade they facilitate the contagion of crises, making them more likely. Nguyen, Castro and Wood (2020) find the probability of financial crises increases when right-wing parties are in office and within 1 year after elections. My results are aligned with these findings.

Finally, I consider the impact of the implementation of an International Monetary Fund (IMF) structuring plan. Since 1992, there have been more than 1,900 different funding plans within the IMF. The decision to engage in these depends on the incentives and constraints faced by national governments, who request and implement them. They are associated with significant policy implementations and can have long term impact in the requesting governments. Abad et al. (2022) show that new governments are more prone to request IMF assistance and implement a structuring plan, as they can more easily attribute the need for assistance to their predecessors, thus reducing the political capital cost of such reforms. I compare the LUI on years where an IMF structuring plan is implemented, but find no difference across mean legal uncertainty. However, when considering the ideological sign of the government I find that a new IMF plan from a right-wing party is associated with an increase in legal uncertainty. Therefore, right-wing governments are associated with more unstable legislation on years where a restructuring program is in place, which suggests that these programs might be more radical under right-wing governments.

TABLE 3
LUI COMPARISONS

	Mean 0	Mean 1	Std. Err	P-Value	N
Full Democracy	0.069	0.051	0.007	0.006	785
Regulatory Qual	0.065	0.045	0.008	0.007	750
Rule Law Index	0.067	0.039	0.008	0.000	750
Parliament Syst	0.072	0.058	0.007	0.031	832
Change in Gov	0.066	0.059	0.009	0.208	831
Left Wing Gov	0.069	0.061	0.008	0.150	677
New Left Gov	0.068	0.039	0.014	0.022	677
IMF Structure	0.063	0.067	0.012	0.349	927
IMF Strt. Right	0.064	0.089	0.015	0.049	677

Notes: Mean LUI by group, standard errors of the difference in means and p-values for the null hypothesis that mean difference is 0 using a t-test. Baseline LUI measure considering forecast error measure a la Jurado, Ludvigson and Ng and AR(3) process. Full Democracy equals 1 if the polity score equals 10 for country i and year t, and 0 otherwise. Regulatory qual equals 1 if country i is in the 25 percentile in the overall score of the Regulatory Quality index from WGI. Rule of Law Index equals 1 if country i is in the 25 percentile in Rule of Law index from WGI. Parliament System equals 1 if country i has a parliamentary system of government. Change in Government equals 1 if there is a new government in country i and year t. Right and Left Wing Governments equal 1 if the government is classified ideologically as right or left. New Right Government and New Left Government are the interactions of Change in Government and Right and Left Wing Government, respectively.

The institutional comparisons presented can only be performed at a yearly level, given the nature of the full dataset obtained for the LUI. To further investigate the LUI and learn more about the behaviour of the index I analyze a specific country. I select Chile as the test run for the index for a variety of reasons. Chile is en emerging economy that has experienced relatively rapid economic growth since 1990 together with an institutional transformation from an authoritarian regime under dictator Augusto Pinochet to a full democracy with party alternation in power. Therefore it serves as a perfect laboratory to study how changes in the institutional environment can impact decision making and the overall economy.

Furthermore, legal texts are readily available for the entire period of 1990-2021, including date of enactment of each law. This enables me to construct the LUI on a monthly basis. Finally, Chile also allows to consider the EPU index as a benchmark to which compare the performance of the LUI.

4.2. Case Study: Chile

The first step in the implementation of the LUI implies running the LDA model on the full corpus of laws enacted in Chile from 1990 to 2021. As a starting point I again consider 18 topics. Then I run the LDA and obtain the topic distribution of each law in each month.

Figure 5 presents the word cloud representation of each topic. This visualization of topics gives some insights on what the legislation in Chile was about in this period. For example, Topic 1 deals with taxation and contracts. Topic 4 is related to public administration while Topic 6 refers to social services including health and education. Topic 10 refers to international agreements, while Topic 12 is related to commercial treaties. Finally Topic 17 deals with public finance while Topic 18 is related to transportation and the environment.

FIGURE 5

TOPIC WORDS FOR CHILE'S LAWS: 1990- 2021

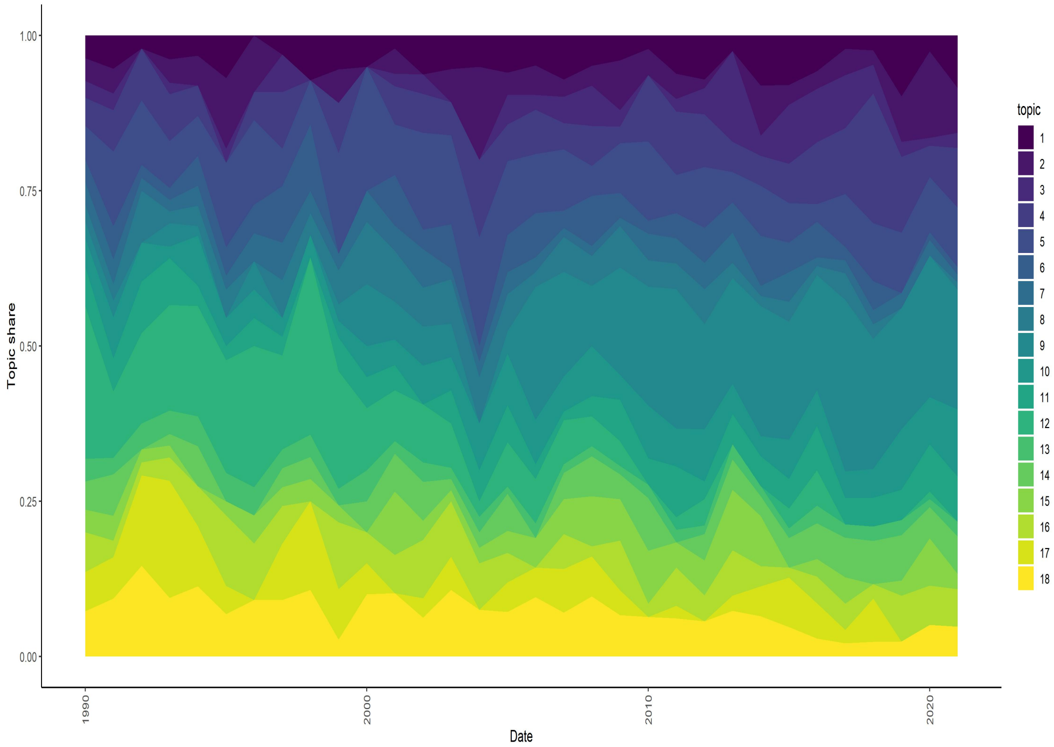
Source: Author based on data from Chile's Library of the National Congress

One quick way to visualize how they change is to plot the average topic shares. Figure 6 shows the evolution of the average topic shares for all the laws approved in Chile in a given year. Even aggregating shares by year we can see some variation in the topics covered by laws in Chile, as the prominence of topics changes every year. For example Topic 8, which deals with access to public information, was almost negligible before 2008. In August of that year

Chile approved Law number 20,285 which regulated the basic principles of transparency in public service. Since then the topic has become very prominent in the country's legislation.

Figure 6

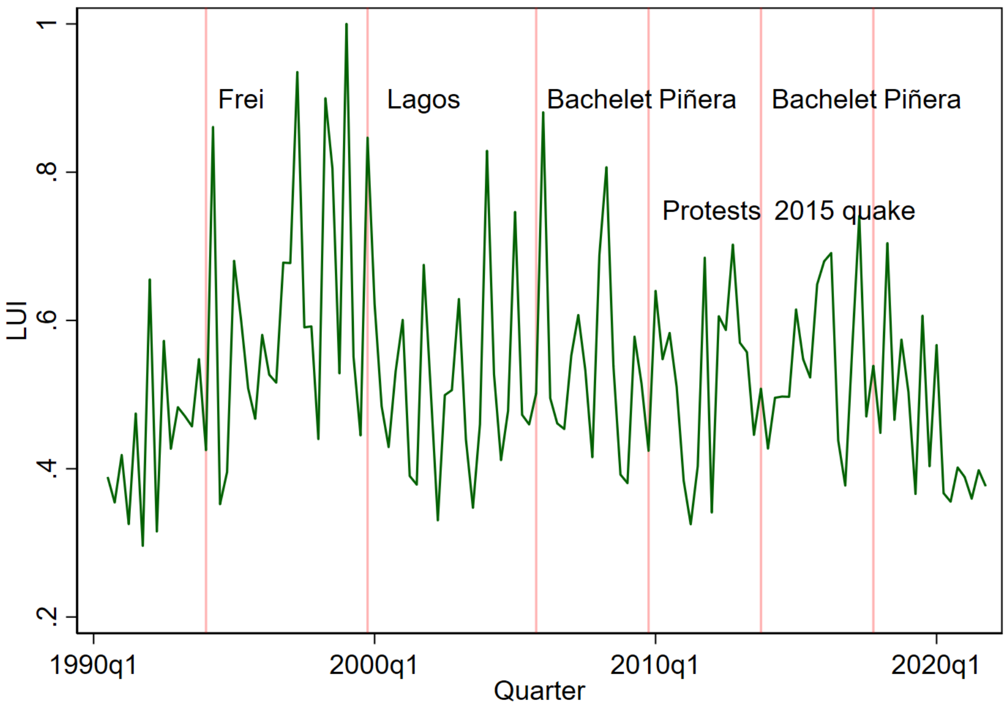
Annual Average Topic Shares for Chile's Laws: 1990- 2021



Source: Author based on data from Chile's Library of the National Congress

With the average topic shares by month I can now construct the LUI. Figure 7 plots the index from January 1990 to December 2021. Results in Figure 7 also show significant jumps in the LUI during each new administration and the massive protests under the first administration of Sebastián Piñera, as well as a spike after the devastating quake in 2015.

Figure 7
LUI for Chile: January 1990 - December 2021



Notes: Legal Uncertainty Index for Chile on a monthly basis. Vertical lines indicate new presidential administration. *Source:* Author based on data from Chile's Library of the National Congress

Chile normally ranks high across most measures of presidential strength, a featured that has been well documented (Alemán and Navia (2009)). The president has proactive powers like agenda setting, a monopoly power on bills regarding budgeting, spending and taxes, decree powers, and the power to introduce amendments at any point in the legislative process. The president also has reactive powers, such as partial, amending, and absolute veto powers (Mimica, Navia and Osorio, 2022). The majority of bills passed in Chile originate at the President's office (Mimica, Navia and Osorio, 2022). Therefore, it is natural to analyze how bill introduction and passage changes under a new presidential administration.

Figure 7 plots the changes in presidential administrations with red vertical lines. In the first quarter of each new administration there is a jump in the LUI index, specially during the center-left *"Concertación"* administrations from 1994 to 2010. The jump moderates from 2010 onward when conservative Sebastián Piñera took office but is still present, particularly in his second term in 2018. Therefore the LUI captures a change in policy implementation with every new administration.

Both Mimica, Navia and Osorio (2022) and Alemán and Navia (2009) show that presidential bills introduced early in the term during the "honeymoon" period were more likely to pass. Therefore, we should expect an increase in legal uncertainty when a new president takes office, as the honeymoon period is ideal for the president to set their agenda. Results in Figure 7 show the new laws differ significantly from those approved at the end of the previous administration. As such, the new legislation generates a jump in legal uncertainty at the start of the new government, given the variance in law content. This finding also reinforces the notion that the political climate enjoyed during the beginning of an administration provides particular advantages for presidents.

The variation in topics captured by the LUI could be susceptible to the number of topics considered. As a robustness test I consider the LUI with 15 and 21 topics respectively. The correlations with these specifications are 0.82 and 0.83 respectively. I conclude that the index is not impacted by the choice of topics in the LDA specification. For the monthly data I assume an AR structure in the behavior of $\hat{\theta}_{k,t+1}$ with 6 lags. As additional robustness tests I consider an AR with 3 and 9 lags respectively. The correlations of the LUI with these specifications are 0.91 and 0.85 respectively. Therefore, results are also unaffected by these changes in AR structure.

5. Legal uncertainty and the economy

5.1. The Chilean Case - VAR Analysis

To investigate whether legal uncertainty matters for economic outcomes, I analyze both the case study of Chile as well as the cross-country dataset, in separate studies.

As discussed in section 1, policy uncertainty makes firm delay investment and hiring (Bernanke, 1983; Bloom, 2014). Policy uncertainties can increase the risk premiums in financial markets, thereby increasing borrowing costs, undercutting productivity, and slowing employment, ultimately resulting in poor economic prospects (Gilchrist, Sim and Zakrajšek, 2014). Therefore, uncertainty has a counter-cyclical relationship with the business cycle: it bottoms out during booms and peaks during recessions (Bloom, 2014). Furthermore, an increase in policy uncertainty will have a long-term impact on capital stocks.

Under the assumption that laws are an instrument of public policy, sudden and unexpected changes in legislation can be considered as an additional source of public policy uncertainty. This leads to two testable hypothesis regarding legal uncertainty:

Hypothesis 1: An increase in legal uncertainty in a country in a given period is associated with a reduction in investment

Hypothesis 2: An increase in legal uncertainty in a country in a given period is associated with an increase in unemployment

For the Chilean case I consider macro data in a VAR analysis, following the methodology by Baker, Bloom and Davis (2016) and Husted, Rogers and Sun (2020). These models exploit time-series variation at the country level, and are useful for characterizing dynamic relationships. VAR models allow to capture potential channels through which legal uncertainty may impact, albeit with limited assurance of causal interpretation.

I start by fitting a VAR to monthly Chilean data. As the main dependent variable I consider national quarterly investment growth, extrapolated monthly. To recover orthogonal shocks, I use a Cholesky decomposition with the order LUI and the quarterly growth in investment. The ordering implies that the contemporaneous values of LUI_t impact investment growth I_t, but not vice-versa. The underlying assumption is that the legislative process is slower than firm decisions, and does not react to investment shocks. Following Baker, Bloom and Davis (2016) the baseline VAR specification includes three lags of all variables and time trends, and also six lags as robustness test. The system of equations fitted is as follows:

$$X_t = \alpha + \Gamma_1 X_{t-1} + \Gamma_2 X_{t-2} + \Gamma_3 X_{t-3} + \epsilon_t \tag{6}$$

with

$$X_t = \begin{pmatrix} I_t \\ LUI_t \end{pmatrix}$$

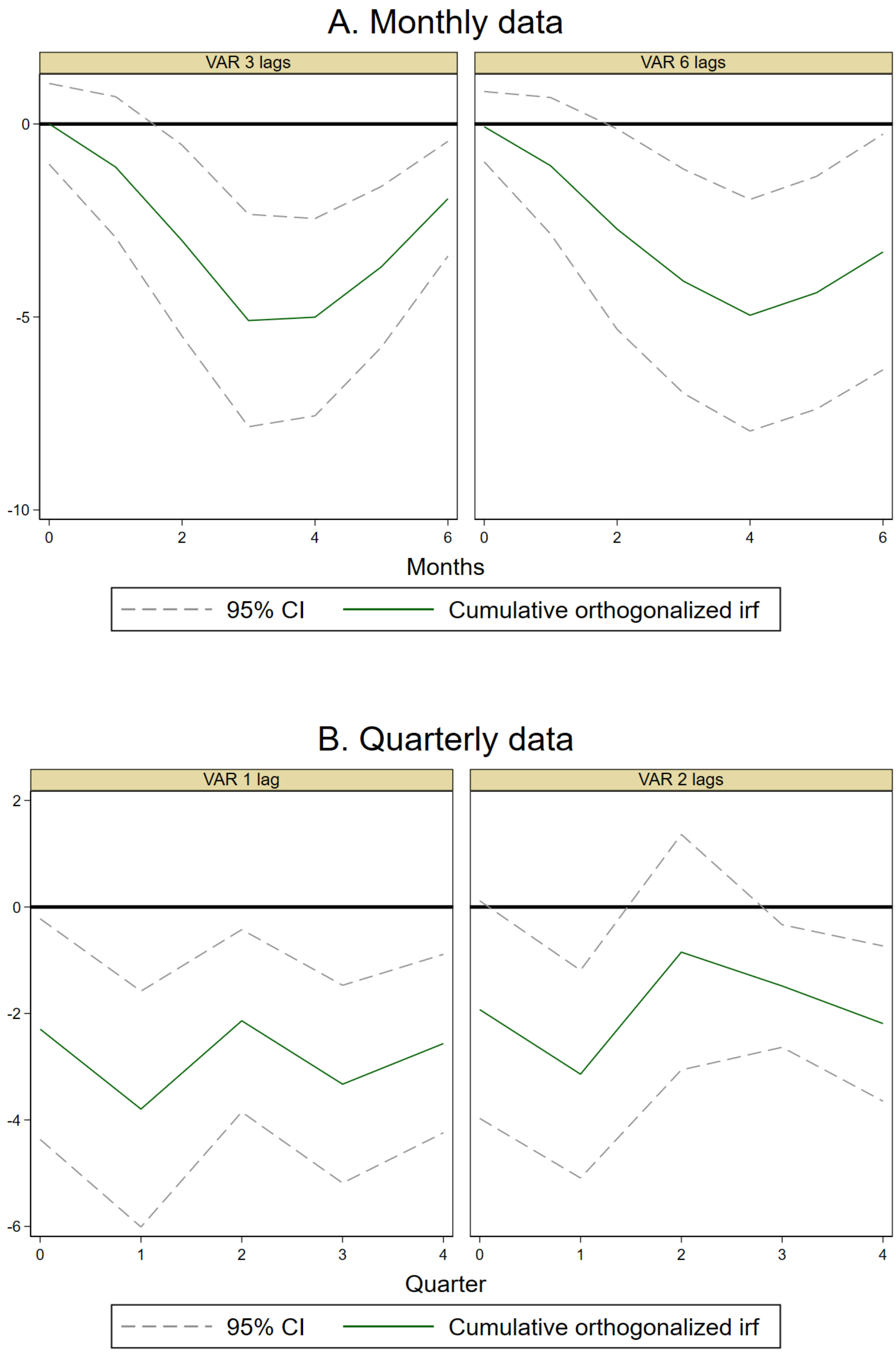
Panel A of Figure 8 depicts the model-implied responses of growth in investment to a one standard deviation positive innovation in the LUI, following the model in equation 6 and monthly data. The response to the LUI shock is negative and statistically significant. Therefore, a positive shock in legal uncertainty reduces investment growth up to the following six months, by a magnitude of about 5% of the mean quarterly growth, which represents a sizable effect. This is the first evidence we have that legal uncertainty impacts economic outcomes by negatively affecting investment decisions.[13]

Alternatively to the monthly extrapolation I also estimated an analogous VAR model on quarterly data, using the same type of Cholesky decomposition to identify shocks, and 1 and 2 lags (equivalent to 3 and 6 lags in the monthly model). Results are robust, as seen in Panel B of Figure 8.

[13]Estimation of the VAR requires that the variables are covariance stationary, as a regression on two random walk processes would yield a significant coefficient in a spurious regression (Granger and Newbold, 1974). An augmented Dickey-Fuller test shows that none of the variables have a unit root at the 1% confidence level. Next, I test for stability. All the eigenvalues lie inside the unit circle, and therefore the VAR satisfies stability conditions. Finally, I test for Granger causality on LUI and investment growth. For these I perform a Wald test on the null hypothesis that the coefficients for the lags of LUI are zero in the equation for investment growth. The p-value is 0.002, so I reject the null and find evidence that the LUI Granger-causes investment growth.

Figure 8

VAR-estimated Impulse Response Functions for Chile - Investment Growth



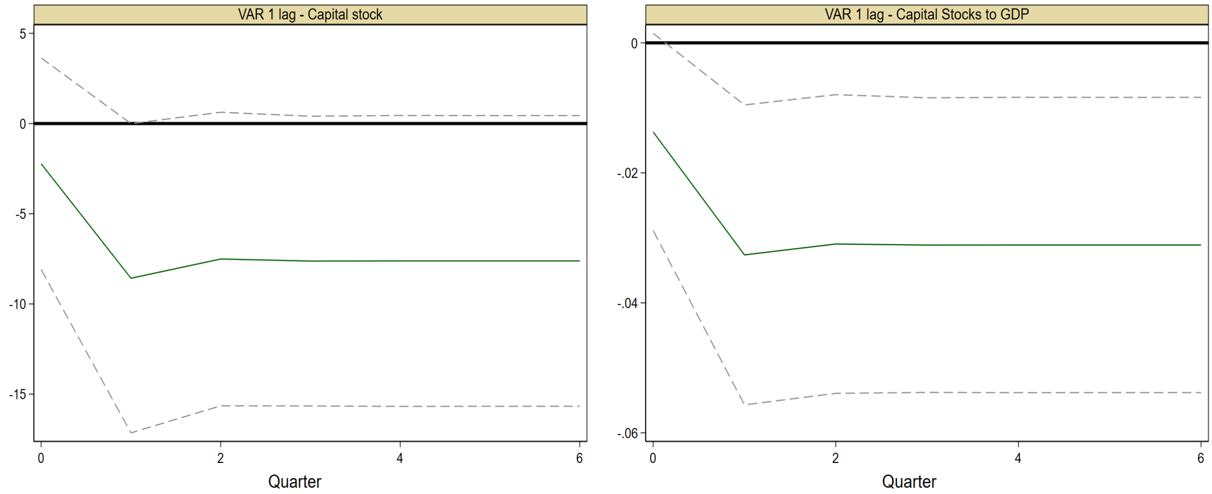
Cumulative IRF for investment growth to an innovation of one standard deviation in the LUI, with 95% confidence intervals. Identification considering three months (left chart) and six months (right chart) lags, time trends and Cholesky decomposition with the following order: LUI and investment growth. Fit to monthly data (Panel A) and quarterly data (Panel B) from January 1990 to December 2021.

As the main robustness test I include the EPU index, available from Cerda, Silva and Valente (2016) on a monthly basis from January 1993. The rationale is to show how the LUI captures variation in uncertainty that is beyond the economic policy uncertainty captured in the EPU, and serves as validity and test of innovation of the LUI. I use a Cholesky decomposition with the order LUI, EPU and the quarterly growth in investment. Results are available in Figure B.2 in the Appendix. Again, the response to the LUI shock is negative and statistically significant while there is no impact from the EPU shock.

For an additional robustness test I consider the LUI with 15 and 21 topics respectively, and implement the same VAR specification as before. The impulse-response functions are available in Figure B.3 in the Appendix. Results are robust to either specification of the LUI, and I conclude that results are not affected by the choice of topics in the LDA. Finally, results are also robust to the LUI specification constructed with AR process with 3 and 9 lags, as seen in Figure B.4 in the Appendix. Therefore, the AR process estimated to compute the forecast error in each period within the LUI calculation does not affect the overall results.

Results in Figure 8 suggest that a shock in legal uncertainty reduces the growth rate of investment, at least temporarily. This implies that capital accumulations should also be curtailed, as the growth rate trajectory of capital changes in response to the uncertainty shock. To analyze this I fit a VAR with quarterly measured capital stocks. To recover orthogonal shocks, I use a Cholesky decomposition with the order LUI and the quarterly capital stocks. Figure 9 presents the response of capital stocks to a one standard deviation positive innovation in the LUI. Panel A considers gross capital stocks and panel B the ratio of capital stocks to GDP. Results show that a 1% positive shock in legislative uncertainty reduces the capital stock by about 0.03% in the capital to GDP ratio, a small yet persistent effect.

Figure 9
VAR-estimated Impulse Response Functions for Chile - Quarterly Capital Stocks



(a) Capital stocks

(b) Capital stocks to GDP

Notes: Cumulative IRF for capital stocks to an innovation of one standard deviation in the LUI index, with 95% confidence intervals. Identification based on 1 quarterly lag, time trends and Cholesky decomposition with the following order: LUI and capital stocks. Fit to quarterly data from January 1990 to December 2021.

Next I estimate a VAR for unemployment, so as to understand an additional channel through which legal uncertainty can impact economic outcomes, following Hypothesis 2. The VAR specification follows equation 6, where y_t is the unemployment rate in month t. Again, to recover orthogonal shocks I use a Cholesky decomposition with the order LUI and unemployment, assuming that the legislative process is slower than firm decisions and does not react to contemporaneous (same month) employment shocks. The baseline VAR includes three lags of all variables and time trends, and also six lags as robustness test.[14]

Results of the estimation are available in Figure 10. The graph shows the impulse-response function to a one-standard deviation innovation in the monthly LUI. The response to the LUI shock is positive and statistically significant. A shock equivalent to one standard deviation of the LUI index is associated with an increase in the monthly unemployment rate of 0.2%, or 2,5% of the mean monthly rate, a modest effect.

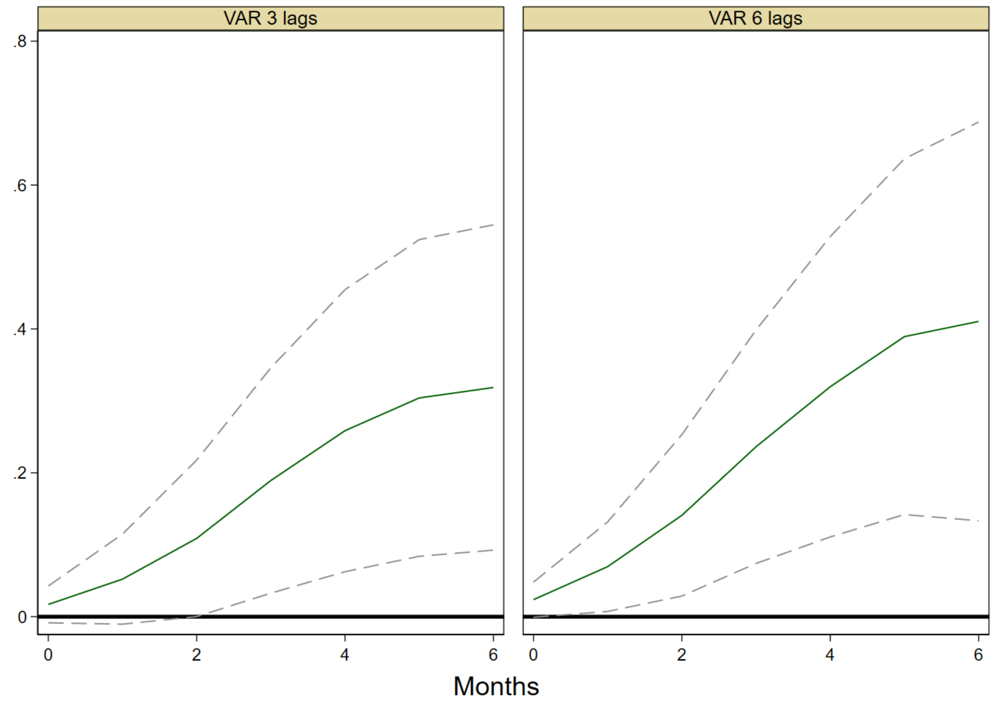
Again, there results are robust to including the EPU index as additional dependent variable (see Figure B.5 in the Appendix). They are also robust to considering the LUI specification constructed with an LDA using 15 and 21 topics respectively (see Figure B.6), and to the LUI computed with the forecast error based on an AR process estimation with 3 and 9 lags, although estimates are more imprecise in this case (see Figure B.7).

One final concern regarding the LUI is that it may be capturing variation that stems from other aspects of legislative texts. The index may be affected by a larger number of laws

[14]Once again, the validity of the VAR model relies on the variables being covariance stationary. The augmented Dickey-Fuller test shows that none of the variables have a unit root at the 5% confidence level or better.

Figure 10

VAR-estimated Impulse Response Functions for Chile - Unemployment



Cumulative IRF for unemployment to an innovation of one standard deviation in the LUI, with 95% confidence intervals. Identification based on three lags (left chart), six lags (right chart) and time trends and Cholesky decomposition with the following order: LUI and unemployment. Fit to monthly data from January 1990 to December 2021

or lengthier and more complex laws being enacted, as there is a greater chance that they will cover more topics compared to fewer and shorter laws. The concern is that more laws and more complex texts may imply bigger or more innovative policy changes. In section B.3 in the Appendix I consider additional controls for text measures on length and complexity. Results for the impact of the LUI on both investment growth and unemployment are robust to the inclusion of these measures.

Based on the results presented I conclude that, in the event of an increase in legal uncertainty, Chilean firms react by slowing down investments and laying off workers and/or not hiring new ones entering the workforce. These results are robust to a wide set of specifications.

5.2. Cross country study - Panel specification

The cross-country variation in investment activity is truly remarkable. For the 30-year period between 1990 and 2020, the rate of gross fixed capital formation worldwide ranged from 1 to 90 percent of GDP, a variance more than two times that of economic growth. Financial development and institutional quality are robust determinants of investment (Lim (2014) and the literature cited therein). For the cross-country analysis I consider a linear model with panel data. I extend the specification by Lim (2014) to consider legal uncertainty as a new measure of institutional quality. For this I estimate the following equation:

$$y_{it} = \alpha + \beta LUI_{it} + \gamma Z_{it} + \eta_i + \epsilon_{it} \quad (7)$$

Where y_{it} is the annual investment growth in country i and year t, LUI_{it} is the Legal Uncertainty Index, Z_{it} are country controls, and η_i are country fixed effects. I estimate equation 7 with ordinary least squares.

Results are available in Table 4. Column 1 presents the results with country fixed effects and no controls. The beta estimate is negative and statistically significant at the 5% level. The LUI has been normalized to 1 so we can interpret the coefficient directly. A 1% increase in the LUI is associated with a 0.100% reduction in the annual national investment growth. Adding year fixed effects to account for time trends has limited impact on the beta estimate of the LUI, as seen in Columns 2. Column 3 includes controls, for which I consider Lim (2014).[15] The beta coefficient estimate of the LUI index is robust to their inclusion. Finally in Column 4, I consider the World Uncertainty index from Ahir, Bloom and Furceri (2022). Results for the LUI are robust to the inclusion of the WUI. The WUI itself enters positively but non statistically significant in the regression.

TABLE 4

EFFECT OF LEGAL UNCERTAINTY ON INVESTMENT GROWTH (FIXED CAPITAL FORMATION)

	(1) Fixed Cap. Form	(2) Fixed Cap. Form	(3) Fixed Cap. Form	(4) Fixed Cap. Form
LUI	-0.100**	-0.080**	-0.114**	-0.090*
	(0.039)	(0.033)	(0.047)	(0.053)
WUI				0.003
				(0.057)
Mean of dep.var.	0.04	0.04	0.03	0.04
R2	0.01	0.19	0.32	0.00
Obs.	828	828	249	778
Country FE	YES	YES	YES	YES
Year FE	NO	YES	NO	NO
Macro Controls	NO	NO	YES	NO

Notes: Cluster-robust standard errors at the country level in parentheses. Baseline LUI measure considering forecast error measure a la Jurado, Ludvigson and Ng and AR(3) process. Investment growth is calculated as the annual variation rate in gross fixed capital formation. For Macro controls I follow Lim (2014) and include annual GDP, GDP growth, trade openness ratio, interest rate, financial freedom index, investment freedom index, property rights index, corruption index, tax revenue as percentage of GDP and credit to private sector as percentage of GDP. WUI index available from Ahir, Bloom and Furceri (2022)

$^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$

[15] See Table B.5 in the Appendix for detailed inclusion of controls

Therefore, I find additional evidence that an increase legal uncertainty correlates with a reduction in investment growth. Under a more uncertain legal framework companies across the economy react by curtailing investment plans, which generates a fall in the national investment aggregate. This is more evidence in support of Hypothesis 1, and provides the first application of the cross-country LUI at a macroeconomic level.

Results are robust considering alternative measure of investment growth. In particular, when considering the growth rate of capital formation, capital formation over GDP and fixed capital formation over GDP we obtain comparable results (see Table B.1 in the Appendix). Moreover, results are also robust to considering alternative specifications of the LUI. I consider both the LUI estimated with an AR(2) and AR(4) process for the forecast error calculation (see Table B.2 in the Appendix) as well as topic modelling with LDA and 15 and 21 topics respectively (see Table B.3).

An additional concern refers to the possibility that the LUI may capture uncertainty due unexpected outcomes in an election, which can potentially yield an unexpected government with a new policy agenda. To control for this I perform two additional controls. First I estimate equation 7 but dropping all observations with years where there is a change in government. As a second test I estimate equation 7 including country-election year fixed effects. Results are available Table B.4 in the Appendix, where I find that the impact of the LUI over investment growth remains negative and statistically significant, albeit smaller. Therefore, although it is plausible that variations in the LUI do capture some uncertainty about election outcomes, they are only partially responsible for the variation seen in the index.

Next I now turn to testing Hypothesis 2, under which an increase in legal uncertainty is associated with an increase in the national unemployment rate. The rationale was that the increase in risk premiums associated with legal uncertainty not only slow down investment plans by companies but also reduce hiring, as firms delay costly decisions until the uncertainty is resolved. The Chilean case showed how a shock to legal uncertainty increased the monthly unemployment rate, and the shock was persistent for up to six months.

For the cross-country analysis I estimate a linear model as in equation 7 with the annual change in the national unemployment rate as dependent variable. I extend the institutional-based specification by Baccaro and Rei (2007) to consider legal uncertainty as an additional determinant of the change in unemployment. Results are available in Table 5. Column 1 presents the results with country fixed effects. The beta estimate is positive and statistically significant at the 1% level. A 1% increase in the LUI is associated with a 0.23% increase in the change in national unemployment. Nonetheless, the effect disappears when considering the macroeconomic controls (Column 3). Therefore, the LUI is only weakly associated with the unemployment rate, and the effect is dampened when considering the macroeconomic determinants of unemployment.

TABLE 5
EFFECT OF LEGAL UNCERTAINTY ON UNEMPLOYMENT

	(1) Unemployment	(2) Unemployment	(3) Unemployment
LUI	0.229***	0.206***	0.055
	(0.082)	(0.076)	(0.110)
Mean of dep.var.	0.02	0.02	0.02
R2	0.00	0.10	0.15
Obs.	875	875	488
Country FE	YES	YES	YES
Year FE	NO	YES	YES
Controls	NO	NO	YES

Notes: Cluster-robust standard errors at the country level in parentheses. Baseline LUI measure considering forecast error measure a la Jurado, Ludvigson and Ng and AR(3) process. Change in unemployment is calculated as the annual variation rate the national unemployment rate. For macroeconomic controls I follow Baccaro and Rei (2007) and include annual GDP growth, inflation rate, interest rate and real exchange rate.
$^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$

6. Discussion

In this paper I provide a new measure of legal uncertainty that can complement other established measures of both policy uncertainty and institutional quality. I argue that the LUI captures variation not currently properly reflected in available measures and can be an additional tool for economic analysis. The proposed methodology incorporates tools from machine learning and NLP that are becoming increasingly popular in social sciences. The methodology is straightforward and can be replicated to other regions and time periods. In this regard, I show how legislative texts can be used to deepen our understanding the impact of the legal environment on economic outcomes.

Results show that legal uncertainty varies significantly by country and year. Enhanced Rule of Law, democracy and parliamentarism are all associated with less legal uncertainty. When considering the economic consequences, I find evidence that legal uncertainty is associated with a reduction in investment growth, under a robust set of specifications, providing evidence in favor of Hypothesis 1. These results provide an additional channel through which policy uncertainty increases the risk premium and forces firms to delay investment decisions. These results are broadly consistent with theories that highlight negative economic effects of uncertainty shocks. The evidence also shows that unemployment increases, although results are more imprecisely estimated and less stable across specifications. This is in line with findings on the limited impact of institutional design on unemployment, as previously noted by Heimberger (2021), and offers only partial evidence in favor of Hypothesis 2. This suggests that firms may also react by curtailing hiring and laying off workers in case of unexpected legal uncertainty shocks.

Future research ought to consider additional elements of the legislative process beyond the texts of enacted laws. Parliamentary discussion records, the early drafts of laws and policy papers are additional important sources of data that can impact economic agents, as they incorporate new information in their decision making. Furthermore, court rulings are also a potential source of data, particularly in countries where jurisprudence sets precedent and has implications beyond specific court decisions. The vast availability of additional data makes for an exciting avenue to pursue research.

References

Abad, Jose María, Vicente J. Bermejo, Felipe Carozzi, and Andrés Gago. 2022. "Blaming Your Predecessor: Government Turnover and External Financial Assistance." *Available at https://www.dropbox.com/s/6vs85d2lqx03r0e/Abad_Gago_et_al.pdf?dl=0*.

Ahir, Hites, Nicholas Bloom, and Davide Furceri. 2022. "The World Uncertainty Index." *NBER Working Paper*, 29763.

Ajina, Aymen, Mhamed Laouiti, and Badreddine Msolli. 2016. "Guiding through the fog: Does annual report readability reveal earnings management?" *Research in International Business and Finance*, 38: 509–516.

Alemán, Eduardo, and Patricio Navia. 2009. "Institutions and the Legislative Success of 'Strong' Presidents: An Analysis of Government Bills in Chile." *The Journal of Legislative Studies*, 15(4): 401–419.

Al-Thaqeb, Saud Asaad, and Barrak Ghanim Algharabali. 2019. "Economic policy uncertainty: A literature review." *The Journal of Economic Asymmetries*, 20.

Ash, Elliott, Massimo Morelli, and Matia Vannoni. 2022. "More Laws, More Growth? Evidence from U.S. States." *Bocconi Working Paper Series*, 178.

Azqueta Gavaldon, Andres. 2017. "Developing news-based Economic Policy Uncertainty index with unsupervised machine learning." *Economics Letters*, 158: 47–50.

Baccaro, Lucio, and Diego Rei. 2007. "Institutional Determinants of Unemployment in OECD Countries: Does the Deregulatory View Hold Water?" *International Organization*, 61: 527–569.

Bachmann, R., S. Elstner, and E. R. Sims. 2013. "Uncertainty and economic activity: Evidence from business survey data." *American Economic Journal: Macroeconomics*, 5(2): 217–249.

Baker, Scott R, Nicholas Bloom, and Steven J Davis. 2016. "Measuring economic policy uncertainty." *The Quarterly Journal of Economics*, 131(4): 1593–1636.

Baker, Scott R., Nicholas Bloom, Steven J. Davis, and Kyle J. Kost. 2019. "Policy news and stock market volatility." *NBER Working Paper Series 25720*.

Bechtel, Michael M. 2009. "The Political Sources of Systematic Investment Risk: Lessons from a Consensus Democracy." *The Journal of Politics*, 71(2): 661–677.

Bernanke, Ben S. 1983. "Irreversibility, uncertainty, and cyclical investment." *The Quarterly Journal of Economics*, 98(1): 85–106.

Blad, Johan, and Karin Svensson. 2020. "Exploring NMF and LDA Topic Models of Swedish News Articles." *Uppsala University Masters Thesis*.

Blei, David M., Andrew Y. Ng, and Michael I. Jordan. 2003. “Latent Dirichlet Allocation.” *Journal of Machine Learning Research*, 3: 993–1022.

Bloom, Nicholas. 2009. “The Impact of Uncertainty Shocks.” *Econometrica*, 77(3): 623–685.

Bloom, Nicholas. 2014. “Fluctuations in uncertainty.” *Journal of Economic Perspectives*, 28(2): 153–76.

Boella, Guido, Luigi Di Caro, and Llio Humphreys. 2011. “Using classification to support legal knowledge engineers in the Eunomos legal document management system.” *In: Fifth international workshop on Juris-informatics (JURISIN)*.

Bonaime, Alice, Huseyin Gulen, and Mihai Ion. 2018. “Does policy uncertainty affect mergers and acquisitions?” *Journal of Financial Economics*, 129: 531–558.

Caldara, Dario, and Matteo Iacoviello. 2022. “Measuring Geopolitical Risk.” *American Economic Review*, 112(4): 257–300.

Calvo-González, Oscar, Áxel Eizmendi, and Germán Reyes. 2018. “Winners Never Quit, Quitters Never Grow: Using Text Mining to Measure Policy Volatility and Its Link with Long-Term Growth in Latin America.” *Policy Research Working Paper*, 68.

Carlsson, Mikael. 2007. “Investment and Uncertainty: A Theory-based Empirical Approach.” *Oxford Bulletin of Economics and Statistics*, 69(5): 603–617.

Cerda, Rodrigo, Álvaro Silva, and José Tomás Valente. 2016. “Economic Policy Uncertainty Indices for Chile.” *Available at: https://www.policyuncertainty.com/media/EPU_Chile.pdf*.

Cruz, Cesi, Philip Keefer, and Carlos Scartascini. 2021. “DPI2020 Database of Political Institutions.” *Inter-American Development BankDepartment of Research and Chief Economist*.

Cruz Tadle, Raul. 2021. “FOMC minutes sentiments and their impact on financial markets.” *Journal of Economics and Business*, 68(3): 106021.

Davis, Kevin. 2017. “The Concept of Legal Uncertainty: Definition and Measurement.” *SSRN Electronic Journal*.

Deffains, Bruno, and Catherine Kessedjian. 2015. “Index of Legal Certainty: Report for the Civil Law Initiative.” *Fondation pour le droit continental*.

Djankov, Simeon, Rafael La Porta, Florencio Lopez de Silanes, and Andrei Shleifer. 2003. “Courts.” *The Quarterly Journal of Economics*, 118(2): 453–517.

Fernandez-Villaverde, Jesus, Pablo Guerron-Quintana, Keith Kuester, and Juan Rubio-Ramirez. 2015. “Fiscal Volatility Shocks and Economic Activity.” *American Economic Review*, 105(11): 3352–3384.

Fountas, S., M. Karanasos, and J. Kim. 2006. “Inflation uncertainty, output growth uncertainty and macroeconomic performance.” *Oxford Bulletin of Economics and Statistics*, 68(3): 319–343.

Gentzkow, Matthew, and Jesse M. Shapiro. 2010. “What Drives Media Slant? Evidence from U.S. Daily Newspapers.” *Econometrica*, 78(1): 35–71.

Gerrish, Sean M., and David M. Blei. 2011. “Predicting Legislative Roll Calls from Text.” *Proceedings of the 28th International Conference on Machine Learning*.

Gilchrist, Simon, Jae W. Sim, and Egon Zakrajšek. 2014. “Uncertainty, Financial Frictions, and Investment Dynamics.” *NBER Working Paper Series 20038*.

Gonzalez, Mario, and Raul Cruz Tadle. 2021. “Monetary Policy Press Releases: An International Comparison.” *Working Papers Central Bank of Chile*, 912.

Granger, C.W.J., and P. Newbold. 1974. “Spurious regressions in econometrics.” *Journal of Econometrics*, 2(42): 111–120.

Griffiths, Thomas L., and Mark Steyvers. 2004. “Finding scientific topics.” *Proceedings of the National Academy of Sciences of the United States of America*, Suppl 1: 5228–5235.

Gulen, Huseyin, and Mihai Ion. 2015. “Policy Uncertainty and Corporate Investment.” *The Review of Financial Studies*, 29(3): 523–564.

Hansen, Stephen, and Michael McMahon. 2016. “Shocking language: Understanding the macroeconomic effects of central bank communication.” *Journal of International Economics*, 99: 114–133.

Hassan, Tarek A, Stephan Hollander, Laurence van Lent, and Ahmed Tahoun. 2019. “Firm-Level Political Risk: Measurement and Effects.” *The Quarterly Journal of Economics*, 134(4): 2135–2202.

Hassett, Kevin A., and Gilbert E. Metcalf. 1997. “Investment with Uncertain Tax Policy: Does Random Tax Policy Discourage Investment?” *Economic Journal*, 109: 372–393.

Heimberger, Philipp. 2021. “What is structural about unemployment in OECD countries?” *Review of Social Economy*, 79(2): 380–412.

Ho, Benjamin, and Peng Liu. 2015. “Herd journalism: Investment in novelty and popularity in markets for news.” *Information Economics and Policy*, 31: 33–46.

Hu, Minqing, and Bing Liu. 2004. “Mining and Summarizing Customer Reviews.” *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*.

Husted, Lucas, John Rogers, and Bo Sun. 2020. “Monetary policy uncertainty.” *Journal of Monetary Economics*, 115: 20–36.

Jens, Candace E. 2017. "Political uncertainty and investment: Causal evidence from U.S. gubernatorial elections." *Journal of Financial Economics*, 124: 563–579.

Jurado, Kyle, Sydney Ludvigson, and Serena Ng. 2015. "Measuring Uncertainty." *American Economic Review*, 105(3): 1177–1216.

Kayam, Orly. 2018. "The readability and simplicity of Donald Trump's language." *Political Studies Review*, 16(1): 73–88.

Kelly, Bryan, L'uboš Pástor, and Pietro Veronesi. 2016. "The Price of Political Uncertainty: Theory and Evidence from the Option Market." *The Journal of Finance*, 71(5): 2417–2280.

Kincaid, J. Peter, Robert Fishburne Jr, Richard Rogers, and Brad Chissom. 1975. "Derivation of new readability formulas (automated readability index, fog count and esch reading ease formula) for navy enlisted personnel." *Technical report, Naval Technical Training Command Millington TN Research Branch,*.

Le Moglie, Marco, and Gilberto Turati. 2019. "Electoral cycle bias in the media coverage of corruption news." *Journal of Economic Behavior and Organization*, 163: 140–157.

Lim, Jamus Jerome. 2014. "Institutional and structural determinants of investment worldwide." *Journal of Macroeconomics*, 41: 160–177.

Linz, Juan. 1994. "Presidentialism or Parliamentarism: Does It Make a Difference?" *in Linz, Juan J. and Arturo Valenzuela, (eds.), The Failure of Presidential Democracy*, 3–90.

Mimica, Nicolás, Patricio Navia, and Rodrigo Osorio. 2022. "Changes in the Rules of the Lawmaking Process and the Success of Presidential Bills: Chile, 1990–2018." *Legislative Studies Quarterly*.

Mueller, Hannes, and Christopher Rauh. 2018. "Reading Between the Lines: Prediction of Political Violence Using Newspaper Text." *American Political Science Review*, 112(2): 1194–1225.

Mullainathan, Sendhil, and Andrei Shleifer. 2005. "The Market for News." *American Economic Review*, 95(4): 1031–1053.

Mumcuoglu, Emre, Ceyhun E. Ozturk, Haldun M. Ozaktas, and Aykut Koc. 2021. "Natural language processing in law: Prediction of outcomes in the higher courts of Turkey." *Information Processing and Management*, 58.

Nguyen, Thanh C., Vítor Castro, and Justine Wood. 2020. "Political environment and financial crises." *International Journal of Finance & Economics*, 27: 417–438.

North, Douglas. 1990. *Institutions, Institutional Change, and Economic Performance.* New York: Cambridge University Press.

Osnabrugge, Moritz, Elliott Ash, and Massimo Morelli. 2021. “Cross-Domain Topic Classification for Political Texts.” *Political Analysis*, 1–22.

Porter, Martin F. 1980. “An algorithm for suffix stripping.” *Program*.

Pástor, L’uboš, and Pietro Veronesi. 2013. “Political uncertainty and risk premia.” *Journal of Financial Economics*, 110(3): 520–545.

Raban, Ofer. 2009. “The Fallacy of Legal Certainty: Why Vague Legal Standards May Be Better for Capitalism and Liberalism.” *Boston University Public Interest Law Journal*, 175(19).

Repetto, Luca, and Maximiliano Sosa Andrés. 2022. “Divided government and polarization: Regression-discontinuity evidence from US states.” *CESifo Working Paper*, 9823.

Rodrik, Dani. 1991. “Policy uncertainty and private investment in developing countries.” *Journal of Development Economics*, 36(2): 229–242.

Röder, Michael, Andreas Both, and Alexander Hinneburgz. 2015. “Exploring the Space of Topic Coherence Measures.” *Proceedings of the 8th ACM International Conference on Web Search and Data Mining*, 399–408.

Špeh, Jaka, Andrej Muhic, and Jan Rupnik. 2013. “Parameter estimation for the latent dirichlet allocation.” *Proceedings of the Conference on Data Mining and Data Warehouses, Ljubljana, Slovenia*, 1–4.

Sulea, Octavia-Maria, Marcos Zampieri, Shervin Malmasi, Mihaela Vela, Liviu P. Dinu, and Josef van Genabith. 2017. “Exploring the Use of Text Classification in the Legal Domain.” *Proceedings of the 2nd Workshop on Automated Semantic Analysis of Information in Legal Texts*.

Trubek, David M. 2008. “Max Weber on Law and the Rise of Capitalism.” *in The Sociology of Law*.

Vannoni, Matia, Elliott Ash, and Massimo Morelli. 2021. “Measuring Discretion and Delegation in Legislative Texts: Methods and Application to US States.” *Political Analysis*, 43–57.

Appendix

A. Methodology

A.1. Latent Dirichlet Allocation

To estimate the latent structure of the corpus, LDA considers the words in each document to determine the data process most likely to have generated the documents. Next I describe the algorithm more formally:

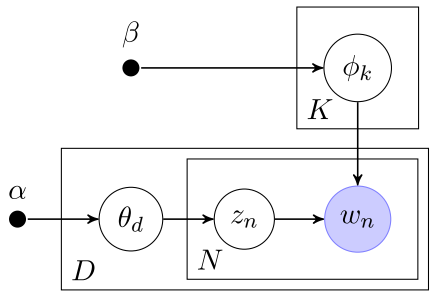
Let θ_d be the topic distribution for document d and ϕ_k the term distribution for topic k. Under LDA we assume a Dirichlet distribution for both θ_d and ϕ_k with parameters α and β respectively. The advantage of the distribution is that if the prior distribution of the multinomial parameters is a Dirichlet distribution, then the posterior is also a Dirichlet distribution.

The α parameter represents the prior belief about the topic distribution of documents, with higher α implying documents are expected to contain a mix of most of the topics, while lower α imply that document contain fewer topics. Conversely, β represents the prior belief about the word distribution of topics, with higher β implying topics are comprised of most words and lower β implying each topic consists of fewer words. Next I describe the document generating process. The following is taken from Blad and Svensson (2020) and Calvo-González, Eizmendi and Reyes (2018):

1. For each topic $k \in \{1, 2, ..., K\}$
   (a) draw a distribution of terms $\phi_k \sim Dir(\beta)$
2. For each document $d \in \{1, 2, ..., D\}$
   (a) draw a vector of topic proportions of terms $\theta_d \sim Dir(\alpha)$
   (b) For each term $w_n \in \{1, 2, ..., N\}$
      i. draw a topic assignment $z_n \sim Multinomial(\theta)$

These latent variables and the chain assumed can be seen in Figure A.1.

FIGURE A.1
LDA DIAGRAM



Note: Plate notation of LDA. Figure reproduced based on Špeh, Muhic and Rupnik (2013).

LDA follows this process by backward iteration, inferring the latent distribution of θ_d, ϕ_k and z_n. LDA computes the posterior distribution of these latent variables given a document and the priors:

$$P(\theta, z, \phi|w, \alpha, \beta) = \frac{P(\theta, z, \phi|\alpha, \beta)}{P(w|\alpha, \beta)} \tag{A.1}$$

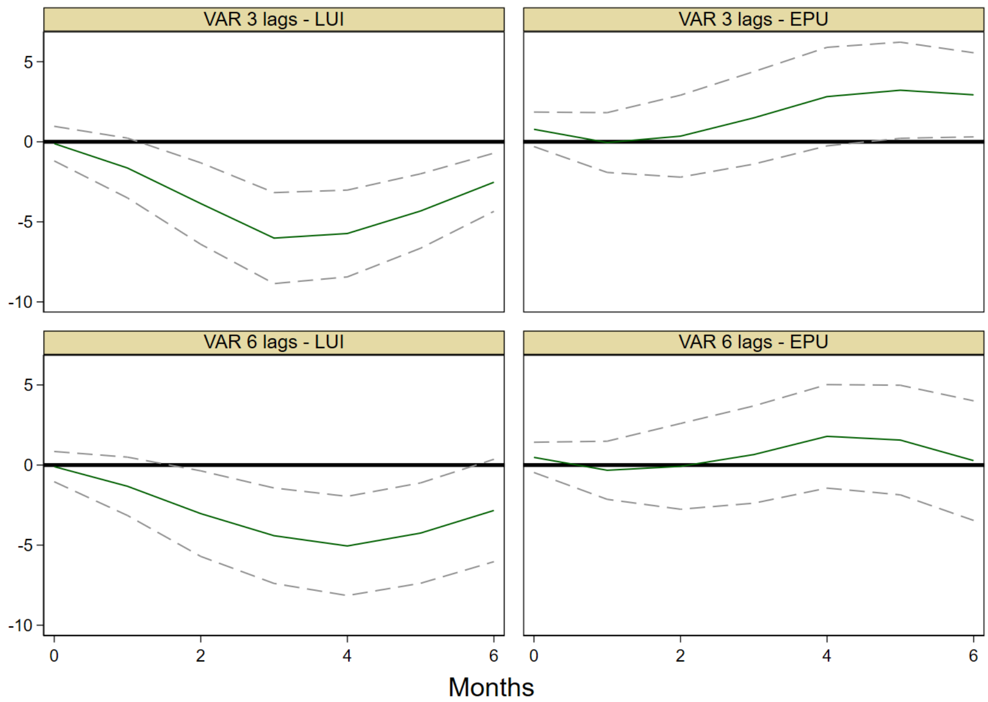
Nonetheless, the posterior needs to be approximated through algorithms as it is computationally intractable (Špeh, Muhic and Rupnik, 2013). A common alternative is to use Collapsed Gibbs sampling (Griffiths and Steyvers, 2004), where we first estimate z fixing θ_d and ϕ_k and then approximate θ_d and ϕ_k with the posterior of z. To implement Gibbs sampling we need the distribution parameters α and β. Following Griffiths and Steyvers (2004) I consider $\alpha = \frac{50}{k}$ and $\beta = 0.1$.

B. Additional results

B.1. Additional VAR specifications - Chile

FIGURE B.2

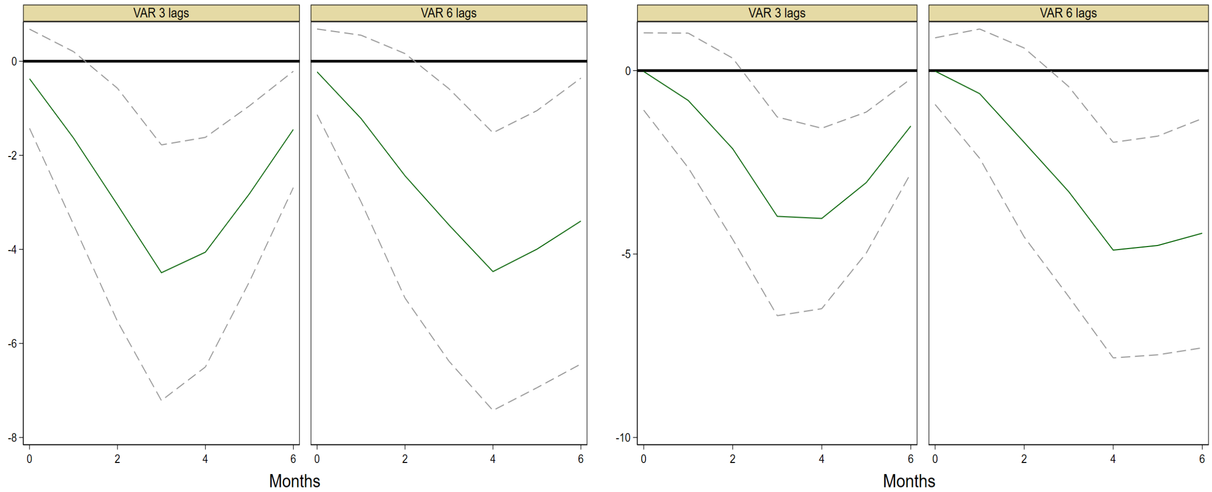
VAR-ESTIMATED IMPULSE RESPONSE FUNCTIONS FOR CHILE - INVESTMENT GROWTH



Cumulative IRF for investment growth to an innovation of one standard deviation in the LUI (left side) and EPU index (right side), with 95% confidence intervals. Identification based on three lags (upper charts), six lags (middle charts), time trends and Cholesky decomposition with the following order: LUI, EPU and investment growth. Fit to monthly data from January 1990 to December 2021.

Figure B.3

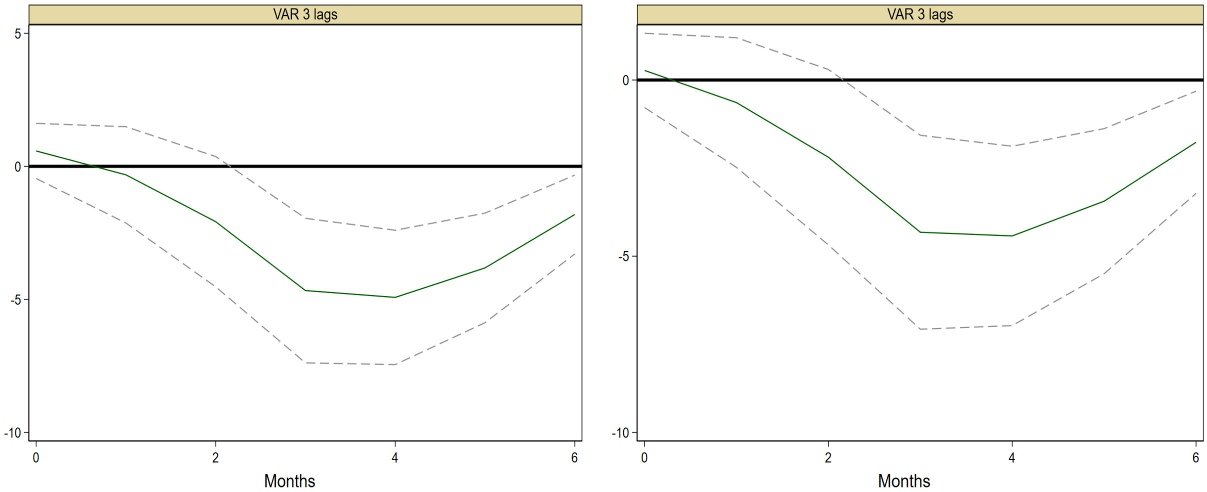
VAR-estimated Impulse Response Functions for Chile - Investment Growth - LUI with Different Number of Topics



(a) LUI - LDA done with 15 topics (b) LUI - LDA done with 21 topics

Notes: Cumulative IRF for investment growth to an innovation of one standard deviation in the LUI index, with 95% confidence intervals. Identification based on three and six lags respectively, time trends and Cholesky decomposition with the following order: LUI and investment growth. Fit to monthly data from January 1990 to December 2021. LUI considering forecast error measure a la Jurado, Ludvigson and Ng considering an LDA with 15 and 21 topics respectively.

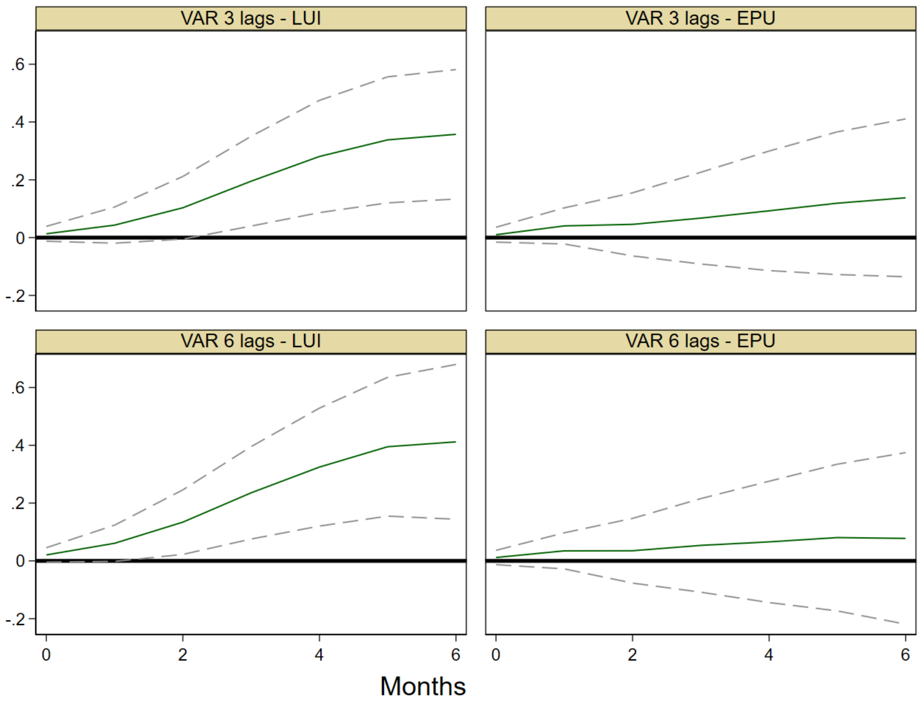
Figure B.4
VAR-estimated Impulse Response Functions for Chile - Investment Growth - Alternative AR process for LUI calculation



(a) LUI - Forecast error calculated with AR 3 (b) LUI - Forecast error calculated with AR 9

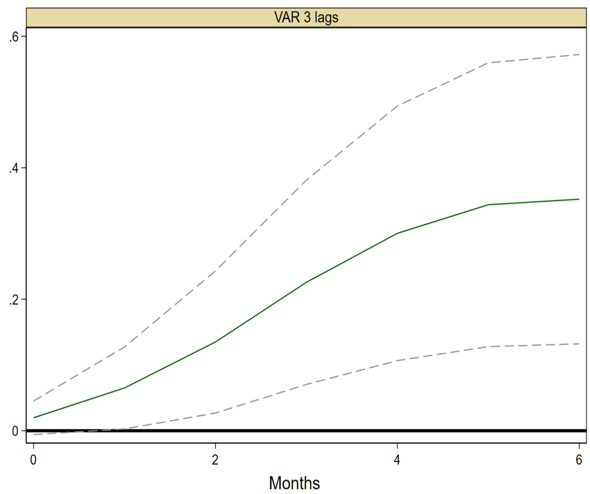
Notes: Cumulative IRF for investment growth to an innovation of one standard deviation in the LUI index, with 95% confidence intervals. Identification based on three lags, time trends and Cholesky decomposition with the following order: LUI and investment growth. Fit to monthly data from January 1990 to December 2021.

FIGURE B.5
VAR-ESTIMATED IMPULSE RESPONSE FUNCTIONS FOR CHILE - UNEMPLOYMENT

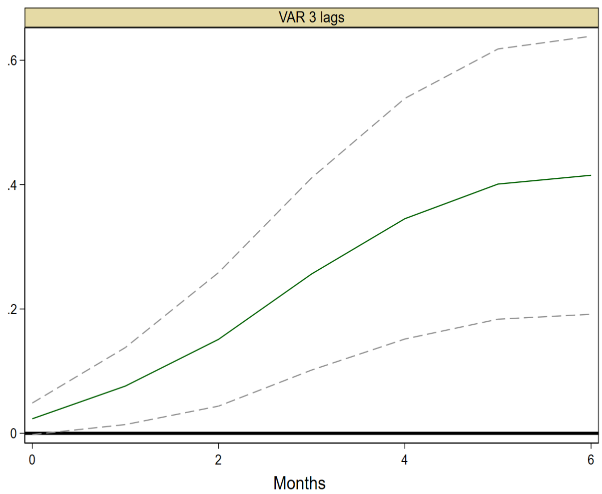


Cumulative IRF for unemployment to an innovation of one standard deviation in the LUI (left side) and EPU index (right side), with 95% confidence intervals. Identification based on three lags (upper charts), six lags (lower charts), time trends and Cholesky decomposition with the following order: LUI, EPU and unemployment. Fit to monthly data from January 1990 to December 2021.

FIGURE B.6
VAR-ESTIMATED IMPULSE RESPONSE FUNCTIONS FOR CHILE - UNEMPLOYMENT - LUI WITH DIFFERENT NUMBER OF TOPICS



(A) LUI - LDA DONE WITH 15 TOPICS

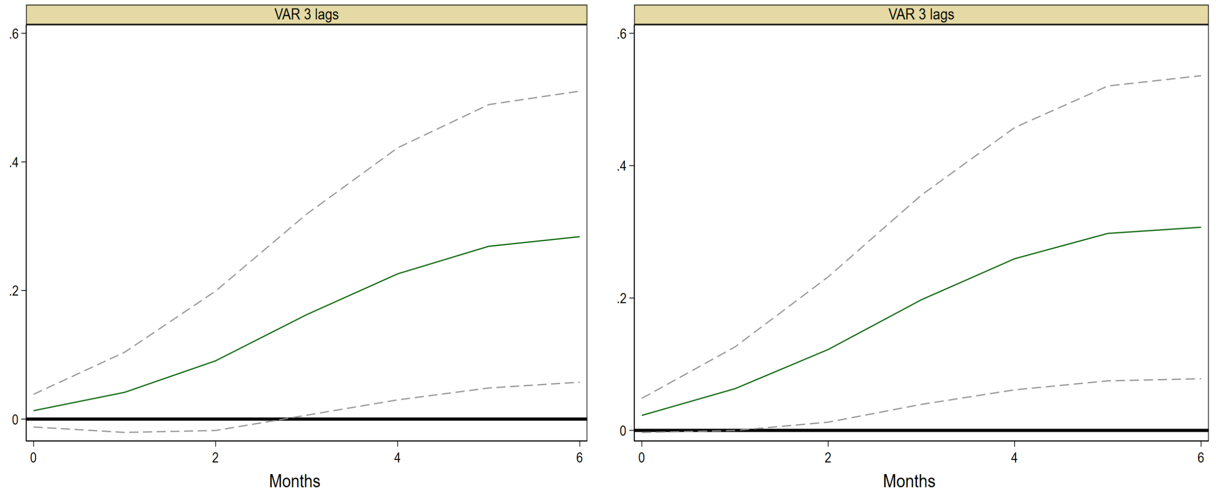


(B) LUI - LDA DONE WITH 21 TOPICS

Notes: Cumulative IRF for unemployment to an innovation of one standard deviation in the LUI index, with 95% confidence intervals. Identification based on three lags, time trends and Cholesky decomposition with the following order: LUI and unemployment. Fit to monthly data from January 1990 6 to December 2021. LUI considering forecast error measure a la Jurado, Ludvigson and Ng considering an LDA with 15 and 21 topics respectively.

Figure B.7

VAR-estimated Impulse Response Functions for Chile - Unemployment - Alternative AR process for LUI calculation



(a) LUI - Forecast error calculated with AR 3 (b) LUI - Forecast error calculated with AR 9

Notes: Cumulative IRF for unemployment to an innovation of one standard deviation in the LUI index, with 95% confidence intervals. Identification based on three lags, time trends and Cholesky decomposition with the following order: LUI and unemployment. Fit to monthly data from January 1990 to December 2021.

B.2. Additional panel data specifications

Table B.1

Effect of Legal Uncertainty on Investment Growth - Alternative investment measure

	(1) Capital Form/GDP	(2) Fixed Capital Form /GDP	(3) Capital Form
LUI	-0.108***	-0.083***	-0.172***
	(0.029)	(0.028)	(0.051)
Mean of dep.var.	1.01	1.01	1.01
R2	0.01	0.01	0.01
Obs.	889	889	889
Country FE	YES	YES	YES

Notes: Cluster-robust standard errors at the country level in parentheses. Investment growth measures considering the annual variation rate capital formation, capital formation over GDP and fixed capital formation over GDP. Baseline LUI measure considering forecast error measure a la Jurado, Ludvigson and Ng and AR(3) process.

$^{***}p < 0.01,^{**}p < 0.05,^{*}p < 0.1$

TABLE B.2

EFFECT OF LEGAL UNCERTAINTY ON INVESTMENT GROWTH - ALTERNATIVE AR PROCESSES TO CONSTRUCT LUI

	(1) Fixed Cap. Form	(2) Fixed Cap. Form
LUI AR4	-0.086***	
	(0.030)	
LUI AR2		-0.057*
		(0.032)
Mean of dep.var.	0.04	0.04
R2	0.01	0.00
Obs.	790	868
Country Fixed Effects	YES	YES

Notes: Cluster-robust standard errors at the country level in parentheses. Investment growth is calculated as the annual variation rate in gross fixed capital formation. LUI considering the forecast error measure a la Jurado, Ludvigson and Ng using an AR(4) and AR(2) process for estimation respectively.

$^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$

TABLE B.3

EFFECT OF LEGAL UNCERTAINTY ON INVESTMENT GROWTH - ALTERNATIVE LUI

	(1) Fixed Cap. Form	(2) Fixed Cap. Form	(3) Fixed Cap. Form
LUI 15 Topics	-0.063		
	(0.067)		
LUI 21 Topics		-0.096**	-0.078**
		(0.038)	(0.033)
Mean of dep.var.	0.04	0.04	0.04
R2	0.00	0.01	0.19
Obs.	828	828	828
Country FE	YES	YES	YES
Year FE	NO	NO	YES

Notes: Cluster-robust standard errors at the country level in parentheses. Investment growth is calculated as the annual variation rate in gross fixed capital formation. LUI2 considering "Earth Movers' Distance" measure

$^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$

Table B.4

Effect of Legal Uncertainty on Investment Growth - Election-Year controls

	(1) Fixed Cap. Form	(2) Fixed Cap. Form
LUI	-0.067*	-0.097**
	(0.033)	(0.040)
Mean of dep.var.	0.04	0.04
R2	0.00	0.06
Obs.	664	796
Country FE	YES	YES
Drop Election Year	YES	NO
Election Year FE	NO	YES

Notes: Cluster-robust standard errors at the country level in parentheses. Investment growth measures considering the annual variation rate in gross fixed capital formation. LUI considering the forecast error measure a la Jurado, Ludvigson and Ng using an AR(3) for estimation per country. Estimation in Column (1) drops all observations with years where there is a change in government. Estimation in Column (2) includes country-election year fixed effects.

$^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$

Table B.5
Effect of Legal Uncertainty on Investment Growth

	(1) Fixed Cap. Form	(2) Fixed Cap. Form	(3) Fixed Cap. Form	(4) Fixed Cap. Form	Fi
LUI	-0.099**	-0.105**	-0.107**	-0.157**	
	(0.042)	(0.048)	(0.046)	(0.063)	
GDP	-0.000**	-0.000*	-0.000*	-0.000*	
	(0.000)	(0.000)	(0.000)	(0.000)	
Interest	-0.002***	-0.002***	-0.002***	-0.001	
	(0.000)	(0.000)	(0.000)	(0.001)	
Trade Open	-0.002**	-0.002*	-0.002*	-0.003	
	(0.001)	(0.001)	(0.001)	(0.002)	
Fin Freedom		0.000	0.000	-0.000	
		(0.000)	(0.000)	(0.001)	
Polity index		-0.007***	-0.007***	0.002	
		(0.003)	(0.003)	(0.009)	
Inv Freedom		-0.001	-0.001	-0.001	
		(0.001)	(0.001)	(0.001)	
Corruption			-0.060	-0.004	
			(0.168)	(0.166)	
Property Rights			0.000	0.000	
			(0.001)	(0.001)	
Tax Revenue				0.005	
				(0.005)	
Credit to private					
GDP growth					
Constant	0.096***	0.178***	0.202**	0.056	
	(0.014)	(0.041)	(0.077)	(0.138)	
Mean of dep.var.	0.04	0.04	0.04	0.04	
R2	0.08	0.07	0.08	0.07	
Obs.	463	403	403	293	
Country Fixed Effects	YES	YES	YES	YES	

Notes: Cluster-robust standard errors at the country level in parentheses. Investment growth is calculated as the annual variation rate in gross fixed capital formation. For controls I follow Lim (2014) and include annual GDP, GDP growth, trade openness ratio, interest rate, financial freedom index, investment freedom index, property rights index, corruption index, tax revenue as percentage of GDP and credit to private sector as percentage of GDP.

$^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$

Table B.6

Descriptive Statistics - Cross Country Sample

	Mean	Std.Dev	Min	Max	N
Investment Growth	0.04	0.12	-0.50	1.00	967
Investment/GDP Growth	1.01	0.09	0.70	1.90	984
Unemployment	8.58	5.68	0.40	33.29	1012
Real Fx Rate	98.99	19.80	9.71	215.52	1007
Inflation Rate	15.10	157.00	-4.48	4734.91	1016
GDP Growth	2.87	4.11	-22.90	25.18	1056
Trade Openess	0.26	8.38	-29.69	50.92	1045
Interest Rate	17.60	22.33	0.99	250.28	614
Credit to private (% GDP)	70.31	46.36	0.00	217.76	817
Stock market trading (% GDP)	31.62	41.38	0.00	355.52	696
Financial Freedom	59.82	16.62	20.00	90.00	876
Investment Freedom	65.24	19.01	10.00	95.00	876
Corruption Index	0.31	0.29	0.01	0.95	959
Property Rights Index	62.51	22.93	10.00	95.00	876
Polity index	7.54	6.09	-88.00	10.00	935
Rule of Law index	0.77	0.25	0.04	1.00	940
New Government	0.20	0.40	0.00	1.00	982
Right Wing Gov	0.43	0.50	0.00	1.00	812
Parliamentary System	0.49	0.50	0.00	1.00	983
IMF Program	0.10	0.30	0.00	1.00	1043
WUI	0.18	0.14	0.00	1.26	978

Notes: Descripitve statistics for full-country sample for period 1990 - 2021 on a yearly basis (depending on availability).

Table B.7

Data description - Chile data

Variable	**Description**	**Availability**	**Source**
Laws	Laws enacted in Chile	Jan 1990 - Dec 2021	National Congress Library
EPU	Economic Policy Uncertainty	Jan 1993 - Dec 2021	Cerda, Silva & Valente (2016)
Investment	Fixed Capital Formation	Jan 1990 - Dec 2021	Ministry of Finance
Inflation	Inflation rate	Jan 1990 - Dec 2021	Central Bank
Unemployment	Unemployment rate	Jan 1990 - Dec 2021	National Statistics Institute
Stock Variation	IPSA index of stock returns	Oct 2009 - Dec 2021	Santiago's stock exchange
Business Trust	Business sentiment index	Jul 2003 - Dec 2021	Universidad del Desarrollo

Table B.8
Data description - Cross country data

Variable	Description	Availability	Source
Laws	Laws enacted	1990-2021	Global Regulation
Investment	Fixed capital formation	1990-2020	WDI
Unemployment	Unemployment rate	1990-2021	WDI
Inflation	Inflation rate	1990-2021	WDI
GDP	Gross Domestic Product	1990-2020	WDI
Trade Openess	(Exports- Imports) / GDP	1990-2020	WDI
Interest Rate	Average annual interest rate	1990-2021	WDI
Credit to private	Credit to private sector (% GDP)	1990-2020	WDI
Stock market	Stock market trading (% GDP)	1990-2020	WDI
Real Fx Rate	Real exchange rate index	1990-2019	Bruegel
Fin Freedom	Financial freedom index	1995-2020	QoG Institute
InvFreedom	Investment freedom index	1995-2020	QoG Institute
Corruption Index	Corruption index	1990-2019	QoG Institute
Property Rights	Property rights index	1995-2020	QoG Institute
Polity Index	Polity index of Democracy	1990-2018	Polity project
Rule of Law	Rule of Law index from WGI	1990-2020	World Bank
Regulatory Qual	Regulatory Quality index from WGI	1990-2020	World Bank
New Government	Indicator of change in goverment	1990-2020	DPI
Right Wing Gov	Indicator of government ideology	1990-2020	DPI
Parliamentary	Indicator of government system	1990-2020	DPI
IMF Program	Indicator of IMF assistance program	1990-2021	IMF
WUI	World Uncertainty Index	1952-2021	Ahir, Bloom and Furceri

B.3. Other Text Analysis: Number of Laws, Length, Sentiment and Complexity

In this section I consider the number of laws enacted in each period as well as their length and text complexity as additional text measures. The number of laws and their length are straightforward measures to capture policy changes through legal texts. Still, we can also consider the complexity of the laws based on their readability. Burrowing from computational linguistics we can calculate an index calibrated to measure the readability of a document in terms of years of education needed to understand it. For example, a document with a score of 12 indicates that it requires 12 years of education to read and understand it. Although these indexes are seldom used in economics, they are more commonly used across social sciences.[16]

I calculate the Kincaid index (Kincaid et al., 1975), which is defined as:

$$Kincaid = 0.39(\frac{words}{sentences}) + 11.8(\frac{syllables}{words}) - 15.59 \tag{B.1}$$

where words and sentences refer to their number in each law. I then aggregate the average Kincaid score on a monthly/yearly basis to obtain a time series, comparable to the LUI.

Finally, I consider an algorithm for sentiment classification that uses a dictionary method with pre-defined sentiment lexicons to determine how may words in a text are positive and how many are negative (Hu and Liu, 2004). A positive word is assigned a positive score of 1, while negative words get the opposite. Then a final score is calculated for each text by adding all scores of each word in the text. A legal text is said to be positive if the final score is above 0, and negative otherwise. I then calculate the percentage of laws deemed as positive on a monthly basis to obtain a time series.

I first estimate a VAR model with the national investment growth as dependent variable, and the LUI index and the additional text measures as endogenous variables. If the variation captured by the LUI stems from either the number of laws or their length and complexity, or the wording sentiment, then the impact of LUI should be smaller compared to the base specification.[17] To recover orthogonal shocks, I use a Cholesky decomposition with the order LUI, Number of Laws, Lenght of Laws, Kincaid index, Percentage of Laws with positive texts and the quarterly growth in investment. I estimate a model equivalent to equation 6 with the following specification of endogenous variables:

[16] Some recent examples include Ajina, Laouiti and Msolli (2016); Kayam (2018), and Gonzalez and Cruz Tadle (2021)

[17] Once again, the VAR requires covariance stationary. The augmented Dickey-Fuller test shows that none of the variables have a unit root at the 1% confidence level.

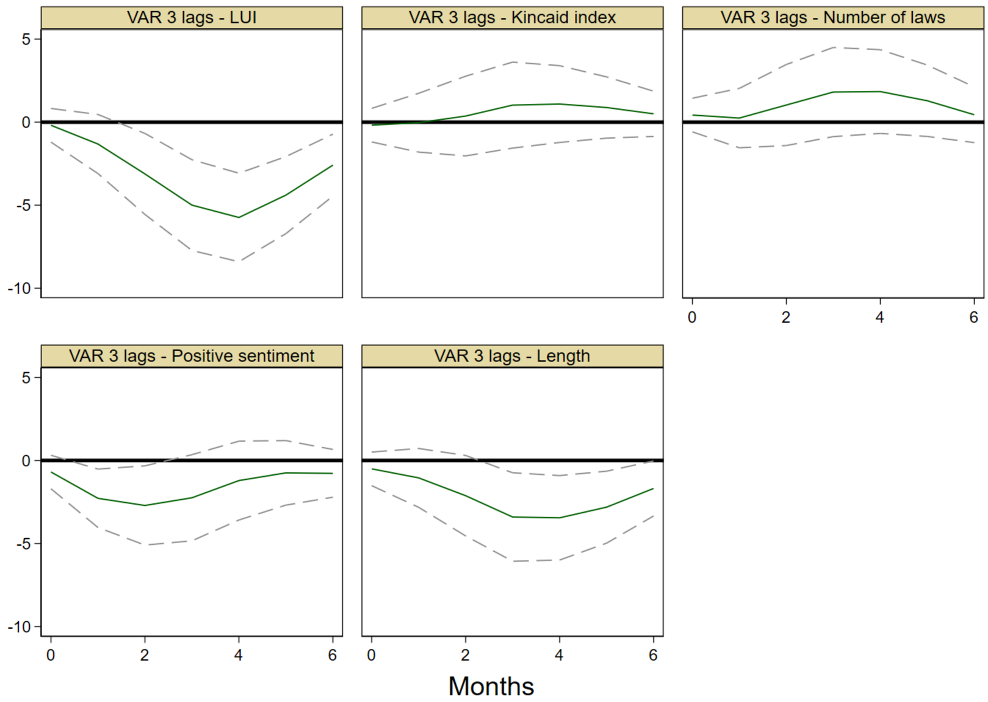
$$X_t = \begin{pmatrix} I_t \\ LUI_t \\ Laws_t \\ AverageLenght_t \\ Kincaid_t \\ \%PositiveText_t \end{pmatrix}$$

The impulse response functions are available in Figure B.8. The baseline VAR specification includes three lags of all variables and time trends.[18] The first graph depicts the model-implied responses of growth in investment to a one standard deviation upward innovation in the LUI. The response to the LUI shock remains negative and statistically significant, albeit of a smaller magnitude compared to the point estimates in Figure 8. There is no statistically significant impact of a shock in the number of laws or their sentiment or complexity as measured by readability. There is a negative and statistically significant impact of a shock in the average length in laws on investment growth. This suggest that more detailed and complex laws also have a negative impact over investment sentiments. Still, the effects are complimentary and not substitutory as the LUI remains statistically significant in the VAR specification that includes both endogenous variables.

Therefore, I find evidence in favor that length rather than the amount of new legislation enacted can scare investors off. This effect comes in addition to the uncertainty generated from changes in legal content.

[18]Results are robust to considering six lags of all variables, as well as the alternative specifications of the LUI (not shown)

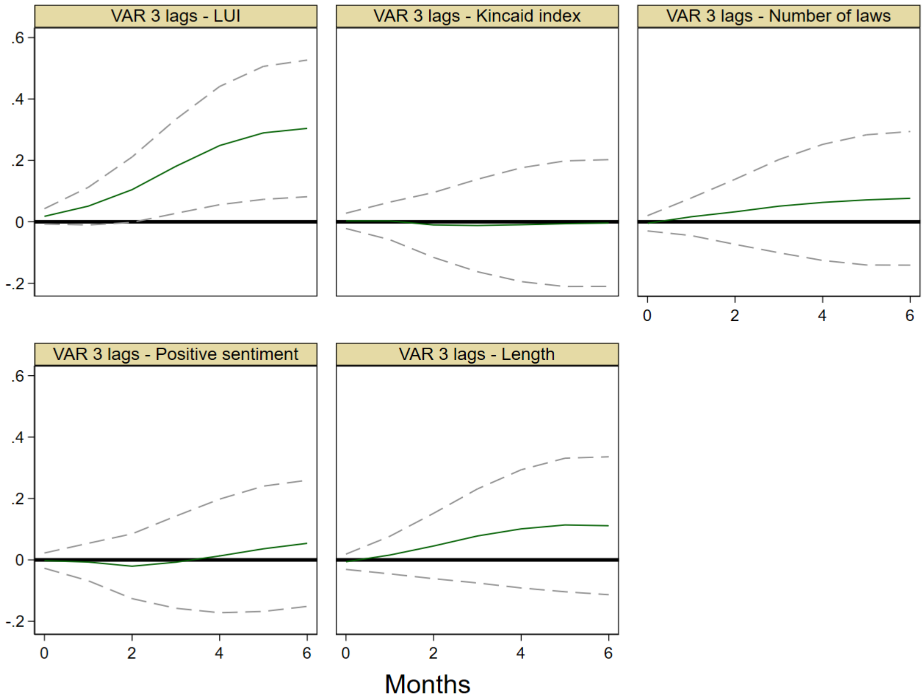
Figure B.8
VAR-estimated Impulse Response Functions for Investment- Additional Text Measures



Notes: Cumulative IRF to an innovation of one standard deviation in the LUI index, Number of Laws, Length of Laws, Kincaid index and Percentage of Positive sentiment laws with 95% confidence intervals. Identification based on three lags and time trends and Cholesky decomposition. Fit to monthly data from January 1990 to December 2021

Finally, I turn to the robustness of the impact of LUI over unemployment. Results are available in Figure B.9. Once again, the response to the LUI shock is positive and statistically significant, albeit marginally at the 10% confidence level. None of the other measures have an impact over the unemployment rate. Therefore, I conclude that the impact of legal uncertainty is not affected by other sources of variation coming from legal texts. The LUI captures the reaction of firms to unexpected changes in policy content regardless of how complex the texts are or how many new laws are implemented.

FIGURE B.9
VAR-ESTIMATED IMPULSE RESPONSE FUNCTIONS FOR UNEMPLOYMENT - ADDITIONAL TEXT MEASURES



Notes: Cumulative IRF for investment growth to an innovation of one standard deviation in the LUI index, Number of Laws, Length of Laws, Kincaid index and Percentage of Positive sentiment laws with 95% confidence intervals. Identification based on three lags and time trends and Cholesky decomposition. Fit to monthly data from January 1990 to December 2021