

A Text-Based Measure of Policy Volatility*

Oscar Calvo-González Axel Eizmendi Germán Reyes

August 2026

Abstract

We propose a text-based measure of policy-agenda volatility, built from the similarity between the policies discussed in a government’s consecutive programmatic statements. Whereas existing policy volatility measures require fiscal statistics available only for recent decades, political text reaches back centuries. We implement the measure on 1,149 presidential addresses delivered in Latin America between 1819 and 2021 and validate it: the policy agenda expressed in presidential addresses shifts when power changes hands, with larger estimated movements following irregular transitions, and larger agenda movements coincide with larger changes in government spending. We use the measure in three applications. First, policy-agenda volatility rises and falls over the two centuries, peaking in the middle of the twentieth century. Second, in the nineteenth century, more democratic countries had more volatile policy agendas, while the region’s entrenched autocracies repeated a nearly fixed agenda for decades; in recent decades, with entrenched autocracies rare, the relationship reverses. Finally, we study the relationship between nineteenth-century volatility and income today, and early democracy’s role in it.

*Calvo-González: World Bank (ocalvogonzalez@worldbank.org), Eizmendi: Tufts University (axel.eizmendi_larrinaga@tufts.edu), Reyes: Middlebury College and IZA (greyes@middlebury.edu). A previous version of this paper circulated under the title “Winners Never Quit, Quitters Never Grow: Using Text Mining to Measure Policy Volatility and Its Link with Long-Term Growth in Latin America” (World Bank Policy Research Working Paper 8310). For helpful comments, we thank Laura Chioda, Alexandru Cojocaru, Bill Maloney, Ambar Narayan, Roy Van der Weide, and participants from the Poverty GP lunch seminar. We acknowledge funding from the World Bank Research Department. Errors and omissions are our own.

1 Introduction

Stable policy is widely viewed as a foundation for economic development. When governments shift priorities erratically, firms postpone investment, public agencies struggle to build capacity, and households cannot form reliable expectations (Rodrik, 1991; Dixit and Pindyck, 1994; Henisz, 2000). Cross-country evidence supports this view: discretionary changes in fiscal policy are associated with lower growth (Fatás and Mihov, 2013), political instability depresses investment (Alesina et al., 1996; Aisen and Veiga, 2013), and spikes in policy uncertainty predict declines in employment and output (Baker et al., 2016).

Any long-run question about policy volatility, however, requires a measure that spans centuries. Existing measures are typically built from data on realized policy (e.g., the composition of government spending, tax rates, or tariff schedules) produced by national statistical agencies and finance ministries. Most countries began collecting these data only in the second half of the twentieth century; many developing economies began decades later. Thus, no such measure exists for the nineteenth century or earlier, the eras whose conditions a large literature links to development today (e.g., Acemoglu et al., 2001; Nunn, 2008; Dell, 2010).

This paper proposes a proxy for realized policy volatility built from the policy agenda: the policies presidents discuss in their annual addresses to Congress. Such addresses are delivered regularly in many countries, and the practice began long before states collected statistics. In Spanish-speaking Latin America, for example, constitutions dating to the early nineteenth century mandate that presidents deliver an annual address to Congress—the regional counterpart of the U.S. State of the Union—reviewing their administration’s accomplishments and laying out its priorities.

Our measure captures how much the government’s policy priorities change from one year’s address to the next over a given period. We build it in three steps, illustrated in Figure 1. First, following the text-as-data literature (Grimmer and Stewart, 2013; Gentzkow et al., 2019), we use a natural-language-processing algorithm, Latent Dirichlet Allocation (LDA), to represent each address as a vector of topic shares: the estimated share of its content that the model attributes to different policies, such as war, infrastructure, or social welfare (Blei et al., 2003; Blei, 2012). Second, we compute the cosine similarity of consecutive addresses’ topic-share vectors, a standard measure of the alignment between two vectors, and define policy-agenda change as one minus this similarity. Policy-agenda change equals zero when two addresses divide their content across policies identically and

approaches one when they emphasize entirely disjoint policies. Third, we define a country’s policy-agenda volatility over a period as the average of these year-over-year changes.

We implement the measure on a corpus of 1,149 presidential addresses delivered in ten countries over two centuries (1819–2021). Our analysis sample comprises eight countries with speech series long and dense enough to compute within-country volatility: Argentina, Chile, Costa Rica, Ecuador, Mexico, Paraguay, Peru, and Venezuela. Nothing in our approach, however, is specific to presidential addresses or to Latin America: the measure applies to any long-running series of programmatic government text, such as State of the Union messages (Rule et al., 2015), throne speeches (John and Jennings, 2010), party manifestos (Budge et al., 2001), or the annual statements each UN member state delivers to the General Assembly (Baturu et al., 2017).

We conduct two validation exercises that show that the measure registers shifts in the policy agenda and that those shifts carry information about realized policy. First, the measure rises when power changes hands, and more so when the change is irregular: relative to consecutive addresses with no transition between them, the agenda moves about twice as much when an electoral handover separates the two addresses, and roughly two and a half times as much when an irregular transition, such as a coup or a revolution, does. Second, larger agenda shifts coincide with larger fiscal changes, as they should if addresses carry some signal about policy implementation rather than rhetoric alone: in annual central-government spending data, years with larger agenda movement show larger changes in total spending.

We then present three applications as proof of concept for the questions the measure can help address. First, we trace policy-agenda volatility over the two centuries and decompose it into two types: within-president volatility, in which a given government shifts its own policy priorities from year to year, and between-president volatility, in which changes of government move the policy agenda. We document a rise and fall of policy-agenda volatility. Volatility remains relatively constant from about 1850 through 1930, triples by about 1950, stays near its peak into the mid-1970s, and falls back after 1980. Both types of volatility rise and fall together, and the surge reflects the size of agenda movements rather than the frequency of turnover. With data beginning in 1980, we would observe only low and slowly declining volatility and would miss the mid-century surge; the two-century series shows the present as a return to nineteenth-century levels.

In our second application we ask what kind of political system exhibits higher policy-agenda volatility. Consistent with modern evidence that democracies and constrained

executives have more stable policy ([Henisz, 2004](#); [Dutt and Mobarak, 2016](#)), we find a negative correlation between policy-agenda volatility and democracy in the twenty-first century. Extending the time period *reverses* this conclusion. In the nineteenth and twentieth centuries the correlation is positive, that is, more democratic countries have *more* volatile agendas. The sign change partly reflects the changing mix of regimes. The least volatile agendas in our sample belong to the region’s long autocracies—Porfirian Mexico, where the Díaz regime emphasized nearly identical developmentalist policies for decades, and postwar Paraguay—and regimes of that kind have become rare. Thus, a stable policy agenda can be the product of entrenched autocracy, and high volatility can partly reflect political competition. Without historical data on policy-agenda volatility, we would miss this shift.

In our final application we examine whether policy-agenda volatility in the nineteenth century predicts development today. We correlate nineteenth-century volatility with log GDP per capita in 2022. Across specifications, outcomes, and samples—including a country-decade panel that tests the relationship within countries—we find no robust evidence that policy-agenda volatility is associated with lower long-run growth. Yet the same design recovers the positive relationship between democracy and income today ([Acemoglu et al., 2019](#)). Why does volatility not predict development? One hypothesis follows from the second application: volatility moves with democracy, so its association with income combines a positive democracy component with a negative component of volatility itself, and the two can cancel. Consistent with this hypothesis, conditioning on nineteenth-century democracy turns the volatility coefficient negative. The association reflects sitting presidents shifting their own priorities from year to year: conditional on nineteenth-century democracy, within-president volatility over 1819–1939 correlates negatively with growth over 1940–2022, while between-president volatility, which arises when power changes hands, is less strongly related.

We contribute to the text-as-data literature in political economy ([Grimmer and Stewart, 2013](#); [Gentzkow et al., 2019](#); [Hansen et al., 2018](#)) by measuring a new object over a longer horizon: the reallocation of stated policy priorities from one executive speech to the next, tracked over two centuries. Existing text-based policy indicators measure perceived uncertainty about future policy at high frequency over recent decades: [Baker et al. \(2016\)](#) build an economy-wide uncertainty index from newspaper coverage, and [Hassan et al. \(2019\)](#) measure firm-level political risk from earnings calls. The closest antecedents are in the comparative-agendas tradition of political science, which codes executive speeches into

issue categories and studies the stability of government agendas in postwar democracies (John and Jennings, 2010; Mortensen et al., 2011; Camargo et al., 2021). Our measure complements both: it captures the reallocation of stated priorities rather than perceived uncertainty, its series begins roughly a century earlier than existing systematic indicators, and we validate it against realized-policy counterparts in the decades where those data exist. Our measure is the policy analogue of the firm-level statistic of Cohen et al. (2020): one minus the cosine similarity between a firm’s consecutive quarterly and annual filings predicts returns, earnings, and bankruptcies.

We also add a long-run, direct measure to the literature relating policy volatility and political instability to growth (Rodrik, 1991; Alesina et al., 1996; Henisz, 2000; Acemoglu et al., 2003; Aisen and Veiga, 2013; Fatás and Mihov, 2013), and to the political economy of policymaking in Latin America (Spiller and Tommasi, 2003; Stein et al., 2006; Acuña et al., 2007; Doyle, 2014). Whereas existing volatility measures in this literature are built from fiscal aggregates or expert assessments available only for recent decades, ours extends the same object back to the early nineteenth century. The long horizon lets us test a channel proposed in the literature on institutions and long-run development (North, 1990; Glaeser et al., 2004; Galiani and Sened, 2014; Dincecco and Katz, 2016; Prados de la Escosura, 2009): whether the stability of policy priorities is a path from early institutions to income today.

The remainder of the paper proceeds as follows. Section 2 describes the speech corpus and the outcome and covariate data. Section 3 constructs the volatility measure and validates it. Section 4 presents the three applications. Section 5 concludes.

2 Data

This section describes the speech corpus, the analysis sample, and the outcome and covariate data.

2.1 Corpus of Presidential Speeches

Our starting point is the corpus of presidential speeches assembled in Calvo-González et al. (2026): 1,149 annual addresses delivered between 1819 and 2021 in ten Spanish-speaking Latin American countries. In these constitutionally mandated speeches, the president reports to Congress on the state of the nation and lays out priorities for the year ahead. The speeches thus form a long, reasonably regular time series of statements of government

priorities, delivered by the head of state in a comparable format across countries and centuries.

The speeches were collected from national congresses, congressional libraries, official online repositories, and digitized volumes held by the U.S. Library of Congress (Appendix B summarizes the sources). Scanned documents were processed with optical character recognition and manually corrected. All text analysis uses the original Spanish. Calvo-González et al. (2026) document the corpus in detail, including a missing-speech analysis showing that availability is largely unrelated to observable country characteristics.

2.2 Analysis Sample

Measuring long-term policy-agenda volatility requires a long series of speeches. Two of the ten corpus countries, Colombia and the Dominican Republic, have speech series that begin only in recent decades, too short to measure within-country volatility. Our analysis sample comprises the remaining eight countries: Argentina, Chile, Costa Rica, Ecuador, Mexico, Paraguay, Peru, and Venezuela, with 1,104 speeches in total.¹

We introduce two further sample restrictions. First, because topic shares estimated from a short address carry sampling error that inflates measured change, we require the shorter address of each pair to have at least 1,000 words, roughly the fifth percentile of address length (Appendix Figure A1 shows the distribution). This restriction eliminates 78 of the 1,096 pairs of consecutive available addresses (7 percent). Second, because longer gaps between consecutive speeches mechanically inflate measured change and cluster in turbulent periods, we keep only pairs of addresses delivered at most three years apart. This restriction eliminates 32 of the remaining 1,018 pairs (3 percent), leaving 986 pairs for analysis. Our results are robust to relaxing both restrictions (Appendix Table A1).

Table 1 reports each country’s summary statistics. Coverage begins between 1819 (Venezuela) and 1881 (Paraguay) and extends to 2021 in every country (Appendix Figure A2 plots cumulative coverage). The average country contributes 138 addresses, and the average address contains about 11,000 words, which corresponds to about an hour and a half of continuous delivery at a typical speaking pace of 130 words per minute.

¹We use the data from Colombia and the Dominican Republic in two places: we estimate the topic model of Section 3 on the full ten-country corpus, and both countries enter the validation analysis of Section 3.3, which requires only recent data.

2.3 Outcome and Covariate Data

Presidential transitions. Transitions and their type (electoral versus irregular) come from Calvo-González et al. (2026), complemented with the Center for Systemic Peace coup list (Marshall and Marshall, 2021).

Government spending. Central-government total expenditure and social expenditure by function (education, health, housing, and social protection), both as shares of GDP, come from the United Nations Economic Commission for Latin America and the Caribbean (ECLAC) and cover 1990–2019. The inputs to the Fatás and Mihov (2013) measure of discretionary fiscal volatility—real government consumption, real GDP, and GDP-deflator inflation—come from the World Development Indicators.

Political institutions. Our democracy measure is the Polity 2 score, available annually from 1800 (Marshall and Gurr, 2020). For robustness we also use Polity’s executive-constraints index and the V-Dem electoral-democracy and liberal-democracy indices (Coppedge et al., 2026).

Income. GDP per capita comes from the Maddison Project Database 2023 (Bolt and van Zanden, 2025).

Appendix B gives coverage details for each source.

3 A Text-Based Measure of Policy-Agenda Volatility

This section describes our measure of policy-agenda volatility. We construct the measure in three steps, illustrated in Figure 1: we represent each address as a vector of topic shares, compute the distance between consecutive addresses as one minus the cosine similarity of their vectors, and average these distances over the period of interest. We then validate the measure with two tests.

3.1 From Speeches to Topic Distributions

We first convert the raw text of each presidential address into a probability distribution over policy topics. We do so with LDA, the standard topic model of the text-as-data literature (Blei et al., 2003; Griffiths and Steyvers, 2004). LDA represents each document as a mixture of topics, in which each topic is a probability distribution over words (Appendix B describes the algorithm in more detail). The model assigns each speech d a vector of topic shares $\theta_d = (\theta_{d,1}, \dots, \theta_{d,K})$ with $\sum_k \theta_{d,k} = 1$. The share $\theta_{d,k}$ measures the fraction of speech

d devoted to topic k . For example, $\theta_{d,1} = 0.2$ means that the model attributes 20 percent of the speech’s content to the first topic. The topic shares summarize the attention an address devotes to different policy areas. Because the annual address is an official statement of the administration’s accomplishments and plans, we interpret θ_d as its expressed policy priorities in that year: which policies it presents as important and how much attention each receives. We do not assume that every stated priority is implemented or that every implemented policy appears in the address.

We estimate the model on the pooled Spanish-language corpus and choose the number of topics ($K = 33$) to minimize held-out perplexity (Appendix Figure B1), a standard measure of how poorly the model predicts the words of documents excluded from its estimation (Appendix B describes the choice of K). Appendix Table A2 lists the topics that average at least one percent of the corpus, with their top defining keywords. Because the corpus is pooled, the set of topics is fixed across countries and years, making the vectors comparable over the full two-century span.

3.2 From Topic Distributions to Policy-Agenda Volatility

The second step defines policy-agenda change between consecutive addresses. We want a notion of distance between the priorities expressed in two addresses: the farther apart the priorities, the larger the measured policy-agenda change. In principle, one could measure the distance between the *entire texts* of two consecutive addresses, for example, by embedding each text and computing the distance between embeddings (Le and Mikolov, 2014; Ash and Hansen, 2023). However, this approach would partly reflect changes in tone, vocabulary, and verbosity. We instead compute the distance between *topic distributions*, which isolate the share of the address devoted to each policy.²

Our benchmark distance is one minus the cosine similarity of the two addresses’ topic-share vectors. We use cosine distance because it is the standard document distance in text-as-data applications, keeping our measure comparable to the existing literature; we discuss alternative notions of distance below. Formally, for a country c with consecutive available speeches in years $t' < t$,

$$\Delta_{c,t} = 1 - \frac{\theta_{c,t} \cdot \theta_{c,t'}}{\|\theta_{c,t}\| \|\theta_{c,t'}\|}, \quad (1)$$

²In practice, both measures are highly correlated. Computing policy-agenda change from multilingual sentence embeddings of the raw text instead of topic distributions yields a pair-level correlation of 0.80 with our measure and a country-level correlation of 0.87 in the nineteenth century.

which equals zero when the two addresses allocate attention identically and approaches one when they emphasize entirely disjoint topics.

This measure has three desirable properties. First, $\Delta_{c,t}$ incorporates information on all the topics discussed in the speeches: a reallocation between any two policies moves the measure. Second, $\Delta_{c,t}$ is cardinal in the size of the reallocation: holding fixed which policies gain and which lose, a larger shift of shares produces a larger $\Delta_{c,t}$. Third, $\Delta_{c,t}$ is quadratic in the share differences: squaring makes the small estimation error that every share carries contribute little—a metric linear in the differences, such as total variation, accumulates that error even between near-identical addresses—and weights the largest reallocations most, so measured changes are mainly driven by movement in the addresses’ main priorities.³

Conceptually, any metric of distance between two probability distributions could serve in this step. To see if this choice matters empirically, we consider eight standard alternatives to cosine distance: total-variation distance, Jensen–Shannon divergence, Kullback–Leibler divergence, Hellinger distance, Euclidean distance, the Bhattacharyya coefficient, the Spearman rank correlation of topic shares, and the R^2 of a regression of one address’s topic shares on the other’s (Appendix B.4 defines each). Appendix Figure A3 reports the pairwise correlations among all nine measures (similarity measures enter as one minus the measure, so that larger values always mean more policy-agenda change). All measures are highly correlated: $1 - R^2$ is nearly identical to the benchmark (correlation = 0.98), and the six distribution-based measures form a tight cluster, correlating 0.85 to 1.00 with one another and 0.89 to 0.96 with the benchmark.⁴

In the third step, we define a country’s policy-agenda volatility over a window of years T as the average policy-agenda change across consecutive addresses at most three years apart in that window:

$$\text{Volatility}_{c,T} = \frac{1}{N_{c,T}} \sum_{t \in T} \Delta_{c,t}, \quad (2)$$

³Not every notion of distance has these properties. For example, a measure based only on the ranking of topics (such as the Spearman-based alternative considered below) fails the first two. If consecutive addresses allocate shares of (0.5, 0.3, 0.2) and (0.6, 0.3, 0.1) across three policies, the ranking is identical in both years and thus a rank-based distance is zero. Doubling the shift to (0.7, 0.3, 0.0) still leaves it at zero, while $\Delta_{c,t}$ rises from 0.02 to 0.06.

⁴The exception is the Spearman-based measure, which correlates only 0.09 to 0.21 with the rest. It depends only on the ranking of topics, and the ranking barely moves between consecutive addresses: a measure that tracks rank reversals shares little variation with measures that track how much probability mass moves.

where the sum runs over the years in T (e.g., the nineteenth century) and $N_{c,T}$ counts those pairs.⁵

3.3 Validation

We next validate the measure in two steps. First, we ask whether it captures movements in the policy agenda by examining whether measured change rises when political power changes hands. Second, we ask whether the measured agenda movements have signal value about realized policy by examining whether they correlate with larger changes in government spending.

Presidential Transitions. Our first test asks how the measure responds when political power changes hands. If the measure captures the volatility of policy priorities, then the change in policy agenda should be greater when power changes hands and it should rise by even more when the transition is irregular than when it is electoral.

We implement the test in two ways. First, in an event-study design, we compute each address’s policy-agenda change relative to the outgoing president’s final address and average it by year relative to the transition, separately for electoral transitions (constitutional handovers following elections) and irregular ones (coups, revolutions, forced resignations, and deaths in office). The analysis sample contains the 304 transitions with at least one address within five years of the change of power. Second, in Table 2 we estimate

$$\Delta_{c,t} = \alpha_c + \gamma_t + \beta_1 \text{Irregular}_{c,t} + \beta_2 \text{Electoral}_{c,t} + \varepsilon_{c,t}, \quad (3)$$

where $\text{Irregular}_{c,t}$ and $\text{Electoral}_{c,t}$ indicate that an irregular or an electoral transition falls between the two addresses that enter $\Delta_{c,t}$, α_c is a country fixed effect, and γ_t is a year fixed effect. Under the first prediction, β_1 and β_2 are both positive; under the second, β_1 exceeds β_2 . We cluster standard errors at the country level. With only eight clusters, asymptotic cluster-robust inference over-rejects (Cameron et al., 2008), so we rely mainly on wild-cluster bootstrap p -values (see Appendix B.5 for implementation details).

⁵An alternative to measuring distance over consecutive addresses would be to average the distance over all pairs of addresses in the window T . The shorter the window, the closer the two constructions stay. Over longer windows they diverge: addresses centuries apart differ partly because priorities shift as states age—from the wars and internal order of new republics to the social services of established states (Calvo-González et al., 2026)—so the all-pairs average mixes this secular reallocation with year-over-year instability. In practice the two are highly correlated: 0.73 across trailing twenty-year country windows and 0.71 across countries in the nineteenth-century benchmark window.

Presidential transitions are accompanied by pronounced shifts in the policy agenda, with larger average movements following irregular transitions. In the event study, policy-agenda change averages only 0.05 during the three years preceding a transition (Figure 2). Over the first three years after a transition, it rises to 0.10 when the handover is electoral (nearly twice the pre-transition level) and to 0.16 when it is irregular (nearly three times that level). Consistent with this, the regression estimates show that both transition coefficients are positive in every specification (no fixed effects, country fixed effects, and two-way fixed effects) and that the irregular point estimate always exceeds the electoral one (though the difference is not significant at the usual levels, $p = 0.232$). In our preferred specification (Table 2, column 3), annual policy-agenda change is $\hat{\beta}_1 = 0.080$ above the no-transition baseline of 0.050 when an irregular transition falls between the two addresses (2.6 times the baseline, wild-cluster bootstrap $p = 0.002$) and $\hat{\beta}_2 = 0.054$ above it when an electoral transition does (2.1 times the baseline, $p = 0.012$).

Government Spending. Our second test asks whether movements in the expressed policy agenda track changes in realized fiscal policy. Intuitively, if speech-based policy-agenda change reflects changes in realized policy and not only in rhetoric, spending should change by more in years when the stated agenda moves by more.

We implement the test in two ways. First, in the cross section, we correlate policy-agenda volatility with two realized-policy counterparts: the volatility of social-spending composition and the volatility of discretionary fiscal policy constructed following [Fatás and Mihov \(2013\)](#).⁶ Second, we estimate

$$\text{FiscalChange}_{c,t} = \alpha_c + \gamma_t + \beta \Delta_{c,t} + \varepsilon_{c,t}, \quad (4)$$

where $\text{FiscalChange}_{c,t}$ is either the change in spending composition or the absolute log change in total spending between years $t - 1$ and t , $\Delta_{c,t}$ is speech-based policy-agenda change between the addresses of years $t - 1$ and t , and α_c and γ_t are country and year fixed effects. If speech-based policy-agenda change tracks realized fiscal change, β should

⁶The test uses 1990–2019, the period the ECLAC spending series cover. Composition volatility is one minus the cosine similarity between consecutive years’ central-government expenditure shares on education, health, housing, and social protection, the same formula as the speech measure. Discretionary fiscal volatility is the standard deviation of residuals from country-by-country regressions of government-consumption growth on output growth—instrumented with its two lags—lagged consumption growth, log gross GDP-deflator inflation and its square, and a linear trend (constant local-currency-unit series from the World Development Indicators).

be positive.

Speech-based policy-agenda change tracks the size of changes in government spending. In the cross section of country means, policy-agenda volatility correlates 0.78 with composition volatility (Figure 3, Panel A) and 0.81 with discretionary fiscal volatility (Panel B). Consistent with this, Table 3 shows that years with larger speech-based policy-agenda change tend to be years with larger absolute changes in total spending. For example, in the two-way fixed-effects specification (column 3), $\hat{\beta} = 0.160$ (wild-cluster bootstrap $p = 0.066$). The positive relationship suggests that the stated agenda and realized spending move with the same underlying policy shifts. The relationship is imperfect, suggesting that our volatility measure also contains information about policy priorities that spending data do not capture.

4 Applications

This section presents three applications of the measure: how policy-agenda volatility has changed over the two centuries, what kind of political system exhibits higher volatility, and whether historical volatility predicts development today.

4.1 The Rise and Fall of Policy-Agenda Volatility

Our first application asks how policy-agenda volatility has changed over the last two centuries and why. To trace the measure over time, we compute each country’s mean pair-level policy-agenda change over a trailing twenty-year window and average the resulting series across countries.

We find a rise and fall of policy-agenda volatility that peaks in the middle of the twentieth century. Average agenda change fluctuates between 0.03 and 0.05 from about 1850 through the early 1930s, at similar levels in the nineteenth century and the early twentieth (Figure 4, Panel A). The average then climbs from 0.05 in 1930 to a peak of 0.14 in 1948, stays between 0.11 and 0.14 into the mid-1970s, and falls to 0.07 by 1980 and back to its nineteenth-century range by the 2010s. The peak decades are the era of populism and recurrent military intervention: Perón (1946–1955) and the coups of 1955, 1962, 1966, and 1976 that followed him in Argentina, the 1945–1958 cycle of revolution, dictatorship, and pacted democracy in Venezuela, and the recurring presidencies of Velasco Ibarra in Ecuador. The decline coincides with the redemocratization wave of the 1980s and 1990s (Huntington, 1991; Mainwaring and Pérez-Liñán, 2013). The pattern is robust to

using alternative distance measures (Appendix Figure A4), to computing agenda change from sentence embeddings of the raw text once adjusted for address length (Appendix Figure A5), and to restricting the analysis to a balanced sample of countries (Appendix Figure A6).

What might explain the rise and fall pattern? In an accounting sense, it could reflect power changing hands more often, or larger agenda movements, within presidencies or across them. Given that each pair of consecutive addresses either falls within a single presidency or spans a change of president, a country’s mean policy-agenda change is the share-weighted average of the two groups’ means:

$$\text{Volatility}_c = \underbrace{(1 - s_c) \bar{\Delta}_{W,c}}_{\text{Within share}} + \underbrace{s_c \bar{\Delta}_{B,c}}_{\text{Between share}}, \quad (5)$$

where s_c is the share of pairs spanning a change of president and $\bar{\Delta}_{W,c}$ and $\bar{\Delta}_{B,c}$ are mean policy-agenda change within and between presidencies. The identity holds in any set of address pairs, so we apply it within each country’s twenty-year window and average each term across countries. Figure 4, Panel B plots the means $\bar{\Delta}_{B,c}$ and $\bar{\Delta}_{W,c}$, which measure the size of agenda movements; Panel C plots the components $s_c \bar{\Delta}_{B,c}$ and $(1 - s_c) \bar{\Delta}_{W,c}$, which sum to total volatility.

The identity attributes the rise and fall to the size of agenda movements (the means $\bar{\Delta}_{W,c}$ and $\bar{\Delta}_{B,c}$) rather than to the frequency of turnover (the share s_c). Mean agenda change across a presidential transition quadruples, from roughly 0.06 before 1930 to 0.27 at its peak in 1969, and falls below 0.10 in the 1990s, with a partial rebound in the 2010s (Panel B). The within-president mean rises from roughly 0.04 to 0.10 before returning to its historical level. The share of pairs spanning a transition, by contrast, stays between 0.14 and 0.37 over the whole period and is lower after 1960 than before, so more frequent turnover explains none of the mid-century surge. Both components rise and fall with the total (Panel C): before 1930 the within-president component accounts for about two thirds of total volatility and the between-president component for the remaining third; over 1940–1975 the split is nearly even, and the two components contribute about equally to the mid-century rise.

In the next two applications, we relate nineteenth-century volatility to political regimes and to development today, so in Table 4 we also apply the identity to that cross-section. Panel A reports the pooled components and Panel B every component by country, splitting between-president pairs by transition type. In the long autocracies, even changes of

president barely moved the agenda: $\overline{\Delta}_B$ is 0.008 in Mexico and 0.004 in Paraguay, below those countries' own within-president means. Turnover frequency ranges from fewer than one pair in ten spanning a change of president (Mexico) to nearly one in two (Venezuela), so similar levels of aggregate volatility can arise from different combinations of turnover and agenda change.

4.2 What Kind of Political System Has a Volatile Agenda?

Our second application asks whether autocracies or democracies have more volatile policy agendas. In the growth literature, the presumption is that volatile policy reflects weak institutions and erratic leadership (Acemoglu et al., 2003; Aisen and Veiga, 2013; Fatás and Mihov, 2013). If the presumption is right, volatility should be highest where rulers face the fewest institutional constraints, that is, in autocracies. To assess this, we correlate each country's policy-agenda volatility with its average Polity 2 score (a regime scale running from full autocracy, -10 , to full democracy, $+10$) over the same window.⁷

Across the two centuries, democratic regimes tend to have *more* volatile policy agendas. On average across countries and all years, the correlation between volatility and the democracy score is 0.54 (Figure 5, Panel A). The average masks large heterogeneity across periods: the correlation is 0.79 in the nineteenth century, 0.53 in the twentieth, and -0.53 in the twenty-first (Panel B). The pattern is robust across constructions of the volatility measure and the democracy measure: the correlation is positive in every construction in the nineteenth and twentieth centuries and negative in every construction in the twenty-first (Appendix Table A3). In the benchmark nineteenth-century window (1819–1899), the correlation with Polity lies between 0.50 and 0.83 across all twenty-two constructions of the volatility measure (Appendix Table A1).⁸

To illustrate the relationship between democracy and policy-agenda volatility, we examine selected historical episodes. Panel C of Figure 5 displays, for each country, the cosine similarity between the topic shares of every pair of its addresses (the same similarity that underlies the volatility measure, computed at all horizons rather than only between

⁷For example, Chile under Michelle Bachelet (2006–2010, 2014–2018) scores $+10$; Venezuela under Hugo Chávez falls from $+7$ in 1999 to -3 by 2009; Chile under Augusto Pinochet (1973–1989) scores -7 for most of the regime; Paraguay under José Gaspar Rodríguez de Francia (1814–1840) scores -9 throughout, as does Mexico in nearly every year of the Porfiriato (1876–1911).

⁸The decomposition of the first application shows that both components carry the correlation between volatility and democracy in the nineteenth century. Democracy correlates 0.72 with the within-president component ($p = 0.084$) and 0.61 with the between-president component ($p = 0.053$). Thus more democratic countries had agendas that moved more, both within presidencies and across changes of government.

consecutive addresses). Long autocracies appear as large uniform blocks of near-identical policy priorities, whereas politically competitive countries display many smaller blocks corresponding to shorter regimes. The two most prominent blocks mark the sample’s two most entrenched regimes. In Mexico, the average similarity between any two of the addresses Porfirio Díaz gave over three and a half decades (1876–1911) is 0.96, and addresses given twenty-five or more years apart still average 0.87, against 0.32 for all same-country pairs that far apart. In Paraguay, the one-party governments that followed the War of the Triple Alliance repeated an almost fixed agenda across successive presidents for more than two decades (average pairwise similarity 0.96). No other country produced a comparable block: even where similarity is high on average, as in nineteenth-century Chile or Costa Rica, it is interrupted by successive administrations changing the emphasis of the agenda.

These findings connect to theory and existing evidence. The nineteenth-century pattern is consistent with theories of entrenched rule. Veto-player theory predicts a stable status quo under a single decision maker, since policy follows the ruler’s preferences for as long as the ruler holds power (Tsebelis, 1995). Olson (1993) predicts stable policies under secure autocrats, particularly where an institutionalized ruling party extends the ruler’s horizon (Gehlbach and Keefer, 2011).

The twenty-first-century correlation reproduces the finding that in recent decades, democracies and constrained executives have more stable fiscal and trade policy (Henisz, 2004; Dutt and Mobarak, 2016), and autocratic budgets stay nearly fixed for long periods and then change in rare, large shifts (Chan and Zhao, 2016; Baumgartner et al., 2017). With data from the twenty-first century alone, we would conclude that democracy stabilizes the policy agenda. However, that conclusion would partly be an artifact of sample composition: entrenched autocracies are rare in recent times, and the multi-decade regimes that hold the agenda fixed appear mainly in the historical record.

4.3 Policy-Agenda Volatility and Long-Run Development

Our third application asks whether historical policy-agenda volatility predicts development today. In the growth literature, volatile policy depresses investment and growth (Rodrik, 1991; Alesina et al., 1996; Fatás and Mihov, 2013). If that relationship compounds over the long run, countries with more volatile nineteenth-century agendas should be poorer today. To examine this prediction, we correlate each country’s nineteenth-century policy-agenda volatility with its log GDP per capita in 2022, trace each country’s income path over the twentieth century, and test whether volatility predicts growth within countries at

the decade level.

We find no evidence that nineteenth-century policy-agenda volatility predicts income today. Figure 6, Panel A shows a weak and statistically insignificant relationship (correlation = -0.08 ; exhaustive permutation $p = 0.85$). This weak relationship holds regardless of the window we use to build our volatility measure: as the end year varies from the early 1880s to 1940, the correlation stays small and never distinguishable from zero (Panel B). Panel C shows the same result over time: since 1900, the income paths of the four countries above the median of nineteenth-century volatility and of the four below are indistinguishable. Each group also spans the income distribution: the high-volatility four contain both Chile, the region’s top performer, and Venezuela, its bottom one; the low-volatility four contain both Mexico, the third richest, and Paraguay, among the poorest. The relationship remains weak across the alternative constructions of the policy-agenda volatility measure (Appendix Table A1) and is robust to dropping countries one at a time (Appendix Figure A7). In Appendix B.6, we also test whether volatility predicts income *within* countries rather than *between* them, and find no evidence that it does.

Motivated by Section 4.2, we next document how the cross-sectional association between nineteenth-century policy-agenda volatility and income today changes after conditioning on nineteenth-century democracy. This comparison is informative because democracy is positively associated with policy-agenda volatility in our sample and has also been associated with later income in previous research (Acemoglu et al., 2019). Table 5 reports unconditional and partial correlations. In each column, we regress income in 2022 on policy-agenda volatility and the indicated variables—log GDP per capita in 1940, average nineteenth-century Polity 2, or both—and rescale the coefficients so that they can be interpreted as partial correlations. Column 1, with no controls, reproduces the near-zero unconditional correlation of Figure 6, Panel A.

Conditioning on nineteenth-century democracy turns the correlation between policy-agenda volatility and income sharply negative. The partial correlation between policy-agenda volatility and income is -0.82 after controlling for average nineteenth-century Polity 2 (exhaustive permutation $p = 0.008$, column 3) and -0.86 after adjusting for both Polity and initial income ($p = 0.070$, column 4). Similar sign changes appear for income relative to the United States and for annualized growth over 1940–2022 (Panels B and C). We emphasize that we interpret these conditional correlations as predictive associations. A causal interpretation would require a strong selection-on-observables identification assumption: that, given nineteenth-century democracy and initial income, countries with

more volatile agendas do not differ in other determinants of income today. This assumption is unlikely to hold in these data. For example, democracy may capture a broader bundle of historical institutions and development trajectories, and its association with agenda volatility may itself be part of the historical process under study. We therefore present the unconditional null and the negative conditional correlations as complementary stylized facts that can motivate future causal work on policy-agenda volatility and long-run development.

Our decomposition further allows us to ask whether a specific component of volatility is more predictive of long-run outcomes. To examine this, we correlate each component, computed over 1819–1939 (before the outcome window), with growth over 1940–2022, conditioning on the nineteenth-century Polity average (Appendix Figure A8). Given nineteenth-century democracy, growth is negatively related to the within-president component (correlation = -0.96 ; exhaustive permutation $p = 0.011$) and less strongly to the between-president component (-0.42). The within-president correlation stays between -0.85 and -0.99 when dropping countries one at a time and is -0.73 for components computed over the full period, 1819–2021. It is smaller in magnitude when the outcome is measured in levels instead of growth rates: -0.57 for log GDP per capita and -0.38 for income relative to the United States, both in 2022. The predictive variation comes from the early twentieth century: components built only on 1900–1939 pairs give nearly the same within-president correlation (-0.92), while components built only on nineteenth-century pairs give a correlation near zero (-0.07). Thus the association reflects governments shifting their own priorities in the early twentieth century rather than in the nineteenth.

5 Conclusion

This paper develops a text-based measure of policy-agenda volatility that extends the study of policy volatility to periods before fiscal statistics existed. Existing empirical measures of policy volatility are generally based on realized policy outcomes, particularly fiscal data available only for recent decades. We instead construct our measure from the annual addresses in which presidents report their priorities to Congress, a practice that began long before states collected such statistics. For each pair of consecutive addresses, we quantify policy-agenda change as one minus the cosine similarity between their topic-share vectors. A country’s policy-agenda volatility over a period is the average of these pairwise agenda movements. Agenda movements rise sharply when political power changes hands, with

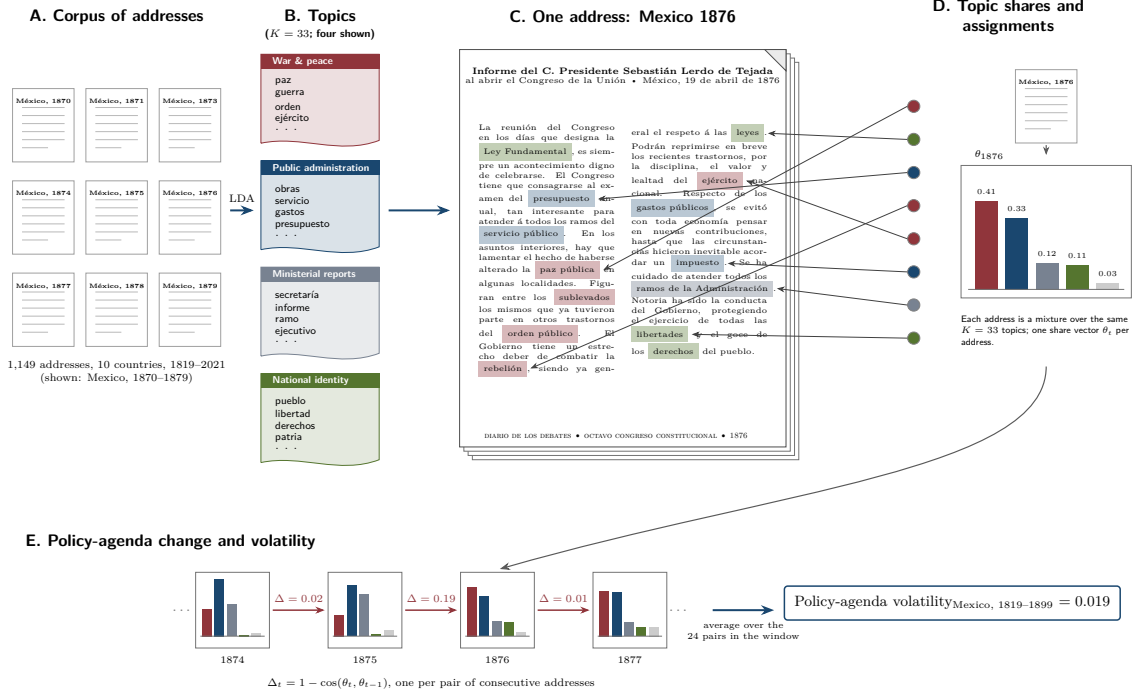
larger point estimates following irregular transitions, supporting the interpretation that the measure captures movements in the expressed policy agenda. Policy-agenda volatility also correlates with government spending volatility, suggesting that the measure contains information about realized policy.

The measure's two-century span allows us to study questions whose relevant variation unfolds across regimes and generations. Our applications illustrate this payoff in three ways. First, policy-agenda volatility rose and fell over the two centuries: in our sample, it stays roughly flat from about 1850 through 1930, triples by about 1950, and falls back after 1980. Second, democracy does not imply a stable policy agenda. Over most of the two centuries, more democratic countries had more volatile agendas, while the most stable agendas belonged to long autocracies whose priorities barely moved either within presidencies or when power changed hands. Thus, agenda stability can accompany political entrenchment, while greater movement can accompany political competition. Third, the unconditional correlation between nineteenth-century volatility and income today is near zero, and it turns negative once nineteenth-century democracy is held constant.

We note that our analysis has several limitations. First, the sample has only eight countries, which limits precision and external validity. Second, none of the applications isolates plausibly exogenous variation in policy-agenda movement, so the relationships we report are descriptive. Third, the measure captures movement in the policy agenda expressed in presidential addresses; it does not establish how much of that movement was implemented or distinguish constructive reform from destabilizing change, responses to shocks, or rhetorical shifts. Future work can distinguish among substantively different kinds of agenda movement and combine the measure with research designs that provide plausibly exogenous variation.

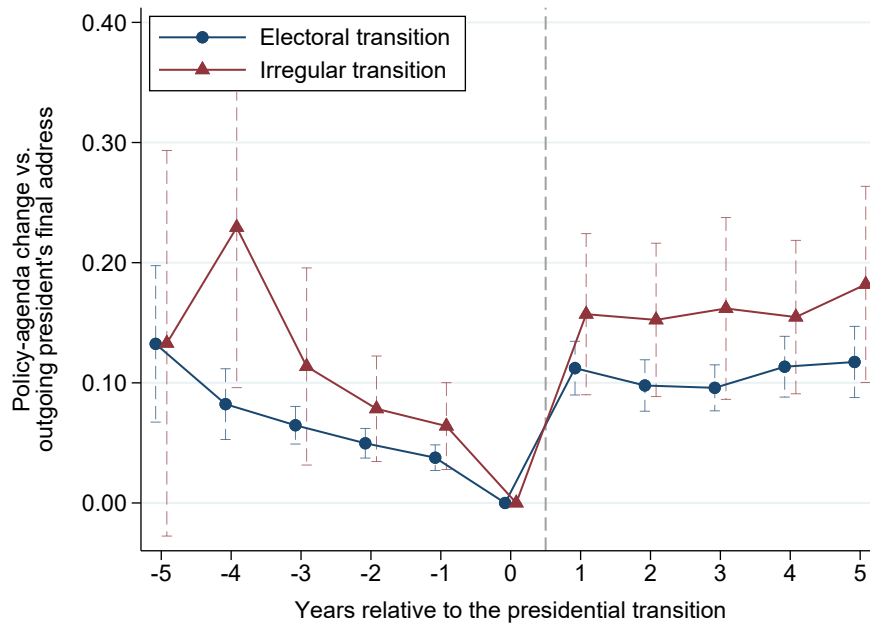
Figures and Tables

Figure 1: Construction of the Policy-Agenda Volatility Measure



Notes: This figure illustrates the construction of the measure with Mexico’s addresses; all values shown are actual estimates. Panel A: the corpus of annual presidential addresses (nine Mexican addresses shown). Panel B: LDA, estimated on the pooled ten-country corpus, represents each of the $K = 33$ topics as a distribution over words; four topics are shown with their top Spanish keywords. Panel C: the address Sebastián Lerdo de Tejada delivered in April 1876, with words characteristic of each displayed topic shaded in that topic’s color. Panel D: the address’s estimated topic-share vector θ_{1876} , with one bar per displayed topic and the remaining topics grouped. Panel E: policy-agenda change between consecutive addresses, $\Delta_t = 1 - \cos(\theta_t, \theta_{t-1})$, shown for 1874–1877, and policy-agenda volatility as the average of Δ_t over all pairs in a window, here Mexico’s nineteenth-century window (1819–1899).

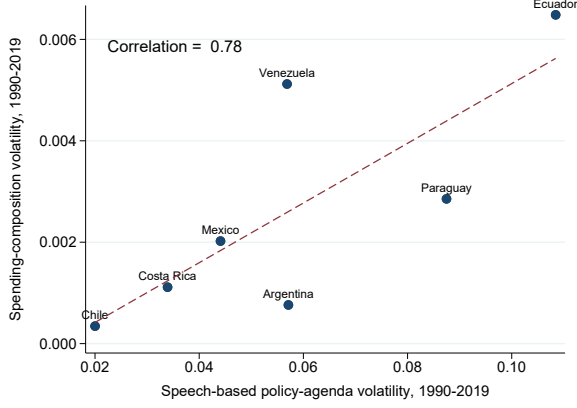
Figure 2: Policy-Agenda Change Around Presidential Transitions



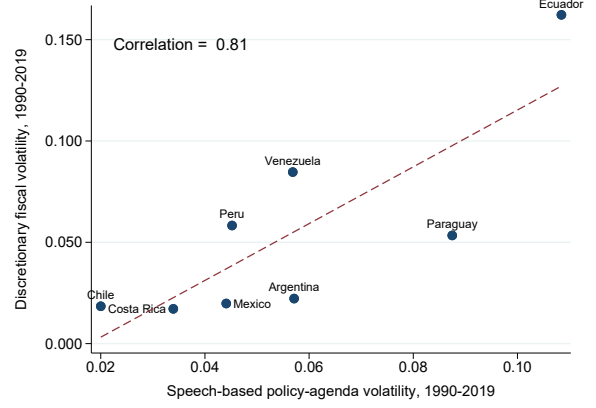
Notes: This figure plots the average policy-agenda change relative to the outgoing president's final address, by year relative to the presidential transition and by transition type. Policy-agenda change between two speeches is one minus the cosine similarity of their topic distributions. Electoral transitions are constitutional handovers following elections; irregular transitions comprise coups, revolutions, forced resignations, and deaths in office. The dashed vertical line marks the transition. Bars are 95 percent confidence intervals around the means. The sample pools all transitions in the eight analysis countries, 1819–2021.

Figure 3: Country-Mean Policy-Agenda Volatility and Realized-Policy Volatility

Panel A. Spending-composition volatility



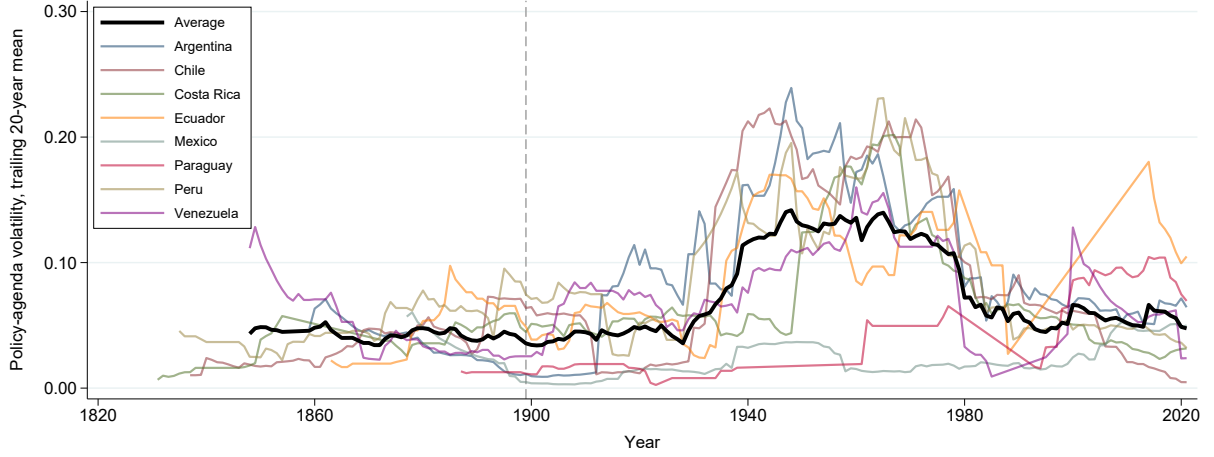
Panel B. Discretionary fiscal volatility



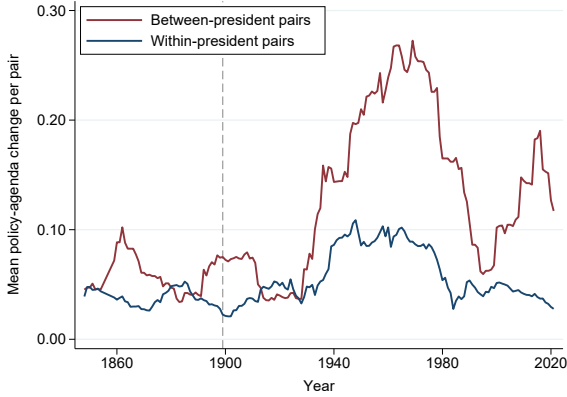
Notes: This figure plots two realized-policy volatility measures against speech-based policy-agenda volatility, all computed by country over 1990–2019. In Panel A, the outcome is spending-composition volatility: one minus the cosine similarity between consecutive years’ shares of central-government social expenditure on education, health, housing, and social protection (ECLAC). Peru is missing from the ECLAC series, leaving seven countries. In Panel B, the outcome is the volatility of discretionary fiscal policy, constructed following [Fatás and Mihov \(2013\)](#) as the standard deviation of residuals from country-by-country regressions of government-consumption growth (World Development Indicators series); all eight countries appear. Speech-based policy-agenda volatility is one minus the cosine similarity between consecutive addresses’ topic distributions, averaged by country over the same window. Lines are least-squares fits.

Figure 4: The Rise and Fall of Policy-Agenda Volatility

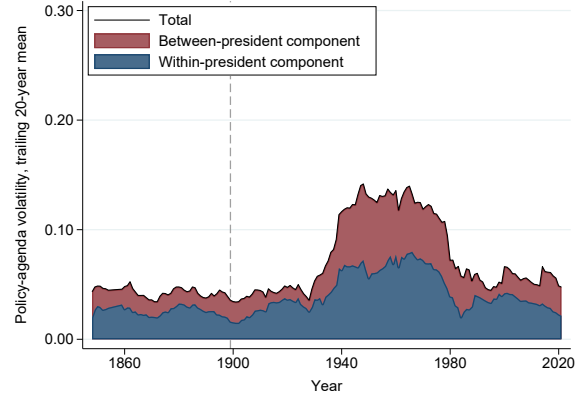
Panel A. Policy-agenda volatility over two centuries



Panel B. Mean policy-agenda change by pair type

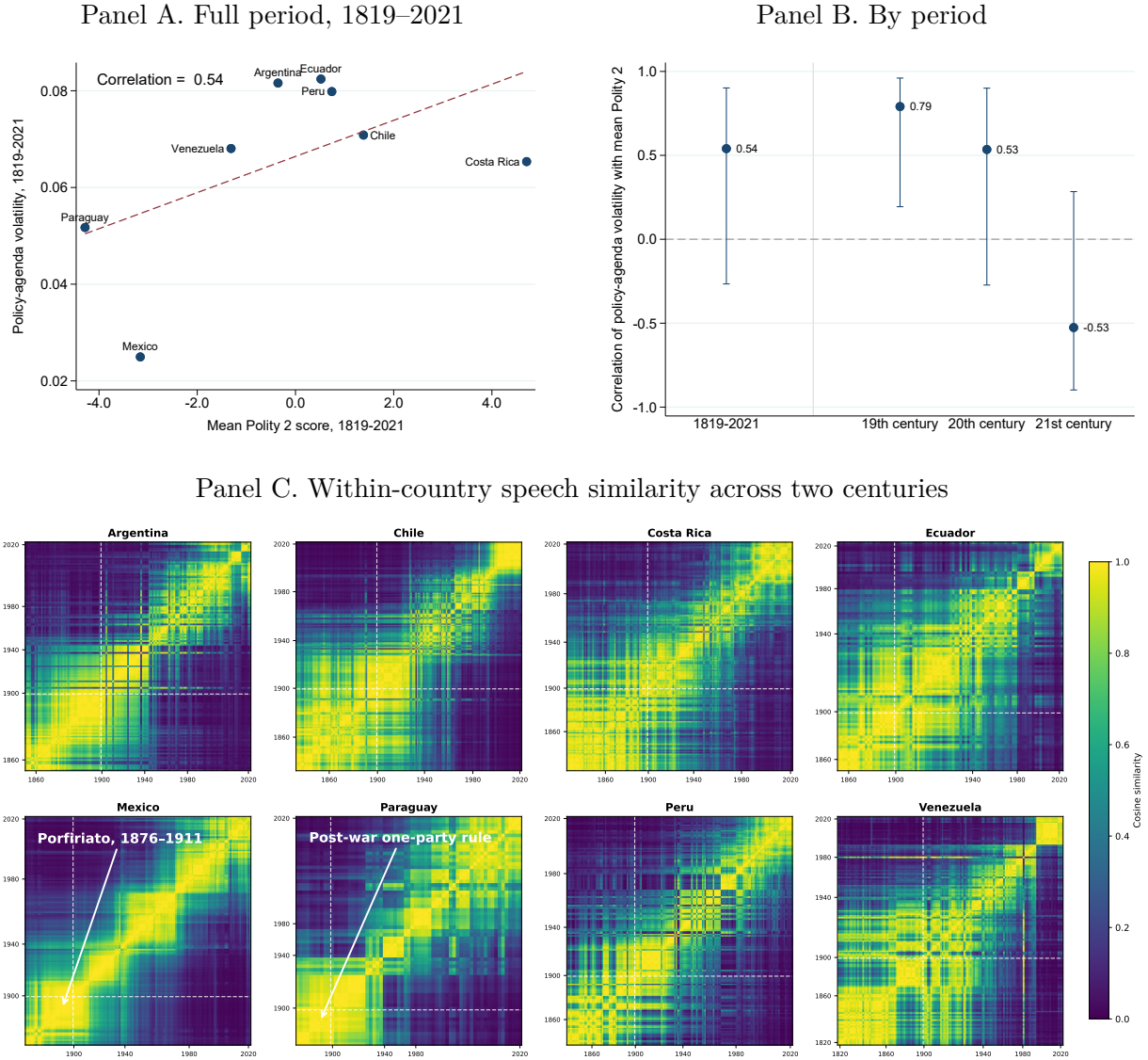


Panel C. The components of the decomposition



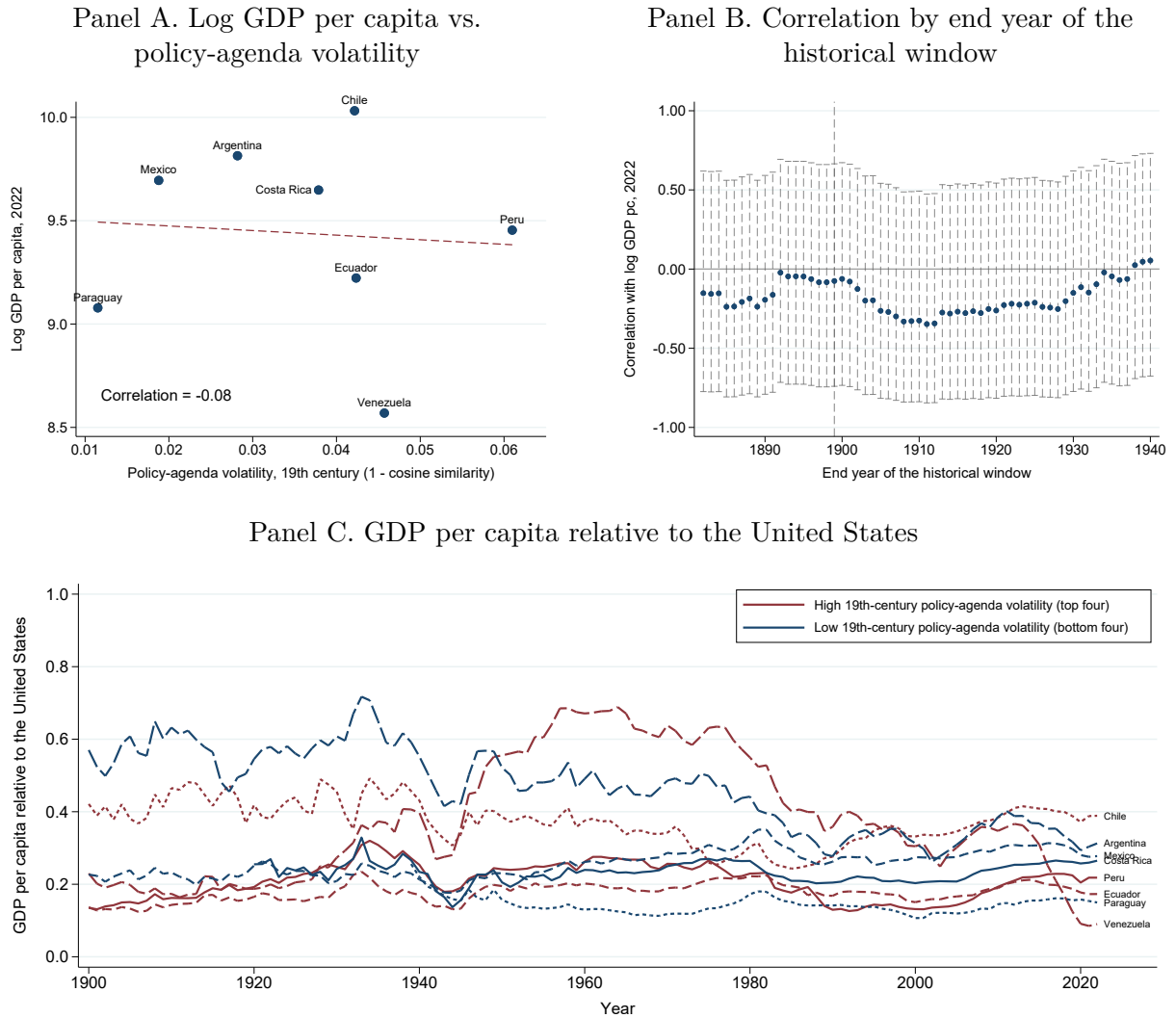
Notes: Panel A plots, for each country, mean policy-agenda change (one minus the cosine similarity between consecutive addresses' topic distributions, pairs at most three years apart and the shorter address at least 1,000 words) over a trailing twenty-year window, shown where the window holds at least five pairs; the bold line is the equal-weight average over the countries whose window qualifies. Panel B plots mean policy-agenda change per pair, separately for pairs spanning a change of president and pairs within a presidency. Panel C splits the average into the between-president and within-president components of equation (5), applied within each country-window and averaged across countries; the components are stacked, so the top of the shaded areas is total policy-agenda volatility. Averaged series are shown for years in which at least four countries qualify and are interrupted over 1855–1858, when fewer do. The dashed vertical line marks the end of the benchmark nineteenth-century window (1899).

Figure 5: Policy-Agenda Volatility and Democracy



Notes: Panel A plots policy-agenda volatility (the average of one minus the cosine similarity between consecutive addresses' topic distributions over 1819–2021, pairs at most three years apart and the shorter address at least 1,000 words) against the average Polity 2 score over the same period; higher scores indicate more democratic institutions, and the line is the least-squares fit. Panel B reports the correlation between the two variables recomputed within each window: the full period and each century (Polity coverage ends in 2018); whiskers are 95 percent confidence intervals based on the Fisher z transformation. Panel C displays each country's matrix of pairwise cosine similarities between all of its addresses, countries in alphabetical order, with both axes listing addresses in chronological order (ticks mark address years); brighter cells indicate more similar topic distributions, dashed white lines mark 1900, and annotations mark the single-regime spells discussed in the text.

Figure 6: Nineteenth-Century Policy-Agenda Volatility and Long-Run Development



Notes: Panel A plots log GDP per capita in 2022 (Maddison Project Database 2023) against nineteenth-century policy-agenda volatility (the average of one minus the cosine similarity between consecutive addresses' topic distributions over 1819–1899, pairs at most three years apart and the shorter address at least 1,000 words); the line is the least-squares fit. Panel B plots the correlation between nineteenth-century policy-agenda volatility and log GDP per capita in 2022 as the end year of the historical window varies from 1882 to 1940; bars are 95 percent confidence intervals from the Fisher z transformation, and the dashed vertical line marks the benchmark window ending in 1899. Panel C plots GDP per capita relative to the United States from 1900 to 2022 for the same eight countries: maroon lines mark the four countries with above-median nineteenth-century policy-agenda volatility, navy lines the four below the median, line patterns distinguish countries within each group, and labels mark each country's 2022 endpoint.

Table 1: Summary Statistics of the Analysis Sample

	Speeches	Coverage	Pairs, 19th c.	Volatility, 19th c.	Volatility, full period	GDP pc 2022	Polity 2, 19th c.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Argentina	137	1854–2021	45	0.028	0.082	18,292	-2.7
Chile	187	1832–2021	66	0.042	0.071	22,741	-0.9
Costa Rica	184	1824–2021	35	0.038	0.065	15,497	-1.9
Ecuador	91	1853–2021	20	0.042	0.082	10,124	-1.0
Mexico	137	1867–2021	24	0.019	0.025	16,235	-4.7
Paraguay	65	1881–2021	8	0.012	0.052	8,763	-7.0
Peru	166	1827–2021	42	0.061	0.080	12,763	-0.2
Venezuela	137	1819–2021	33	0.046	0.068	5,267	-4.6

Notes: This table summarizes the eight-country analysis sample. Column 1 reports the number of speeches; column 2 the first and last speech year. Column 3 reports the number of nineteenth-century benchmark pairs: consecutive available speeches at most three years apart whose shorter address has at least 1,000 words (Section 2.2). Columns 4 and 5 report policy-agenda volatility (the average of one minus the cosine similarity between consecutive addresses' topic distributions, over benchmark pairs) over 1819–1899 and over the full period. Column 6 reports GDP per capita in 2022 in 2011 dollars (Maddison Project Database 2023); Venezuela's figure reflects its post-2014 collapse. Column 7 reports the average Polity 2 score over 1819–1899.

Table 2: Policy-Agenda Change and Presidential Transitions

	Outcome: policy-agenda change		
	(1)	(2)	(3)
Span includes irregular transition	0.079*** (0.014)	0.075*** (0.016)	0.080*** (0.015)
Span includes electoral transition	0.059*** (0.009)	0.059*** (0.009)	0.054*** (0.013)
Mean policy–agenda change, no transition	0.050	0.050	0.050
Wild–bootstrap p : irregular	0.004	0.009	0.002
Wild–bootstrap p : electoral	0.003	0.002	0.012
Wild–bootstrap p : equality	0.257	0.372	0.232
Country FE	No	Yes	Yes
Year FE	No	No	Yes
N (pairs)	986	986	986

Notes: This table estimates equation (3) on the 986 benchmark pairs over the full period, 1819–2021. The outcome is pair-level policy-agenda change; the regressors indicate that the pair of consecutive addresses spans an irregular transition (coup, revolution, forced resignation, or death in office) or an electoral transition. A pair spans a transition when its two addresses are delivered by different presidents, with the type taken from the dated transition records for the years the pair covers; mixed spans are classified as irregular. Pairs with a change of president but no coded electoral or irregular transition (appointments, interim administrations, and residual types) are counted in the no-transition baseline. Columns differ in fixed effects as indicated; year effects enter as year indicators. Standard errors clustered by country in parentheses; reported p -values are from the wild-cluster restricted bootstrap with Webb weights and 9,999 replications, which is the inference we rely on with eight clusters. The equality row tests $H_0 : \beta_1 = \beta_2$. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3: Annual Policy-Agenda Changes and Annual Fiscal Changes

	Outcome: absolute log change in spending, year t		
	(1)	(2)	(3)
Speech-based policy-agenda change (annual pairs)	0.163*** (0.048)	0.131** (0.051)	0.160** (0.064)
Wild-bootstrap p	0.010	0.077	0.066
Country FE	No	Yes	Yes
Year FE	No	No	Yes
N (pair-years)	176	176	176
Countries	9	9	9

Notes: This table estimates equation (4) under the same three specifications as Table 2, on matched country-years, 1990–2019, for the nine corpus countries with usable overlapping data (Colombia and the Dominican Republic included; Peru’s central-government series is unavailable, Appendix B). Policy-agenda change is one minus the cosine similarity between consecutive years’ addresses (gap of one year, minimum-length restriction). Total-spending change is the absolute log change in total central-government expenditure (ECLAC). Columns differ in fixed effects as indicated; year effects enter as year indicators. Standard errors clustered by country in parentheses; reported p -values are from the wild-cluster restricted bootstrap with Webb weights and 9,999 replications, which is the inference we rely on with nine clusters. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 4: Decomposing Volatility into Turnover and Conditional Agenda Change

		Share of pairs spanning			Mean policy-agenda change		Volatility _c (7)
Pairs		new president s_c	electoral	irregular	within $\bar{\Delta}_{W,c}$	between $\bar{\Delta}_{B,c}$	
(1)	(2)	(3)	(4)	(5)	(6)		
Panel A. Pooled							
All countries	273	0.25	0.13	0.12	0.031	0.066	0.040
Panel B. By country							
Argentina	45	0.22	0.16	0.07	0.023	0.045	0.028
Chile	66	0.14	0.12	0.02	0.034	0.096	0.042
Costa Rica	35	0.17	0.09	0.09	0.034	0.055	0.038
Ecuador	20	0.35	0.15	0.20	0.037	0.053	0.042
Mexico	24	0.08	0.00	0.08	0.020	0.008	0.019
Paraguay	8	0.25	0.25	0.00	0.014	0.004	0.012
Peru	42	0.43	0.10	0.31	0.034	0.097	0.061
Venezuela	33	0.45	0.27	0.18	0.040	0.053	0.046

Notes: This table reports the ingredients of the decomposition in equation (5) for nineteenth-century benchmark pairs (gap at most three years, minimum-length restriction). Panel A pools all pairs across countries; Panel B reports each country separately. Column 2 is the share of pairs spanning a change of president; columns 3 and 4 split that share by the transition type coded in the span (electoral, versus coups, revolutions, forced resignations, and deaths in office); spans with a change of president but no coded electoral or irregular transition in the years the pair covers (appointments, interim administrations, and residual types) count in column 2 but in neither column 3 nor column 4. Columns 5 and 6 report mean policy-agenda change within and between presidencies; column 7 is country-level policy-agenda volatility, which equals $(1 - s)\bar{\Delta}_W + s\bar{\Delta}_B$ exactly.

Table 5: Nineteenth-Century Policy-Agenda Volatility and Long-Run Development

	Control added:			
	None	Log GDP pc, 1940	Polity 2, 19th c.	Both
	(1)	(2)	(3)	(4)
Panel A. Outcome: log GDP per capita, 2022				
Policy-agenda volatility	−0.08 (0.41)	−0.10 (0.44)	−0.82** (0.26)	−0.86** (0.26)
Log GDP pc, 1940		0.23 (0.44)		0.50 (0.43)
Polity 2, 19th c.			0.86** (0.23)	0.89** (0.23)
Permutation p -value	0.847	0.787	0.008	0.070
Panel B. Outcome: GDP per capita relative to the U.S.				
Policy-agenda volatility	−0.04 (0.41)	−0.10 (0.45)	−0.72* (0.31)	−0.83** (0.28)
Log GDP pc, 1940		0.39 (0.41)		0.66 (0.37)
Polity 2, 19th c.			0.79** (0.28)	0.87** (0.25)
Permutation p -value	0.905	0.800	0.011	0.069
Panel C. Outcome: annualized growth, 1940–2022				
Policy-agenda volatility	−0.15 (0.40)	−0.10 (0.44)	−0.73* (0.31)	−0.86** (0.26)
Log GDP pc, 1940		−0.55 (0.37)		−0.81** (0.29)
Polity 2, 19th c.			0.76** (0.29)	0.89** (0.23)
Permutation p -value	0.686	0.787	0.204	0.070

Notes: This table reports partial correlations between each panel’s outcome and each row variable across the eight benchmark countries. The outcomes are log GDP per capita in 2022 (Panel A), GDP per capita relative to the United States in 2022 (Panel B), and annualized growth of GDP per capita over 1940–2022, in percent (Panel C). Each cell residualizes the outcome and the row variable on the column’s other variables—none (column 1), log GDP per capita in 1940 (column 2), the average Polity 2 score over 1819–1899 (column 3), or both (column 4)—and correlates the residuals, so column 1 reports unconditional correlations. Standard errors in parentheses are $\sqrt{(1 - r^2)/(6 - k)}$, where k counts the column’s controls, with stars from the implied t -test. Permutation p -values for the volatility row are exhaustive and use the HC3-studentized statistic from the regression of the outcome on the column’s variables, with Freedman–Lane residual permutation in columns with controls (Appendix B.5). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

References

- Acemoglu, D., S. Johnson, and J. A. Robinson (2001). The colonial origins of comparative development: An empirical investigation. *American Economic Review* 91(5), 1369–1401.
- Acemoglu, D., S. Johnson, J. A. Robinson, and Y. Thaicharoen (2003). Institutional causes, macroeconomic symptoms: Volatility, crises and growth. *Journal of Monetary Economics* 50(1), 49–123.
- Acemoglu, D., S. Naidu, P. Restrepo, and J. A. Robinson (2019). Democracy does cause growth. *Journal of Political Economy* 127(1), 47–100.
- Acuña, C. H., S. Galiani, and M. Tommasi (2007). Understanding the political economy of structural reform: The case of Argentina. In J. M. Fanelli (Ed.), *Understanding Market Reforms in Latin America: Similar Reforms, Diverse Constituencies, Varied Results*, pp. 31–72. New York: Palgrave Macmillan.
- Aisen, A. and F. J. Veiga (2013). How does political instability affect economic growth? *European Journal of Political Economy* 29, 151–167.
- Alesina, A., S. Özler, N. Roubini, and P. Swagel (1996). Political instability and economic growth. *Journal of Economic Growth* 1(2), 189–211.
- Ash, E. and S. Hansen (2023). Text algorithms in economics. *Annual Review of Economics* 15, 659–688.
- Baker, S. R., N. Bloom, and S. J. Davis (2016). Measuring economic policy uncertainty. *Quarterly Journal of Economics* 131(4), 1593–1636.
- Baturo, A., N. Dasandi, and S. J. Mikhaylov (2017). Understanding state preferences with text as data: Introducing the UN General Debate corpus. *Research & Politics* 4(2), 2053168017712821.
- Baumgartner, F. R., M. Carammia, D. A. Epp, B. Noble, B. Rey, and T. M. Yildirim (2017). Budgetary change in authoritarian and democratic regimes. *Journal of European Public Policy* 24(6), 792–808.
- Blei, D. M. (2012). Probabilistic topic models. *Communications of the ACM* 55(4), 77–84.

- Blei, D. M., A. Y. Ng, and M. I. Jordan (2003). Latent Dirichlet allocation. *Journal of Machine Learning Research* 3, 993–1022.
- Bolt, J. and J. L. van Zanden (2025). Maddison-style estimates of the evolution of the world economy: A new 2023 update. *Journal of Economic Surveys* 39(2), 631–671.
- Budge, I., H.-D. Klingemann, A. Volkens, J. Bara, and E. Tanenbaum (2001). *Mapping Policy Preferences: Estimates for Parties, Electors, and Governments 1945–1998*. Oxford: Oxford University Press.
- Calvo-González, O., A. Eizmendi, and G. Reyes (2026). From warfare to welfare: Leaders, eras, and the evolution of policy attention. Working paper.
- Camargo, C. Q., P. John, H. Z. Margetts, and S. A. Hale (2021). Measuring the volatility of the political agenda in public opinion and news media. *Public Opinion Quarterly* 85(2), 493–516.
- Cameron, A. C., J. B. Gelbach, and D. L. Miller (2008). Bootstrap-based improvements for inference with clustered errors. *Review of Economics and Statistics* 90(3), 414–427.
- Chan, K. N. and S. Zhao (2016). Punctuated equilibrium and the information disadvantage of authoritarianism: Evidence from the People’s Republic of China. *Policy Studies Journal* 44(2), 134–155.
- Cohen, L., C. Malloy, and Q. Nguyen (2020). Lazy prices. *Journal of Finance* 75(3), 1371–1415.
- Coppedge, M., J. Gerring, C. H. Knutsen, S. I. Lindberg, J. Teorell, et al. (2026). V-Dem country-year dataset v16. Varieties of Democracy (V-Dem) Project.
- Dell, M. (2010). The persistent effects of Peru’s mining Mita. *Econometrica* 78(6), 1863–1903.
- Dincecco, M. and G. Katz (2016). State capacity and long-run economic performance. *Economic Journal* 126(590), 189–218.
- Dixit, A. K. and R. S. Pindyck (1994). *Investment under Uncertainty*. Princeton, NJ: Princeton University Press.

- Doyle, D. (2014). The political economy of policy volatility in Latin America. *Latin American Politics and Society* 56(4), 1–21.
- Dutt, P. and A. M. Mobarak (2016). Democracy and policy stability. *International Review of Economics & Finance* 42, 499–517.
- Fatás, A. and I. Mihov (2013). Policy volatility, institutions, and economic growth. *Review of Economics and Statistics* 95(2), 362–376.
- Freedman, D. and D. Lane (1983). A nonstochastic interpretation of reported significance levels. *Journal of Business & Economic Statistics* 1(4), 292–298.
- Galiani, S. and I. Sened (Eds.) (2014). *Institutions, Property Rights, and Economic Growth: The Legacy of Douglass North*. Cambridge: Cambridge University Press.
- Gehlbach, S. and P. Keefer (2011). Investment without democracy: Ruling-party institutionalization and credible commitment in autocracies. *Journal of Comparative Economics* 39(2), 123–139.
- Gentzkow, M., B. Kelly, and M. Taddy (2019). Text as data. *Journal of Economic Literature* 57(3), 535–574.
- Glaeser, E. L., R. La Porta, F. Lopez-de-Silanes, and A. Shleifer (2004). Do institutions cause growth? *Journal of Economic Growth* 9(3), 271–303.
- Griffiths, T. L. and M. Steyvers (2004). Finding scientific topics. *Proceedings of the National Academy of Sciences* 101(Supplement 1), 5228–5235.
- Grimmer, J. and B. M. Stewart (2013). Text as data: The promise and pitfalls of automatic content analysis methods for political texts. *Political Analysis* 21(3), 267–297.
- Hansen, S., M. McMahon, and A. Prat (2018). Transparency and deliberation within the FOMC: A computational linguistics approach. *Quarterly Journal of Economics* 133(2), 801–870.
- Hassan, T. A., S. Hollander, L. van Lent, and A. Tahoun (2019). Firm-level political risk: Measurement and effects. *Quarterly Journal of Economics* 134(4), 2135–2202.
- Henisz, W. J. (2000). The institutional environment for economic growth. *Economics & Politics* 12(1), 1–31.

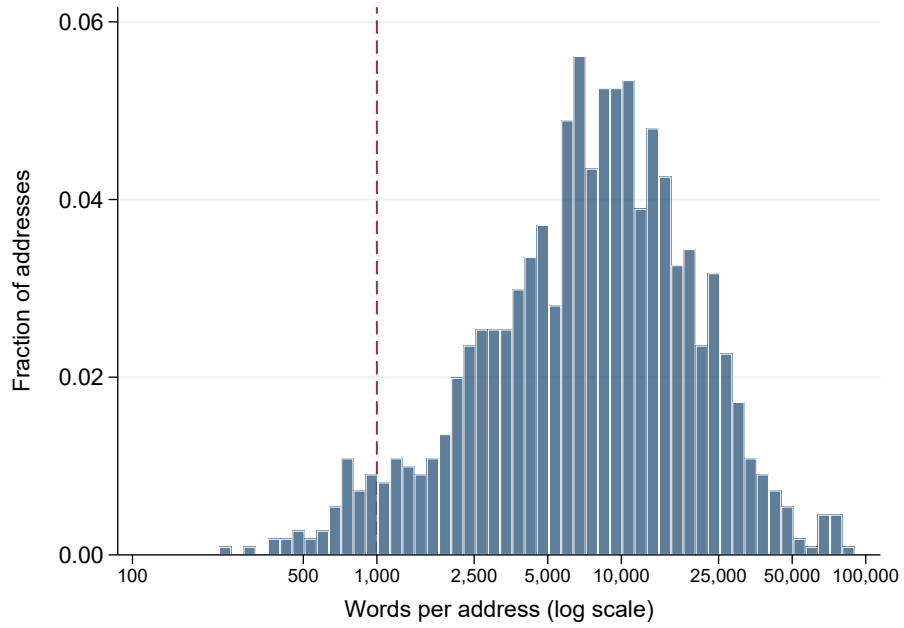
- Henisz, W. J. (2004). Political institutions and policy volatility. *Economics & Politics* 16(1), 1–27.
- Huntington, S. P. (1991). *The Third Wave: Democratization in the Late Twentieth Century*. Norman, OK: University of Oklahoma Press.
- John, P. and W. Jennings (2010). Punctuations and turning points in British politics: The policy agenda of the Queen’s speech, 1940–2005. *British Journal of Political Science* 40(3), 561–586.
- Le, Q. and T. Mikolov (2014). Distributed representations of sentences and documents. In *Proceedings of the 31st International Conference on Machine Learning*, pp. 1188–1196.
- Mainwaring, S. and A. Pérez-Liñán (2013). *Democracies and Dictatorships in Latin America: Emergence, Survival, and Fall*. Cambridge: Cambridge University Press.
- Marshall, M. G. and T. R. Gurr (2020). Polity5: Political regime characteristics and transitions, 1800–2018, dataset users’ manual. Technical report, Center for Systemic Peace.
- Marshall, M. G. and D. R. Marshall (2021). Coup d’état events, 1946–2021, codebook. Center for Systemic Peace.
- Mortensen, P. B., C. Green-Pedersen, G. Breeman, L. Chaqués-Bonafont, W. Jennings, P. John, A. M. Palau, and A. Timmermans (2011). Comparing government agendas: Executive speeches in the Netherlands, United Kingdom, and Denmark. *Comparative Political Studies* 44(8), 973–1000.
- North, D. C. (1990). *Institutions, Institutional Change and Economic Performance*. Cambridge: Cambridge University Press.
- Nunn, N. (2008). The long-term effects of Africa’s slave trades. *Quarterly Journal of Economics* 123(1), 139–176.
- Olson, M. (1993). Dictatorship, democracy, and development. *American Political Science Review* 87(3), 567–576.
- Prados de la Escosura, L. (2009). Lost decades? economic performance in post-independence Latin America. *Journal of Latin American Studies* 41(2), 279–307.

- Rodrik, D. (1991). Policy uncertainty and private investment in developing countries. *Journal of Development Economics* 36(2), 229–242.
- Rule, A., J.-P. Cointet, and P. S. Bearman (2015). Lexical shifts, substantive changes, and continuity in State of the Union discourse, 1790–2014. *Proceedings of the National Academy of Sciences* 112(35), 10837–10844.
- Spiller, P. T. and M. Tommasi (2003). The institutional foundations of public policy: A transactions approach with application to Argentina. *Journal of Law, Economics, and Organization* 19(2), 281–306.
- Stein, E., M. Tommasi, K. Echebarría, E. Lora, and M. Payne (2006). *The Politics of Policies: Economic and Social Progress in Latin America, 2006 Report*. Washington, DC: Inter-American Development Bank and David Rockefeller Center for Latin American Studies.
- Tsebelis, G. (1995). Decision making in political systems: Veto players in presidentialism, parliamentarism, multicameralism and multipartyism. *British Journal of Political Science* 25(3), 289–325.

Appendix

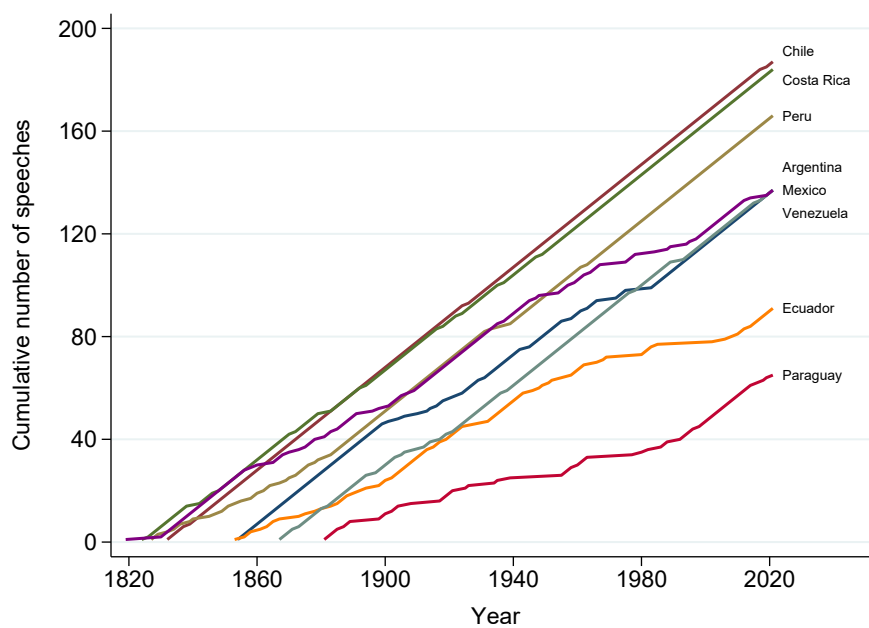
A Appendix Figures and Tables

Figure A1: The Distribution of Address Length



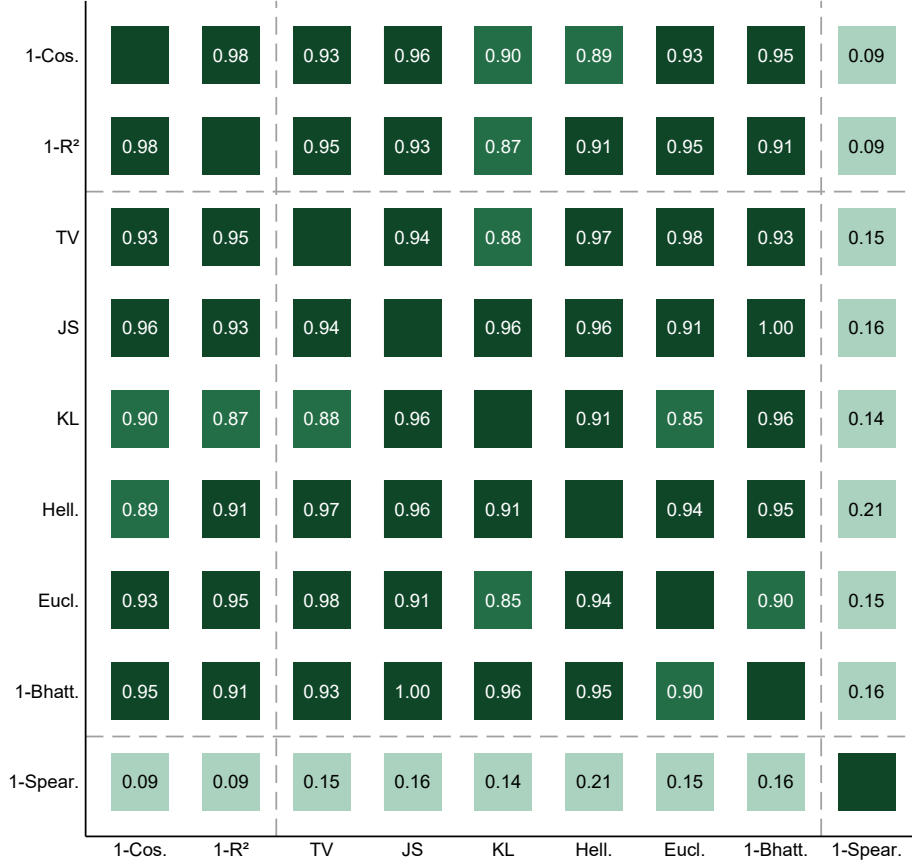
Notes: This figure plots the distribution of address length across the 1,104 addresses delivered in the eight analysis countries, 1819–2021. Length is the address’s word count, shown on a logarithmic axis. The dashed line marks the 1,000-word minimum of the benchmark measure: a pair of consecutive addresses enters the benchmark sample only if both addresses have at least 1,000 words (Section 2.2).

Figure A2: Cumulative Speeches by Country



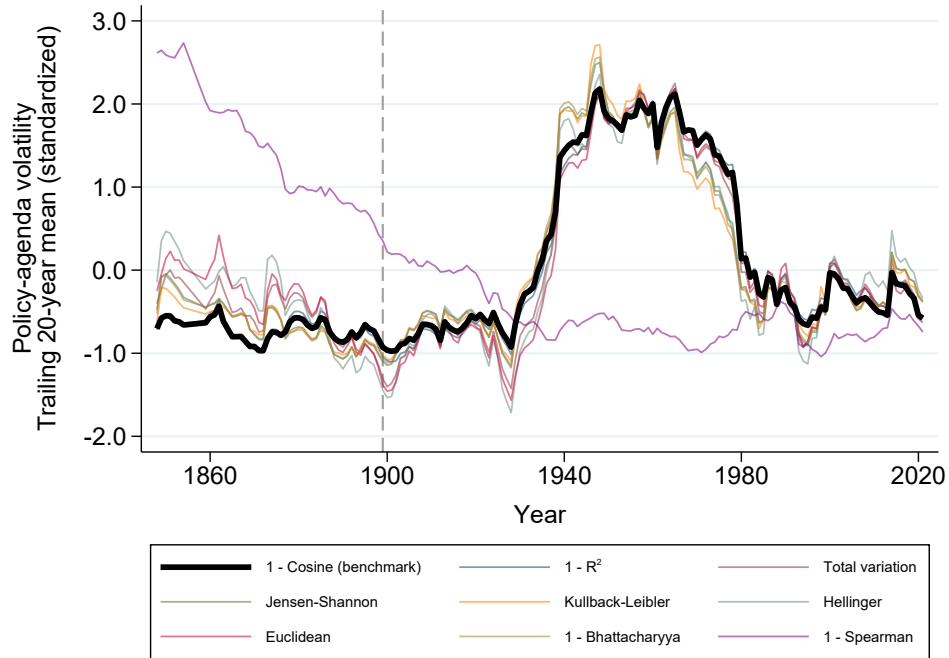
Notes: This figure plots the cumulative number of speeches in the analysis sample by country over time.

Figure A3: Correlation Between Alternative Distance Measures



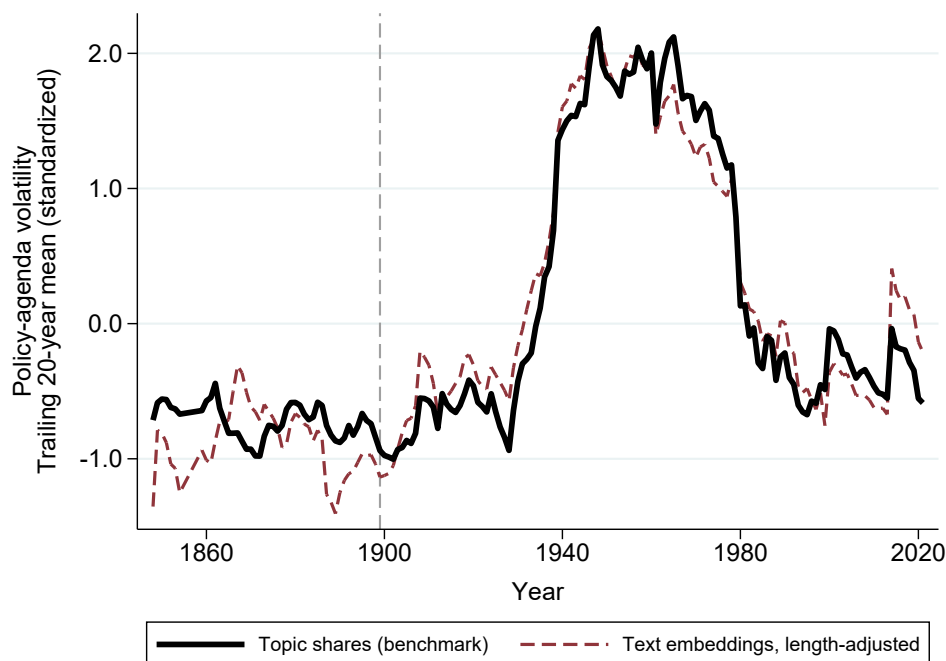
Notes: This figure reports pairwise correlations between nine pair-level measures of policy-agenda change, computed across the 986 benchmark pairs of consecutive addresses. All measures are oriented so that larger values indicate more policy-agenda change: one minus cosine similarity (the benchmark), one minus the R^2 of a regression of one address's topic shares on the other's, total-variation distance, Jensen–Shannon divergence, Kullback–Leibler divergence, Hellinger distance, Euclidean distance, one minus the Bhattacharyya coefficient, and one minus the Spearman rank correlation of topic shares. Appendix B.4 defines each measure. Darker green indicates stronger positive correlation. Dashed lines separate the benchmark and its fit-based twin, the distribution-based measures, and the rank-based measure.

Figure A4: The Rise and Fall of Policy-Agenda Volatility by Distance Measure



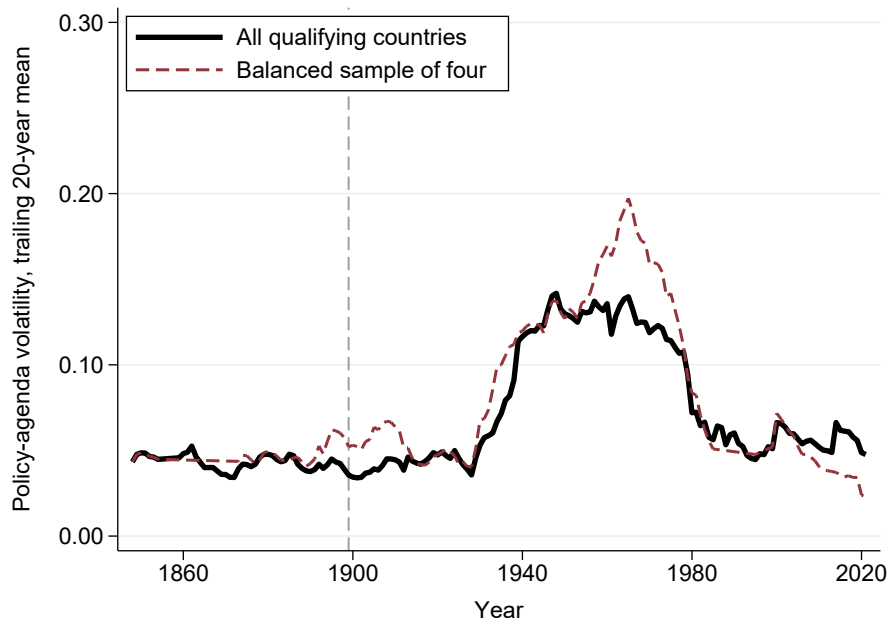
Notes: This figure recomputes the equal-weight cross-country average of the trailing twenty-year window mean from Figure 4, Panel A using each of the nine pair-level distance measures of Appendix B.4, and standardizes each series to mean zero and unit variance over 1848–2021. The benchmark, one minus cosine similarity, is the bold line. Windows hold at least five pairs, and averaged series cover years in which at least four countries qualify. The dashed vertical line marks the end of the benchmark nineteenth-century window (1899).

Figure A5: The Rise and Fall of Policy-Agenda Volatility from Text Embeddings



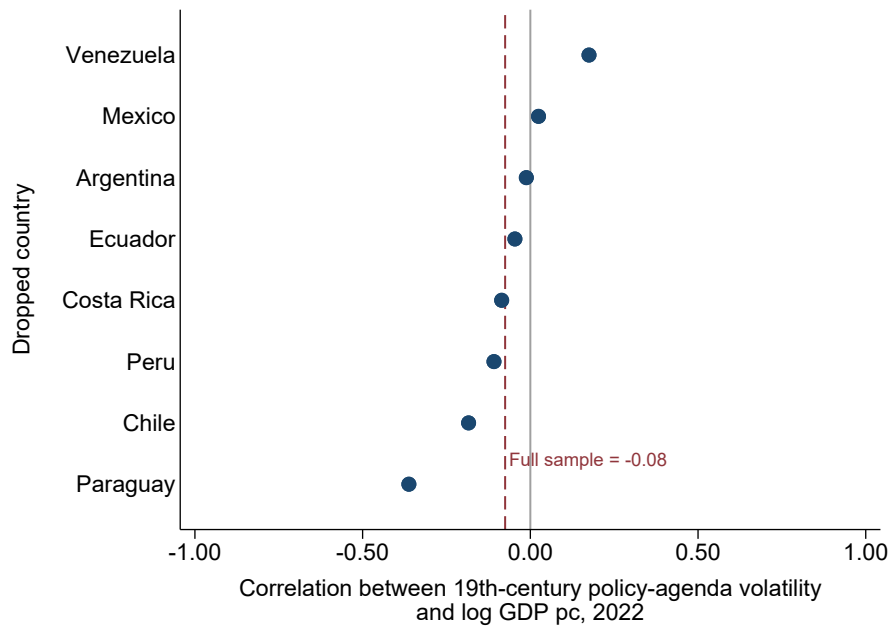
Notes: This figure compares the trailing twenty-year window series of Figure 4, Panel A with the same series computed from an embedding-based measure of policy-agenda change: each address's raw text is embedded with a multilingual sentence-embedding model, and policy-agenda change is one minus the cosine similarity between consecutive addresses' document vectors. Because the embedding-based measure is strongly related to address length, the length-adjusted series residualizes it on the log word count of the shorter address in each pair. Both series are computed on the 984 benchmark pairs with raw text available for both addresses, averaged across countries with equal weights, and standardized to mean zero and unit variance over 1848–2021. Windows hold at least five pairs, and averaged series cover years in which at least four countries qualify. The dashed vertical line marks the end of the benchmark nineteenth-century window (1899).

Figure A6: The Rise and Fall of Policy-Agenda Volatility in a Balanced Sample



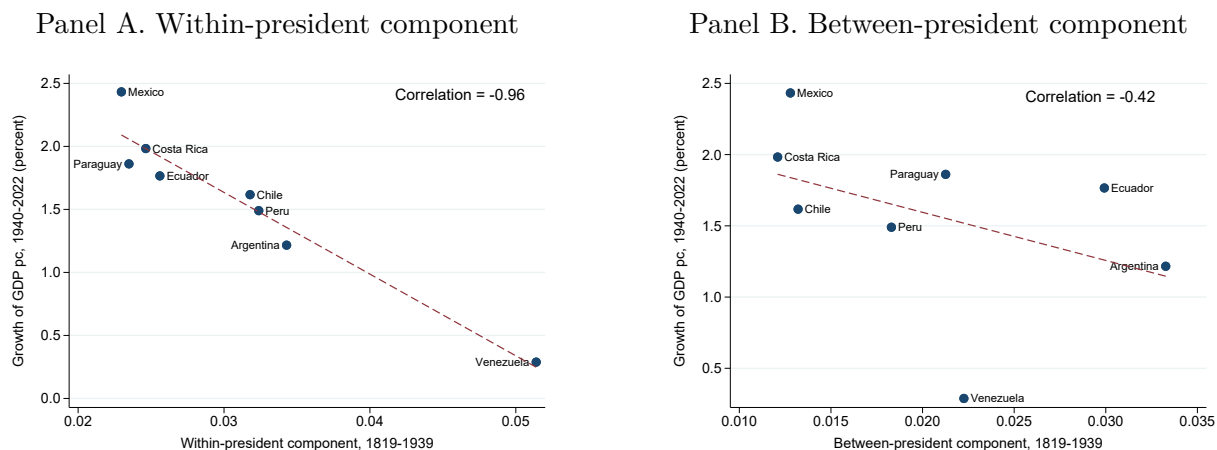
Notes: This figure compares the equal-weight average of the trailing twenty-year window mean over all qualifying countries (Figure 4, Panel A) with the average over a balanced sample: the four countries whose windows qualify from 1848 (Chile, Costa Rica, Peru, and Venezuela), averaged only in years where all four qualify. Windows hold at least five pairs. The dashed vertical line marks the end of the benchmark nineteenth-century window (1899).

Figure A7: Leave-One-Out Correlations



Notes: This figure plots the correlation between nineteenth-century policy-agenda volatility and log GDP per capita in 2022 when each country is dropped in turn. The dashed line marks the full-sample correlation.

Figure A8: Long-Run Growth and the Components of Pre-1940 Policy-Agenda Volatility,
Conditioning on Nineteenth-Century Democracy



Notes: This figure plots annualized growth of GDP per capita over 1940–2022, from the Maddison Project Database 2023, against the two components of policy-agenda volatility from equation (5), computed over 1819–1939, before the outcome window. In each panel, growth and the component are residualized on the average nineteenth-century Polity 2 score and recentered at their means, so the annotated correlation is the partial correlation between the two variables given nineteenth-century democracy. Lines are least-squares fits.

Table A1: Alternative Constructions of the Volatility Measure

	Correlation with benchmark (1)	Correlation with Polity 2, 19th c. (2)	Correlation with log GDP pc, 2022 (3)
Panel A. Benchmark and length restriction			
1 – cosine similarity (benchmark)	1.00	0.79	-0.08
No minimum-length restriction	0.87	0.77	0.04
Noise-corrected (1/length residual)	0.34	0.64	0.38
Panel B. Gap between addresses			
Adjacent years only (gap = 1)	0.81	0.52	0.02
Gap ≤ 2	0.98	0.75	-0.03
Gap ≤ 5	1.00	0.77	-0.08
Annualized: (1 – cosine)/gap	0.88	0.68	0.14
Residualized on gap length	0.88	0.66	0.09
Panel C. Distance measure			
Total-variation distance	0.88	0.68	-0.28
1 – Pearson correlation	1.00	0.80	-0.06
Jensen–Shannon divergence	0.92	0.70	-0.27
1 – R^2 of topic-share regression	1.00	0.79	-0.10
Kullback–Leibler divergence	0.93	0.67	-0.29
Hellinger distance	0.84	0.65	-0.30
Euclidean distance	0.95	0.67	-0.27
1 – Bhattacharyya coefficient	0.92	0.70	-0.27
1 – Spearman rank correlation	0.74	0.64	-0.15
Panel D. Topic model			
Pooled LDA, K = 13	0.96	0.80	-0.02
Pooled LDA, K = 23	0.96	0.79	0.08
Pooled LDA, K = 43	1.00	0.80	-0.03
Six core policy domains	0.99	0.83	-0.02
Country-specific LDA	0.74	0.50	0.07

Notes: This table recomputes nineteenth-century policy-agenda volatility under the alternative constructions of Table A3, Panel A, and under additional distance measures. Panel A varies the minimum-length restriction and its noise correction; Panel B varies the maximum gap between paired addresses, where the annualized variant divides each pair’s policy-agenda change by the gap length and the residualized variant regresses pair-level policy-agenda change on gap-length indicators (pooled across all pairs) and averages the recentered residuals; Panel C varies the distance measure (Appendix B.4); Panel D varies the topic model, including the six core policy domains of Calvo-González et al. (2026) (shares renormalized) and country-specific LDA models, in which each country’s speeches are represented in their own topic space. Columns report each construction’s correlation with the benchmark measure (column 1), with the average Polity 2 score over 1819–1899 (column 2), and with log GDP per capita in 2022 (column 3), across the eight benchmark countries.

Table A2: Topics and Their Top Defining Keywords

Topic	Share (%)	Top-10 keywords
1	23.8	works, service, national, public, jobs, administration, trade, number, expenditures, budget
2	13.6	peace, peoples, war, administration, national, homeland, nation, order, public, army
3	10.7	development, national, social, program, sector, policy, system, resources, process, growth
4	10.9	the people, social, policy, life, nation, right, national, freedom, homeland, economic
5	10.8	national, works, production, construction, social, plan, activities, services, education, ministry
6	8.0	health, family, education, person, quality, program, poverty, security, project, social
7	1.9	law, general, very, senate, project, today, fellow citizens, regime, provinces, there is
8	1.9	federal, hectares, works, district, city, construction, tons, production, kilometers, credit
9	1.7	the people, national, world, Chavez, social, oil, revolution, people, mission, plan
10	1.6	school, coffee, legislative, executive, road, Limón, bank, Puntarenas, roads, expenditures
11	1.6	department, having been, river, school, period, road, irrigation, region, jobs, Puno
12	1.5	executive, united, federal, secretariat, district, period, branch, kilometers, report, Veracruz
13	1.4	secretariat, department, federal, executive, district, report, having been, commission, treasury, public
14	1.3	federal, Maracaibo, revenue, Cabello, district, ministry, annual report, debt, port, Guaira
15	1.2	province, nation, provinces, capital, Córdoba, pesos fuertes, gold, territories, immigration, Corrientes
16	1.1	executive, nation, currency, provinces, exercise, national, period, federal, resources, works
17	1.1	copper, treasury, Valparaíso, seven hundred, escudos, number, project, three hundred, Carabineros, Arica
18	1.0	province, provinces, Cuenca, Esmeraldas, road, ministry, Loja, police, Guayas, report

Notes: This table lists the 18 of the pooled LDA model's 33 topics whose average share within the eight-country analysis sample is at least 1 percent. Topics are numbered by their frequency rank in the ten-country corpus of Calvo-González et al. (2026); column 2 reports the average share within the analysis sample, so the two orderings differ slightly. Column 3 lists the ten keywords most strongly associated with each topic (translated from Spanish; non-lexical tokens are dropped). Topics 1–6 are the six core policy domains of Calvo-González et al. (2026): public administration, war and peace, economic development, national identity, infrastructure, and social welfare.

Table A3: Robustness of the Volatility–Democracy Correlation

		By century:		
	1819–2021	19th c.	20th c.	21st c.
	(1)	(2)	(3)	(4)
Panel A. Constructions of the measure (correlation with Polity 2)				
1 – cosine similarity (benchmark)	0.54	0.79	0.53	-0.53
No minimum-length restriction	0.59	0.77	0.50	-0.53
Noise-corrected (1/length residual)	0.39	0.64	0.49	-0.53
Adjacent years only (gap = 1)	0.55	0.52	0.54	-0.34
Gap ≤ 2	0.56	0.75	0.56	-0.53
Gap ≤ 5	0.46	0.77	0.48	-0.52
Annualized: (1 – cosine)/gap	0.62	0.68	0.58	-0.42
Residualized on gap length	0.57	0.66	0.55	-0.51
Total-variation distance	0.51	0.68	0.65	-0.37
1 – Pearson correlation	0.54	0.80	0.54	-0.52
Jensen–Shannon divergence	0.51	0.70	0.51	-0.58
Pooled LDA, K = 13	0.55	0.80	0.54	-0.51
Pooled LDA, K = 23	0.53	0.79	0.55	-0.56
Pooled LDA, K = 43	0.58	0.80	0.57	-0.59
Six core policy domains	0.42	0.83	0.49	-0.96
Country-specific LDA	0.63	0.50	0.58	-0.97
Panel B. Alignment, sample, and coverage (benchmark measure)				
Polity averaged over pair years only	0.45	0.62	0.56	-0.58
Excluding Paraguay	0.48	0.63	0.40	-0.59
Controlling for log speeches in window	0.44	0.64	0.42	0.00
Controlling for mean speech length	0.37	0.78	0.41	-0.82
Panel C. Alternative democracy measures (benchmark measure)				
Polity 2 score	0.54	0.79	0.53	-0.53
Executive constraints	0.34	0.53	0.30	-0.52
V-Dem electoral democracy	0.40	0.38	0.56	-0.57
V-Dem liberal democracy	0.35	0.25	0.55	-0.66

Notes: This table reports associations between policy-agenda volatility and window-mean democracy within the indicated window: the full period (column 1) and the nineteenth, twentieth, and twenty-first centuries (columns 2–4). Policy-agenda volatility and the democracy measures are recomputed within each window, and Polity coverage ends in 2018. Panel A reports correlations between each construction of policy-agenda volatility and the average Polity 2 score. In Panel B, the first row reports correlations using Polity averaged only over years covered by usable speech pairs, and the second excludes Paraguay. The remaining two rows report the standardized coefficient on policy-agenda volatility from a regression of the Polity 2 score on volatility and the listed control, with all variables standardized within the window. Panel C reports correlations between the benchmark measure and each alternative democracy measure. Leave-one-out correlations for the nineteenth-century benchmark range from 0.63 to 0.93, with a worst-case permutation p -value of 0.16.

B Empirical Appendix

B.1 Presidential Speeches

The speech corpus is documented in [Calvo-González et al. \(2026\)](#); we summarize the features relevant for this paper. Sixteen Spanish-speaking Latin American countries have a constitutional mandate requiring the president to report annually to the legislature. The corpus covers the ten with at least two decades of obtainable speeches. Speeches were collected from national congresses and congressional libraries, official online repositories, and, for Venezuela, digitized volumes at the U.S. Library of Congress complemented with newspaper archives; [Calvo-González et al. \(2026\)](#) detail each country’s sources. Scanned documents were processed with optical character recognition and manually corrected by a team member.

Because the volatility measure requires long within-country series, we exclude Colombia and the Dominican Republic from the eight-country analysis sample; their speech series begin in 1999. Both countries enter the pair-year validation of Section 3.3, which uses only post-1990 annual pairs.

B.2 Outcome and Covariate Data

Transitions and coups. Presidential transitions and their classification (election, coup, revolution, resignation, death, appointment, interim, or a residual other) were hand-coded for [Calvo-González et al. \(2026\)](#) from the Georgetown Political Database of the Americas and country sources. We classify coups, revolutions, resignations, and deaths as irregular transitions. The Center for Systemic Peace coup list ([Marshall and Marshall, 2021](#)) corroborates the coding where the two overlap: it records a coup or attempted coup within a year of 59 percent of our post-1945 irregular transitions, consistent with the list covering coups but not the resignations and deaths that our irregular category also includes.

Government spending. Central-government social expenditure by function (education, health, housing and community amenities, and social protection) and total central-government expenditure, both as percentages of GDP, come from CEPALSTAT (ECLAC) and cover 1990–2019. Peru is missing from both series (its social spending is reported only under general-government coverage). Composition shares are computed within the four social functions and require all four to be reported in a year.

Institutions. Polity 2 and executive constraints (XCONST) come from the Polity5 an-

nual data, 1800–2018 (Marshall and Gurr, 2020); we set the special interregnum codes to missing for executive constraints and use the standard Polity 2 treatment otherwise. The V-Dem electoral-democracy (v2x_POLYARCHY) and liberal-democracy (v2x_LIBDEM) indices come from V-Dem version 16 (Coppedge et al., 2026); nineteenth-century values are missing for some early post-independence years, and window averages use available years.

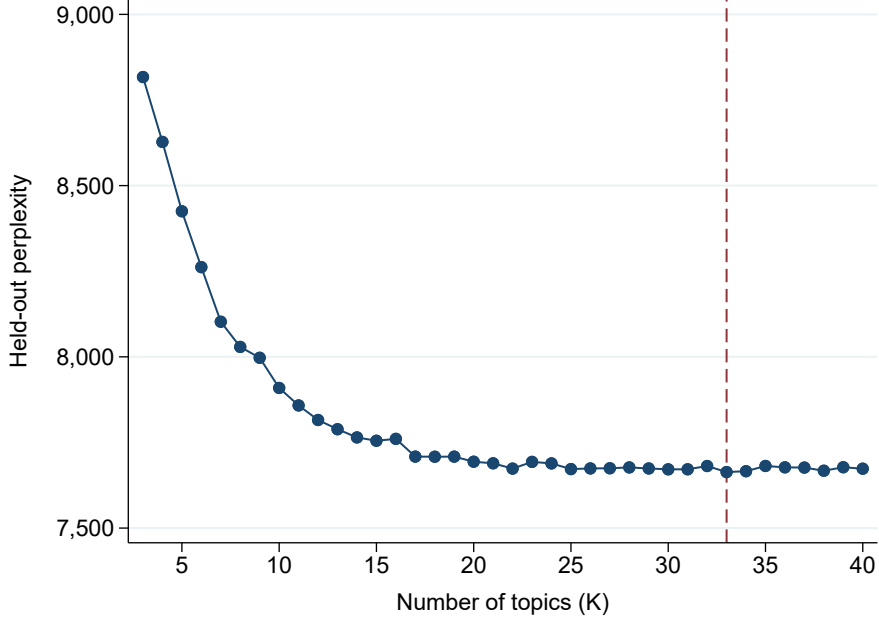
GDP per capita. We use the Maddison Project Database 2023 (Bolt and van Zanden, 2025), in 2011 international dollars. Maddison reports scattered pre-modern benchmarks for Mexico (from 1550) and Peru (from 1595); coverage relevant for our analysis begins in 1800 for Argentina, Chile, and Venezuela, in 1870 for Ecuador, in 1920 for Costa Rica, and in 1939 for Paraguay, so 1940 is the first round year available for all eight countries; we use it as the initial-income control. The country-decade growth panel uses decade-endpoint observations where available.

B.3 Latent Dirichlet Allocation

LDA is a generative model of documents (Blei et al., 2003; Blei, 2012). Each topic $k \in \{1, \dots, K\}$ is a distribution β_k over the vocabulary; each document d has a distribution θ_d over topics. Each word in document d is generated by drawing a topic $z \sim \text{Multinomial}(\theta_d)$ and then a word $w \sim \text{Multinomial}(\beta_z)$. The Dirichlet priors on θ and β regularize the estimates.

Calvo-González et al. (2026) estimate the model on the pooled Spanish-language corpus by collapsed Gibbs sampling (Griffiths and Steyvers, 2004), after removing stopwords and consolidating term variants, and select the number of topics by minimizing perplexity—the inverse geometric mean of the predictive likelihood—on a held-out ten percent of documents. Perplexity is minimized at $K = 33$ (Appendix Figure B1). Our robustness checks re-estimate the volatility measure from models with $K = 13, 23$, and 43 topics, from country-specific models, and from the 33 topics aggregated into six core policy domains.

Figure B1: Held-Out Perplexity and the Number of Topics



Notes: This figure plots the held-out perplexity of the pooled LDA model as a function of the number of topics K . The dashed line marks the minimum at $K = 33$.

B.4 Alternative Distance Measures

Let t and p denote the topic-share vectors of the later and earlier address of a pair, with components indexed by k . Appendix Figure A3 compares the benchmark cosine distance with eight alternatives, each oriented so that larger values mean more policy-agenda change.

Total-variation distance. The share of probability mass that must move to turn one distribution into the other:

$$\text{TV}(t, p) = \frac{1}{2} \sum_k |t_k - p_k|.$$

Jensen–Shannon divergence. A symmetric and bounded variant of the Kullback–Leibler divergence, computed against the mixture $m = (t + p)/2$ with logarithms base 2, so it is bounded between zero and one:

$$\text{JS}(t, p) = \frac{1}{2} \text{KL}(t \parallel m) + \frac{1}{2} \text{KL}(p \parallel m).$$

Kullback–Leibler divergence. The information lost when the earlier address’s distribu-

tion is used to approximate the later one:

$$\text{KL}(t \parallel p) = \sum_k t_k \ln(t_k/p_k).$$

Hellinger distance. A bounded distance based on the overlap of the two distributions:

$$\text{H}(t, p) = \sqrt{1 - \sum_k \sqrt{t_k p_k}}.$$

Euclidean distance. The ordinary distance between the two share vectors:

$$\text{E}(t, p) = \sqrt{\sum_k (t_k - p_k)^2}.$$

Bhattacharyya coefficient. A similarity bounded between zero and one; it enters as one minus the coefficient:

$$\text{BC}(t, p) = \sum_k \sqrt{t_k p_k}.$$

Spearman rank correlation. The correlation between the two addresses' rankings of topics; it enters as one minus the correlation:

$$\rho_S(t, p) = \text{corr}(\text{rank}(t), \text{rank}(p)).$$

R^2 . The squared correlation between the two share vectors, equivalently the R^2 of a regression of one address's shares on the other's; it enters as one minus the R^2 :

$$R^2(t, p) = \text{corr}(t, p)^2.$$

B.5 Inference

With eight countries, asymptotic inference is unreliable, so all cross-sectional p -values in the paper are exhaustive permutation p -values computed under the null of exchangeability. For a bivariate association, we enumerate all $8! = 40,320$ permutations of the outcome across countries, recompute the test statistic under each, and report the share of permutations whose statistic is at least as large in absolute value as the observed one (Freedman and Lane, 1983). The statistic is the HC3-studentized coefficient, so the permutation test matches the heteroskedasticity-robust standard errors reported in the tables and guards

against single-country leverage. For regressions with a control, we use the Freedman–Lane procedure: we regress the outcome on the control, permute the residuals, add them back to the fitted values, and re-estimate the full model under each permutation, again enumerating all permutations of the estimation sample. Enumeration is complete, so no simulation error enters.

The country-decade panel and the transition regressions cluster standard errors by country, with only eight clusters. Asymptotic cluster-robust inference over-rejects with so few clusters (Cameron et al., 2008), so we report p -values from the wild-cluster restricted bootstrap with Webb six-point weights and 9,999 replications.

B.6 Volatility and Income Within Countries

We test whether policy-agenda volatility predicts income *within* countries rather than *between* them. We build a country-decade panel of 137 country-decades, compute policy-agenda volatility within each country-decade, and estimate

$$Y_{c,t} = \alpha_c + \gamma_t + \beta \text{Volatility}_{c,t} + \delta X_{c,t} + \varepsilon_{c,t}, \quad (\text{B1})$$

where $Y_{c,t}$ is an income outcome of country c in decade t , $\text{Volatility}_{c,t}$ is policy-agenda volatility computed within the decade, α_c and γ_t are country and decade fixed effects, and $X_{c,t}$ is a control included in some specifications.

We measure income in three ways—log GDP per capita, GDP per capita relative to the United States, and annualized growth of GDP per capita—with each outcome measured in the same decade t or the next ($t+1$), and three specifications for each: no further controls, controlling for log income at the start of the decade, and controlling for the decade’s average Polity 2 score. We report each estimate as a partial correlation, residualizing the outcome and volatility on the fixed effects and any controls, so the estimates are on the same scale as the cross-sectional correlations of Table 5. Appendix Table B1 shows that volatility predicts none of the three: across horizons and specifications, no correlation is statistically distinguishable from zero (smallest wild-cluster bootstrap $p = 0.21$; Panels A–C).

Table B1: Policy-Agenda Volatility and Development Outcomes: Country-Decade Panel

	Decade t			Decade $t + 1$		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Outcome: log GDP per capita						
Policy-agenda volatility in decade t	0.05 (0.10)	-0.16 (0.13)	-0.01 (0.10)	-0.00 (0.08)	-0.06 (0.08)	-0.03 (0.07)
Log GDP pc, start of decade		0.95*** (0.01)			0.53*** (0.08)	
Polity 2, decade average			0.43* (0.18)			0.33 (0.18)
Wild-bootstrap p -value	0.682	0.303	0.917	0.995	0.423	0.638
N (country-decades)	122	122	115	118	114	118
Panel B. Outcome: GDP per capita relative to the U.S.						
Policy-agenda volatility in decade t	0.03 (0.09)	-0.11 (0.10)	-0.01 (0.07)	-0.01 (0.08)	-0.09 (0.07)	-0.05 (0.07)
Log GDP pc, start of decade		0.89*** (0.08)			0.59*** (0.12)	
Polity 2, decade average			0.45* (0.22)			0.35* (0.18)
Wild-bootstrap p -value	0.730	0.267	0.842	0.849	0.222	0.417
N (country-decades)	122	122	115	118	114	118
Panel C. Outcome: annualized growth of GDP per capita						
Policy-agenda volatility in decade t	-0.09 (0.07)	-0.07 (0.08)	-0.09 (0.07)	-0.00 (0.08)	0.05 (0.08)	0.01 (0.09)
Log GDP pc, start of decade		-0.24** (0.08)			-0.45*** (0.09)	
Polity 2, decade average			0.02 (0.07)			-0.12 (0.08)
Wild-bootstrap p -value	0.208	0.403	0.229	0.957	0.529	0.942
N (country-decades)	114	114	114	109	106	109

Notes: This table reports partial correlations between each panel's outcome and policy-agenda volatility in decade t (the average of one minus the cosine similarity between consecutive addresses at most three years apart). The outcomes are the decade means of log GDP per capita (Panel A) and GDP per capita relative to the United States (Panel B), and annualized GDP per capita growth in log points $\times 100$ (Panel C), measured in the same decade (columns 1–3) or the next calendar decade (columns 4–6). Each cell residualizes the outcome and the row variable on country and decade fixed effects and the column's control—none (columns 1 and 4), log GDP per capita at the decade's first year (columns 2 and 5), or the decade's average Polity 2 score (columns 3 and 6)—and correlates the residuals, so estimates are on the same scale as the correlations in Table 5. Standard errors in parentheses are the underlying regression's cluster-robust (by country) standard errors rescaled to the correlation scale, leaving t -statistics unchanged; reported p -values are from the wild-cluster bootstrap with Webb weights and 9,999 replications. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.