

# Winners Never Quit, Quitters Never Grow

Using Text Mining to Measure Policy Volatility  
and Its Link with Long-Term Growth in Latin America

*Oscar Calvo-González*

*Axel Eizmendi*

*Germán Reyes*



**WORLD BANK GROUP**

Poverty and Equity Global Practice Group

January 2018

## Abstract

Although there is wide recognition of the negative consequences of policy volatility for countries' long-term economic growth, there is limited empirical work on this subject. One of the reasons is the difficulty of measuring policy volatility over long periods of time, especially in developing countries. This paper contributes to this literature by constructing a proxy for policy volatility that exploits the information content of the priorities conveyed in presidential speeches.

The study creates a policy volatility measure using a Latent Dirichlet Allocation algorithm on a novel data set of 953 presidential speeches in 10 Latin American countries and Spain. The paper shows that the proxy for policy volatility is negatively correlated with long-term growth over 1940–2010. The results are robust to a large set of changes in the construction of the proxy for policy volatility.

---

This paper is a product of the Poverty and Equity Global Practice Group. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at <http://econ.worldbank.org>. The authors may be contacted at [ocalvogonzalez@worldbank.org](mailto:ocalvogonzalez@worldbank.org).

*The Policy Research Working Paper Series disseminates the findings of work in progress to encourage the exchange of ideas about development issues. An objective of the series is to get the findings out quickly, even if the presentations are less than fully polished. The papers carry the names of the authors and should be cited accordingly. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the International Bank for Reconstruction and Development/World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent.*

# Winners Never Quit, Quitters Never Grow: Using Text Mining to Measure Policy Volatility and Its Link with Long-Term Growth in Latin America\*

Oscar Calvo-González<sup>†</sup>, Axel Eizmendi<sup>‡</sup>, Germán Reyes<sup>§</sup>

JEL Classification: E60, H11, O11, O54

Keywords: Policy stability, Growth, Text Mining, Presidential Speeches, Latin America

---

\* The authors would like to thank Laura Chioda, Alexandru Cojocaru, Peter Hakim, Luis Felipe Lopez-Calva, Bill Maloney, Ambar Narayan, Mariano Tommasi, Roy Van der Weide, and participants at an internal seminar for fruitful discussions and suggestions. We also thank Sandra Pastrán, Evangelina Cabrera, Yalile Uarac, and Cesar Polit and Martha Narvaez, for compiling and forwarding to us multiple speeches from Argentina, Paraguay, Chile and Ecuador, respectively. Natalia García-Peña Bersh helped compile data from Colombia. We are grateful to the Research Support Budget from the World Bank for financial support.

<sup>†</sup> Poverty and Equity Global Practice, World Bank. Mail: [ocalvogonzalez@worldbank.org](mailto:ocalvogonzalez@worldbank.org)

<sup>‡</sup> University of Oxford, Oxford, UK. Mail: [axel.eizmendi.larrinaga@gmail.com](mailto:axel.eizmendi.larrinaga@gmail.com)

<sup>§</sup> Cornell University, Ithaca, NY, USA. Mail: [gjr66@cornell.edu](mailto:gjr66@cornell.edu)

# 1. Introduction

Frequent policy changes are widely thought to affect economic growth negatively. One channel through which policy volatility has such negative impact is by affecting investment decisions. This argument goes back at least to the work on investment under uncertainty (Bernanke, 1983; Dixit and Pindyck, 1994). Attention to economic policy uncertainty has gained interest in recent years. While this literature shares an interest in measuring the extent and impact of uncertainty, the approaches to do so vary significantly: from extracting the information content of asset price volatility (Bloom, 2009, and Bryan et al., 2016), to exploiting a broad range of macroeconomic time series (Jurado et al., 2015) or a narrower set of fiscal time series (Fernández-Villaverde et al., 2015), to developing a new index of economic policy uncertainty based on newspaper coverage frequency (Bloom, 2014, and Baker et al. 2016). Regardless of the method, another commonality in the recent literature is their shared view that the impact of policy uncertainty is large.

We add to this literature by providing a measure of the volatility of policy priorities over the long run. We do so by extracting the information content of the priorities conveyed in presidential speeches at the annual ‘state-of-the-union’ addresses over a 70-year period for Latin American countries and Spain. As in the case of Bloom (2014) and Baker et al. (2016), we use texts as our main data, but in contrast with that literature we focus on texts that capture the evolution of policy priorities. As the empirical literature on policy uncertainty mentioned above shows, there is no definitive single measure of policy uncertainty. Our contribution is in this spirit, by providing one additional way to capture policy uncertainty that is not meant to substitute for but to complement other measures.

Our focus on policy priorities as captured through presidential speeches has three additional characteristics that make it of interest. First, it allows us to take the analysis back in time over long periods of time for countries where other data are not available. While it is widely held that policy continuity is positive for countries’ long-term economic growth (Commission on Growth and Development, 2008:3), there is surprisingly limited empirical evidence of such long-run effects. We try to address this limited evidence by mobilizing a type of data, namely the text of presidential speeches, that has previously not been exploited in a quantitative way. Recently, Shiller has argued that economics should be “expanded to include serious quantitative study of changing popular narratives” (Shiller 2017:967). In a similar vein, we expand the analysis of policy uncertainty to study the narratives of policy-makers by quantitatively analyzing how the topics about which they talk change over time.

Second, it can help us zero in on fundamental shifts in policy. In our view, it may be informative to distinguish different types of uncertainty about economic policy. Think of the following two scenarios. The first one refers to a country where there is an increase in uncertainty as to whether the independent central bank may raise interest rates by 25 basis points given a strong economic recovery. The second scenario refers to a country where a

new government has announced plans to nationalize large swaths of economic activity. By design our measures of policy volatility would not detect the increase in uncertainty observed in the first scenario but may be a good way to capture the essence of what is happening in the second scenario. Arguably the uncertainty created by the type of issues exemplified in the second scenario is particularly relevant in a developing country context.

The third reason why our measure of the volatility of policy priorities is of interest takes us beyond the issue of uncertainty. In addition to increasing uncertainty, frequent changes in policy priorities can also negatively impact the performance of the public sector. As the World Bank's World Development Report 2017 on governance and the law argues, the result of a proliferation of competing interests in the policy arena may be that public institutions become "overloaded with multiple pressures, undermining the coherence and effectiveness of public policies" (WDR 2017:24). Because our measure of policy volatility is at the level of priorities, we capture this source of ineffectiveness.

Our measure of policy priorities has also disadvantages. Annual presidential speeches are not the only vehicle for conveying policy priorities to the public. In addition, any methodology that extracts and codifies information from texts is a form of a simplification. Moreover, policy makers can express priorities but fail to deliver on them. This limitation, however, is less of an issue when addressing the question of our paper, since we are interested in the volatility of policy priorities.

The message on the importance of policy uncertainty is particularly relevant in Latin America – a region where analysts and policy makers have in the last couple of decades been searching for the reasons why the policy reforms of the 1980s and 1990s did not deliver the expected results. Analyses of the experience with policy reform in Latin America have pointed out the importance of credibility (or lack of it) of reforms, and explored the institutional settings which gave rise to credible reforms (Tommasi, 2006). Our analysis can help inform this literature on the political economy of reform in Latin America by providing a proxy measure of policy volatility, a factor which is hypothesized in the literature to be a contributor to the limited credibility of reforms.

The rest of the paper is organized as follows. Section two presents the data set we built as well as some descriptive statistics. The third section discusses the topic modeling methodology – using a Latent Dirichlet Allocation (LDA) algorithm – and implements this approach to the data set of presidential speeches to obtain the evolution of topics over time. In section four we produce a new measure of the volatility of policy priorities based on the topics uncovered using the LDA algorithm. In section five, we explore how the volatility of policy priorities correlates with long-term growth. We document that higher volatility of policy priorities appears to be associated with lower growth in the long-run and check that this correlation is robust to alternative measures of policy volatility. Section six concludes.

## 2. Presidential speeches data and descriptive statistics

One of the contributions of our paper is the compilation and use of a novel data set of presidential speeches. The data set is comprised of 953 speeches delivered during the period 1819-2016 in 10 Latin American countries (Argentina, Chile, Colombia, Costa Rica, Dominican Republic, Ecuador, Mexico, Paraguay, Peru, and the República Bolivariana de Venezuela) and Spain.

We selected these countries because their constitutions mandate that the president must give an annual speech that gives an overview of the work the government has performed in each legislative session, as well as an outline of the policy goals and priorities for the future (Dulanto 2006).<sup>1</sup> One can think of these speeches as constituting the respective countries' closest equivalent of the United States' "State of the Union Address." These speeches, therefore, provide a general reflection of the political priorities of each government.

In the case of Spain, no such mandate exists, so we use speeches that closely approximate the programmatic content of the constitutionally-mandated speeches. For the period 1939-1974, we use Franco's end-of-year speeches and for the period 1979-1982, we use the inaugural speech each president delivered during the assumption of the presidency. Finally, throughout the 1983-2015 period we rely on a mix of two types of speeches. In years in which a presidential election took place, we use the president's inaugural speech, whereas in years with no presidential election, we use the speeches presidents deliver during the yearly "debates about the state of the nation,"—a political event that did not become a tradition until 1983. All these speeches embody the programmatic nature of the constitutionally-mandated speeches in the rest of our sample.

The compilation of the data set consisted of a two-stage process: collecting the speeches and processing them. The sources used for speech collection vary by country. Access to most of the speeches from Argentina, Ecuador and Paraguay was through the respective National Congresses and their librarians. In the case of Chile, Colombia, Costa Rica, Dominican Republic, and Mexico, the speeches were collected through a variety of online sources. Most of the República Bolivariana de Venezuela's speeches were scanned from books containing compilations of presidential speeches available at the US Library of Congress. In Appendix A we list the source of each of the speeches.

The second stage involved processing the speeches to enable text analysis. For the speeches that were already in a digitized format, this meant converting each file into text format and removing any text that did not form part of the speech, such as the title or date. For scanned speeches—i.e., those in "image" format—the text in every page was extracted by performing Optical Character Recognition (OCR), the conversion of images of text into machine-encoded, editable and searchable text. While OCR is certainly more efficient than

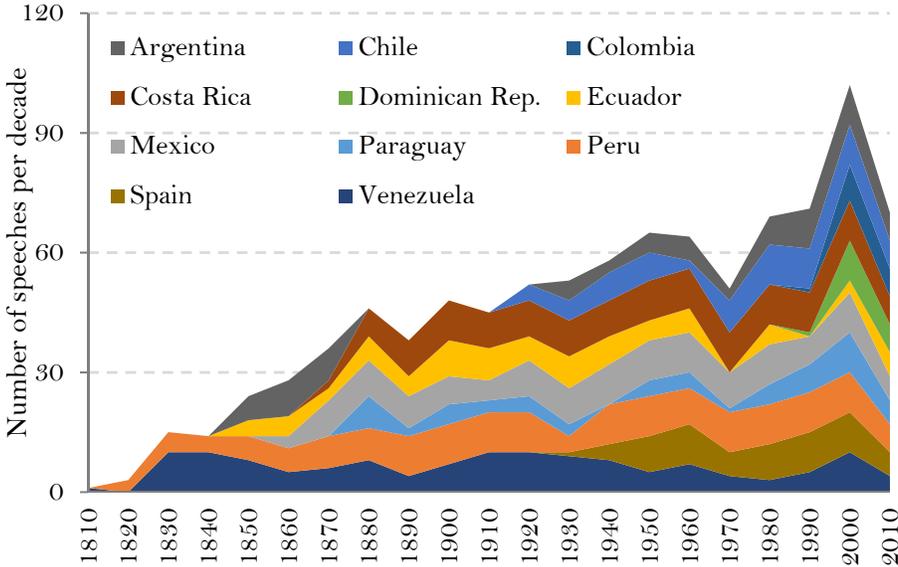
---

<sup>11</sup> In Appendix A we list the Spanish-speaking Latin American countries that have this constitutional mandate, as well as the specific articles in the constitution that establish this requirement.

transcribing text images manually, some inaccuracies are inevitable, especially when the quality of the original text is not optimal. This was the case for some of the older speeches for which the typewritten text was not as clear. To ensure the quality of the data, the vast majority of inaccuracies in the processed texts were corrected manually.

Figure 1 below shows the distribution of speeches by country and by decade, while Appendix Table A1 details the precise number of observations on each decade by country.

Figure 1. Number of presidential speeches per decade by country



Source: Authors' calculations.

Due to variability in data availability, both online and in the respective Congressional libraries, not every country nor every decade is represented equally in the data set. Costa Rica, Mexico, Peru and the República Bolivariana de Venezuela are the most highly represented countries in the data set, each with between 14% and 17% of total speeches. Argentina, Chile, Ecuador, and Paraguay have a moderate representation, between 6% and 9% of total speeches. Colombia and the Dominican Republic are equally underrepresented in the data set, with only 2% of total speeches, respectively. Representation in terms of time periods is also asymmetric. Most speeches (60%) correspond to the 20<sup>th</sup> century, 22% to the 19<sup>th</sup> century and the remaining 18% to the 21<sup>st</sup> century.

In this paper, we use a subset of this data set for which we have a relatively balanced panel of speeches for most countries. Because our research question is concerned with policy volatility in the long-run, we exclude Colombia and the Dominican Republic—countries for which we have no speeches prior to the year 2000. Similarly, since we are also interested in analyzing long-term growth, we start our analysis in 1940, as GDP estimates of previous

years are not as reliable. This reduces our sample size from 953 presidential speeches to 550. In the next section, we show what is the topic content of those speeches.

### 3. Topic models

Topic models are statistical models that machine learning researchers have recently developed to extract the main themes contained in large, unstructured collections of documents (Blei, 2012). A significant advantage that these algorithms offer is the fact that they do not require the researcher to specify a set of arbitrary topics into which the documents are classified. Instead, topic models use modeling assumptions and properties of the texts to automatically “discover” a set of topics and simultaneously assign documents to those topics.

Topic models have been used for a wide range of purposes, including, among others, the analysis of biological and medical data sets (Chen et al., 2014), modeling of online user reviews (McDonald and Titov, 2008), and the analysis of social media (He et al. 2011; Davison and Hong 2010). Topic modeling has also been applied to analyze political texts. Wang and McCallum (2006), for example, introduce the “topics over time” model and apply it to over two centuries of US State of the Union addresses. Quinn et al. (2010) use the “dynamic multitopic model” to model the daily attention allocated to a variety of topics in US Senate floor speeches, while Grimmer (2010) relies on his “expressed agenda model” to measure how senators allocate their attention to different topics, as reflected by their press releases.

The most widely used topic model is Latent Dirichlet Allocation (LDA), first introduced by Blei et al. (2003). Given a collection of documents, LDA “discovers” the primary topics in each document as well as the degree, or proportion, in which each document exhibits those topics. A particularly attractive feature of this model is the fact that the only restriction, or arbitrary decision, imposed by the researcher is choosing the number of topics to be extracted. The discovery of topics is then performed using words in the documents as the *only* observable variables. To do so, LDA assumes that documents are random probability distributions over topics, and that topics are random probability distributions over words. In other words, documents contain a mixture of topics with varying probabilities and topics contain words with different probabilities. Another important assumption is the “exchangeability” or “bag-of-words” assumption, which means that the order in which the words appear in the document is not important; LDA relies on term frequencies instead.

The fundamental assumption of LDA is that the observed documents were generated through a probabilistic generative process. However, the parameters or “recipe” of this generative process are hidden, or “latent.” This defines the key inferential task of LDA: estimating the “latent” structure—the topics and the topic composition of each document.

LDA performs this task by working through the generative process in reverse. That is, it uses the observed words in each document to estimate the parameters of the generative process that are most likely to have generated the observed collection of documents. The generative process can be described as follows:

- 1- For each topic, decide what words are likely.
- 2- For each document:
  - a. Decide what proportions of topics should be in the document, say 20% topic A, 50% topic B, 30% topic C.
  - b. For each word:
    - i. Choose a topic. Based on the topic proportions from step 2.a., topic A is more likely to be chosen.
    - ii. Given this topic, choose a likely word (generated in step 1).

We can describe this process more formally using the model parameters and corresponding probability distributions. After specifying a number of topics  $k$ :

- 1- For each topic  $k$ , draw a distribution over words  $\varphi_k$  according to a Dirichlet distribution  $\sim \text{Dir}(\beta)$ , where  $\beta$  is the parameter of the Dirichlet prior on the per-topic word distribution.<sup>2</sup>
- 2- For each document  $D$ :
  - a. Draw a vector of topic proportions  $\theta_d$  according to a Dirichlet distribution  $\sim \text{Dir}(\alpha)$ , where  $\alpha$  is the parameter of the Dirichlet prior on the per-document topic distribution.<sup>3</sup>
  - b. For each of the  $N$  words  $w_n$ :
    - i. Draw a topic assignment  $z_n$  according to a multinomial distribution  $\sim \text{Multinomial}(\theta)$  according to the topic proportion  $\theta_d$
    - ii. Choose a word  $w_n$  from  $p(w_n | z_n, \varphi)$ , a multinomial probability conditioned on the topic  $z_n$ .

The key inferential task of LDA consists in performing this assumed generative process in reverse. That is, using the observed documents and words, LDA works backwards to infer the “latent structure”—the distribution of the parameters  $\theta$ ,  $z$ , and  $\varphi$ —that are most likely to have generated the documents in the sample. Where  $z$  represents the per-word topic assignments and  $\theta$  gives the topic distribution of each document, which indicates the extent to which each document belongs to each topic;  $\varphi$  gives the distribution of words in

---

<sup>2</sup> The beta parameter represents the “prior” belief about the per-topic word distributions. A high beta value means that each topic is likely to be made up of most of the words in the corpus, whereas a low beta means that each topic will have fewer words.

<sup>3</sup> The alpha parameter represents the prior belief about the per-document topic distributions. A high alpha-value means that each document is likely to contain a mixture of most of the topics, and not any single topic specifically, whereas a low alpha-value means that each document is likely to contain fewer topics.

topic  $k$ , which is used to define the semantic content of each topic. The objective of LDA consists in computing the posterior distribution of these hidden variables given a document and the Dirichlet priors:

$$P(\theta, z, \varphi | w, \alpha, \beta) = \frac{P(\theta, z, \varphi | \alpha, \beta)}{P(w | \alpha, \beta)}$$

Estimating the maximum likelihood of the model and the distributions of the hidden variables requires marginalizing over the hidden variables to obtain the model’s probability for a given corpus  $w$  and priors  $\beta$  and  $\alpha$ .

$$p(w | \alpha, \beta) = \frac{\Gamma(\sum_i \alpha_i)}{\prod_i \Gamma(\alpha_i)} \int \left( \prod_{i=1}^k \theta_i^{\alpha_i - 1} \right) \left( \prod_{n=1}^N \sum_{i=1}^k \prod_{j=1}^V (\theta_i \beta_{ij})^{w_n^j} \right) d\theta$$

These distributions are intractable to compute, requiring the use of other approximate inference algorithms. Although in the first introduction of LDA, Blei et.al. (2003) relied on a Variational Bayes approximation of the posterior distribution, we use Collapsed Gibbs sampling as our inference technique—a commonly used alternative introduced by Griffith and Steyvers (2004).

The Collapsed Gibbs sampling algorithm is a common Markov Chain Monte Carlo (MCMC) algorithm that is used to approximate posterior distributions when these cannot be directly computed. The idea is to iteratively generate posterior samples by looping through each variable to sample from its conditional distribution, while retaining the values of all other variables fixed in each iteration (Yildirim, 2012). Essentially, we simulate posterior samples by sweeping through all the posterior conditionals, one random variable at a time. Because we initiate the algorithm with random values, the samples simulated at the early iterations are likely not close to the true posteriors. However, the process eventually “converges” at the point where the distribution of the samples closely approximates the distribution of the true posteriors.

In LDA, the variables we want to approximate are the “latent” variables  $\theta$  and  $\varphi$ . This is achieved by generating a sequence of samples of topic assignments  $z$  for each word  $w$ . As mentioned above, for each iteration, Gibbs Sampling requires retaining the values of all variables except for one fixed (see Griffiths and Steyvers, 2004). Therefore, because words are the only observed variables in LDA, at each iteration, the topic assignment of only one word is updated, while the topic assignments for all other words are assumed to be correct (i.e., remain unchanged). Samples from the posterior distribution  $p(z | w)$  are obtained by sampling from:

$$P(z_i = K | w, z_{-i}) \propto \frac{n_{-i,K}^{(j)} + \beta n_{-i,K}^{(di)} + \alpha}{n_{-i,K}^{(\cdot)} + V\beta n_{-i,\cdot}^{(di)} + k\alpha}$$

Where  $z_{-i}$  is the vector of current topic assignments of all words except the  $i$ th word  $w_i$ . The index  $j$  indicates that  $w_i$  is the  $j$ th term in the entire vocabulary of words in the corpus ( $V$ );  $n_{-i,K}^{(j)}$  indicates how often the  $j$ th term of the vocabulary is currently assigned to topic  $K$  without the  $i$ th word. The dot “.” indicates that summation over this index is performed;  $d_i$  denotes the document in the corpus to which word  $w_i$  belongs;  $\beta$  and  $\alpha$  are the hyperparameters of the prior distribution explained earlier (Grun and Hornik 2011).

Intuitively, in the first place, the algorithm goes through each document and randomly assigns each word in the document to one of the  $K$  topics. Because these assignments are random, however, they are poor and must be improved on. To improve these topic assignments, for each word  $i$  in document  $d$ , (for each  $w_{d,i}$ ) and for each topic  $k$ , two values are computed: 1)  $p(\text{topic } k \mid \text{document } d)$ , or the proportion of words in document  $d$  that are currently assigned to topic  $k$ , and 2)  $p(\text{word } w \mid \text{topic } k)$ , or the proportion of assignments to topic  $k$  over all documents that come from this word  $w$ . Then, these two proportions are multiplied to get  $p(\text{topic } t \mid \text{document } d) * p(\text{word } w \mid \text{topic } t)$ , which in the context of LDA’s generative process, gives the probability that topic  $k$  generated word  $w$ . Finally, word  $w$  is reassigned to a new topic based on this probability. To put it simply, for each word, its topic assignment is updated based on two criteria: 1) How prevalent is that word across topics? 2) How prevalent are topics in the document?

As in any Gibbs Sampling algorithm, the above steps are repeated a large number of times. After a large number of iterations, the algorithm converges to a steady state where the topic assignments of each word are close approximations of the true values. At this point, we can finally use these topic assignments to estimate the “latent” variables—the posterior of  $\theta$  and  $\varphi$ —given the observed words  $w$  and topic assignments  $z$ :

$$\hat{\theta}_K^d = \frac{n_K^{(d)} + \alpha}{n^{(d)} + K\alpha} \quad \hat{\varphi}_K^j = \frac{n_K^{(j)} + \beta}{n_K^{(\cdot)} + V\beta} \quad \text{for } j = 1 \dots V \text{ and } d = 1, \dots, D$$

With  $\theta$  and  $\varphi$  estimated, the objective of LDA—extracting topic representations of each document—is achieved.

### 3.1 LDA IMPLEMENTATION

As a first step before running LDA, to improve the discovery of topics, we follow the standard practice of cleaning our collection of documents. Specifically, we remove all punctuations and numbers, as well as “stop words”—terms such as articles, conjunctions and pronouns that are semantically meaningless for defining a topic. Because we are interested in discovering topics that are common across countries, we remove country-specific terms such as, “Peruvians,” “Peru” or “Lima,” which could otherwise bias the topics towards country-specific, rather than subject-specific topics.

We rely on Collapsed Gibbs Sampling for the iterative process of topic inference. This approach requires the specification of values for the parameters of the prior distributions—

$\beta$  (beta) for the per-topic term distributions and  $\alpha$  (alpha) for the per-document topic distributions. Following Griffiths and Steyvers (2004) we select the commonly used values of  $\alpha = 50/t$  (the number of topics) and  $\beta = 0.1$ .

One final, yet crucial parameter that must be specified is the number of topics to be discovered. To determine the optimal number of topics we rely on a measure known as *perplexity* that is often used in information theory and natural language processing to evaluate how well a model can predict the data, with lower perplexity indicating a better model (Blei et.al. 2003). In LDA, this involves, three steps: 1) as is common in machine learning, randomly partitioning our sample into a training set (90% of speeches) and a held-out test set (10% of speeches), 2) iteratively implementing LDA on the training set using a different number of topics  $t \in \{1, \dots, 30\}$  in each iteration, and 3) computing the perplexity of each model, which amounts to evaluating how “perplexed” or surprised each “trained” model is when presented with the previously unseen test set. The lower the perplexity, the better the model can predict the data. More formally, for a test set of  $M$  documents perplexity is defined as:

$$\text{Perplexity}(D_{\text{test}}) = \exp \left\{ -\frac{\sum_{d=1}^M \log p(W_d)}{\sum_{d=1}^M N_d} \right\}$$

Where  $W_d$  represents the words in document  $d$  and  $N_d$  the number of words.

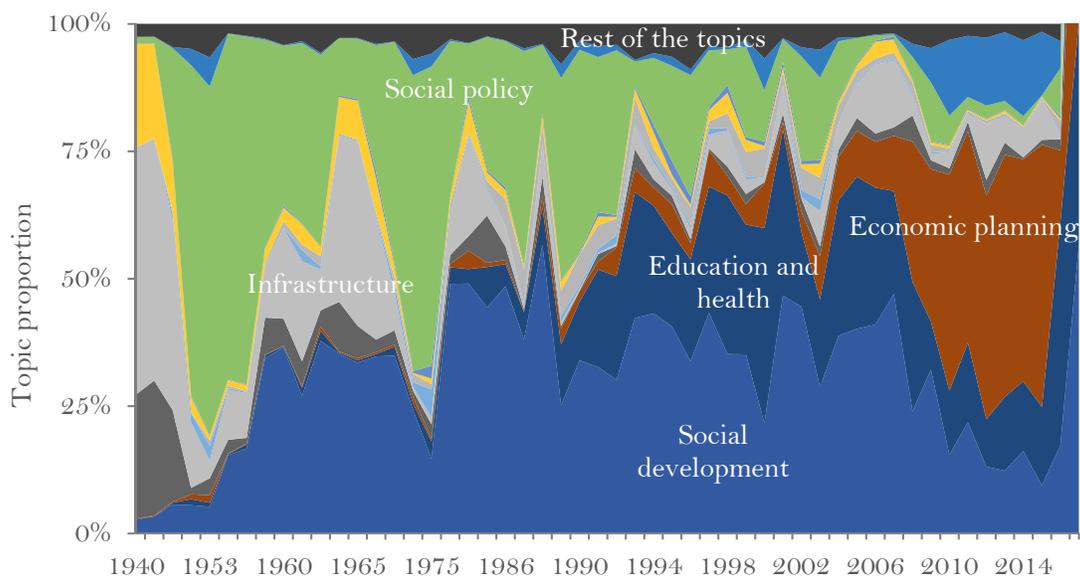
### 3.2 TOPICS OF THE PRESIDENTIAL SPEECHES

We begin by running the LDA model on the collection of speeches of all countries we presented in section 2. We choose to perform the LDA pooling the speeches of all countries to increase cross-country comparability. When performing LDA on each country separately, the number of topics extracted is not the same for every country and the topics themselves—i.e., the words that define them—are also different. This means that the results LDA produces may not be ideal for cross-country comparison. If we pool the speeches from all countries together and perform LDA in the resulting collection of speeches, the issue of comparability across countries can be reduced because every speech is now classified into the same set of topics. A downside of this approach is that the topics are not as country-specific, which diminishes some of the accuracy. Nevertheless, we choose comparability over goodness of fit and use the output from the pooled version of LDA throughout our main analysis. The country-specific LDA output is used to perform robustness checks.

To select the number of topics, we employ the criteria of perplexity minimization, which suggests 23 is the optimal number of topics across countries—a point at which perplexity begins to flatten out. Once LDA has “discovered” the topics contained in the presidential speeches of each country, we can plot the topic proportions ( $\theta$ ) to examine how the content of the speeches has evolved over time. We focus on the period 1940-2010, for which the number of speeches available is generally higher.

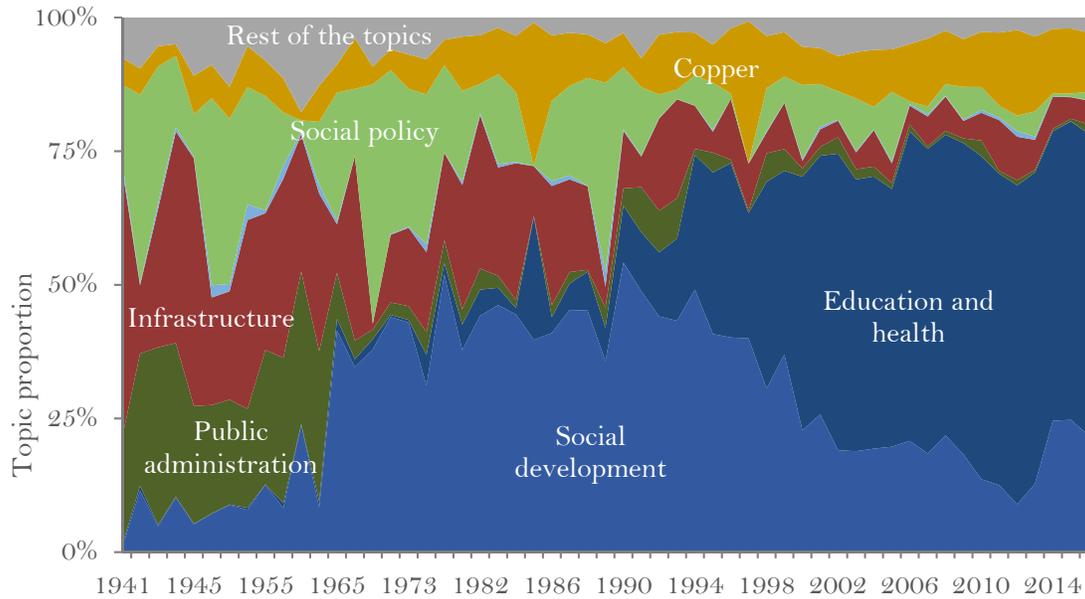
A final point that we must explain is our decision to include Spain in a sample of strictly Latin American countries. We do this because we wish to obtain a benchmark country that has experienced strong long-term economic growth, and which has been relatively politically stable over time. This allows us to put the results for LAC countries in context. With other types of analysis, we could have chosen any country that fulfills the political stability and sustained growth criteria, but for modeling the content of speeches, language is a crucial variable. Because speeches from Spain and LAC have a common language, we can pool the speeches together into one unified database and obtain LDA results that are comparable across all countries. To illustrate the results of performing LDA on our pooled sample of speeches, we present the topics discovered and their evolutions for Argentina, Chile and the República Bolivariana de Venezuela in Figures 2, 3 and 4. The vector of terms that define the topics is presented in Table 1, although for the sake of presentation, we do not provide the terms defining the “copper,” “Chavismo” “oil industry” topics.

Figure 2. Evolution of speech content in Argentina



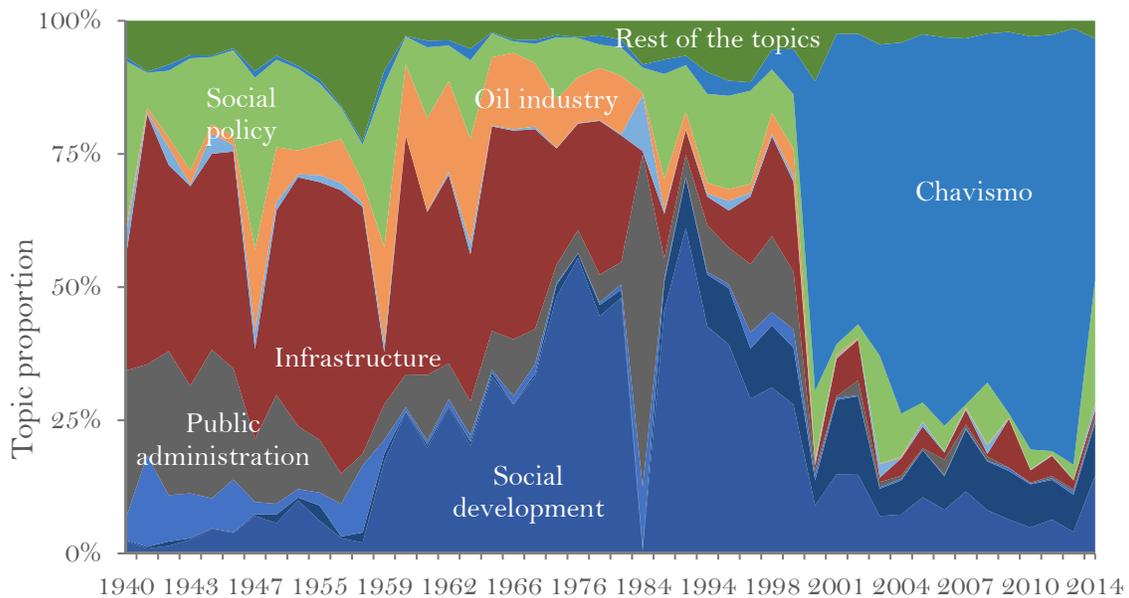
Note: This Figure shows the topic probabilities estimated with LDA. The topic probabilities can be thought of as representing the proportion of each document that belongs to each topic. Only topics whose average document probability exceeds 10% in at least one decade are presented, the rest are grouped as a residual.

Figure 3. Evolution of speech content in Chile



Note: This Figure shows the topic probabilities estimated with LDA. The topic probabilities can be thought of as representing the proportion of each document that belongs to each topic. Only topics whose average document probability exceeds 10% in at least one decade are presented, the rest are grouped as a residual.

Figure 4. Evolution of speech content in the República Bolivariana de Venezuela



Note: This Figure shows the topic probabilities estimated with LDA. The topic probabilities can be thought of as representing the proportion of each document that belongs to each topic. Only topics whose average document probability exceeds 10% in at least one decade are presented, the rest are grouped as a residual.

Table 1. Top probability terms defining the topics

“Public administration”	“Social policy”	“Infrastructure”	“Social development”	“Education and health”	“Economic planning”
public	policy	national	development	education	sector
administration	social	public works	social	health	provinces
national	life	production	national	family	bank
services	national	construction	program	program	growth
interests	nation	education	policy	poverty	province
right	homeland	services	sector	quality	debt
executive	world	Plan	system	security	crisis
public	freedom	social	process	social	plan
trade	right	service	resources	right	economy
order	economic	development	growth	world	money

Note: This table shows the words that have the highest probability of belonging to each topic. Because LDA defines topics as distributions over words, the lists of terms define the subject of each topic.

Figures 2 to 4 illustrate two important points that are relevant for our analysis. First, LDA detects important shifts in the political context of a country. In the case of Argentina, the advent of the Kirchner administration, as well as the arrival of the debt crisis are captured by the fact that beginning in the mid-2000s, the “social development” and “education and health” topics are quickly replaced by the “economic planning” topic. This latter topic is concerned with “debt” “crisis” and “banks,” which clearly reflects the economic realities Argentina faced at the time. Furthermore, this result is consistent with the country-specific LDA results for Argentina. LDA is able to detect an important shift in political context for the República Bolivariana de Venezuela as well. Although a topic concerned with social development prevailed until the late 1990s, it was quickly replaced by a new topic concerned with Chaves and his objectives. Clearly, the coming to power of Chaves constituted an important change in the political context of the República Bolivariana de Venezuela, and LDA captures this shift. In addition, it is worth noting that LDA also detects policy areas that are rather country-specific. For example, a topic about the oil industry is present in the República Bolivariana de Venezuela, and in Chile, a topic about the copper industry remains relevant across a long stretch of time, reflecting the economic significance the industry has had throughout Chile’s history.

The second point that the figures illustrate is that the stability or consistency of the topics—which we take as reflecting policy priorities—varies by country. In Chile, social development was an important topic from early on and it remained that way for decades. Furthermore, the education and health topic has continued to be extremely prevalent since the 1990s, much more so than in any other country. These observations capture the fact that Chile has prioritized social development and education for decades, and that these priorities have been stable over time. On the other hand, such stability was not present in Argentina

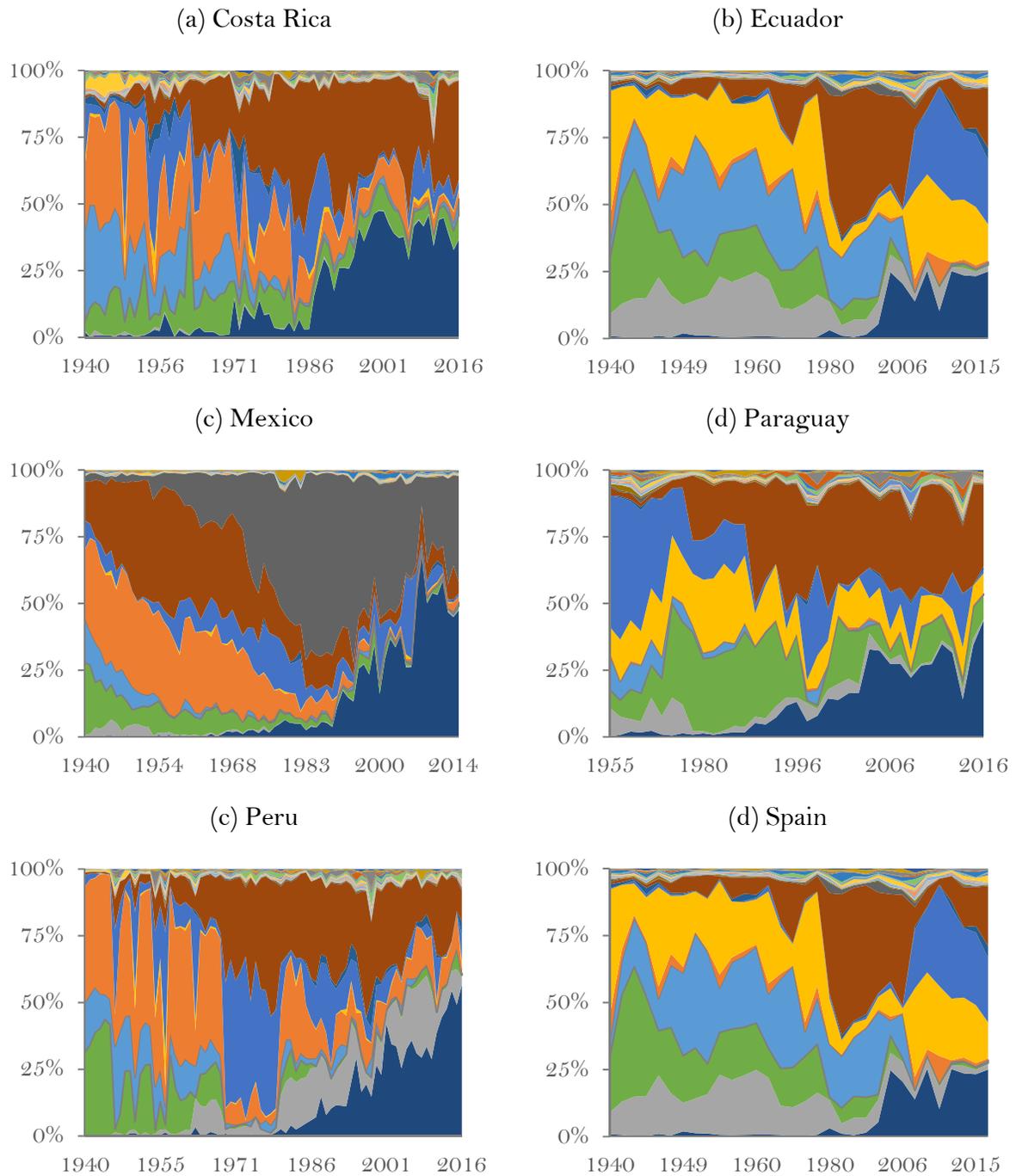
and the República Bolivariana de Venezuela, where, as we mentioned earlier, changing political contexts induced greater volatility in the evolution of the topics. In short, the figures illustrate that LDA is capturing relevant topics that reflect, at least to an extent, the policy priorities of governments (e.g., education and development in Chile), as well as the shifts in the political contexts of these countries (Chavismo in the República Bolivariana de Venezuela, and the debt crisis and Kirchner administration in Argentina). Visually, it becomes clear that the content of presidential speeches has evolved differently for each country, with some exhibiting greater volatility than others.

Indeed, if we look at the remaining six countries in our pooled sample of speeches (Figure 5), it becomes clear that the volatility of topics varies greatly across countries.

In the Appendix B we show two examples of the country-specific LDA (Argentina and Peru) to show the results are qualitatively similar. Admittedly, the results are not the exact same as if we use LDA on the pooled data, since the topics are more tailored to the context of each country (e.g., the country-specific LDA discovers “Peronism” as one of the topics in Argentina).

These cross-country differences in the volatility of the topics over time suggest that the topic proportions could be used to construct a measure of policy stability. We propose such measures in the following section.

Figure 5. Evolution of speech content in Costa Rica, Ecuador, Mexico, Paraguay, Peru, and Spain, 1940-2016



Note: This Figure shows the topic probabilities estimated with LDA. The topic probabilities can be thought of as representing the proportion of each document that belongs to each topic.

## 4. A new measure of policy volatility

The results presented above show that LDA can detect changes in political discourse that reflect broader shifts in the political context and objectives of a country. They also illustrated visually that the stability of this discourse—captured by the topic evolutions discovered by LDA—varies by country. To construct our measure of policy volatility, we exploit this variation in topic prevalence over time and across countries. In other words, we transform the visual representation of the topic evolutions presented earlier into one unique indicator of policy stability.

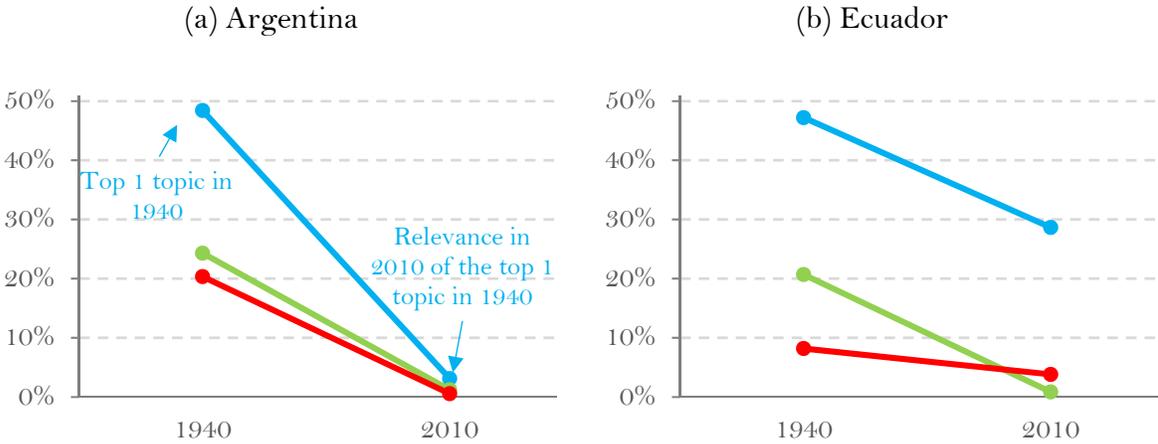
The underlying assumption behind our measure is that the content of presidential speeches reflects the policy objectives and priorities of their respective governments. Consequently, one might expect that in countries where presidents have tended to provide stability to the policies of their predecessors, with relatively few disruptions to the policy agendas of preceding years, the thematic content of presidential speeches will also tend to remain relatively constant. Conversely, in countries where the policy areas prioritized by the government have not remained consistent over time, one might expect that the content of presidential speeches will be characterized by greater volatility. From this set of assumptions, we can construct different measures of policy stability based on how the thematic content of presidential speeches has evolved over time. Specifically, in countries with greater policy stability, the topics covered in presidential speeches are expected to remain relatively constant, whereas in countries with low policy stability, the topics discussed are likely to be more erratic.

Given its ability to extract the topics and topic composition of documents, LDA provides an adequate method to measure how “continuous” the content of presidential speeches has remained over time. Our measures of policy volatility are based on the topic proportions of each speech ( $\theta$ ), which give an estimate of the proportion of the speech that discusses each topic. As a first step in constructing these measures, we perform a process of “deuration” by removing those topics for which the topic proportion did not exceed 10% in any speech. The assumption is that if the prevalence of a topic never exceeds 10% of the speech for any year, that topic does not convey a strong signal of the policy priorities of the speech and thus has limited value from an analytical point of view.

The first of our measures—the baseline policy volatility measure—is the one we use in the main exposition of our results, while the other alternative measures are meant to act as robustness checks. For each country, the baseline policy volatility measure consists of the percent change between 1940 and 2010 of the share of the three topics that were most prevalent in the year 1940. In other words, we identify the three topics with the highest topic proportion ( $\theta$ ) in the speeches of 1940, and compute how prevalent these three topics are in the speeches of 1940 and 2010 by calculating the sum of their topic proportions. We then take the percent change of these sums. The underlying concept is that in countries with low policy volatility, the content of the speeches—which we take as a reflection of policy—should remain relatively constant over time, and so the difference between the topics discussed in 2010 and 1940 should be smaller in countries with lower policy volatility.

As a visual illustration of this measure, in Figure 6 we show, for Argentina and Ecuador, the speech proportion in 2010 of the top three topics in 1940. We can readily observe that policy volatility was much higher in Argentina. In 1940, a topic concerned with the development of infrastructure dominated speech content with a topic proportion of 48%. However, by 2010, this same topic had nearly disappeared. The fall in prevalence is similarly pronounced for the second and third most prevalent topics of 1940. On the other hand, policy volatility in Ecuador was considerably lower. In 1940, a topic discussing a wide range of subjects, such as social policy, the economy and justice, had a speech prevalence of 47%. Although the importance of the topic had fallen by 2010, it remained a significant topic with a proportion of 29%. In other words, policy, as reflected in presidential speeches, was less discontinuous in Ecuador than in Argentina. These are precisely the cross-country differences our measure attempts to capture.

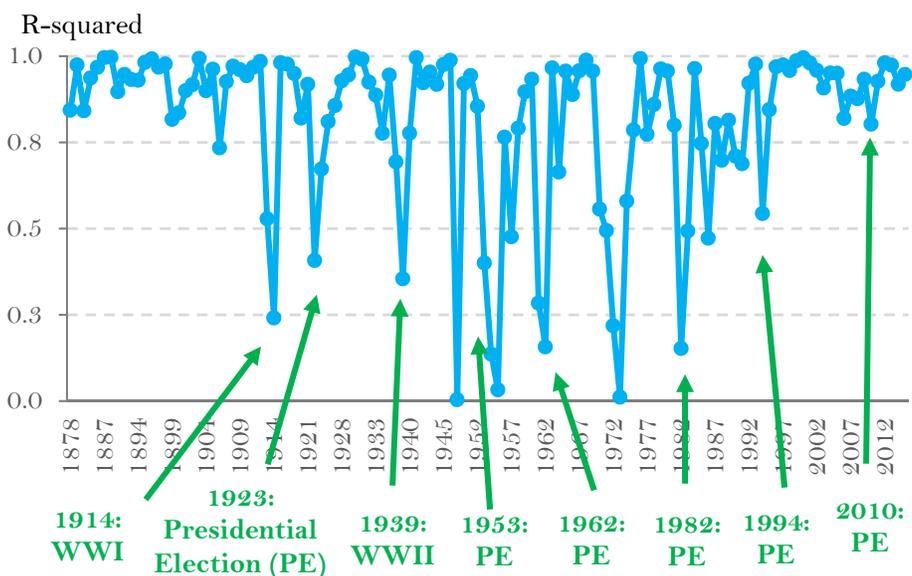
Figure 6. Speech proportion in 2010 of the top three topics in 1940 in Argentina and Ecuador



Note: This Figure shows the speech proportion in 2010 of the top three topics of the presidential speech in 1940. The topics were selected by the LDA using pooled data of all the countries.

One of our alternative measures of policy volatility is based on the R-squared obtained by regressing  $\theta_t$ —the topic proportions of the speech from year  $t$ —on  $\theta_{t+1}$ —the topic proportions of the speech from the preceding year. The R-squared of these regressions provides an indicator of how similar the topic distributions in a speech from a given year  $t$  are compared to the topic distributions of the speech from the preceding year  $t+1$ , with a high R-squared indicating a high degree of similarity between the two years. In the context of our research question, a high R-squared suggests high policy stability. To illustrate this measure, in Figure 7, we show how the R-squared measure captures policy volatility in Costa Rica throughout the 1878-2015 period.

Figure 7. Policy volatility in Costa Rica using the R-squared measure



Note: This Figure shows the R-squared obtained by regressing  $\theta_t$ —the topic proportions of the speech from year  $t$ —on  $\theta_{t+1}$ —the topic proportions of the speech from the preceding year.

Although for many of the years policy volatility is low (R-squared is close to one), it is interesting to note that in periods in which a significant exogenous “event,” such as a presidential election or the two World Wars took place, the R-squared falls abruptly. This result makes intuitive sense, as a change of government often brings policy changes, which are likely reflected in the content of speeches. Similarly, one would expect that a global event of such magnitude as a World War would induce presidents to devote attention to this subject in their speeches, which would represent a clear deviation from the issues discussed in the previous year.

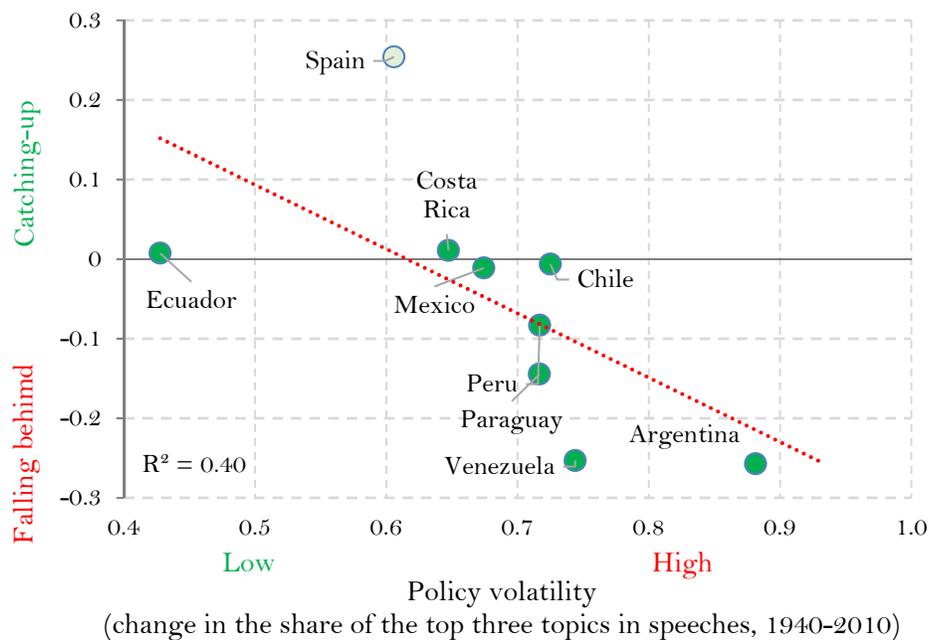
To synthesize Figure 7 into an index, for each country we calculate both the average R-squared, and the coefficient of variation of the R-squares. Another alternative measure of policy volatility consists in computing the average coefficient of variation of the topic proportions ( $\theta$ ) across all topics, a measure of variability in the relevance of the topics in each speech. Finally, we also look at the percentage of speeches in which the main topic—the topic with the highest  $\theta$ —changes, a measure of volatility in policy prioritization.

Although these indicators are useful for measuring the degree of policy stability, we must note that they suffer from one important limitation: they fail to differentiate between the continuity of good and bad policies. Clearly, the type of policy continuity that exerts a positive effect on economic growth is the continuity of *good* policies. An authoritarian government that has pursued the same set of poorly-planned policies for decades has a high degree of policy continuity, but logically, this continuity is not beneficial for growth. We acknowledge that our measures are imperfect because they fail to capture these qualitative differences, but we believe that they are still useful as general indicators of policy continuity.

## 5. Policy volatility and long-term growth

As noted in the introduction, the literature argues that policy continuity ought to be an important driver of economic growth in the long-run. In this section, we provide suggestive evidence that this is indeed the case.<sup>4</sup> In Figure 8 we show that policy volatility is negatively correlated with long-term growth during the 1940-2010 period. Long-term growth is measured by the percentage point change of the GDP per capita of each country (relative to the GDP per capita of the US). Positive values indicate countries that were catching-up with the US, while economies with negative values diverged from the US. Policy volatility is measured as the percentage point change between 1940 and 2010 of the share of the top three topics in the presidential speech of the initial year. Intuitively, this measure captures to what extent the top priorities circa 1940 were still relevant seven decades later. As mentioned previously, we pool the speeches from all countries and perform LDA in the collection of speeches to “discover” the topics in each speech. We then use a 10% minimum threshold to discard topics that appear infrequently in each country.

Figure 8. Policy volatility and convergence with the US GDP per capita, 1940-2010 (using topic modeling on pooled data of speeches from all countries)



Note: This Figure shows the cross-country correlation between long-term growth and policy volatility. Long-term growth is measured as the percentage point change of the GDP per capita of each country (relative to the GDP per capita of the US). Policy volatility is measured as the percentage point change of the share of the top three topics in the presidential speech of the initial year. The topics were selected by the LDA using pooled data of all the countries, and using a 10% minimum threshold to discard topics with low frequencies.

<sup>4</sup> GDP growth data come from The Maddison-Project, <http://www.ggd.net/maddison/maddison-project/home.htm>, 2013 version.

On average, countries with higher policy volatility were the ones that fell-behind the most in terms of long-term growth. The linear correlation between both measures is  $-0.63$  (and statically significant at the 10% level). In fact, Argentina and the República Bolivariana de Venezuela, the two countries with the lowest values of policy stability, were the ones that diverged the most with the US economy. On the other hand, Spain, the country with the second highest policy stability, was the one that converged the most with the US. On average, policy volatility explains about 40% of the cross-country variability in long-term growth, as measured by the R-squared.

It is important to stress that the interpretation is not necessarily causal, but rather suggestive. The relationship between policy stability and long-term growth can go both ways, or could also be driven by a third (omitted) variable. On the one hand, more policy-stability can lead to higher average levels of growth. But as countries grow, policy stability can also be affected through several channels (e.g., better institutions).

In the next subsection, we show that the strong correlation between policy volatility and growth is robust to different decisions underlying the creation of our policy volatility variable, and to different measures of policy volatility and long-term growth.

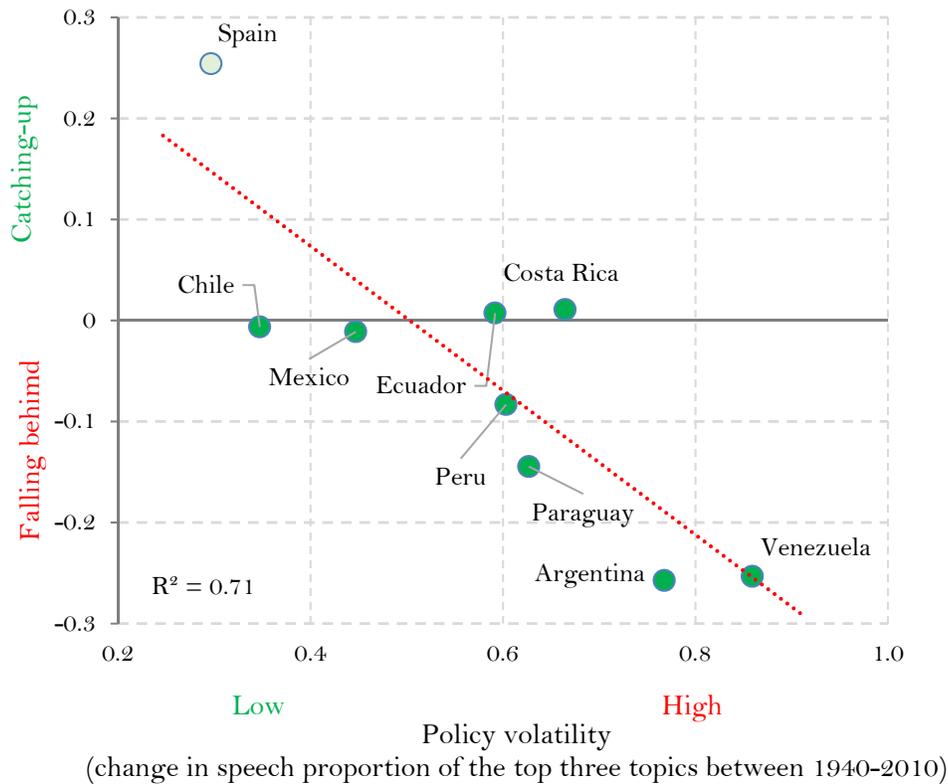
## 5.1 ROBUSTNESS CHECKS

In the process of calculating our policy volatility variable we took several arbitrary decisions: how to define the topics of each speech; whether we should omit topics with a low frequency in the process of creating our variables; and what should be the relevant time-frame of analysis. We now show that our main result is robust to a wide variation of these parameters.

In our baseline measure of policy volatility, we used the LDA output obtained by running the model in the pooled sample of speeches. This guarantees that the number of topics, and the content of the topics are constant between countries—increasing the cross-country variability—while the topic probabilities vary both across countries and across speeches within countries. This methodology assumes there are ‘common’ topics in the speeches across countries of the region, and that the relative relevance of each of those common topics varies across countries (as measured by the topic probabilities).

As a robustness check, we use an alternative approach. Instead of using LDA on the pool of speeches, we use it on a country by country basis. This implies that the optimal number of topics used to run the model, and the content of each of those topics vary across countries. This makes cross-country comparisons more difficult, but it improves how well the topics selected fit the speeches of each country. Figure 9 shows the cross-country correlation using this alternative way of creating our policy volatility measure.

Figure 9. Policy volatility and convergence with the US GDP per capita, 1940-2010

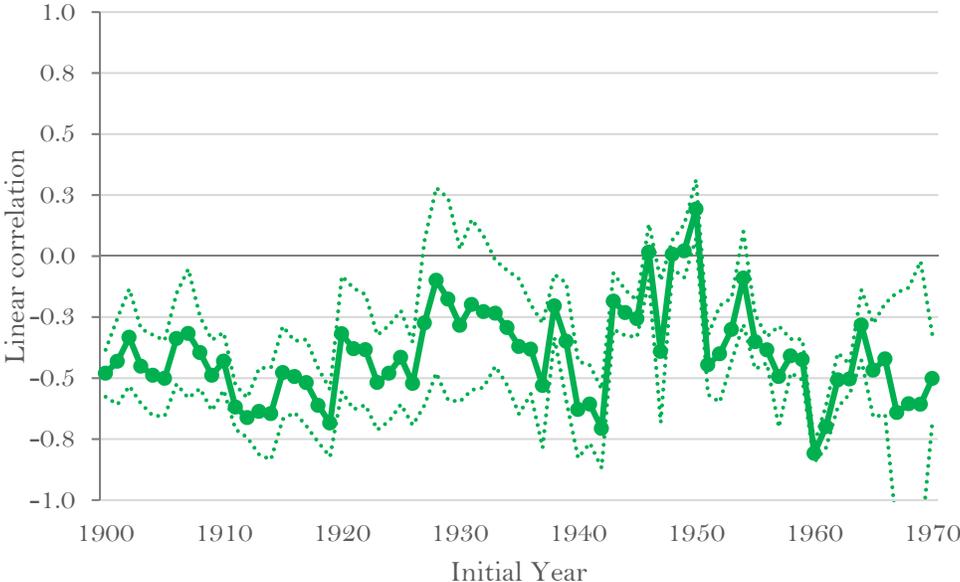


Note: This Figure shows the cross-country correlation between long-term growth and policy volatility during the 1940-2010 period. Long-term growth is measured as the percentage point change of the GDP per capita of each country (relative to the GDP per capita of the US). Policy volatility is measured as the percentage point change of the share of the top three topics in the presidential speech of the initial year. The topics were selected by the LDA on a country by country basis, and using a 10% minimum threshold to discard topics with low frequencies.

Consistent with the previous results, Figure 9 shows a strong association between policy stability and growth. Although there is some minor re-ranking of countries in terms of policy-volatility (e.g., Argentina and the República Bolivariana de Venezuela switch places at the bottom of the ‘policy volatility ranking’), we find the same qualitative results. In this case, the linear correlation is stronger (-0.84) and statistically significant at the 1% level, perhaps reflecting the better fit of the selected topics to each country. Again, the countries with the highest volatility (Argentina and the República Bolivariana de Venezuela) were the ones that fell behind the most, while the countries that caught up the most with the US (Chile and Spain) were the ones with the highest stability. In this case, policy volatility explains, on average, 70% of the cross-country variability of the long-term growth. The linear correlation between the policy volatility variable calculated in the two different ways is 0.35, while the Spearman correlation between the ranking of countries in terms of policy volatility is 0.53.

A second important decision is the year in which we start the analysis, as the initial year defines the top three topics in the initial speech (and therefore, the values of our policy variable). While in the baseline specification we start our analysis in 1940 to ensure reliable data, in Figure 10 we show the linear correlation between policy volatility and long-term growth, starting from different initial years, from 1900 to 1970.

Figure 10. Cross-country correlation between policy volatility and growth using different initial years of analysis

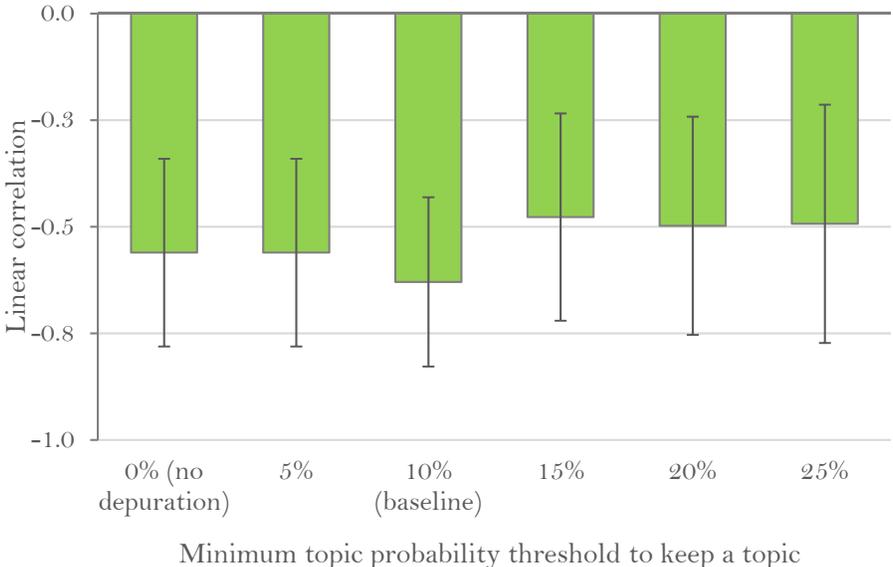


Note: This Figure shows the cross-country correlation between long-term growth and policy volatility starting in different initial years. Dashed lines show 95% confidence intervals. Long-term growth is measured as the percentage point change between different initial years and 2010 of the GDP per capita of each country (as a percentage of the GDP pc of US). Policy volatility is measured as the change between different initial years and 2010 of the topic probabilities of the top three topics of the presidential speech of the initial year. The topics were selected by the LDA using pooled data of all the countries, and using a 10% minimum threshold to discard topics with low frequencies.

Figure 10 shows that the cross-country correlation coefficients are remarkably stable across different initial years. The average linear correlation in the 1900-1970 period is -0.41, with some variability across time. The correlation remained stable at around -0.50 in the 1900-1940 period, with most of the confidence intervals below zero. Coefficients become more unstable by the end of the 1940s and beginning of the 1950s, with some coefficients taking positive values (although not statistically different from zero), then returning to values around -0.50 by the end of 1950s and 1960s.

Next, we show that our results are robust to choosing different thresholds to deplete irrelevant topics (i.e., topics with a very low topic probability).<sup>5</sup> In the baseline specification we use a 10% minimum threshold. This means that, if a topic does not have a probability of 10% or higher in any of the speeches of a country, then we exclude that topic from the analysis. The topic probability of 10% can be interpreted as the proportion of the speech devoted to a topic. Under this interpretation, we drop topics discovered by LDA that were not discussed in any speech of a country in a proportion higher than 10% of the total speech. In Figure 11 we show how the correlation between policy volatility and long-term growth varies depending on the depletion threshold.

Figure 11. Cross-country correlation between policy volatility and growth using different topic depletion minimum thresholds, 1940-2010



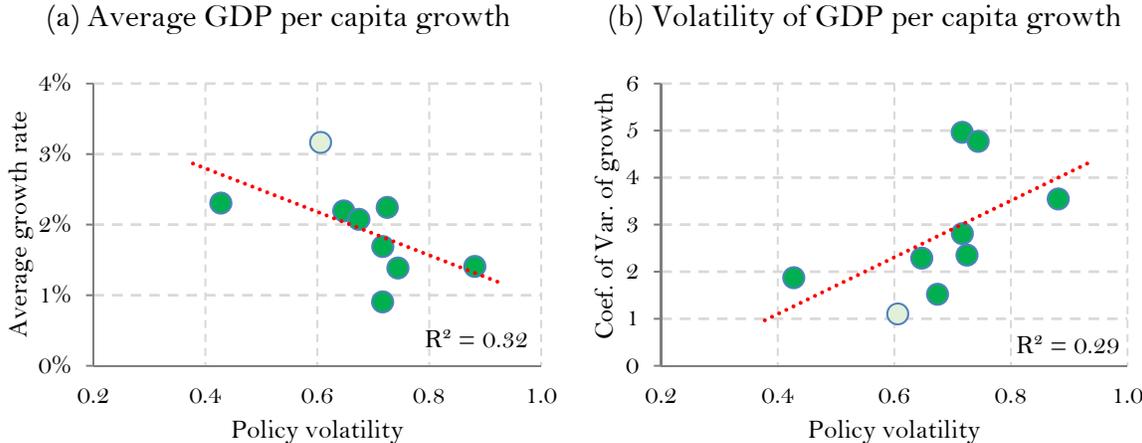
Note: This Figure shows the cross-country correlation between long-term growth and policy volatility during the 1940-2010 period, using different minimum threshold to discard topics with low frequencies. Long-term growth is measured as the percentage point change of the GDP per capita of each country (as a percentage of the GDP per capita of the US). Policy volatility is measured as the percentage point change of the share of the top three topics in the presidential speech of the initial year. The topics were selected by the LDA using pooled data of all the countries.

Figure 11 shows that the threshold choice makes only a marginal difference in quantitative terms. The correlation coefficients range from -0.49 using a 25% threshold, to -0.63 using a 10% threshold (the average correlation across thresholds is -0.53).

<sup>5</sup> In the appendix we show how the number of topics discussed in every speech varies by country depending on the depletion minimum threshold.

Next, we assess how the association between policy volatility and long-term growth varies when we use alternative measures of each of these variables. While in our baseline specification, we measure long-term growth based on whether the countries converged or diverged with the US economy in terms of GDP per capita, in Figure 11 we introduce two alternative measures of growth. In Figure 11a, we correlate our baseline measure of policy volatility with the average GDP per capita growth rate of countries in the 1940-2010 period, while in Figure 11b we correlate it with the volatility of growth during the same period, as measured by the coefficient of variation of the GDP per capita growth.

Figure 12. Correlation between policy volatility and different outcome variables, 1940-2010



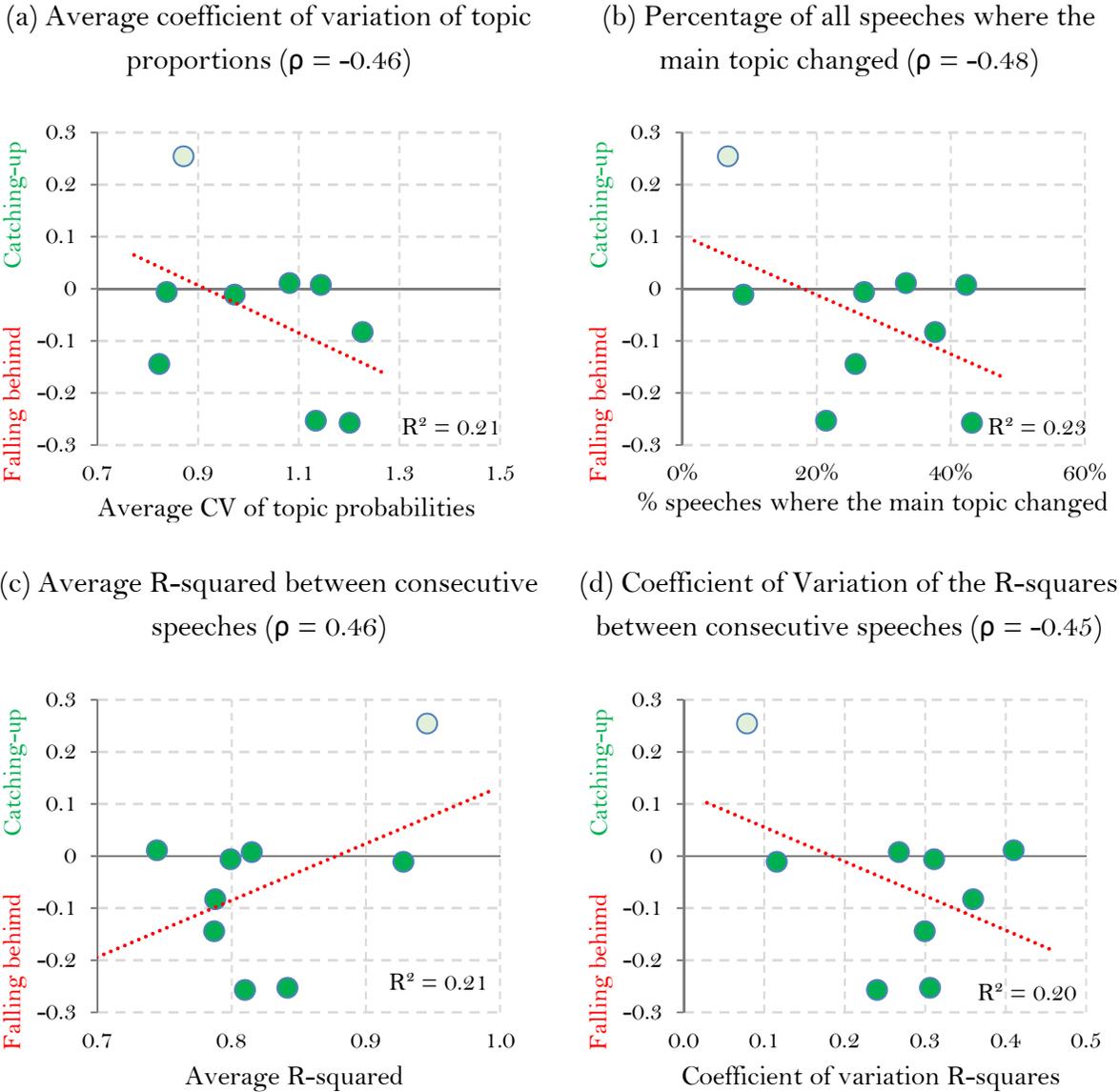
Note: This Figure shows the cross-country correlation between policy volatility and two measures of growth during the 1940-2010 period. In Panel (a) we correlate policy volatility with the average GDP per capita growth rate, while in panel (b) we do so with the coefficient of variation of the GDP per capita growth rate. Policy volatility is measured as the percentage point change of the share of the top three topics in the presidential speech of the initial year. The topics were selected by the LDA using pooled data of all the countries, and using a 10% minimum threshold to discard topics with low frequencies.

Figure 12 shows policy volatility is also associated with these two alternative outcomes. On the one hand, countries with lower average growth rates tend to be the ones with higher policy volatility (the correlation coefficient is  $-0.57$ ). On the other hand, countries with more volatile growth tend to also be the ones with higher policy volatility (the correlation coefficient is  $-0.54$ ), suggesting volatile policies drive volatile growth rates. On average, a third of the cross-country variability on the outcome variable is explained by our policy volatility variable.

So far, all our results have been focused on a single, intuitive measure of policy volatility: whether the top priorities of the countries remain the same across time. As a final robustness check, we assess how our results vary when we use alternative measures of policy volatility. We present four alternative measures of policy volatility. First, the average coefficient of

variation of topic probabilities, a measure of variability in the topic content. Second, the percentage of speeches where the topic with the highest probability changed (in relation with the previous speech), a measure of policy prioritization over time. Third, the average R-squared between consecutive speeches. Fourth, the coefficient of variation of the R-squares between consecutive speeches. Figure 13 shows the main results.

Figure 13. Correlation between different policy volatility measures and long-term growth, 1940-2010



Note: This Figure shows the cross-country correlation between four alternative measures of policy volatility and long-term growth during the 1940-2010 period. In Panel (a) we correlate long-term growth with the average coefficient of variation of topic probabilities, in panel (b) with the percentage of speeches where the topic with the highest probability changed, in panel (c) with the average r-squared between consecutive

speeches, and in panel (d) with the coefficient of variation of the r-squares between consecutive speeches. Long-term growth is measured as the percentage point change of the GDP per capita of each country (as a percentage of the GDP per capita of the US). The topics were selected by the LDA using pooled data of all the countries, and using a 10% minimum threshold to discard topics with low frequencies.

Figure 13 shows that our alternative measures of policy volatility yield consistent results with those previously shown. All our four alternative measures of policy stability have the expected correlation with long-term growth, and a similar explanatory power. On average, countries with higher policy stability have higher long-run growth rates. In this case, the goodness of fit of our measures is lower. On average, about a fifth of the cross-country variability of the long-term growth is explained by any of our policy volatility variables.

## 6. Conclusion

Frequent shifts in economic policies can increase uncertainty, resulting in lower investment and economic growth over the long run. This is an important but particularly difficult premise to test empirically. To begin with, how to measure policy volatility over long periods of time has proven to be elusive.

The first contribution of our paper is to provide a new proxy for policy volatility that exploits the information content of the priorities conveyed in presidential speeches. In most Latin American countries presidents are constitutionally mandated to give a ‘state-of-the-union’-type of speech annually. This is typically a programmatic speech which lays out achievements and priorities of the government in question. We built a new data set with 953 such presidential speeches from 10 Latin American countries and Spain. We then applied a common technique used in the literature on topic modeling, the Latent Dirichlet Allocation algorithm, to identify different topics and identify the share of each speech that is devoted to each topic. This allows us to observe the evolution of topics over time and detect significant differences across countries.

The second contribution of the paper is to provide suggestive evidence of the link between policy volatility and long-term growth. We show that our proxy of policy volatility is negatively correlated with long-term growth over the 1940-2010 period. Our results are robust to a large set of changes on how we construct our proxy of policy volatility.

More broadly, as in the case of Baker et al. (2016) who use the words in newspapers, our contribution demonstrates that drawing on text as data can help deepen our understanding of broad economic, political, and historical development.

## Appendix A. About the Data

The presidential speeches were obtained from the following sources:

**Argentina:** 1854-1868; 1872-1879; 1935-1940; 1942 ; 1946-1947; 1952-1955; 1958-1961; 1963-1966; 1973-1975; 1983-1994 (Provided by Chamber of Deputies); 1995-2001 ([Casa Rosada](#)); 2002 ([lavo.com](#)); 2003-2007 ([Casa Rosada](#)) 2008 ([wikisource](#)); 2009 ([cfkargentina](#)); 2010 ([parlamentario/noticias](#)); 2011 ([cfkargentina](#)); 2012 ([elloropolitico](#)); 2013 ([wikisource](#)); 2014 ([elloropolitico](#)); 2015 ([cfkargentina](#)); 2016 ([lanacion](#)).

**Chile:** 1926-1929; 1932-1933; 1935-1936; 1939; 1941-1945; 1947; 1949; 1950-1953 1955-1956; 1959; 1961; 1965; 1970-1976; 1980-2016; ([Biblioteca de Congreso Nacional](#)).

**Colombia:** 1999-2001 ([ideaspaz.org](#)); 2003 ([Presidencia Colombia](#)); 2004 ([Portal Colombia.com](#)); 2005 ([historico.presidencia.gov.co/discursos/](#)); 2006 ([El tiempo](#)); 2007 ([alvarouribevelez.com.co](#)); 2008 ([elespectador.com](#)); 2009 ([eltiempo.com](#)); 2010 ([historico.presidencia.gov.co](#)); 2011 ([wsp.presidencia.gov.co](#)); 2012 ([e.exam-10.com/ekonomika/13771/index.html](#)); 2013 ([wsp.presidencia.gov.co](#)); 2014 ([semana.com](#)); 2015 ([wp.presidencia.gov.co](#)); 2016 ([elheraldo.co](#)).

**Costa Rica:** 1878-1879; 1883-1892; 1894-1916; 1918-1922; 1924-1935; 1937-1947; 1949-2016. (<https://sites.google.com/site/mensajepresidencialcr/>).

**Dominican Republic:** 1999 ([leonelfernandez.com](#)); 2000 ([leonelfernandez.com](#)); 2001 ([Gobierno Hipolito](#)); 2002 ([Gobierno Hipolito](#)); 2003 ([Gobierno Hipolito](#)); 2004 ([Gobierno Hipolito](#)); 2005 ([diariolibre.com/](#)); 2006 ([leonelfernandez.com](#)); 2007 ([listindiario.com](#)); 2008 ([pld.org.do](#)); 2009 ([leonelfernandez.com](#)); 2010 ([leonelfernandez.com](#)); 2011 ([diariodominicano.com](#)); 2012 ([elnacional.com.do](#)); 2013 ([listindiario.com](#)); 2014 ([listindiario.com](#)); 2015 ([tirapiedras.com](#)); 2016 ([almomento.net](#)).

**Ecuador:** 1853; 1856-1858; 1861; 1863-1865; 1867; 1873; 1875; 1880; 1883; 1885; 1886-1888; 1890; 1892; 1894; 1898-1900; 1902-1924; 1928; 1932-1943; 1946; 1948-1949; 1951-1952; 1955; 1958-1962; 1966; 1968-1969; 1980-1983; 1985; 2002; 2006; 2008; 2010-2012; 2014-2016 (Provided by National Assembly of Ecuador).

**Mexico:** 1867-1871; 1873-1894; 1897-1903; 1905-1906; 1909; 1913-1914; 1917-1919; 1921-1936; 1938-1976; 1978-1989; 1993-2013; ([Biblioteca Garay](#)); 2014 ([Gobierno de México](#)); 2015 ([Gobierno de México](#)).

**Paraguay:** 1881-1889; 1891; 1898-1900; 1902-1904; 1908-1909; 1917-1921; 1925-1926; 1928; 1934-1935; 1939; 1955-1958; 1960-1963; 1976-1977; 1980; 1982; 1986-1988; 1992-1996; 1998-2014 (Provided by Congressional Library of Paraguay); 2016 ([mre.gov.py](#)).

**Peru:** 1829; 1832; 1834-1836; 1839-1840; 1845; 1847; 1849-1851; 1853; 1855; 1858-1860; 1862-1864; 1867; 1869-1870; 1872-1876; 1878-1879; 1881; 1883-1918; 1920-1928; 1930; 1940-1961; 1963-1966; 1968- 2016 ([Museo del Congreso](#)).

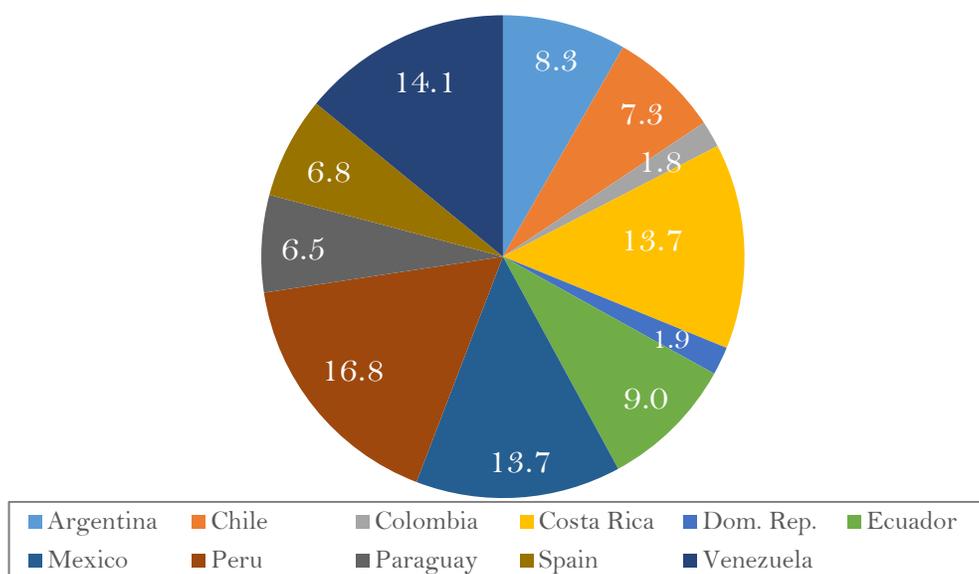
**Spain:** 1939; 1946-1974; 1979; 1981-2015

(<http://www.congreso.es/portal/page/portal/Congreso/Congreso/Diputados/Historia>).

**República Bolivariana de Venezuela:** 1819; 1830-1856; 1858; 1860; 1865-1868; 1870; 1873; 1875; 1876-1877; 1881-1883; 1885-1891; 1896; 1898; 1901-1905; 1907; 1909-1945; 1947-1948; 1954-1957; 1959-1969 (Antonio Moreno (1972), "Mensajes Presidenciales"); 1975 (Rafael Caldera, "Metas de Venezuela") 1976-1978 (Carlos Andres Perez, "Manos a la Obra"); 1984; 1988;1989 (Jaime Lusinchi, "Papeles del Presidente"); 1994-1995; 1997-1998(Rafael Caldera, "Compromiso Solidario); 1999 ([Formación PSUV](#)); 2000 ([Formación PSUV](#)); 2001 ([Formación PSUV](#)); 2002 ([Formación PSUV](#)); 2003 ([Formación PSUV](#)); 2004 ([Formación PSUV](#)); 2005 ([Formación PSUV](#)); 2006 ([Formación PSUV](#)); 2007 ([Todo Chavez](#)); 2008 ([Formación PSUV](#)); 2009 (aristobulo.psuv.org.ve); 2010-2012 ([Todo Chavez](#)); 2014 ([vtv.gob.ve/](#)).

Considering all the speeches mentioned above, the different countries are represented in our database according to Figure A1.

Figure A1. Share of total speeches in each country (in %)



Note: Authors' elaboration.

The distribution of speeches by decade and country is shown in Table A1.

### About the missing speeches

Although we made efforts to compile as many presidential speeches as possible, there are a number of countries for which the speeches from certain years are missing. In what follows, we attempt to outline what we believe are the most probable causes.

For some years, speeches are missing because no speech was delivered. For example, in the case of Paraguay, the president did not deliver a speech throughout the period 1940-1948. Political turmoil, such as a coup or an ongoing revolution, is another probable cause, particularly for missing speeches from the 1970s. Looking at Figure 1 in the first section, which presents the number of speeches in each country by decade, the graph dips visibly in the 1970s, reflecting a lower number of speeches for that decade. Considering the multiple coups and revolts taking place in the region during this time, some of the missing speeches are likely associated with this turbulent political context.

For other missing speeches, it is hard to establish a cause with certainty. In some cases, the congressional libraries were unable to locate the physical copies, in other cases the quality of the original copies was too low for proper digitalization. At times, particularly in the case of the República Bolivariana de Venezuela, we were unable to find publications containing the compilations of the presidents' speeches for certain years.

Table A1 Countries with constitutionally mandated presidential speeches

<b>Country</b>	<b>Article</b>
Argentina	99, subsection 8
Bolivia	96, subsection 10
Colombia	189, subsection 12
Costa Rica	139, subsection 4
Chile	24
Ecuador	171, subsection 7
Guatemala	183, subsection i
Honduras	245, subsection 8
Mexico	69
Nicaragua	150, subsection 15
Panama	178, subsection 5
Paraguay	238, subsection 8
Peru	118, subsection 7
Venezuela, RB	237
Dominican Republic	55, subsection 22
Uruguay	168, subsection 5

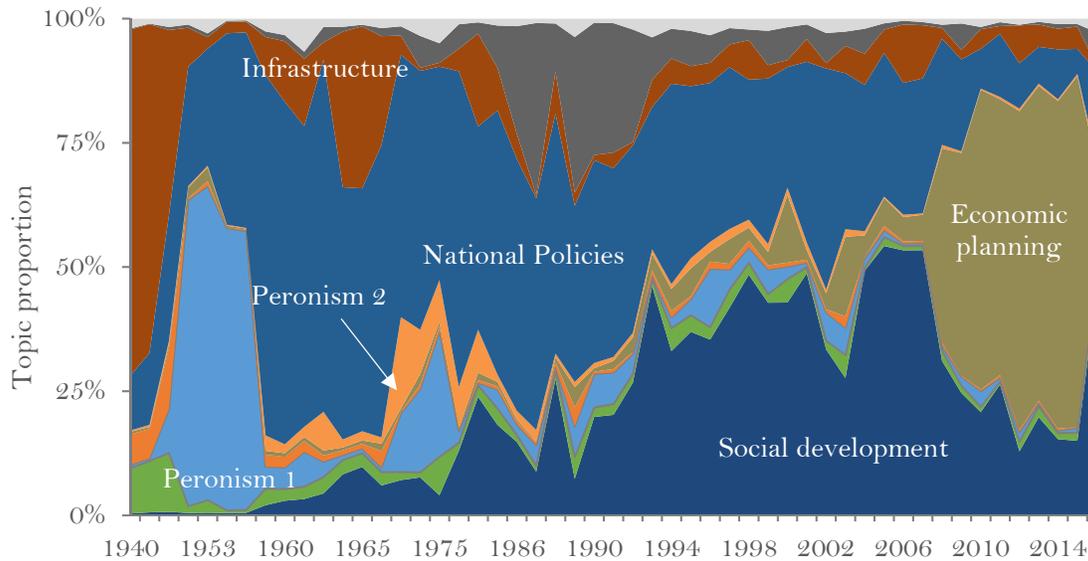
## Appendix B. Country-specific topic discovery

As an illustrative example of the country-specific LDA, we present results for Argentina and Peru in Figures B1 and B2, respectively. The respective topics were labeled based on the top 10 most probable words in each topic, which we present in Tables B1 and B2, respectively.

The figures illustrate two important points that are relevant to our analysis. First, LDA detects important shifts in the political context of a country with considerable accuracy. In the case of Argentina, the emergence of the Peronist doctrine during Juan Domingo Perón's presidency (1946-1955) is clearly reflected in the pronounced increase in prevalence of the "Peronism 1" topic that is defined by terms related to Peronism, such as "doctrine" or "Peronist." Moreover, a second topic related to Peronism—which we label "Peronism 2"—emerges during his 1973-1974 presidency. This could suggest that LDA not only detected the rise of the clearly distinct political discourse of Perón, but that it also detected a change within this discourse. LDA's ability to detect important shifts in the political context can also be observed in the case of Peru. In 1968, a topic concerned with leftist revolutionary discourse emerges abruptly and dominates speech content until 1980, rapidly falling in prevalence thereafter. This period coincides very closely with the revolutionary military government of Juan Velasco Alvarado, who came to power in 1968 and was eventually overthrown in 1975.

The second point that the figures below illustrate is that the stability or consistency of the topics varies by country. For example, in Peru, some of the topics, such as the "infrastructure" topic, or the "public services" topic, exhibit significant volatility from one year to the next. They are also not consistent over long stretches of time, as these topics almost disappear after the 1980s. On the other hand, the topics discovered for Argentina are more consistent over time. For example, the "national policies" topic has high prevalence in the 1940s, and continues to do so well into the early 2000s.

Figure B1. Topic content of the presidential speeches of Argentina using country-specific LDA, 1940-2016



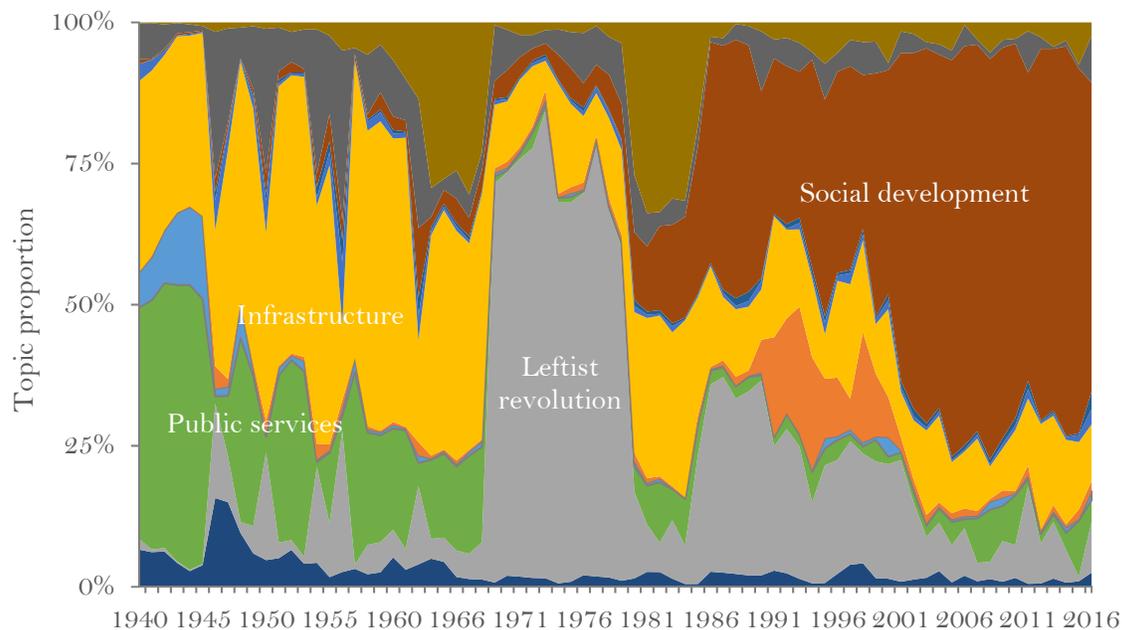
Note: This Figure shows the topic probabilities estimated with LDA for the presidential speech given each year. The topic probabilities can be thought of as representing the proportion of each document that belongs to each topic. Only topics whose average document probability exceeds 10% in at least one decade are presented, the rest are grouped as a residual.

Table B1. Top probability terms defining the topics in Argentina  
(In descending order of probability of term within the topic)

“Infrastructure”	“Peronism 1”	“Peronism 2”	“National policies”	“Social development”	“Economic planning”
Public works	The people	liberation	national	social	sector
executive	social	Peron	policy	growth	plan
national	doctrine	reconstruction	social	system	bank
number	economic	Justicialista	nation	education	world
currency	men	science	development	investment	people
services	justice	area	The people	program	economy
construction	nations	educational	life	development	province
increase	economy	banks	production	quality	justice
exercise	homeland	university	right	security	workers
plan	Peronist	rural	effort	policies	growth

Note: This Table shows the words that have the highest probability of belonging to each topic. Because LDA defines topics as distributions over words, the lists of terms define the subject of each topic.

Figure B2. Topic content of the presidential speeches of Peru using country-specific LDA, 1940-2016



Note: This Figure shows the topic probabilities estimated with LDA for the presidential speech given each year. The topic probabilities can be thought of as representing the proportion of each document that belongs to each topic. Only topics whose average document probability exceeds 10% in at least one decade are presented, the rest are grouped as a residual.

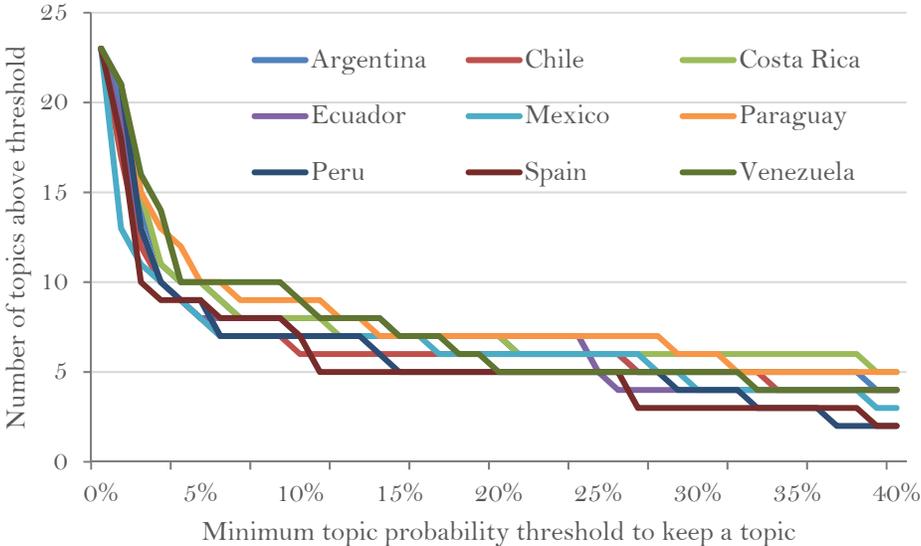
Table B2. Top probability terms defining the topics in Peru  
(In descending order of probability of term within the topic)

“Public services”	“Infrastructure”	“Revolutionary left”	“Social development”
service	national	revolution	national
national	Public works	The people	program
Public works	education	policy	investment
jobs	production	development	social
school	plan	social	health
department	social	process	resources
direction	construction	national	Public works
object	activities	economic	growth
studies	ministry	revolutionary	project
services	bank	companies	security

Note: This Table shows the words that have the highest probability of belonging to each topic. Because LDA defines topics as distributions over words, the lists of terms define the subject of each topic.

# Appendix C. Additional Figures and Tables

Figure C1. Number of topics above minimum topic probability threshold by country



This Figure shows the number of topics in each presidential speech that have a topic probability above each threshold in any of the years.

## References

- Baker, S.R., N. Bloom, and S.J. Davis (2016). “Measuring economic policy uncertainty.” *Quarterly Journal of Economics* 131(4):1593–1636. doi:10.1093/qje/qjw024.
- Bernanke, B.S. 1983. “Irreversibility, Uncertainty, and Cyclical Investment.” *Quarterly Journal of Economics* 98(1):85–106.
- Blei, D.M., A. Y. Ng, and M. I. Jordan (2003). “Latent dirichlet allocation.” *Journal of Machine Learning Research* 3:993–1022.
- Blei, D. M. (2012). “Probabilistic Topic Models.” *Communications of the ACM* 55(4):77–84.
- Blei, D., Lafferty, J. (2006). “Dynamic topic models.” *International Conference on Machine Learning ACM*, 113–120.
- Bloom, N. (2009). “The Impact of Uncertainty Shocks.” *Econometrica* 77 (3): 623–85.
- Bloom, N. (2014). “Fluctuations in uncertainty.” *Journal of Economic Perspectives* 28(2):153–176.
- Bryan, K., L. Pastor, and P. Veronesi (2016), “The Price of Political Uncertainty: Theory and Evidence from the Option Market,” *Journal of Finance* 71:2417—2480.
- Chute, C., D.C. Li, T. Thermeau, and H. Liu (2014). “Discovering associations among diagnosis groups using topic modeling.” *AMIA Summits on Translational Science Proceedings*, 43.
- Commission on Growth and Development (2008). *The Growth Report. Strategies for Sustained Growth and Inclusive Development*. World Bank on behalf of the Commission on Growth and Development chaired by Professor Michael Spence.
- Dixit, A.K., and R.S. Pindyck (1994). *Investment Under Uncertainty*. Princeton, NJ: Princeton University Press.
- Dulanto, F., (2006). “Mensajes presidenciales y otros documentos para la historia política del Perú.”
- Fernández-Villaverde, J., P. Pablo Guerrón-Quintana, K. Kuester, and J. Rubio-Ramírez (2015), “Fiscal volatility shocks and economic activity,” *American Economic Review* 105(11): 3352–3384. <http://dx.doi.org/10.1257/aer.20121236>
- Griffiths, T. L. and M. Steyvers (2004). “Finding scientific topics.” *Proceedings of the National Academy of Sciences of the United States of America* 101(Suppl. 1):5228–5235.
- Grimmer, J. (2009). “A bayesian hierarchical topic model for political texts: Measuring expressed agendas in senate press releases.” *Political Analysis* 18(1): 1–35.

- He, J., W. X., Zhao, J., Jiang, J., Weng, E.P., Lim, H., Yan Jiang and X., Li (2011). "Comparing twitter and traditional media using topic models." *European Conference on Information Retrieval*: 338-349.
- Hornik, K. and Grün, B. (2011). "topicmodels: An R package for fitting topic models." *Journal of Statistical Software* 40(13): 1-30.
- Jurado, K., S. Ludvigson, and S. Ng (2015). "Measuring Uncertainty," *American Economic Review* 105, 1177–1216.
- McDonald, R.T. and I. Titov (2008). "A Joint Model of Text and Aspect Ratings for Sentiment Summarization." *ACL*, 8: 308-316.
- Quinn, K. M., Monroe, B. L., Colaresi, M., Crespín, M. H. and Radev, D. R (2010), "How to Analyze Political Attention with Minimal Assumptions and Costs." *American Journal of Political Science* 54: 209–228.
- Shiller, R.J. (2017). "Narrative economics," *American Economic Review* 107(4):967-1004.
- Tommasi, M. (2006), "The institutional foundations of public policy," *Economía* 6(2):1-36  
10.1353/eco.2006.0022
- Wang, X. and A. McCallum (2006). "Topics over time: a non-Markov continuous-time model of topical trends." *Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining (KDD '06)*. 424-433.
- World Bank (2016). Governance and the Law. World Development Report 2017
- Yildirim, I. (2012). "Bayesian inference: Gibbs sampling." Technical Note, University of Rochester