More Laws, More Growth?
Evidence from U.S. States

Elliott Ash, Massimo Morelli, Matia Vannoni*

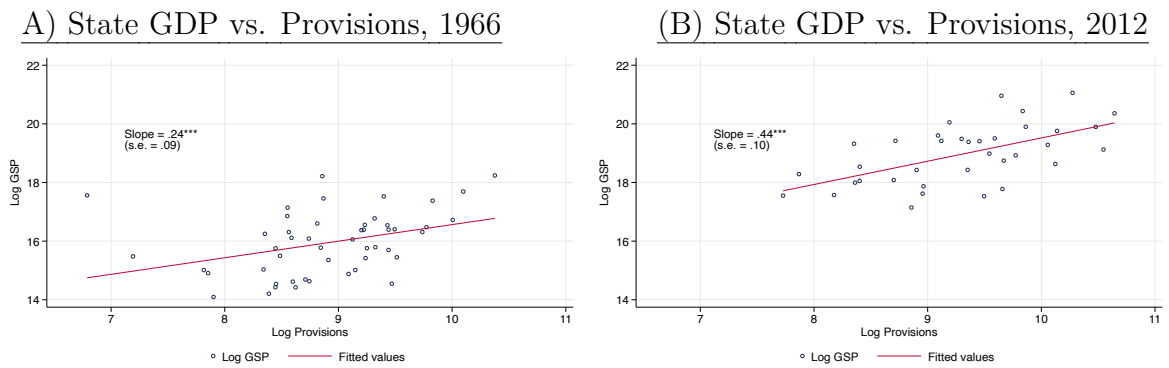
July 14, 2023

Abstract

This paper analyzes the conditions under which more legislation contributes to economic growth. In the context of U.S. states, we apply natural language processing tools to measure legislative flows for the years 1965-2012. We implement a novel shift-share design for text data, where the instrument for legislation is leave-one-out legal-topic flows interacted with pre-treatment legal-topic shares. We find that at the margin, higher legislative output causes more economic growth. Consistent with more complete laws reducing ex post hold-up, we find that the effect is driven by the use of contingent clauses, is largest in sectors with high relationship-specific investments, and is increasing with local economic uncertainty.

*Elliott Ash, ETH Zurich, ashe@ethz.ch; Massimo Morelli, Bocconi University and CEPR, massimo.morelli@unibocconi.it; Matia Vannoni, King's College London, matia.vannoni@kcl.ac.uk. We acknowledge financial support from the European Research Council Advanced Grant 694583 and the Framework for Attraction and Strengthening of Excellence Grant. We also acknowledge support from the Baffi CAREFIN and IGIER centers. We thank Massimo Anelli, Benjamin Arold, Pierpaolo Battigalli, Kirill Borusyak, Quentin Gallea, Livio di Lonardo, Xavier Jaravel, Giovanni Maggi, and Piero Stanig for helpful comments. We also wish to thank seminar participants at EUI, Autonoma de Barcelona, Columbia Political Economy Conference, ETH Political Economy Workshop, King's College London, and Stockholm School of Economics for helpful comments. We thank David Cai, Claudia Marangon, and Emiliano Rinaldi for helpful research assistance.

Figure 1: State GDP and Legislative Output, 1966 and 2012



Notes. Scatter-plots for the relationship between (log) provisions and (log) state GDP at the beginning of our time period (1966) and the end (2012).

1 Introduction

In the cross section, jurisdictions with larger, more complex legal systems also tend to have larger, more productive economies. The correlation between legislative output and GDP in U.S. states, illustrated in Figure 1, provides a clear example of this empirical regularity. A fundamental question is whether these correlations reflect causal links.

While a larger economy could lead to more laws mechanically (as, for example, more industries need to be regulated), it could also be that more legislation (if well-designed) causes economic growth. Consider the introduction of detailed property rights protections, for example, or establishment of the rule of law (Dam, 2007). These institutions could help markets run more efficiently, encourage investment, and increase growth. In particular, a more complete "legislative contract" could lead to more investment by making laws more enforceable and reliable (e.g. Williamson, 1979, Hart and Moore, 1988). Conversely, legislation might instead consist mainly of favors to special interest groups, coming at the cost to overall growth and welfare (e.g. Grossman and Helpman, 2001). Even with good-natured legislators in charge, excessive lawmaking could hinder economic growth by increasing compliance costs (Niskanen, 1971, Botero et al., 2004). Given the current state of the economy, one might postulate an optimal level of legal detail, with movement toward the optimum from either side leading to growth.

Motivated by these debates, we explore empirically the relationship between legislative detail and economic output in U.S. states, 1965-2012. Our research question is whether and how laws impact the economy. In brief, we show that increasing le-

gal detail does lead to more growth. On mechanisms, we report supporting evidence that higher legislating reduces legal uncertainty, leading to more business activity via relationship-specific investments.

The first step is to measure detail in legislation. For each state and biennium, we produce a measure of legislative output from the text of state laws. The measure draws on recently developed methods in computational linguistics to detect *provisions,* legally relevant requirements in statutes (Vannoni et al., 2019). These provisions extract more information than coarser measures based on words or phrases. Further, we use a topic model to measure the allocation of provisions across legal categories (Blei et al., 2003).

Our empirical strategy introduces a novel text-based shift-share instrument for legal detail (e.g. Bartik, 1994). Analogous to standard shift-share instruments that use *sector-specific economic flows* interacted with *pre-period sector shares*, we construct our instrument using *topic-specific legislative flows* interacted with *pre-period topic shares*. For identification, we assume that topic-specific national legislative flows are exogenous to each particular state, in line with recent econometric work by Adao et al. (2019) and Borusyak et al. (2022). Intuitively, we think of states as borrowing language on policies enacted in other states (the shifters), especially when they have relatively low existing detail on that topic (the shares). Our design passes a number of checks recently developed by econometricians for probing the exogeneity of shift-share instruments (Adao et al., 2019, Goldsmith-Pinkham et al., 2020, Borusyak et al., 2022).

The main empirical finding is that more state-level legislation due to the shift-share instrument tends to boost the state economy on average. This effect is robust to a range of alternative choices for text processing, regression specification, and statistical inference. We can rule out a number of alternative channels for the results besides an effect of legislation on the economy, including changes in population, taxes/expenditures, intervening political conditions, and changes in the actions of state regulatory agencies or state courts.

To understand why more laws lead to more growth, consider the following essential feature of our setting and design: The instrument captures state-to-state diffusion on legal topics where borrowing states have not yet legislated. As publication of laws in other states reduces the cost of enacting similar laws, that expands the legislative choice set over beneficial laws while allowing legislators to disregard unsuitable ones. Hence, the instrument is likely to bring efficiency-enhancing laws under minimal assumptions on the benevolence of state legislators. These assumptions seem reasonable in light of previous work showing that states tend to borrow the more successful policies from

each other (e.g. Volden, 2006, Pacheco, 2012, Butler et al., 2017).

What is the economic mechanism underlying the law's boost to growth? In an extensive mechanisms analysis, we focus on an "incomplete contracts" interpretation of legislating and the economy. To summarize, we think of state governments as creating legislation that businesses must abide by, which necessitates specialized investments and supply chain structures. Ambiguous or incomplete legislation can generate uncertainty about enforcement, which deters businesses from making investments in the first place. More comprehensive and clear legislation reduces this hold up problem and encourages such specific investments, thus stimulating economic growth.

We report a series of empirical findings in support of legislative completeness and hold-up reduction as a key mechanism. First, laws that regulate the economy (e.g. land/property rights) have a larger effect on growth than laws regulating social issues (e.g. family law). Second, provisions that produce contingencies – actions or outcomes conditioned on events, reducing uncertainty over those events – are more helpful than non-contingent provisions (Battigalli and Maggi, 2002, 2008). Third, there is decreasing economic benefit to legislating when the existing stock is already relatively detailed. Fourth, the law's effect on growth is focused on sectors relying more on customized inputs requiring relationship-specific investments (Nunn, 2007). And fifth, increasing legal detail is most beneficial under higher local economic uncertainty (Baker et al., 2016).

Further, the estimated effect coefficients are economically meaningful. The regressions suggest that a 10 percent increase in legislative flows due to borrowing increases the state GDP growth rate (per capita) by about 0.15 percentage points, relative to a mean of 3.1 percentage points. For contingent provisions, the estimated effects are even larger – a 10 percent increase in the flow of contingencies would lead to about a 0.6 percentage-point boost to the growth rate. For comparison, achieving a .15 percentage-point increase in growth through fiscal stimulus would require about a 0.1 percentage-point increase in net-of-tax government spending (e.g. Nakamura and Steinsson, 2014, Chodorow-Reich, 2019).

These results add to a long-running debate on how laws and regulations influence growth prospects (e.g. Parker and Kirkpatrick, 2012). Overall, the results are consistent with the "positive view" that legislation is needed to regulate externalities, define the tax base, and allocate government expenditures; and usually, it helps the economy to grow (e.g. Dam, 2007). Empirical work documenting a positive correlation between legislative output and growth includes Mulligan and Shleifer's (2005) cross-sectional

comparison of U.S. states, Fukumoto's (2008) time series comparison in Japan, and Kirchner's (2012) time series comparison in Australia.[1] Conversely, our evidence goes against the "negative view" of public choice theory that regulation is usually excessive and tends to hinder economic growth (e.g. Niskanen, 1971, Davis, 2017); in particular, the effect of compliance costs in hindering business formation, competition, innovation, and skill acquisition (Fonseca et al., 2001, Nicoletti and Scarpetta, 2003, Ciccone and Papaioannou, 2007, Braunerhjelm and Eklund, 2014). Related empirical work emphasizing compliance costs includes Botero et al.'s (2004) cross-country comparison on the regulation of labor, Di Vita's (2017) comparison of regulatory complexity across Italian regions, and Gratton et al. (2021), also in Italy, pointing to electoral incentives as a mechanism for overproduction of laws and negative impacts on economic growth. Finally, more nuanced models describing conditions under which higher legal complexity helps or hurts includes Kawai et al. (2018) and Foarta and Morelli (2022).

Two closely related papers, both in terms of the topic and the method, have used text analysis to assess federal regulations in the United States. Dawson and Seater (2013) show that in the U.S. time series, the number of pages in the Code of Federal Regulations is negatively related to overall national growth. Coffey et al. (2020) produce a panel dataset across industries since 1980, and find that stricter industry-specific regulation is associated with lower industry growth. There could be many reasons for these different results, including the discussion earlier on how our legislation instrument is driven by borrowing of successful policies among state legislators.

The mechanisms analysis around legislative completeness and investment is relevant to the large literature in labor economics and contract theory on hold up and the theory of the firm (e.g. Williamson, 1979, Grossman and Hart, 1986, Hart and Moore, 1988). Here, we extend that idea to the legislative "social contract." The public finance literature on tax compliance has made a similar connection to incomplete laws as incomplete social contracts (Weisbach, 2002, Holtzblatt and McCubbin, 2003, Givati, 2009); consistent with our results, more detailed tax legislation can help the economy by reducing legal uncertainty (Slemrod, 2005, Graetz, 2007). Another related paper is Nunn (2007), who finds that industries relying on inputs requiring relationship-specific investments tend to cluster in countries with more effective legal institutions and better contract en-

[1] A number of papers have used indexes for regulatory quality and shown a positive correlation with economic growth across countries or over time (Gørgens et al., 2004, Loayza et al., 2005, Djankov et al., 2006, Jalilian et al., 2007, Jacobzone et al., 2010).The different indices include measures of regulatory burden from OECD surveys, World Bank's Doing Business, World Bank Governance Indicators, Amadeus database, UNIDO 3-IndStat, and Fraser Institute's Economic Freedom Index.

forcement. The most recent empirical work on legal uncertainty and economic activity includes Bamieh et al. (2023) and Giommoni et al. (2023).

The paper is organized as follows. Sections 2, 3, and 4 describe the data, text analysis methods, and empirical approach, respectively. Section 5 reports the main results on laws and growth, while Section 6 reports the extensive mechanisms analysis on what what types of laws matter, and under what economic conditions the effect is largest. Section 7 concludes.

2 Data Sources

This section describes the data and provides summary statistics. The variables can be roughly divided into three categories: data on economic output and growth, statute text data and legislative output, and control variables. The main summary statistics are reported in Appendix Table A.1. A full list of variables with descriptions is shown in Appendix Table A.2. Additional summary statistics are shown in Appendix Table A.3.

The dataset for our empirical analysis ranges from 1965 through 2012. This period is determined by the beginning of the economic growth variables (in 1965) and the ending of the legislative text variables (in 2012). The data are constructed by biennium (two-year periods), since many states publish their compiled statutes once every two years.

Economic Activity. We have a rich array of variables on the economic conditions by year in each of the 50 states. These data are assembled from the Bureau of Economic Analysis Regional Accounts, County Business Patterns, Klarner (2013), and Ujhelyi (2014).

As our empirical analysis looks at how legal flows impact economic growth, the key variable Y_{st} is local growth, measured by the change in log per capita Gross State Product (GSP) in state s between year $t-2$ and year t (as the data are at the biennium level). Appendix Figure A.1 shows the evolution of this variable over the time period of the data. The data on the numerator (total real GSP) and the denominator (total population) will also be used separately. All economic variables denominated in dollars are deflated to 2007 values using the state-level CPI.

We have a number of additional measures of economic activity. On the worker side, we have labor income and employment. On the firm side, we have number of

establishments and profits.

State Session Laws Corpus. The dataset on legislation includes the full text of U.S. state session laws. This corpus consists of the statutes enacted by each state legislature during each session. The statutes modify the text in the state's compiled legislative code. As mentioned, the laws are published annually or biennially. To ensure consistency, the dataset is built biennially, with the data point for even year t including the laws from t and $t-1$.

The statutes can include new laws, amendments to existing laws (revisions), and repeal of existing laws (deletions). Ideally, one could distinguish the effects of amending and repealing provisions in terms of their effect on the stock of laws. In particular, repeal of clauses usually has a negative effect on the stock of laws, while amending of clauses could have a negative, neutral, or positive effect depending on what they replace. Unfortunately, our corpus does not provide a machine-readable indication of the original text that is being amended or repealed, so we cannot precisely determine the size of removals.[2] Hence, our main measure of legislative volume includes all types of provisions and does not distinguish amendments or repeals. Through qualitative inspection of samples from the corpus, however, we could determine that amendments and repeals are a relatively small share of the text in the state session laws. Quantitatively, we proxy for the share of amendments and repeals by scanning for associated text signifiers ("amend*" or "repeal*"). Appendix Figure A.6 shows the time series for these shares over time, and they are relatively infrequent (about 3 percent repealing and less than 1 percent amending). In any case, the presence of amendments and repeals is not a problem for our empirical analysis as long as their frequency is not confounded with the instrument. Appendix Figure A.15 shows that, reassuringly, the instrument is not correlated with the share of either type of clause.[3]

The next issue is that the text from the state session laws corpus is produced from optical character recognition (OCR) applied to printed laws. From inspecting samples, the OCR is high quality. Appendix Figure A.5 shows the scanned copy of a page from a

[2]Similarly, we cannot cleanly distinguish clauses that add regulations or remove them. So some of our estimated effects could be due to clauses that deregulate rather than regulate. An example of such a "deregulating" law is Texas Utilities Code Title 2.C Ch. 65, "Deregulation of certain incumbent local exchange company markets", enacted in 2005. While that law is taking away power from a telecommunications regulator, it still contains a number of quite detailed provisions. See `https://statutes.capitol.texas.gov/Docs/UT/htm/UT.65.htm`.

[3]Appendix Table A.23 shows that we can also control for these variables in our regressions and it makes no difference.

statute enacted in the Texas Legislature for the 1889 session. As can be seen, although the statute is old, the quality of the digitized version is quite good.

Still, as with any historical digitized corpus there are a significant number of OCR errors. To investigate this, we computed a proxy for OCR as the misspelling rate for common (non-proper) nouns. Appendix Figure A.6 shows the time series in the misspelling rate and it is low (about 3 percent) and smooth over time. These misspellings could add measurement error to the legislative output measure. Again, this is not a major problem for our empirical analysis as long as the OCR error rate is not correlated with either the outcome or the instrument for legislating. Fortunately, the instrument is not correlated with the misspelling rate (Appendix Figure A.14).[4]

Demographics. We link the data on economics and law to demographic data at the state level. Besides population, we use census information on the age distribution, the fraction of urban population, and the share of foreign-born population.

State government finances. We use a set of data on government revenues and expenditures from the state government finances census. These include total government expenditures (in 1000s current dollars), and legislative expenditures (in 1000s current dollars).

Politics. Next, we use measures of state political conditions. For each state and year, we have a measure of Democratic Control, which is the number of governing bodies (lower chamber, upper chamber, and governor) controlled by Democrats. This ranges from zero to three.

Relationship specificity. We measure the importance of relationship-specific investments using the data from Nunn (2007). Building on Rauch (1999), Nunn first identifies inputs that are sold on an organized exchange or reference-priced in trade publications. The idea is that exchange goods or reference-priced goods have an elastic supply on the global market, and they can be purchased in flexible quantities without established relationships. Other goods (neither on exchanges nor price lists) are more specialized and depend on private relationships – and relationship-specific investments. Nunn (2007) scores industry sectors by the relative share of inputs from the latter category – that is, inputs requiring relationship-specific investments. We match those scores with Bureau

[4]In addition, controlling for OCR error rate in our main results does not change them.

of Economic Analysis data on yearly state GDP by sector, using Autor et al's (2013) crosswalk files to match different industry classifications. The resulting dataset has information on relationship-specificity for 30 sectors (3-digit NAICS). Those sectors, ordered by relationship-specificity, are listed in Appendix Table A.4. Intuitively, the lowest-scoring industries are oil/gas extraction and primary metal manufacturing, while the highest-scoring industries are computers/electronics and publishing.

Local economic policy uncertainty. Finally, we have information on state-year-level economic policy uncertainty, constructed from the text of newspaper articles based on the approach from Baker et al. (2016). For this purpose, we use the searchable local newspaper archive `newspapers.com`, which can programmatically provide counts by state and year for articles meeting search criteria. We count the number of articles mentioning the phrase 'economic uncertainty' in a state in a given biennium and then construct a frequency by taking this count divided by the total number of news articles. Appendix Figure A.4 shows that our measure is highly correlated with a state-level measure recently developed by Baker et al's team for recent years (rank correlation coefficient = 0.41).

3 Text Analysis Methods

This section summarizes our methods for extracting useful measures from the statute texts.

3.1 Measuring Legislative Output

Using the digitized text of the state session laws, we start by segmenting the text for each biennium into statutes. Roughly speaking, a "statute" is a singular, coherent enacted bill or policy. It usually corresponds to a "chapter" in the compiled legislative code, which is the second level of organization beneath titles. Appendix Figure A.5 Panel A shows the distribution of the number of statutes by biennium. Panel B shows the distribution of the number of words per statute. Panels C and D respectively show the time series for the number of statutes, and number of words per statute, over time.

Next, the statutes are segmented into sentences using a sentence tokenizer. For each sentence, we extract legally relevant statements following the method in Vannoni et al.

(2019) and Ash et al. (2020). The method works as follows, with more detail provided in Appendix B.2.

We apply a syntactic dependency parser to construct data on the grammatical relations among words in each sentence (Dell'Orletta et al., 2012, Montemagni and Venturi, 2013), as illustrated in Appendix Figure A.7. The dependency parse identifies the main verb in a sentence segment, along with the associated subject, object, helping verb, and information on negation.

To extract legally relevant statements, we define a set of legislative provision types (also called legal frames), including delegations, constraints, and so on (Soria et al., 2007, Saias and Quaresma, 2004). We extract dependency tags associated with each legislative provision type (van Engers et al., 2004, Lame, 2003); for instance, a constraint is characterized by three potential structures: a negative structure with a modal, such as 'the Agent shall not'; a negative structure with a permission verb, such as 'the Agent is not allowed'; or a positive structure with a constraint verb, such as 'the Agent is prohibited from'. The set of provision types, with tagging rules, are listed in Appendix Table A.6. Vannoni et al. (2019) and Ash et al. (2020) use this method to count provisions across different agent types. Here, the aim is less targeted – we count the number of legal provisions by state and over time.

Our measure of legislative output W_{st} is the number of legal provisions counted in the session laws for a state at biennium t. To assess proportional changes in provisions, we use the log of the counts. The evolution of this measure, by year, is illustrated in Figure A.1. Counting provisions should provide a cleaner measure of the flow of legal requirements than would be obtained by a coarser measure, such as word counts or page counts. Word or page counts would be noisier because they include a lot of non-legislative or otherwise less informative content. Vannoni et al. (2019) provide some validation against human annotations that our parser-based measure does a better job than simpler measures in identifying legally relevant statements. Appendix Figure A.2 shows that provision counts and word counts are correlated. Appendix Table A.22 explores variations on our analysis using word counts or page counts.

3.2 Allocating Laws to Topics

An essential ingredient in our analysis is to assign statutes to topics. For this purpose, we apply the Latent Dirichlet Allocation (LDA) model described in Blei et al. (2003). This algorithm, by now well-known in the literature on text data in political economy

(Grimmer and Stewart, 2013, Hansen et al., 2018), assumes that every document is a distribution over topics, which in turn is a distribution over words and phrases. A document is generated by drawing topic shares, and then the words of the document are drawn from those topics.

We trained LDA on our corpus at the statute level using the Mallet wrapper from the Python gensim package. The main tunable hyperparameter in LDA is the number of topics K. Starting with $K = 6$ topics, we increased the number by multiples of six (12, 18, ..., etc) to find the topic count that maximized the topic coherence score. This score was maximized at $K = 42$. We also inspected the topics subjectively, and we agreed that the specification with $K = 18$ topics was a good balance for a relatively small number of intuitive, coherent topics. After producing our main empirical results for all topic counts $K \in \{6, 12, ..., 48\}$, we found that the instrument constructed with $K = 18$ topics (more details below) generates the most consistent estimate across specifications with different sets of predetermined covariates. Therefore, we have two preferred LDA models: 18 topics and 42 topics. For our main results, however, the topic number choice is not important. In Appendix Table A.21 we show consistent results for all LDA models produced ($K \in \{6, 12, ..., 48\}$). The baseline specification for the main text uses the LDA model with $K = 18$ topics. The list of 18 topics is reported in Table 1, sorted by most to least frequent in the state session laws corpus. The model produces clearly interpretable topics for vehicle regulation, licensing, courts, project funding, childcare services, trusts and estates, employment law, taxes, land regulation, retirement regulation, etc. These are the types of legal policy areas that one would expect to arise in the business of U.S. state government.[5]

The 42-topic LDA model is mainly used to flesh out our results by policy type. These more granular topics were easier than the 18-topic model to divide into broader policy areas: economic regulation, social regulation, fiscal policy, and procedural. To make this assignment to policy groups, all three of the co-authors annotated the topics and we assigned the majority annotation, with some discussion under disagreement. The list of topics, with broader category assignments, is reported in Appendix Table A.7. Appendix Figure A.8 shows the legislation shares across these four categories over time.

Using the trained models, we assign to every statute a distribution over topics based on the words and phrases in that statute. For each state-biennium, the number

[5]Some example sentences with topic tags are listed in Tables A.8 and A.9.

Table 1: List of Topics, 18-Topic Specification

Label	**Frequency**	**Most Associated Words**
Courts	0.0724	court judgment attorney case appeal civil petition sheriff trial circuit_court district_court such_person complaint counsel brought circuit warrant paid
Pensions	0.0653	paid benefit rate payment equal death age credit pay total life pension premium calendar_year loss account case per_cent event membership excess maximum
Local Projects	0.0645	development local project budget government cost grant research center local_government data transfer governor is_the_intent develop urban review biennium
Procurement	0.0621	director contract work review civil labor contractor attorney_general bureau final perform audit receipt status exempt panel government firm bid prepared
Elections	0.0612	district town petition charter special ballot mayor voter township precinct cast referendum census elector case town_council said_district such_district
Banking	0.0604	loan trust bank agent partnership institution foreign stock mortgage deposit surplus interest merger credit_union partner case credit gift branch transact
Licensing	0.0593	license fee dealer sale food sold holder sell valid fish agent distributor milk liquor product such_license livestock game card retail misdemeanor fine
Real Estate	0.0576	real interest sale owner contract claim lien payment transfer instrument seller holder issuer debtor claimant buyer pay broker settlement receipt money
Bonds	0.0574	interest bond payment commonwealth cost sale paid pay project power thereon sold debt pledge local_law event hereof proper said_board real port sell therefrom
Expenditures	0.0569	fund account money paid special pay tile payment transfer for_the_fiscal_year excess trust_fund so_much_thereof deposit state_general_fund auditor tie
Bureaucracy	0.0551	governor council government chief fire appoint personnel compact conflict perform shall_consist invalid parish successor volunteer membership head travel
Healthcare	0.0546	health care treatment health_care physician home human patient mental mental_health drug social condition public_health medicaid dental client review institution
Child Custody	0.0535	child court minor children parent age probation crime victim parole guardian adult petition placement youth case social legal child_support obligor home
Taxes	0.0522	tax paid gross credit return net rate exempt assessor case refund equal sale total calendar_year payment fuel portion sold price retail zone pay such_tax
Land & Energy	0.0512	land water owner control site air solid gas tenant oil park airport forest coal plant environment prevent underground power soil portion landlord condition
Education	0.0474	school school_district state_board district student institution higher teacher special aid pupil children school_year tuition high_school school_board
Traffic 1	0.0423	motor highway driver owner traffic plate test vessel accident weight special sect trailer railroad state_highway stricken feet fine alcohol aircraft carrier
Traffic 2	0.0267	street road feet island river run tract team great highway township center_line park center corner lake beach more_or_less san honor creek high_school

Notes. This table shows the 18 topics, along with their frequency and the most associated keywords. As it can be seen, the distribution is rather dispersed and no topic is predominant. The most frequent topics across states and years are Courts, Pension and Local Projects, whereas the least frequent are Education, Traffic 1 and Traffic 2.

of provisions by topic is computed by the sum of provisions in that state-biennium's statutes, weighted by the topic share of each statute. Formally, let L_{st} be the set of laws in state s time t. Each statute $i \in L_{st}$ has a provision count w_i and a distribution over topics $\vec{v} \ni v_i^k, \forall k \in \{1, ..., K\}$, where $v_i^k \geq 0$ and $\sum_k v_i^k = 1$. Then define legislative flows for topic k in state s during t as

$$W_{st}^k = \sum_{i \in L_{st}} v_i^k w_i.$$

This process results in a dataset with the number of provisions by topic for the legislation of a state in a biennium.

3.3 Measuring Contingency in Legal Language

Contingencies are a prominent feature of legal language because they impose more precise conditions under which legal actions will be made (Crawford and Ostrom, 1995, Frantz and Siddiki, 2022).[6] We measure contingency using a simple lexicon-based approach. Starting with several lists developed by linguists to indicate contingency, we searched for examples in the statutes to check which words almost always indicated contingency. After this inspection process, we settled on a relatively short list of words that were distinctive of contingent clauses. Formally, a provision is contingent if one of the following words (or phrases) appears in the same sentence: {*if*, *in case*, *where*, *could*, *unless*, *should*, *would*, *as long as, so long as*, *provided that*, *otherwise*, *supposing*}.

To see what this distinction looks like in context, Appendix Table A.8 shows examples of contingent provisions, while Appendix Table A.9 shows examples of non-contingent provisions. These are randomly sampled from the corpus to represent different states, years and topics. One can clearly see that non-contingent clauses impose rigid requirements, while contingent clauses depend on some environmental factor.

Let W_{st}^C be the number of contingent provisions in the statutes from state s in year t. Let $W_{st}^N = W_{st} - W_{st}^C$ be the number of non-contingent provisions. Following the same procedure as in Subsection 3.2, we also compute topic-specific counts of contingent and non-contingent provisions by state-biennium.

Summary statistics related to contingency are reported in Appendix B.4. About one-fifth of provisions are contingent. Appendix Table A.10 shows the changes in contingency across decades, showing that the share of contingent clauses has decreased

[6]For a more detailed discussion of this issue, see Appendix B.4.

slightly over time, from 19.3% in the 1970s to 18.6% in the 2000s. Appendix Figure A.10 shows the time series for the share of contingencies by the four policy categories. Economic-regulation clauses have usually had the highest degree of contingency.

4 Empirical Approach

This section outlines the main features of our research design for estimating a causal effect of legislative output on economic growth. We use a shift-share instrumental variables design based on topics in legal texts. We describe how the instrument is constructed and present evidence and checks on its validity.

4.1 Linear Regression Specification

Our dataset is at the state-biennium level, for each state s and biennium t. The main research objective is to test whether legislative output W_{st} increases or decreases economic growth Y_{st}. More formally, let W_{st} equal the number of legal provisions enacted, and $\Delta \log Y_{st}$ equal the log change in real per capita GDP, in s during t. We assume a linear model

$$\Delta \log Y_{st} = \alpha_s + \alpha_t + \alpha_s \cdot t + \rho \log W_{st} + X'_{st}\beta + \varepsilon_{st} \tag{1}$$

where α_s includes state fixed effects, α_t includes time (biennium) fixed effects, and $\alpha_s \cdot t$ includes state-specific time trends. When estimated by ordinary least squares (OLS), this is a standard two-way fixed effects model. X_{st} includes a set of additional covariates, for example pre-period state characteristics interacted with the time fixed effects, for use in robustness specifications.[7]

Under strong identification assumptions, OLS estimates for ρ would procure a causal effect of legislative output on growth. The key assumption is that there are no unobserved factors (time-varying at the state level) correlated with both $\log W_{st}$ and

[7]The set of variables included differs by specification. For example, in the main 2SLS results (Table 3), we report results with pre-period economic covariates interacted with biennium effects (initial growth, initial GSP, and initial GSP per capita); controls for initial sector shares interacted by biennium; demographic characteristics (share of urban, foreign, and population) measured in the pre-treatment period interacted with biennium fixed effects; topic share controls; lagged government expenditures; and the lagged dependent variable. The specific variables in each column are listed in the respective table notes. Descriptions of the covariates with data sources are shown in Appendix Table A.2.

$\Delta \log Y_{st}$. This assumption is unrealistic, given that there could be unobserved shocks (e.g., the rise of a new industry) that affect both economic output and legislative output. Our empirical strategy is designed to address these confounders.

4.2 Shift-Share Instrument for Legislative Diffusion

The baseline OLS model (1) is likely produce biased estimates given confounders and reverse causality. We address this issue using an instrumental-variables approach that isolates exogenous variation in legislative output due to sharing of legal texts across U.S. states. This section describes the source of identifying variation and formalizes the construction of the instrument.

A classic motivation for federalism institutions is that the constituent units – e.g. U.S. states – can act as laboratories of democracy in the discovery and adoption of good policies (Burgess et al., 2016). When states adopt good policies through legislation, legislators in other states can learn from that example and adopt similar policies in their laws. As discussed in previous work, one of the main drivers of policy diffusion across U.S. states is policy success; successful policies are more likely to spread across states (Volden, 2006, Pacheco, 2012, Shipan and Volden, 2014, Butler et al., 2017). For example, Yu et al. (2020) find that in the case of drunk driving laws, only policies that reduce the total fatality rate tend to spread across states. Souza et al. (2019) document tax competition between states, while DellaVigna and Kim (2022) find policy diffusion between neighboring states and states with similar political makeups.

As these policies are embodied in legislative text, state-to-state policy diffusion typically consists of borrowing of text (Burgess et al., 2016). Other things equal, it is cheaper to use previously used text than draft something from scratch. This borrowing mechanism is likely to be strengthened in the case of U.S. states because the state legislatures are relatively resource-constrained (e.g. Malhotra, 2006). Burgess et al. (2016) document borrowing of text, showing that in recent years some of that is driven by organizations publishing model legislation (see also Hansen and Jansa, 2021).

To summarize, the previous work shows that state legislators borrow legal provisions from other states. That borrowing is due in part to resource constraints, suggesting that there are drafting costs for new legislation. Hence, we would expect more borrowing on topics where a given state has not yet legislated in much detail. Further, there is a tendency to selectively borrow successful provisions. Hence, the legislation that diffuses by borrowing is positively selected in terms of its impact on society and the economy.

These ideas motivate the construction of our instrument. Formally, we adapt a shift-share instrumental-variables design, often attributed to Bartik (1991, 1994) but popularized by Blanchard and Katz (1992). The original application of the approach was meant to address the endogeneity between employment growth and economic growth; that is, more economically prosperous regions tend to attract more labor. To address this problem, one can instrument local employment growth with the interaction between pre-treatment local employment shares by sector and national employment growth rates by sector. The Bartik approach therefore isolates changes in employment growth due to these labor demand shocks (rather than due to local supply side responses).

While the use in economic growth and employment is still the classic example, more recent applications include migration effects on labor markets (Card, 2001, Basso and Peri, 2015), imports and economic growth (Autor et al., 2013, 2016), market size and drug innovation (Acemoglu and Linn, 2004), small business lending and economic growth (Greenstone et al., 2020), effects of democracy on growth (Acemoglu et al., 2019), and effects of the China shock on nationalism (Colantone and Stanig, 2018) and populism (Autor et al., 2020). In tandem with this diversity of applications, a recent and active literature in econometrics has produced useful results and guidance on how to use these estimators (Borusyak et al., 2022, Jaeger et al., 2018, Adao et al., 2019, Goldsmith-Pinkham et al., 2020).

To link our setting to that of more traditional shift-share designs, let's conceive the flow of legislative provisions as analogous to the flow of workers or flow of migrants. Analogous to economic sectors (which supply workers) and origin countries (which supply migrants), we have legal policy topics (which supply legislative text). The instrument consists of a "share" factor and a "shift" factor, to be described in turn. As above, W_{st} represents the total number of legislative statements in state s at biennium t, while W_{st}^k represents the number of statements on topic k in s at t.

The local "shares" are a state's pre-period stock of legislative output on each topic, analogous to pre-period employment shares across sectors, or pre-period immigrant population shares across origin countries. Formally, we construct the pre-treatment legislative topic shares as the average of topic shares over the decade prior to our analysis (1955-1964), represented as period zero: $\frac{W_{s0}^k}{W_{s0}}$.[8]

[8]Note that we divide a state's cumulative statutes by topic, W_{s0}^k, by the total statute output in those years, W_{s0}. This is needed to normalize variation by state, such that instrument variation is driven by topic variation within state. We cannot use W_{s0}^k itself as the shares (the level, rather than the share) due to the large level differences across states, which then produces a very noisy instrument.

The global "shifter" in our case is nationwide growth in topic-specific legislating, analogous to nationwide growth in employment in a particular sector, or growth in immigration from a particular origin country. Formally, this is the leave-one-out average log change in legislation to topic k in other states, $\frac{1}{49}\sum_{r\neq s}\Delta \log W_{rt}^k$, where r indexes the other 49 states. Borusyak et al. (2022) note that the assumptions for identification are relaxed with the leave-one-out specification for the shifter.

Now we combine the "shifters" and the "shares." The instrument for legislative output is the weighted sum, by topic, of the leave-one-out average legislative flow on that topic in other states, multiplied by this state's pre-treatment topic share:

$$Z_{st} = \sum_{k=1}^{K} \underbrace{\frac{W_{s0}^k}{W_{s0}}}_{\text{shares}} \underbrace{\sum_{r\neq s} \frac{\Delta \log W_{rt}^k}{49}}_{\text{shifts}}. \tag{2}$$

To assist interpretability of the first-stage and reduced-form estimates, Z_{st} is standardized to mean zero and variance one.[9] The first stage equation for legislative output is

$$\log W_{st} = \alpha_s + \alpha_t + \alpha_s \cdot t + \psi Z_{st} + X_{st}'\beta + \eta_{st} \tag{3}$$

where Z_{st} is given by (2). The other items are the same as Equation (1). Reduced form estimates are produced by

$$\Delta \log Y_{st} = \alpha_s + \alpha_t + \alpha_s \cdot t + \gamma Z_{st} + X_{st}'\beta + \epsilon_{st}, \tag{4}$$

that is, regressing the outcome directly on the instrument.

4.3 First Stage

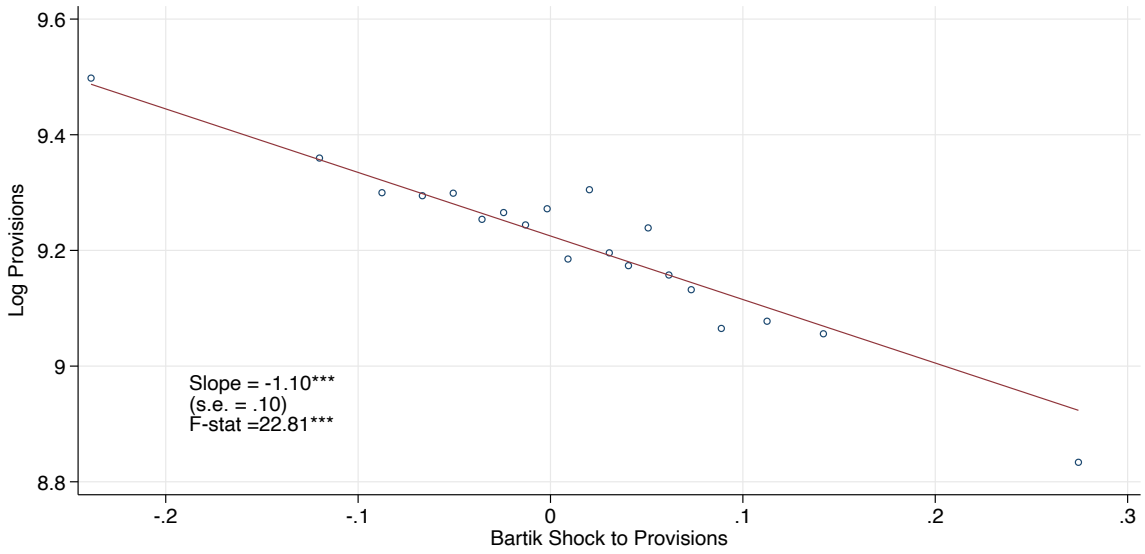
In the classic shift-share instrument, it is expected that the first-stage effect of Z_{st} on the endogenous regressor (e.g. employment), ψ, is positive, as having high previous

In line with that, the shift-share instrument based on levels does not get a strong enough first stage (F-stat < 2.5). The estimated 2SLS coefficients are the same sign and similar in magnitude ($\hat{\rho} \approx 0.03$), but noisy and not statistically significant.

We include all topics in constructing the instrument, as recommended by Borusyak et al. (2022), relative to a situation where only a subset of shares is used for the instrument (as in Autor et al., 2016). Moreover, the use of pre-treatment shares is advisable in situations where shocks are serially correlated and shares are affected by lagged shocks.

[9]See Appendix Table A.11 for summary statistics on the instrument and endogenous regressor by decade.

Figure 2: First Stage: Impact of Shift-Share Legislative Shock on Legislative Output



Notes. Binned scatterplot for the first-stage relationship (Equation 3) between the shift-share instrument (horizontal axis) and the log number of provisions (vertical axis). State and year fixed effects absorbed.

shares of sectors that are increasing nationally will tend to get pulled upward. In our setting, however, the effect of previous shares go in the other direction. As outlined in the discusson above, it is states with relatively low previous legislating on a topic that will be pushed most by national upward shifts on that topic. Hence, we expect $\psi < 0$.

Figure 2 illustrates the first-stage relationship, which is statistically significant ($p = .003$) and produces a Kleibergen-Paap first-stage F-statistic of 22.8 in the baseline specification. As expected, the first-stage relation between legislative flow and the instrument is negative. When a state had initially low detail on a topic, then it is more likely to increase legislating in response to national trends on that topic. Consistent with this interpretation, the "shift" term of the instrument is positively correlated with the endogenous regressor $\log W_{st}$, while the "shares" term is negatively correlated (Appendix Figure A.11).

4.4 Exogeneity and Exclusion

There are two approaches to identification in shift-share designs. In the first approach, one assumes that the pre-period shares are conditionally exogenous (Jaeger et al., 2018, Goldsmith-Pinkham et al., 2020). In this view, the exclusion restriction hinges on the

fact that the shares (normally, sectoral composition, but in our case, topic shares) are as good as randomly assigned conditional on the fixed effects and controls (see Borusyak et al., 2022). In our case, this assumption could be formally stated as

$$\mathbb{E}\{\frac{W_{s0}^k}{W_{s0}} \cdot \epsilon_{st} | \vec{\alpha}_{st}, X_{st}\} = 0, \forall k \tag{5}$$

where $\vec{\alpha}_{st}$ gives the vector of fixed effects. Using the definition of the instrument (2), Equation (5) implies instrument exogeneity. Equation (5) is a relatively strong requirement in most empirical contexts, however. In our case, this would mean that pre-period legislative topic shares are uncorrelated with subsequent trends in economic growth during the treatment period. This assumption is difficult to justify, since the pre-period legislation could be drafted in expectation of future growth trends. For example, the proportion of legislation on taxes or employment regulation in the 1950s could be correlated with growing more or less quickly in the 1960s or 1970s. Still, we show that we can pass the checks proposed by Goldsmith-Pinkham et al. (2020) and Jaeger et al. (2018) in the framework that assumes exogeneity of pre-treatment shares. Appendix Table A.14 shows that the instrument is uncorrelated with pre-treatment state characteristics. Appendix Figure A.13 shows that pre-treatment topic shares are uncorrelated with subsequent growth trends. These statistics lend support to the "exogeneity of shares" assumption, which would suffice for instrument validity.

A second approach to identification, taken by Adao et al. (2019) and Borusyak et al. (2022), relies on different, arguably weaker, assumptions. In these frameworks, the exclusion restriction follows from the conditional exogeneity of the current-period shifters, rather than from the pre-treatment shares. No assumption is needed with respect to the pre-treatment shares, and instead this approach assumes that the global shocks are uncorrelated with the exposure-weighted average of potential outcomes. In the case of Autor et al. (2016), for example, the identification assumption is that average unobserved determinants of economic growth across states must be unrelated to flows of Chinese imports. With panel data (as in our context), the assumption can be further relaxed. Formally, we have

$$\mathbb{E}\{\sum_{r \neq s} \frac{\Delta \log W_{rt}^k}{49} \cdot \epsilon_{st} | \vec{\alpha}_{st}, X_{st}\} = 0, \forall k \tag{6}$$

where the terms and technicalities are as above. With the inclusion of state and time

fixed effects, shocks are allowed to be correlated with exposure-weighted averages of state and time-invariant unobservables, or linearly varying within state given the inclusion of state-time trends (Borusyak et al., 2022).[10]

In line with Adao et al. (2019) and Borusyak et al. (2022), we take a number of steps to assess the validity of Z_{st} as an instrument for $\log W_{st}$ (see Appendix C). First, to check that the relevance of the shift-share instrument is driven by a majority of topics, we regress the increase in provisions related to a topic in a state on the increase in the total provisions related to that topic in other states and the increase in all legal provisions in that state, for every topic (including state and year fixed effects and clustering standard errors by state). We find that topic growth is statistically significant in the great majority of topics, as shown in Appendix Figure A.12. Second, we use the test for weak instruments, robust to heteroscedasticity, serial correlation, and clustering, proposed by Olea and Pflueger (2013). A rule of thumb for 2SLS is to reject the null hypothesis of a weak instrument when the effective F is greater than 23.1. In our data, the effective F statistic equals 132.8 and we reject the weak instrument null at 5 percent significance. Third, Appendix Table A.13 reports the following placebo test: we regress economic growth on future values of the legislative-growth instruments. The estimates are not statistically significant.[11] Fourth, we run a balance test by regressing the instrument on some potential confounders. Appendix Table A.15 shows the instrument is not correlated with current or lagged values for relevant state characteristics.

[10]There are two additional identification issues that should be discussed. First, there is the issue of shared economic or political shocks across multiple states, which drive both legislation and economic growth. Economic crises like the Great Recession and the Covid Pandemic would be examples of such events. A similar issue is there for the classic Bartik (1991) instrument for employment and growth. All of our validity checks, for example the placebos for time and other variables, are designed to support our assumption that such joint shocks are second order, once integrated into the constructed instrument.

A second issue is how the instrument, and the resulting nudge to detail, impacts other neighboring states. There could be positive spillovers in the outcome due to gains from trade, for example, or negative spillovers due to migration of labor or capital. There may also be spillovers in the effect on legislative output. We assume that to the extent that these spillovers exist, they are second-order to the main effect of the instrument. Further exploring such spillovers is an important area for future work, as discussed in the conclusion below.

[11]See also Appendix Table A.18 for additional results on leads and lags of the effect.

Table 2: First Stage, OLS, and Reduced Form

	(1)	(2)	(3)	(4)	(5)	(6)
	Effect on Provisions		Effect on Real GDP Growth Per Capita			
	FS	FS	OLS	OLS	RF	RF
Legislative Output			0.0146+	0.0152		
			(0.00832)	(0.0123)		
Instrument (Z_{st})	-1.099**	-1.221**			-0.0200*	-0.0205*
	(0.230)	(0.259)			(0.00883)	(0.00940)
Observations	1,183	1,183	1,182	1,182	1,182	1,182
R-squared	0.813	0.9	0.431	0.446	0.420	0.440
State FE	X	X	X	X	X	X
Time FE	X	X	X	X	X	X
State-Specific Trends		X		X		X

Notes. Columns 1 and 2 show the estimates for the First Stage (FS) (Equation 3). Columns 3 and 4 show the results for OLS estimates of Equation (1). Columns 5 and 6 give the Reduced Form (RF) specification (Equation 4), regressing the outcome (growth per capita) directly on the instrument . All specifications include state and biennium fixed effect, with a second column including state-specific trends. All standard errors clustered by state. **p<.01; *p<.05; +p<.1.

5 Main Results

This section reports the main empirical results. We start in Section 5.1 with an empirical test for whether greater legal output causes greater growth, or lower growth, at the margin.. We then report some robustness checks and supporting results on the main effect in Section 5.2.

5.1 Main Results on Legislation and Growth

Table 2 presents the first results for legislative output and growth. Columns 1 and 2 show estimates for the first stage Equation (3), illustrating a negative and significant effect of the instrument on log provisions. In Columns 3 and 4, we see that OLS estimates of the second stage Equation (1) are positive, but not robustly significant. Columns 5 and 6 show a significant reduced-form effect of the instrument on growth, from Equation (4). As previously discussed, the reduced-form coefficient is negative, reflecting that lower pre-treatment detail on a topic is associated with a positive shock to legislative output. Additional specifications for OLS and the reduced form are shown in Appendix Table A.16 and Appendix Table A.17, respectively.

2SLS estimates for ρ, the effect of legislative output on growth, are reported in Table

Table 3: Effect of Legislative Output on Economic Growth (2SLS)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Effect on Growth Rate Per capita						
Legislative Output	0.0182*	0.0168+	0.0152*	0.0134+	0.0116+	0.0222*	0.0094+
	(0.00903)	(0.00863)	(0.00704)	(0.00687)	(0.00602)	(0.0106)	(0.00507)
First Stage F-stat	22.86	22.19	23.11	22.92	44.51	19.69	27.30
Observations	1,182	1,182	1,182	1,182	1,134	1,182	1,086
Time FE	X	X	X	X	X	X	X
State FE	X	X	X	X	X	X	X
State Trends		X					X
Econ Vars × Time			X				X
Sector Shares × Time				X			X
Demog Vars × Time					X		X
Topic Shares						X	X
Lagged Govt Expend							X
Lagged Dep. Var.							X

Notes. Results for the 2SLS model (Second Stage 1 and First Stage 3). All specs include state and biennium fixed effects. Column 2 adds state-specific linear trends. Column 3 adds a set of pre-period economic covariates interacted with biennium effects (initial growth, initial GSP, and initial GSP per capita). Column 4 controls for initial sector shares interacted by biennium, and Column 5 adds demographic characteristics (share of urban, foreign, and population) measured in the pre-treatment period interacted with biennium fixed effects. Column 6 includes topic share controls. Column 7 includes all covariates and adds lagged government expenditures and the lagged dependent variable. Descriptions of the covariates with data sources are shown in Appendix Table A.2.Standard errors clustered by state. **p<.01; *p<.05; +p<.1.

3. Column 1 gives the baseline 2SLS estimate with state fixed effects and biennium fixed effects. It is positive and statistically significant, similar in magnitude to the OLS, meaning that at the margin an exogenous shift in legislative output due to nationwide text flows is associated with increased economic growth. The rest of the columns provide an array of robustness checks. Column 2's state-specific linear time trends do not change things. Nor do the set of pre-treatment controls, interacted with fully saturated time effects, added in Columns 3 through 5. The results are not sensitive to controls for current-period topic shares (Column 6). Finally, we can take everything together and add the lagged dependent variable (Column 7), still producing a positive and statistically significant coefficient, but with a slightly smaller magnitude when including endogenous controls.

These robust positive effects of legislative output on growth are economically meaningful in their magnitudes. The estimates suggest that a 10 percent increase in borrowed legislation – the approximate change triggered by a one-standard-deviation change in the residualized instrument[12] – would increase the per capita economic growth rate by 0.1 to 0.2 percentage points, relative to a mean of 3.1 percentage points. To give some intuition for this magnitude, recent work in empirical macro suggests that achieving a .15 percentage-point increase in the growth rate through a fiscal stimulus would require about a 0.1 percentage-point increase in net-of-tax government spending (e.g. Nakamura and Steinsson, 2014, Chodorow-Reich, 2019). Hence, a 10 percent increase in statute flows is equivalent in its growth effect to a 0.1 percent increase in government spending flows.

5.2 Robustness Checks on Main Results

This section enumerates a series of checks to assess robustness of the specification and evaluate alternative channels for the results. First, Appendix Table A.18 reports regression estimates for leads and lags of the growth effect of increased legislative output. As with the main regressions, the current-period effect is positive and significant. The placebo lead (effect of next biennium's legislating) is a precisely estimated zero (Columns 1-3). Meanwhile, the lagged effect is positive, suggesting an additional delayed effect in the subsequent biennium. The lagged effect is not statistically sig-

[12]More precisely, after residualizing out state and biennium fixed effects, a one-standard-deviation change in the instrument is 0.12. Multiplying that by the first stage coefficient 1.1 generates a predicted change of 0.13, or 13 percent increase in legislative output.

nificant in 2SLS (Columns 4-7) but significant at the 10 percent level in the reduced form (Column 8).[13] The zero estimate for the pre-shock placebo lead and a positive estimate for the post-shock lag provide some additional reassurance on the validity of the instrumental-variables design.

Next, we report a number of robustness checks in regard to the topics. In line with Borusyak et al. (2022), we show that results are robust to the inclusion of topic share controls, both in levels and in changes (see Appendix Table A.20). The results are not driven by any single topic (Appendix Figure A.16). Next, the results are not sensitive to the number of topics used in the construction of the instrument. Appendix Table A.21 shows results for 6, 12, 24, 30, 36, 42, and 48 topics. The main results hold regardless of the topic count for constructing the instrument.

To check further that our syntax-based measure of legislative provisions is capturing the legally and economically relevant component of legislative text, we run our analysis using alternative measures of legislative detail for the instrument and treatment. Appendix Table A.22 shows that when using number of words, rather than number of provisions, as the endogenous regressor (and for constructing the instrument), we obtain positive 2SLS estimates that are not statistically significant (Columns 1 and 2). This result supports our argument from above that our NLP method is needed to extract legally relevant information from the statute texts. In line with this idea, our main result is robust to including as a control the number of pages in the published statutes (Column 4).

To better understand the effect on growth, Appendix D.3 provides supporting results for the effect of legislative detail on some alternative outcomes. Appendix Table A.25 shows that the estimated coefficient is identical (yet less precise) when using growth in GDP as the outcome rather than growth in GDP per capita (Column 1), as there is no effect on population (Column 2). Further, we can rule out that the effects on growth are driven by changes in employment (Column 3) or the number of establishments (Column 6). Meanwhile, there are positive effects on other signifiers of economic expansion, including profits (Column 4) and wages (Column 5). Looking to other government activities besides legislation, Appendix Table A.27 shows there is no effect on total government expenditures, expenditures on legislative expenses, taxes, or party control (Democrat/Republican) of state government. That there is no effect on government spending or taxes suggests that the effect on growth is not driven by a fiscal shock,

[13] In particular, we note that in the specification with additional controls, the p-value for the lead is 0.86 (Column 3), while the p-value for the lag is 0.27 (Column 7).

where new legislation mechanically causes new spending. That there is no effect on legislative spending suggests that the growth effect is not driven by confounding effects on the legislative process, for example increased quality of policymaking procedure. The null effect on party control means that there does not appear to be intervening effects in the state political environment.

So far, our analysis has left out some potentially important additional sources of laws: bureaucratic regulations and the courts. Appendix D.4 provides a detailed analysis of the relevance of these alternative legal sources. We built auxiliary corpora of state regulations and state court cases to assess their relevance for our instrumental-variables analysis. We find that our instrument does not have a direct effect on these other legal sources, and that our main results hold when controlling for the volume of text from these other sources. Thus, we can rule out that our effects are driven by regulations or caselaw.

Finally, Appendix Table A.19 reports the baseline specification with alternative clustering of standard errors. The results are robust to not clustering (Columns 1 and 2) as well as two-way clustering by state and year (Columns 3 and 4). Following Adao et al. (2019), we apply k-means clustering on the pre-period topic share vectors to group states according to their initial topic shares. We then cluster standard errors on 12, 16, and 20 initial-topic groups, and results are still robustly significant (Columns 5 to 10).

6 Legal and Economic Mechanisms

This section provides additional evidence to unpack the legal and economic factors that are most relevant to the impact of laws on growth. First, Section 6.1 lays out our conceptual framework for analyzing laws and growth, founded in how more complete laws can increase economic activity through more relationship-specific investments. We then report evidence on the resulting predictions. These include heterogeneity by type of policy (Section 6.2), by contingent versus non-contingent clauses (Section 6.3), by pre-existing level of detail (Section 6.4), by relationship specificity across industrial sectors (Section 6.5), and by levels of economic policy uncertainty (Section 6.6).

6.1 Conceptual Framework

These supporting analyses are motivated by the idea that more complete legislation can increase location-specific and relationship-specific investments by reducing ex post

hold-up. The "hold-up model" has a long tradition in the economic literature, showing that contract detail is important for relationship-specific investments (Williamson, 1979, 1985, Grossman and Hart, 1986). The main idea is input suppliers need to make specific investments to customize the input for the needs of the final good producer. Hence they need more protection, namely more detailed and enforceable contracts. If the contracts are not well enforced ex post, because of the lack of details, there will be less investment ex ante (Klein et al., 1978, Hart and Moore, 1990, Nunn, 2007).

Applied to state legislatures, we start with the notion that state government creates legislation that businesses need to comply with. This legislation might require businesses to specialize their investments and structure their supply chains to this legal context, meaning those investments have less value if moved to other jurisdictions. If legislation is ambiguous or incomplete, then businesses face uncertainty about how the rules will be enforced by regulators or courts. If a business makes relationship-specific investments based on the incomplete legislation, but then regulators or courts fail to enforce them as expected, the business can be "held up" where its investments become less valuable or even worthless. Hence, in states with incomplete legislation, businesses will be more hesitant to make full, optimal investments in the first place. That under-investment limits economic growth, and provides a mechanism by which increased completeness in legislation can lead to more growth at the margin.

An implicit assumption in this hold-up interpretation is that the adopted laws are helpful to businesses on average. That is, the laws mostly make the economic environment more stable for commerce, rather than do harm to the economy due to mistake or rent-seeking. In the case of U.S. states, some notable institutional factors contribute to a beneficial effect of more lawmaking. There is competition between states to attract businesses, and there is social learning between states about reforms adopted in other states (Souza et al., 2019). In particular, our instrument is likely to be driven by efficiency-enhancing laws and regulations, as it is constructed based on laws that are borrowed from other jurisdictions. The publication of laws in other states reduces the writing costs of enacting those laws. If those laws are helpful, they can be adopted; if they are not a good fit, they can be ignored. Hence our instrument captures an expansion of the choice set; under minimal benevolence assumptions, the legislative detail triggered by the instrument is likely to increase completeness in the legislative social contract, and to help businesses on average. As mentioned above, previous empirical work on state-to-state policy diffusion suggests that successful policies are more likely to diffuse (Volden, 2006, Pacheco, 2012, Shipan and Volden, 2014, Butler et al., 2017,

Yu et al., 2020).

These ideas and empirical predictions are put on a more formal footing in Appendix E. First, we give more formal detail to a writing costs approach to legislating, based on Battigalli and Maggi (2002, 2008). Second, we present an alternative decision theory framework, which models the legislator's choice when to legislate. Both models undergird and complement the hold-up model based on relationship-specific investments and contract completeness.

Now we outline a number of additional testable predictions that arise from a hold-up model of legislative detail. These can then be taken to data in the subsequent sections.

Heterogeneity across legislative policy topics. If our results have to do with business investments, then a first expectation is that the clauses that matter most should be those on policies about regulating business. Policies that are less related to business should have less of an effect on growth.

Relative effect of contingent clauses. A key feature of complete contracts is the inclusion of contingencies, which condition actions and outcomes on the state of the world (Battigalli and Maggi, 2002). Contingencies do more than non-contingencies to split up the state space and leave less ambiguity for regulators and courts in the interpretation of laws. Contingencies are especially valuable in long-term relationships that are more likely to involve specific investments (Battigalli and Maggi, 2008). Hence, if those contingencies come through borrowing from other jurisdictions, they are even more likely to promote growth than non-contingencies. Non-contingent laws impose rigid requirements, or else give discretion to enforcers. Therefore non-contingent laws may even tend to inhibit business activity.

Concavity in existing legal detail. In any model of contract completeness (e.g. Battigalli and Maggi, 2002), one can rank the topics or clause types by their relative legal and economic importance. The contract designer will write the most important clauses first, and as one moves down the ranking, there is a decreasing marginal benefit of adding clauses. Hence, we expect heterogeneity by pre-existing detail in response to an exogenous increase in clauses. Starting at a relatively low level of detail, there should be a larger effect on specific investments and economic growth.

Relationship specificity of sector inputs. If our model is right, we would see an increase in relationship-specific investments between firms in response to an increase in legal detail. But relationship-specific investments cannot be observed directly. As a proxy, following Nunn (2007), we can assess their importance indirectly by looking for heterogeneous effects across sectors based on relationship specificity of the intermediate

inputs in that sector. We expect that effects of laws on growth are concentrated among the sectors relying more on relationship-specific investments.

Economic policy uncertainty. A key ingredient of the hold-up model is uncertainty about the future. When a rare event (not covered by the contract) becomes more likely due to increased uncertainty, then the expected costs of that event increase, and in turn the benefits of describing that event in the contract increase. Adapting that to our legislative context, we expect that an exogenous increase in legal detail would have a larger positive impact on growth when economic uncertainty is higher. More specifically, as discussed in Battigalli and Maggi (2008), we expect that under higher uncertainty more contingencies is more beneficial for growth, as the benefit to conditioning legal outcomes on the state increases.

6.2 Heterogeneity Across Legislative Policy Topics

Our instrument identifies an average effect that combines many factors across many different types of legislative texts. Here, we check what types of policies are pivotal for the effect. We expect the effect to be driven by policies related to specific business investments. In particular, we would expect the largest effect from policies that regulate economic activity (e.g. contracts, licensing, property rights), with less of an effect from other policies, such as those regulating social issues (e.g. family law, criminal justice).

As described in Section 3.2, we divide the LDA topics into the four more interpretable categories: economic regulation, social regulation, fiscal policy, and procedural. Thus we have four separate endogenous regressors W_{st}^l, representing the log number of provisions in state s at biennium t allocated to topics in policy category l. In turn, we produce separate shift-share instruments for each of the four categories. The calculation is the same as in subsection 4.2, except that rather than summing over all topics K, we sum over the subset of topics K_l within each respective policy category. We therefore get a separate instrument Z_{st}^l for each policy. We then estimate the baseline 2SLS system (Equations 3 and 1) separately for each of the four categories l, where the category-specific endogenous regressor W_{st}^l and instrument Z_{st}^l are appropriately slotted in.

The effects across policy categories are reported in Table 4. We can see, first, that there is a positive and significant ($p < 10\%$) effect of economic regulations (Column 1), and no effect at all of social regulations (Column 2). That is consistent with the investment hypothesis, where clearer rules about economic issues reduce hold up and

Table 4: What Policies are Driving the Effect of Lawmaking on Growth?

	(1)	(2)	(3)	(4)
	Effect on Real GDP Growth Per Capita			
Policy Category	*Economic Regulation*	*Social Regulation*	*Fiscal*	*Procedural*
Legislative Output	0.0125+	-0.0006	0.0220*	0.0009
	(0.00697)	(0.0097)	(0.0107)	(0.009)
First Stage F-stat	42.53	13.42	18.68	49.12
Observations	1,182	1,182	1,181	1,182
Time FE	X	X	X	X
State FE	X	X	X	X

Notes. Results for the 2SLS model (Second Stage 1 and First Stage 3), where the instruments and endogenous regressors are constructed separately by the four larger policy categories. Columns give the respective policy category. All specifications include time and state fixed effects. **p<.01; *p<.05; +p<.1.

lead to more economic activity, while clearer rules about social issues have less of an effect.

Further, we find that fiscal policy rules are impactful for growth (Column 3). That also makes sense from an investment view given that many place-based policies are implemented through taxes and public spending. Consistent with that view, recall that the effect of laws on growth is not driven by changes in government expenditures (Appendix Table A.27). That is, the fiscal-policy effect is driven not by a spending multiplier, but rather through legal changes in how money is collected or spent (e.g. targeted tax exemptions or subsidies). Finally, procedural rules (e.g. electoral administration) are not as important for economic growth (Column 4).

6.3 Relative Effect of Contingent Clauses

In the context of making an optimal set of rules or encouraging relationship-specific investments, contingencies are pivotal in moving legislation toward a more complete contract. In Battigalli and Maggi (2002), for example, it is optimal for the most important contract topics to have more contingent clauses (see also Battigalli and Maggi, 2008). A more complete contract helps reduce legal uncertainty, and reduction in legal uncertainty generates more stable relationships within and across firms, thereby allowing for better economic outcomes.

As described in Section 3.3, we produce separate counts for contingent provisions

(W_{st}^C) and non-contingent provisions (W_{st}^N). We estimate variants of the 2SLS system (3) and (1), but using the contingent and non-contingent measures of laws as joint endogenous regressors. The second stage is

$$\Delta \log Y_{st} = \alpha_s + \alpha_t + \alpha_s \cdot t + \rho_C \log W_{st}^C + \rho_N \log W_{st}^N + X_{st}'\beta + \varepsilon_{st} \tag{7}$$

where now we have two endogenous regressors, with the associated causal effects of interest for contingencies (ρ_c) and non-contingencies (ρ_N).

With two endogenous regressors, we need at least two instruments. To that end, we compute two variants of the shift-share instrument using the same formula (2), but where all provisions counts are replaced with contingent provision counts and non-contingent provision counts, respectively. Let Z_{st}^C give the contingency instrument and let Z_{st}^N give the non-contingency instrument. The first stage equations are

$$\log W_{st}^C = \alpha_s + \alpha_t + \alpha_s \cdot t + \psi_C Z_{st}^C + \psi_N Z_{st}^N + X_{st}'\beta + \eta_{st}^C \tag{8}$$

$$\log W_{st}^N = \alpha_s + \alpha_t + \alpha_s \cdot t + \psi_C Z_{st}^C + \psi_N Z_{st}^N + X_{st}'\beta + \eta_{st}^N \tag{9}$$

where all terms are as above. Appendix D.5 reports additional checks and results for these instruments.

In addition to the joint treatment, we estimate an alternative specification using as a single endogenous regressor the log difference between contingency and non-contingency, $\log W_{st}^C - \log W_{st}^N$. The second stage is

$$\Delta \log Y_{st} = \alpha_s + \alpha_t + \alpha_s \cdot t + \rho_{CN}(\log W_{st}^C - \log W_{st}^N) + X_{st}'\beta + \varepsilon_{st} \tag{10}$$

where the causal effect of interest is ρ_{CN}, giving the effect of contingencies relative to non-contingencies. We use both contingency instruments in the first stage:

$$(\log W_{st}^C - \log W_{st}^N) = \alpha_s + \alpha_t + \alpha_s \cdot t + \psi_C Z_{st}^C + \psi_N Z_{st}^N + X_{st}'\beta + \eta_{st} \tag{11}$$

which gave a higher first-stage F-statistic than computing a single differenced instrument. We will report first stage statistics for all specifications along with the 2SLS estimates.

The 2SLS regression estimates for contingency are reported in Table 5, with the different specifications analogous to those from Table 3. Columns 1 and 2 provide the estimates for the second stage (7) with two endogenous regressors (contingent and non-

Table 5: Effect of Contingent and Non-Contingent Clauses on Economic Growth

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Effect on Real GDP Growth Per Capita						
Contingent Provisions	0.0638**	0.0590**					
	(0.0226)	(0.0215)					
Non-Contingent Provisions	-0.0559*	-0.0511*					
	(0.0242)	(0.0228)					
Contingent -			0.0752**	0.0697**	0.0501*	0.0379*	0.0773**
Non-Contingent			(0.0242)	(0.0229)	(0.0219)	(0.0158)	(0.0219)
First Stage F-stat	22.27	36.82	22.83	36.60	15.13	31.68	23.86
Observations	1,182	1,182	1,182	1,182	1,182	1,182	1,134
Time FE	X	X	X	X	X	X	X
State FE	X	X	X	X	X	X	X
State Trends		X		X			
Econ Vars × Time					X		
Sector Shares × Time						X	
Demog Vars × Time							X

Notes. Results for the 2SLS model of contingencies. Column 1 and 2 show results for contingent and non-contingent clauses together (Second Stage 7 and First Stages 8 and 9), adding state specific trends in the second column. Columns 3-7 show the results for the difference between contingent and non-contingent clauses (Second Stage 10) and First Stages 11). Column 4 adds state specific trends, Column 5 adds pre-period economic variables interacted by year, Column 6 interacts initial sector shares by biennium, Column 7 initial demographic characteristics interacted by biennium. All specifications include controls for state and biennium fixed effects. **p<.01; *p<.05; +p<.1. Standard errors clustered by state.

contingent), instrumented by first stages (8) and (9). We can see in both columns that the 2SLS effect of contingent clauses is positive, while the 2SLS effect of non-contingent clauses is negative.

Next, Columns 3 through 7 show the estimates for the differenced (contingent minus non-contingent) second stage (10) with first stage (11). Consistent with the separate-treatments specification, there is a large positive effect of relative use of contingency. The effect is robust to including state trends or including pre-treatment characteristics interacted with time fixed effects.

The magnitude of the coefficients on contingency clauses are also notable. They are much larger than that for total provisions – three to four times as large. Overall, these results support the view that contingent clauses are most important for promoting investment and growth.[14]

Appendix D.5 reports a number of supporting results. Appendix Table A.35 reports additional specifications with the differenced treatment variable, showing that it is robust to inclusion of other variables. Appendix Table A.36 shows the effects for other intermediate economic outcomes. Appendix Table A.37 shows the results when using contingency and non-contingency counts by themselves as the endogenous regressor.

6.4 Concavity in Existing Legal Detail

Here we assess a potential concave relationship of legislative detail and economic growth. Since there would be decreasing marginal benefits in completing a legislative contract, the effect of adding laws should be larger in contexts with a relatively low pre-existing stock of laws. We take that intuitive idea to the data.

Historical records on the stock of legislation (the annotated code) are not available. Instead, we proxy for the stock using recent levels of the flows – in particular the number of provisions issued in the state over the last five bienniums (ten years). The idea is that, at any given point, the ranking of states by the historical flow of provisions can proxy for the ranking of states by the total stock of provisions.

Correspondingly, we rank the state-biennium observations by recent detail and then split the sample into three terciles by that ranking. We then estimate the baseline 2SLS system (Equations (3) and (1)), but subsetting by the three terciles. We also look at

[14]From Appendix Table A.10 and Appendix Figure A.9, we see that the log difference in contingencies and non-contingencies is actually slightly decreasing over time. So the overall aggregate predicted change in output due to changes in legislative volume over this period may be negative.

concavity in the effect of contingent clauses by estimating the 2SLS system for the effect of the difference in contingencies and non-contingencies (Equations (11) and (10)). We would expect a larger effect of new laws in the sample with lowest previous detail.

Appendix Table A.38 reports the estimates. Consistent with a concave relationship, we find that the effect of new laws on economic growth is stronger for states with low recent legal volume (Columns 1-3) compared to states with medium detail (Columns 4-5) or high detail (Columns 6-7). The effect for low-detail states is robust to state trends (Column 2), and also holds for the effect of contingencies (Column 3). Appendix Section D.6 provides additional specification checks for the concavity analysis. In particular, Appendix Table A.40 shows that we get similar results when the concavity thresholds are computed after residualizing on the state and year fixed effects.

6.5 Sectoral Relationship Specificity

Our framework takes a relatively broad view of specific investments in our empirical context, where for example an investment could be specific to a state subsidy or banking regulation. That said, firm-to-firm investments are key and do depend on the legal environment. While relationship-specific investments between firms cannot be measured directly, we can test for their importance indirectly by assessing heterogeneity across sectors that vary by relationship specificity. Specifically, we expect that the effect of additional clauses will be larger in those sectors where there are more goods with intermediate inputs that require relationship-specific investments.

For each industry, we have a proxy from Nunn (2007) on the proportion of intermediate inputs that are relationship-specific. That is measured as the proportion of inputs neither sold on global exchanges nor reference-priced. We calculate state-biennium GDP growth but limited to the sectors with high and low relationship-specificity, respectively. We then estimate the 2SLS regressions from above using the separate outcomes. We would expect a larger effect of new laws in the sectors with high relationship specificity.

Table 6 reports the results on heterogeneity by sectoral relationship specificity. First, Columns 1-3 show results for sectors with low relationship specificity – that is, sectors like fossil fuels and primary metals where inputs are purchased on global exchanges at publicly referenced prices. Exogenous increases in legal detail, overall (Column 1) or through contingencies (Columns 2 and 3), have no effect on output in those sectors.

Next, Columns 4-9 report the estimates for sectors with high relationship specificity – sectors like electronics and publishing where firms have special relationships with

Table 6: Hetereogeneous Effects by Relationship-Specific Investments

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Effect on Real GDP Growth by Sector Group								
Relationship Specificity	Low			High					
Legislative Output	-0.00854			0.0381+	0.0350+				
	(0.0238)			(0.0199)	(0.0193)				
Contingent Provisions		0.0857				0.135+	0.138*		
		(0.106)				(0.0736)	(0.0685)		
Non-Contingent Provisions		-0.105				-0.12	-0.124		
		(0.121)				(0.0822)	(0.0761)		
Contingent -			0.0556					0.159*	0.158*
Non-Contingent			(0.0926)					(0.0677)	(0.0630)
First Stage F-stat	22.83	18.2	19.26	22.83	21.74	18.2	34.4	19.26	33.42
Observations	1133	1133	1133	1133	1133	1133	1133	1133	1133
Time FE	X	X	X	X	X	X	X	X	X
State FE	X	X	X	X	X	X	X	X	X
State Trends					X		X		X

Notes. Results for the 2SLS model (Second Stage 1 and First Stage 3), where GDP growth (not per capita) is constructed from sectors with below-median relationship specificity (Columns 1-3), or from those with above-median relationship specificity (Columns 4-9). Relationship specificity scores are constructed using data from Nunn (2007) on the share of inputs that are neither traded on public exchanges nor reference-priced. Also includes estimates for the differenced contingency detail measure, as indicated. All specifications include state and biennium fixed effects and results for state specific trends are also reported (as indicated). **p<.01; *p<.05; +p<.1.

suppliers to provide customized inputs that are not sold at publicly referenced prices. We can see here that, in contrast to the low-specificity sectors, there is a positive and statistically significant effect of laws on growth. That holds for overall legislative output (Columns 4-5) as well as contingencies (Columns 6-9), and it is robust to inclusion of state trends. These results are consistent with relationship specific investments being an essential mechanism in the effect of laws on growth.

6.6 Economic Policy Uncertainty

A final supporting analysis is on the moderating role of uncertainty in the economic environment. When uncertainty increases, the benefits from a greater completeness of the law typically increase. Rare events become more frequent and therefore need to be covered by contingencies to avoid hold-up.

To measure such uncertainty, we adapt the validated measure of economic policy uncertainty (EPU), constructed and explored by Baker et al. (2016) in the context of the U.S. national economy. We use the state-level annual measure of local EPU described in Subsection 2, and rank the state-biennium observations by uncertainty. We then split the sample into three terciles based on the uncertainty ranking.

Table 7 reports 2SLS estimates for each tercile in EPU, looking at the baseline results with total provisions, as well as the contingency analysis using contingent and non-contingent clauses. Columns 1 and 2 include estimates for low uncertainty, Columns 3 and 4 with medium uncertainty, and Columns 5 through 10 with high uncertainty. The specifications are the same as those reported in Table 3 (baseline) and Table 5 (contingency).

First consider Columns 1 through 4, with low or medium uncertainty. These are all zeros, regardless of the specification. The coefficients are all relatively small in magnitude, and none are statistically significant. Note that the first stage is sometimes weak, however.

In contrast, consider Columns 5 through 10, focusing on the highest-uncertainty tercile. Columns 5 and 6 show a positive and significant effect of legislative output, about twice in magnitude to the full-sample estimate from Table 3. A similar magnified effect is seen for contingency in Columns 7 through 10. Contingent clauses have a relatively large positive effect on economic growth under high uncertainty. Meanwhile, the computed first-stage F-statistics are consistent with a sufficiently strong first stage for all of these regressions. Overall, these estimates provide support for the view that the

Table 7: Effect of Laws on Growth by the Level of Economic Policy Uncertainty

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Effect on Real GDP Growth Per Capita									
Economic Uncertainty	Low		Medium		High					
Legislative Output	0.00448		0.00699		0.0373*	0.0391*				
	(0.0111)		(0.0111)		(0.0153)	(0.0176)				
Contingent Provisions							0.145*	0.170*		
							(0.0560)	(0.0672)		
Non-Contingent Provisions							-0.137*	-0.163*		
							(0.0624)	(0.0775)		
Contingent -		0.0823		0.000182					0.164**	0.189**
Non-Contingent		(0.0692)		(0.0310)					(0.0465)	(0.0568)
First Stage F-stat	65.92	4.251	5.389	12.03	46.50	108.2	10.24	9.433	10.65	10.34
Observations	345	345	373	373	377	377	377	377	377	377
Time FE	X	X	X	X	X	X	X	X	X	X
State FE	X	X	X	X	X	X	X	X	X	X
State Trends						X		X		X

Notes. Results for the 2SLS model (Second Stage 1 and First Stage 3), splitting up the data by terciles in economic policy uncertainty (Baker et al., 2016). Also includes estimates for the differenced contingency detail measure, as indicated. Columns 1-2 show results for states with lowest tercile uncertainty. Columns 3-4 report results for those with median uncertainty while Columns 5-10 states with uncertainty in the higher tercile. All specifications include state and biennium fixed effects, while for High Uncertainty states, results controlling for state specific trends are also included (as indicated). **p<.01; *p<.05; +p<.1.

effects of law on growth are moderated by higher or lower economic policy uncertainty.

Appendix Section D.6 provides additional specification checks for the uncertainty analysis. In particular, Appendix Table A.39 shows that concavity and uncertainty recover independent dimensions in the dataset. In addition, Appendix Table A.41 shows similar results when the uncertainty variable is residualized on state and year fixed effects before the ranking and division into terciles.

Appendix Table A.42 shows that the uncertainty effect is robust to the inclusion of lagged growth per capita, suggesting that it is not driven just by the EPU measure picking up the business cycle. Also consistent with this point: Appendix Table A.43 shows that if we split up the sample based on recent growth (rather than recent detail or current economic policy uncertainty), we see effects of legislative output on growth in both the top and bottom tercile. Overall, these checks suggest the effect heterogeneity from high EPU is not driven by confounding business cycle trends.

7 Conclusion

This paper explores what makes legislative output matter for growth. In the empirical setting of the U.S. states for the years 1965 through 2012, we find that more legislation tends to boost the economy, although that average result conceals important heterogeneity. That heterogeneity is revealed by additional empirical analysis motivated by the mechanism that we consider most important for the results: making a more complete legislative contract, which reduces ex post hold up and increases ex ante relationship-specific investments. Consistent with this mechanism, we indeed find that the positive impact on growth is driven by economic rather than social regulations, is higher when the additional legislation is in the form of contingent clauses, is larger starting from lower legislative completeness, is strongest for sectors that rely more on relationship-specific inputs, and is concentrated in periods of greater economic policy uncertainty.

Methodologically, we build on the empirical literature in economics through novel use of legal text data in a causal framework. First, we introduce a new measure of legislative output from the text of state laws based on tools from computational linguistics. Second, we implement a text-based shift-share instrumental variables strategy that isolates exogenous variation in legislative output. These methods could be useful in other contexts with borrowing of texts between units, for example diffusion of

technologies into patent filings, or diffusion of source code between software projects, or diffusion of narratives on social media. Such explorations could use simulations and other methods to better understand the robustness of text-based instruments, and in particular their sensitivity to different pre-processing or featurization steps.

Substantively, it could be interesting to extend the approach to allow for spillover effects of laws on neighboring states (Souza et al., 2019, DellaVigna and Kim, 2022). The economic policies identified in our study could have both positive spillovers – for example through gains from trade – and negative spillovers – for example through displacement of labor and capital. Understanding these spillovers would give a fuller picture of the welfare consequences of legislative borrowing. For example, one could use county-level economic output data at state borders to help get at this question, potentially using data on job-to-job transfers and cross-state commuting.

The external validity of our empirical results is an open question, and it would be interesting and useful to seek similar evidence in other federal systems, such as Canada, Switzerland, or the European Union. The theoretical mechanisms that we have explored could apply more broadly, however, and could help guide future empirical work. In particular, external validity to other contexts would depend on different institutional frameworks. As shown in Gratton et al. (2021), for example, signaling incentives can have a strong effect on the quantity and quality of laws. In a system with strong signaling incentives and a large stock of legislation – e.g. Italy – a reduction in legal writing costs may have a very different impact from the case of U.S. state legislators, who have weaker signaling distortions and face competitive pressures tending toward efficiency. Other relevant factors include professionalism among state legislators, the quality of laws in other states, and specialized agencies to support legislative drafting (Bendor, 1995). Foarta and Morelli (2022) suggest a theory of complexity and reforms that reconciles some of the different empirical findings. We hope that a combination of more theory and data analysis could bring about a broader understanding of the applicability of these results across different institutional contexts.

References

Acemoglu, D. and Linn, J. (2004). Market size in innovation: theory and evidence from the pharmaceutical industry. *The Quarterly Journal of Economics*, 119(3):1049–1090.

Acemoglu, D., Naidu, S., Restrepo, P., and Robinson, J. A. (2019). Democracy does cause growth. *Journal of Political Economy*, 127(1):47–100.

Adao, R., Kolesar, M., and Morales, E. (2019). Shift-Share Designs: Theory and Inference*. Technical Report 4.

Ash, E., Jacobs, J., MacLeod, B., Naidu, S., and Stammbach, D. (2020). Unsupervised extraction of workplace rights and duties from collective bargaining agreements. *MLLD*.

Autor, D., Dorn, D., Hanson, G., and Majlesi, K. (2020). Importing political polarization? the electoral consequences of rising trade exposure. *American Economic Review*, 110(10):3139–83.

Autor, D. H., Dorn, D., and Hanson, G. H. (2013). The china syndrome: Local labor market effects of import competition in the united states. *American Economic Review*, 103(6):2121–68.

Autor, D. H., Dorn, D., and Hanson, G. H. (2016). The china shock: Learning from labor-market adjustment to large changes in trade. *Annual Review of Economics*, 8:205–240.

Autor, D. H., Kerr, W. R., and Kugler, A. D. (2007). Does employment protection reduce productivity? evidence from us states. *The Economic Journal*, 117(521):F189–F217.

Baker, S. R., Bloom, N., and Davis, S. J. (2016). Measuring economic policy uncertainty. *The quarterly journal of economics*, 131(4):1593–1636.

Bamieh, O., Coviello, D., Ichino, A., and Persico, N. (2023). Effect of business uncertainty on turnover.

Bartik, T. J. (1991). Who benefits from state and local economic development policies?

Bartik, T. J. (1994). The effects of metropolitan job growth on the size distribution of family income. *Journal of Regional Science*, 34(4):483–501.

Basso, G. and Peri, G. (2015). The association between immigration and labor market outcomes in the united states. *IZA DP No. 9436*.

Battigalli, P. and Maggi, G. (2002). Rigidity, discretion, and the costs of writing contracts. *American Economic Review*, 92(4):798–817.

Battigalli, P. and Maggi, G. (2008). Costly contracting in a long-term relationship. *The*

RAND Journal of Economics, 39(2):352–377.

Bendor, J. (1995). A model of muddling through. *American Political Science Review*, 89(4):819–840.

Blanchard, O. and Katz, L. (1992). Regional evolutions. *Brookings Papers on Economic Activity*, 23(1):1–76.

Blei, D. M., Ng, A. Y., and Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of machine Learning research*, 3(Jan):993–1022.

Bombardini, M. and Trebbi, F. (2020). Empirical models of lobbying. *Annual Review of Economics*, 12:391–413.

Borusyak, K., Hull, P., and Jaravel, X. (2022). Quasi-experimental shift-share research designs. Technical Report 1.

Botero, J. C., Djankov, S., Porta, R. L., Lopez-de Silanes, F., and Shleifer, A. (2004). The regulation of labor. *The Quarterly Journal of Economics*, 119(4):1339–1382.

Braunerhjelm, P. and Eklund, J. E. (2014). Taxes, tax administrative burdens and new firm formation. *Kyklos*, 67(1):1–11.

Burgess, M., Giraudy, E., Katz-Samuels, J., Walsh, J., Willis, D., Haynes, L., and Ghani, R. (2016). The legislative influence detector: Finding text reuse in state legislation. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 57–66.

Butler, D. M., Volden, C., Dynes, A. M., and Shor, B. (2017). Ideology, learning, and policy diffusion: Experimental evidence. *American Journal of Political Science*, 61(1):37–49.

Card, D. (2001). Immigrant inflows, native outflows, and the local labor market impacts of higher immigration. *Journal of Labor Economics*, 19(1):22–64.

Chodorow-Reich, G. (2019). Geographic cross-sectional fiscal spending multipliers: What have we learned? *American Economic Journal: Economic Policy*, 11(2):1–34.

Ciccone, A. and Papaioannou, E. (2007). Red tape and delayed entry. *Journal of the European Economic Association*, 5(2-3):444–458.

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

Colantone, I. and Stanig, P. (2018). Global competition and brexit. *American political science review*, 112(2):201–218.

Crawford, S. E. and Ostrom, E. (1995). A grammar of institutions. *American political science review*, 89(3):582–600.

Dam, K. (2007). *The Law-Growth Nexus: The Rule of Law and Economic Development*. Brookings Institution Press.

Davis, S. J. (2017). Regulatory complexity and policy uncertainty: headwinds of our own making. *Becker Friedman Institute for Research in economics working paper*, (2723980).

Dawson, J. W. and Seater, J. J. (2013). Federal regulation and aggregate economic growth. *Journal of Economic Growth*, 18(2):137–177.

DellaVigna, S. and Kim, W. (2022). Policy diffusion and polarization across us states. Technical report, National Bureau of Economic Research.

Dell'Orletta, F., Marchi, S., Montemagni, S., Plank, B., and Venturi, G. (2012). The splet-2012 shared task on dependency parsing of legal texts. *Proceedings of the 4th Workshop on Semantic Processing of Legal Texts*.

Di Vita, G. (2017). Institutional quality and the growth rates of the italian regions: The costs of regulatory complexity. *Papers in Regional Science*.

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

Foarta, D. and Morelli, M. (2022). The common determinants of legislative and regulatory complexity.

Fonseca, R., Lopez-Garcia, P., and Pissarides, C. A. (2001). Entrepreneurship, start-up costs and employment. *European Economic Review*, 45(4):692–705.

Frantz, C. K. and Siddiki, S. (2022). *Institutional Grammar: Foundations and Applications for Institutional Analysis*. Springer Nature.

Fukumoto, K. (2008). Legislative production in comparative perspective: Cross-sectional study of 42 countries and time-series analysis of the japan case. *Japanese Journal of Political Science*, 9(1):1–19.

Gennaioli, N. and Shleifer, A. (2007). The evolution of common law. *Journal of Political Economy*, 115(1):43–68.

Giommoni, T., Guiso, L., Michelacci, C., and Morelli, M. (2023). The economic cost of ambiguous laws.

Givati, Y. (2009). Resolving legal uncertainty: The unfulfilled promise of advance tax rulings. *Virginia Tax Review*, 29:137 at 169–175.

Goldsmith-Pinkham, P., Sorkin, I., and Swift, H. (2020). Bartik instruments: What, when, why, and how. *American Economic Review*, 110(8):2586–2624.

Gørgens, T., Paldam, M., Würtz, A., et al. (2004). *How does public regulation affect growth?* University of Aarhus.

Graetz, M. J. (2007). Tax reform unraveling. *The Journal of Economic Perspectives*, 21(1):69–90.

Gratton, G., Guiso, L., Michelacci, C., and Morelli, M. (2021). From weber to kafka: Political instability and the overproduction of laws. Technical report.

Greenstone, M., Mas, A., and Nguyen, H.-L. (2020). Do credit market shocks affect the real economy? quasi-experimental evidence from the great recession and "normal" economic times. Technical Report 1.

Grimmer, J. and Stewart, B. M. (2013). Text as data: The promise and pitfalls of automatic content analysis methods for political texts. *Political Analysis*, 21(3):267–297.

Grossman, G. M. and Helpman, E. (2001). *Special interest politics*. MIT press.

Grossman, S. J. and Hart, O. D. (1986). The costs and benefits of ownership: A theory of vertical and lateral integration. *Journal of political economy*, 94(4):691–719.

Hansen, E. R. and Jansa, J. M. (2021). Complexity, resources and text borrowing in state legislatures. *Journal of Public Policy*, 41(4):752–775.

Hansen, S., McMahon, M., and Prat, A. (2018). Transparency and deliberation within the fomc: a computational linguistics approach. *The Quarterly Journal of Economics*, 133(2):801–870.

Hart, O. and Moore, J. (1988). Incomplete contracts and renegotiation. *Econometrica: Journal of the Econometric Society*, pages 755–785.

Hart, O. and Moore, J. (1990). Property rights and the nature of the firm. *Journal of political economy*, 98(6):1119–1158.

Holtzblatt, J. and McCubbin, J. (2003). Whose child is it anyway? simplifying the definition of a child. *National Tax Journal*, pages 701–718.

Honnibal, M., Johnson, M., et al. (2015). An improved non-monotonic transition system for dependency parsing. *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing, Lisbon, Portugal, 17-21 September 2015*, pages 1373–1378.

Jacobzone, S., Steiner, F., Ponton, E. L., and Job, E. (2010). Assessing the impact of regulatory management systems: preliminary statistical and econometric estimates.

Jaeger, D. A., Ruist, J., and Stuhler, J. (2018). Shift-share instruments and the impact of immigration. Technical report, National Bureau of Economic Research.

Jalilian, H., Kirkpatrick, C., and Parker, D. (2007). The impact of regulation on economic growth in developing countries: A cross-country analysis. *World development*, 35(1):87–103.

Kawai, K., Lang, R., and Li, H. (2018). Political kludges. *American Economic Journal: Microeconomics*, 10(4):131–58.

Kirchner, S. (2012). Federal legislative activism in australia: a new approach to testing wagner's law. *Public Choice*, 153(3):375–392.

Klarner, C. (2013). State economic data. *IQSS Dataverse Network [Distributor]. Retrieved from*.

Klein, B., Crawford, R. G., and Alchian, A. A. (1978). Vertical integration, appropriable rents, and the competitive contracting process. *The journal of Law and Economics*, 21(2):297–326.

Lame, G. (2003). Using text analysis techniques to identify legal ontologies' components. *ICAIL 2003 Workshop on Legal Ontologies /& Web Based Legal Information Management*, pages 50–61.

Loayza, N. V., Oviedo, A. M., and Serven, L. (2005). *Regulation and macroeconomic performance*. World Bank Washington, DC.

MacLeod, W. B. and Nakavachara, V. (2007). Can wrongful discharge law enhance employment? *The Economic Journal*, 117(521):F218–F278.

Malhotra, N. (2006). Government growth and professionalism in us state legislatures. *Legislative Studies Quarterly*, 31(4):563–584.

McLaughlin, P. and Sherouse, O. (2017). Dodd-frank is one of the biggest regulatory events ever. Technical report, George Mason University, Mercatus Center.

McLaughlin, P. A. and Sherouse, O. (2016). The impact of federal regulation on the 50 states. *Available at SSRN 2939260*.

Montemagni, S. and Venturi, G. (2013). Natural language processing and legal knowledge extraction. *Summer School LEX2013 - 5th September 2013*.

Mulligan, C. B. and Shleifer, A. (2005). The extent of the market and the supply of regulation. *The Quarterly Journal of Economics*, 120(4):1445–1473.

Nakamura, E. and Steinsson, J. (2014). Fiscal stimulus in a monetary union: Evidence from us regions. *American Economic Review*, 104(3):753–792.

Nicoletti, G. and Scarpetta, S. (2003). Regulation, productivity and growth: Oecd evidence. *Economic policy*, 18(36):9–72.

Niskanen, W. A. (1971). *Bureaucracy and representative government*. Aldine, Chicago. (William Arthur) William A. Niskanen.

Nunn, N. (2007). Relationship-specificity, incomplete contracts, and the pattern of trade. *The quarterly journal of economics*, 122(2):569–600.

Olea, J. L. M. and Pflueger, C. (2013). A robust test for weak instruments. *Journal of*

Business & Economic Statistics, 31(3):358–369.

Pacheco, J. (2012). The social contagion model: Exploring the role of public opinion on the diffusion of antismoking legislation across the american states. *The Journal of Politics*, 74(1):187–202.

Parker, D. and Kirkpatrick, C. (2012). Measuring regulatory performance. *The economic impact of*.

Rauch, J. E. (1999). Networks versus markets in international trade. *Journal of international Economics*, 48(1):7–35.

Saias, J. and Quaresma, P. (2004). Using nlp techniques to create legal ontologies in a logic programming based web information retrieval system. In *Workshop on Legal Ontologies and Web based legal information management of the 9th International Conference on Artificial Intelligence and Law, Edinburgh, Scotland*.

Shipan, C. R. and Volden, C. (2014). When the smoke clears: expertise, learning and policy diffusion. *Journal of Public Policy*, 34(3):357–387.

Slemrod, J. (2005). The etiology of tax complexity: Evidence from us state income tax systems. *Public Finance Review*, 33(3):279–299.

Soria, C., Bartolini, R., Lenci, A., Montemagni, S., and Pirrelli, V. (2007). Automatic extraction of semantics in law documents. In *Proceedings of the V Legislative XML Workshop*, pages 253–266.

Souza, P. C., Rasul, I., and de Paula, Á. (2019). Identifying network ties from panel data: theory and an application to tax competition.

Ujhelyi, G. (2014). Civil service rules and policy choices: evidence from us state governments. *American Economic Journal: Economic Policy*, 6(2):338–380.

van Engers, T. M., van Gog, R., and Sayah, K. (2004). A case study on automated norm extraction. *Legal knowledge and information systems, Jurix*, pages 49–58.

Vannoni, M., Ash, E., and Morelli, M. (2019). Legislative information extraction: Methods and application to us state session laws. *working paper*.

Volden, C. (2006). States as policy laboratories: Emulating success in the children's health insurance program. *American journal of political science*, 50(2):294–312.

Weisbach, D. A. (2002). An economic analysis of anti tax avoidance doctrines. *American Law and Economics Review*, 4(1):88–115.

Williamson, O. E. (1979). Transaction-cost economics: the governance of contractual relations. *The journal of Law and Economics*, 22(2):233–261.

Williamson, O. E. (1985). Assessing contract. *The Journal of Law, Economics, and Organization*, 1(1):177–208.

Yu, J., Jennings, E. T., and Butler, J. (2020). Lobbying, learning and policy reinvention: An examination of the american states drunk driving laws. *Journal of Public Policy*, 40(2):259–279.

Table A.1: Summary Statistics

Variables	(1) Obs	(2) Mean	(3) SD	(4) Min	(5) Max
Economic Output Variables					
Log Real GSP per Capita	1,250	3.652	0.281	2.803	4.844
Log Real GSP	1,250	17.79	1.436	14.09	21.54
Log Real GSP per Capita Growth	1,249	0.031	0.050	-0.174	0.332
Log Employment Growth	823	0.057	0.064	-0.151	0.930
Log Number of Establishments Growth	823	0.045	0.058	-0.146	0.409
Log Establishment Profit Growth	550	0.163	0.109	-0.403	0.818
Statute Text Variables					
Log Provisions (Legislative Output)	1,183	9.211	0.887	2.996	11.42
Log Contingent Provisions	1,183	7.528	0.983	0.405	9.859
Log Non-Contingent Provisions	1,183	8.908	0.893	2.890	11.03
Covariates					
Log Population	1,250	14.94	1.029	12.51	17.47
Democratic Control	1,127	1.802	1.057	0	3
Log Income	1,250	3.479	0.267	2.563	4.144
Log Govt. Expenditure	1,250	15.57	1.471	11.89	19.46
Log Legis. Expenditure	1,250	9.410	1.384	5.176	12.73

Notes. Summary statistics for the main variables. The different number of observations is due to the availability of different years in the different datasets/sources we use.

Online Appendix

A Data Appendix

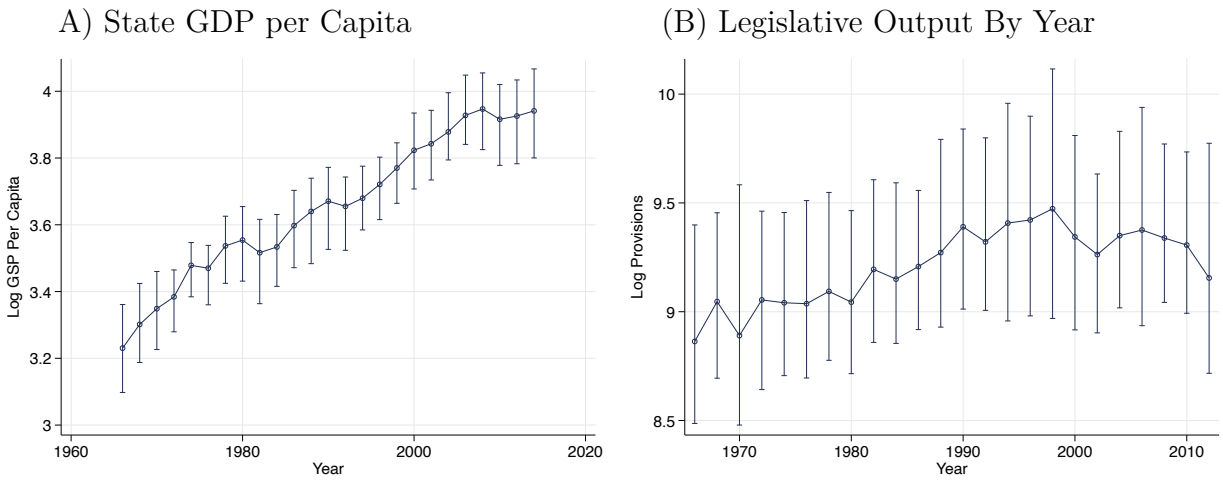
Table A.2: List of Variables with Source and Description

Variable	Data Source	Description
Economic Outcomes		
Log Real GSP	BEA Regional Accounts	Logged gross State Product deflated ti 2007 values using state CPI
Log Real GSP Growth	BEA Regional Accounts	Biennial growth in log real GSP
Log Real GSP Per Capita	BEA Regional Accounts	Log of GSP divided by population
Log Real GSP Per Capita Growth	BEA Regional Accounts	Biannial growth in Log Real GSP per capita
Growth Number Estabilishments	County Business Patterns	Growth of logged number of establishments
Log Establishment Profit Growth	County Business Patterns	Growth of logged establishment profits
Log Employment Growth	County Business Patterns	Growth of logged employment
Legislation		
Shock to Provisions	State session laws	Instrument for logged number of legal provisions in state statutes
Shock to Contingent Provisions	State session laws	Instrument for logged contingent legal provisions in state statutes
Shock Non-Contingent Provisions	State session laws	Instrument for logged non-contingent legal provisions in state statutes
Log Provisions	State session laws	Logged number of legal provisions in state statutes
Log State Statute Words	State session laws	Logged number of words in state statutes
Log Contingent Provisions	State session laws	Logged number of contingent legal provisions in state statutes
Log Non-Contingent Provisions	State session laws	Logged number of non-contingent legal provisions in state statutes
Log Share Amend Sentences	State session laws	Logged share of sentences that contain 'amend'
Log Share Repeal Sentences	State session laws	Logged share of sentences that contain 'repeal'
Log Statute Misspelling Rate	State session laws	Logged OCR error rate in in state statutes
Covariates		
Log Population	Klarner (2013)	Logged population
Urban Population	Ujhelyi (2014)	Logged urban population
Democratic Control	Klarner (2013)	Number of bodies under democratic control
Log Income	Klarner (2013)	Logged labour income
Log Expenses	Klarner (2013)	Logged government expenditures (in 1000s current dollars)
Log Legislative Expenses	Klarner (2013)	Logged legislative expenditures (in 1000s current dollars)
Log State News Uncertainty	newspapers.com	Logged number of articles mentioning the phrase 'economic uncertainty'
Log Real Tax Per Capita	Klarner (2013)	Logged per capita taxation deflated to CPI
Log Real General Exp p. Capita	Klarner (2013)	Logged per capita expenses deflated to CPI
Log Campaign Contributions	opensecrets.org	Logged dollars spent in campaign contributions
Campaign Finance Regulation	Book of the States	Openness of campaign finance contribution regulation
Regulation and Courts		
Log Federal Regulation	QuantGov	Federal Regulation and State Enterprise (FRASE) index
Log State Regulation Words	LexisNexis	Log word count in state regulations, from the Regulation Report database
Log State Court Opinion Words	LexisNexis	Log word count in state court opinions, from Court Opinions database

Table A.3: Extended Summary Statistics

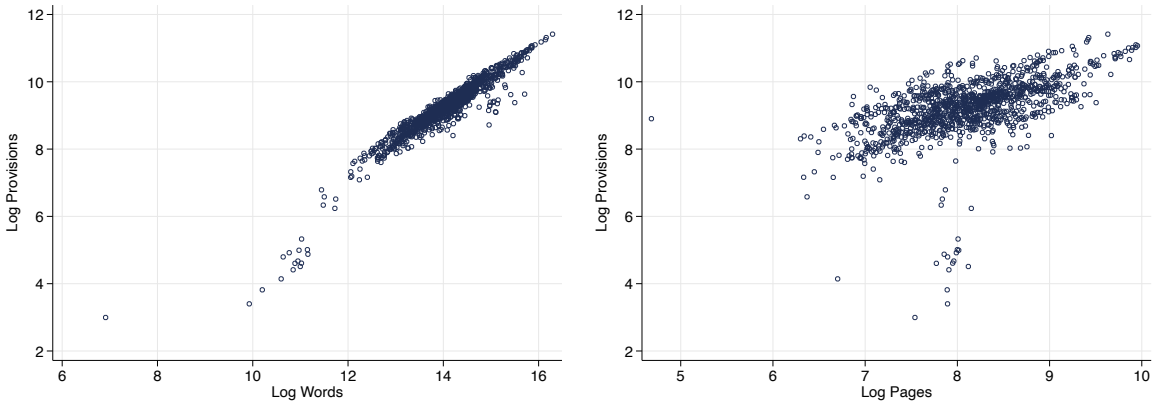
Variable	N	mean	std dev	min	max
Economic Outcomes					
Log Real GSP	1,250	17.79	1.436	14.09	21.54
Log Real GSP Growth	1,250	0.134	0.0701	-0.0870	0.665
Log Real GSP Per Capita	1,250	3.652	0.281	2.803	4.844
Log Real GSP Per Capita Growth	1,249	0.0314	0.0502	-0.174	0.332
Growth Number Estabilishments	823	0.0451	0.0584	-0.146	0.409
Log Establishment Profit Growth	550	0.163	0.109	-0.403	0.818
Log Employment Growth	823	0.0568	0.0643	-0.151	0.930
Legislation					
Shock to Provisions	1,183	0.0131	1.025	-2.191	2.563
Shock to Contingent Provisions	1,183	0.00371	1.023	-2.035	2.803
Shock Non-Contingent Provisions	1,183	0.0100	1.031	-2.218	2.629
Log Provisions	1,183	9.211	0.887	2.996	11.42
Log State Statute Words	1,183	14.03	0.833	6.912	16.29
Log Contingent Provisions	1,183	7.528	0.983	0.405	9.859
Log Non-Contingent Provisions	1,183	8.908	0.893	2.890	11.03
Log Share Amend Sentences	1,159	-3.619	0.576	-7.321	-2.098
Log Share Repeal Sentences	1,159	-5.496	0.897	-11.79	-2.240
Log Statute Misspelling Rate	1,183	0.0306	0.00699	0.0131	0.0649
Covariates					
Log Population	1,250	14.94	1.029	12.51	17.47
Urban Population	1,248	0.635	0.144	0.359	0.887
Democratic Control	1,127	1.802	1.057	0	3
Log Income	1,250	3.479	0.267	2.563	4.144
Log Expenses	1,250	15.57	1.471	11.89	19.46
Log Legislative Expenses	1,250	9.410	1.384	5.176	12.73
Log State News Uncertainty	1,208	-4.628	0.348	-6.044	-3.479
Log Real Tax Per Capita	1,250	0.721	0.370	-0.466	2.715
Log Real General Exp p. Capita	1,250	1.348	0.418	-0.0226	2.879
Log Campaign Contributions	659	16.32	1.674	9.179	20.26
Campaign Finance Regulation	1,199	0.990	0.712	0	2
Regulation and Courts					
Log Federal Regulation	450	13.75	0.184	13.24	14.43
Log State Regulation Words	444	14.55	1.757	0	16.84
Log State Court Opinion Words	1,183	14.77	1.049	12.335	18.267

Figure A.1: State-Level Economic Output and Legislative Output By Year



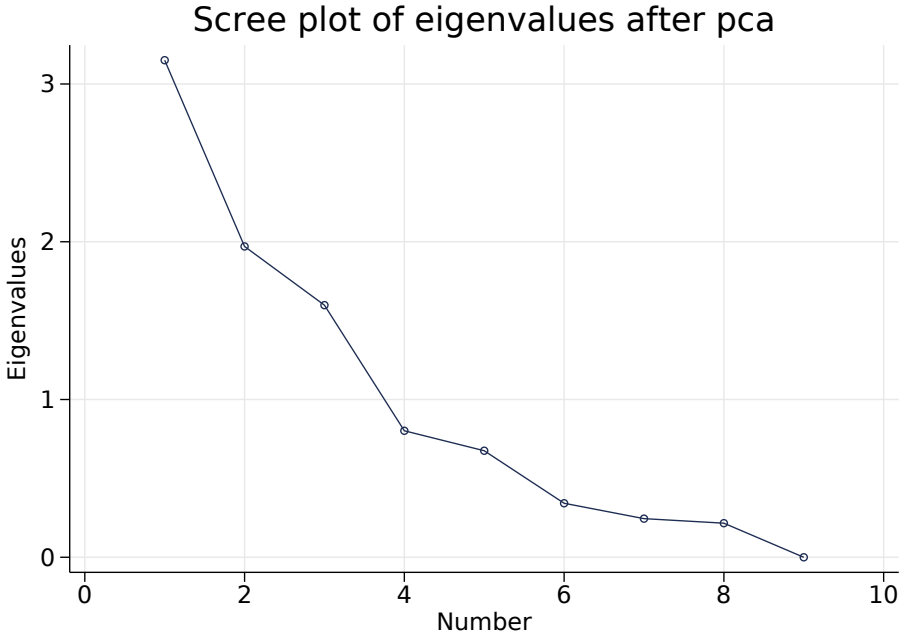
Notes. Line graphs showing the mean of (log) private GDP (Panel A) and (log) provisions (Panel B) across states over time. Error spikes give 90% confidence intervals from standard errors of the mean.

Figure A.2: Scatter Plots of Provisions vs. Word Counts and Page Counts



Notes. The figure shows a scatter plot for the relationship between (logged) words in the left panel and (logged) pages in the right panel and (logged) provisions, respectively.

Figure A.3: First Principal Components of Initial Economic Sectors



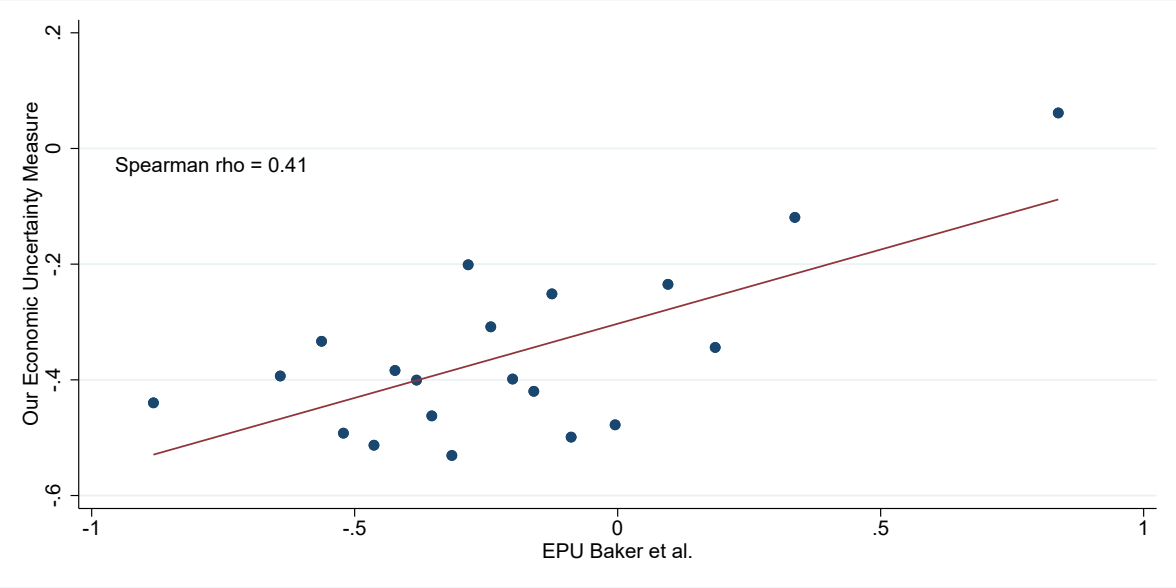
Notes. Screeplot for the principal components in economic sectors. We include the first four, or first six, components, interacted with year, as exogenous covariates in the regression analysis.

Table A.4: Relationship Specificity Scores, by Industry

NAICS Code	Industry Description	Relationship Specificity Score	GSP Share (%)
211	Oil & Gas Extraction	0.171	4.89
331	Primary Metal Manufacturing	0.201	2.05
324	Petroleum and Coal Products	0.246	2.89
112	Animal Production	0.271	5.12
311	Food Manufacturing	0.305	6.34
325	Chemical Manufacturing	0.322	2.13
313	Textile Mills	0.34	0.59
327	Nonmetallic Mineral Products	0.368	1.49
322	Paper Manufacturing	0.373	2.16
111	Crop Production	0.374	5.12
326	Plastics and Rubber Products	0.393	2.20
212	Mining (not Oil/Gas)	0.396	3.55
115	Agriculture/Forestry Support	0.405	1.07
314	Textile Product Mills	0.456	0.75
332	Fabricated Metal Manufacturing	0.457	2.10
113	Forestry and Logging	0.482	0.95
312	Beverage/Tobacco Manufacturing	0.509	6.34
114	Fishing & Hunting	0.521	2.12
337	Furniture Manufacturing	0.536	1.01
213	Mining Support	0.539	0.87
315	Apparel Manufacturing	0.561	0.52
335	Electrical Equipment and Appliances	0.567	1.71
339	Miscellaneous Manufacturing	0.585	3.14
323	Printing	0.614	1.37
321	Wood Product Manufacturing	0.639	1.19
316	Leather Manufacturing	0.66	0.64
333	Machinery Manufacturing	0.699	1.75
336	Transportation Equipment Manufacturing	0.787	3.98
334	Computers and Electronics	0.81	6.61
511	Publishing	0.83	7.54

Notes. List of industries matched from Nunn (2007) to our US BEA data on state-year sectoral output. Rows are 3-digit NAICS 1997 sectors. Sorted by third column – the relationship specificity score from Nunn, measuring the proportion of inputs not sold on an exchange, and not reference-priced. Fourth column gives the average GSP share (in percent) out of this set of sectors, across all states and years.

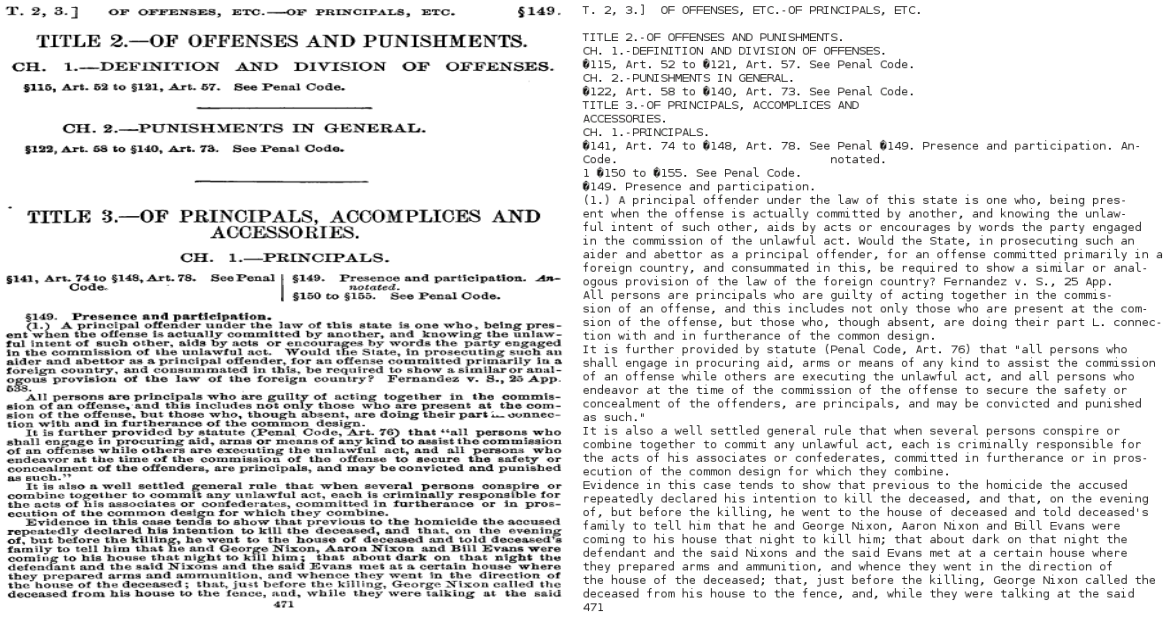
Figure A.4: Economic Policy Uncertainty Index



Notes. Binned scatterplot showing the cross-sectional relationship between our measure of economic policy uncertainty and the state-level one from policyuncertainty.com, available since the 1990s for most states. The regression model represented includes year FE. The Spearman rank correlation coefficient is 0.41 (rejecting the null hypothesis that the ranks of the two variables are independent).

B Details on Text Features

Figure A.5: State Session Laws Corpus



T. 2, 3.] OF OFFENSES, ETC.—OF PRINCIPALS, ETC. §149.

TITLE 2.—OF OFFENSES AND PUNISHMENTS.

CH. 1.—DEFINITION AND DIVISION OF OFFENSES.

§115, Art. 52 to §121, Art. 57. See Penal Code.

CH. 2.—PUNISHMENTS IN GENERAL.

§122, Art. 58 to §140, Art. 73. See Penal Code.

TITLE 3.—OF PRINCIPALS, ACCOMPLICES AND ACCESSORIES.

CH. 1.—PRINCIPALS.

§141, Art. 74 to §148, Art. 78. See Penal Code.

§149. Presence and participation. *Annotated.*

§150 to §155. See Penal Code.

§149. **Presence and participation.**

(1.) A principal offender under the law of this state is one who, being present when the offense is actually committed by another, and knowing the unlawful intent of such other, aids by acts or encourages by words the party engaged in the commission of the unlawful act. Would the State, in prosecuting such an aider and abettor as a principal offender, for an offense committed primarily in a foreign country, and consummated in this, be required to show a similar or analogous provision of the law of the foreign country? Fernandez v. S., 25 App. 658.

All persons are principals who are guilty of acting together in the commission of an offense, and this includes not only those who are present at the commission of the offense, but those who, though absent, are doing their part in connection with and in furtherance of the common design.

It is further provided by statute (Penal Code, Art. 76) that "all persons who shall engage in procuring aid, arms or means of any kind to assist the commission of an offense while others are executing the unlawful act, and all persons who endeavor at the time of the commission of the offense to secure the safety or concealment of the offenders, are principals, and may be convicted and punished as such."

It is also a well settled general rule that when several persons conspire or combine together to commit any unlawful act, each is criminally responsible for the acts of his associates or confederates, committed in furtherance or in prosecution of the common design for which they combine.

Evidence in this case tends to show that previous to the homicide the accused repeatedly declared his intention to kill the deceased, and that, on the evening of, but before the killing, he went to the house of deceased and told deceased's family to tell him that he and George Nixon, Aaron Nixon and Bill Evans were coming to his house that night to kill him; that about dark on that night the defendant and the said Nixons and the said Evans met at a certain house where they prepared arms and ammunition, and whence they went in the direction of the house of the deceased; that, just before the killing, George Nixon called the deceased from his house to the fence, and, while they were talking at the said

471

T. 2, 3.] OF OFFENSES, ETC.-OF PRINCIPALS, ETC.

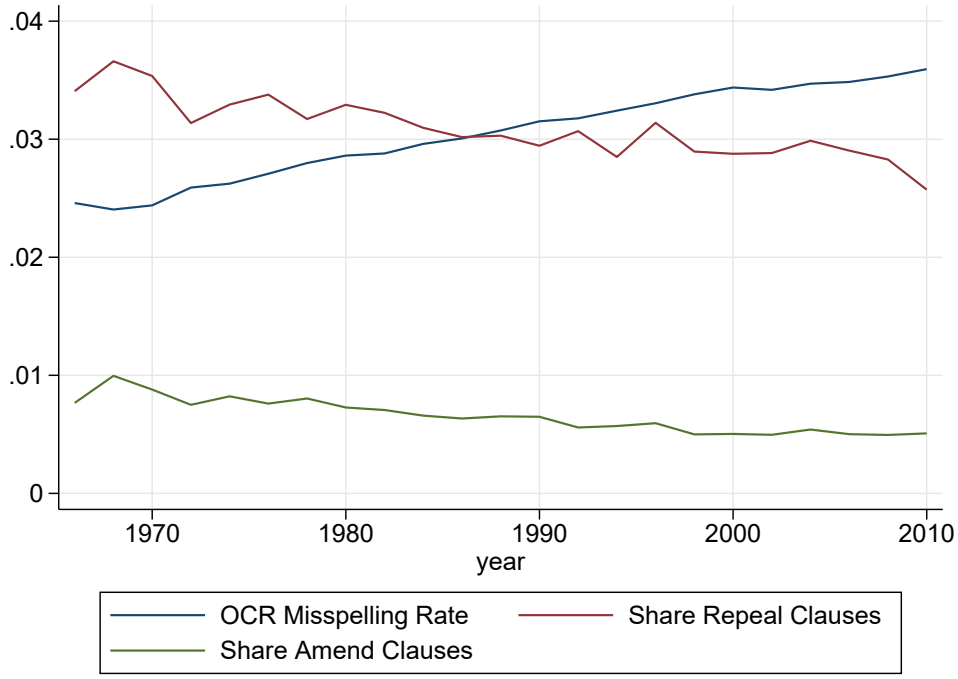
TITLE 2.-OF OFFENSES AND PUNISHMENTS.
CH. 1.-DEFINITION AND DIVISION OF OFFENSES.
§115, Art. 52 to §121, Art. 57. See Penal Code.
CH. 2.-PUNISHMENTS IN GENERAL.
§122, Art. 58 to §140, Art. 73. See Penal Code.
TITLE 3.-OF PRINCIPALS, ACCOMPLICES AND ACCESSORIES.
CH. 1.-PRINCIPALS.
§141, Art. 74 to §148, Art. 78. See Penal Code. §149. Presence and participation. Annotated.
§150 to §155. See Penal Code.
§149. Presence and participation.
(1.) A principal offender under the law of this state is one who, being present when the offense is actually committed by another, and knowing the unlawful intent of such other, aids by acts or encourages by words the party engaged in the commission of the unlawful act. Would the State, in prosecuting such an aider and abettor as a principal offender, for an offense committed primarily in a foreign country, and consummated in this, be required to show a similar or analogous provision of the law of the foreign country? Fernandez v. S., 25 App.
All persons are principals who are guilty of acting together in the commission of an offense, and this includes not only those who are present at the commission of the offense, but those who, though absent, are doing their part L. connection with and in furtherance of the common design.
It is further provided by statute (Penal Code, Art. 76) that "all persons who shall engage in procuring aid, arms or means of any kind to assist the commission of an offense while others are executing the unlawful act, and all persons who endeavor at the time of the commission of the offense to secure the safety or concealment of the offenders, are principals, and may be convicted and punished as such."
It is also a well settled general rule that when several persons conspire or combine together to commit any unlawful act, each is criminally responsible for the acts of his associates or confederates, committed in furtherance or in prosecution of the common design for which they combine.
Evidence in this case tends to show that previous to the homicide the accused repeatedly declared his intention to kill the deceased, and that, on the evening of, but before the killing, he went to the house of deceased and told deceased's family to tell him that he and George Nixon, Aaron Nixon and Bill Evans were coming to his house that night to kill him; that about dark on that night the defendant and the said Nixons and the said Evans met at a certain house where they prepared arms and ammunition, and whence they went in the direction of the house of the deceased; that, just before the killing, George Nixon called the deceased from his house to the fence, and, while they were talking at the said
471

Notes. Scanned image and associated OCR for example page from State Session Laws corpus.

B.1 Segmenting Statutes

The first step is to merge and process this raw text. A script serves to append pages, remove headers, footers, tables of contents, indexes, and other non-statute material. Then, it segments the text into individual bills, acts, and resolutions using text markers for the start of new statutes. These include indicators for new Chapters, Articles, or Titles, such as a line with CHAPTER followed by a Roman numeral. Some states have their own standard indicators, such as P.A followed by a number to indicate a new Public Act. The script also uses common text for the beginning of a statute preamble (e.g., An act to...) and for enacting clauses (e.g., Be it enacted that...). Research assistants checked samples of the statute segmenter for each state-year to make sure it worked well.

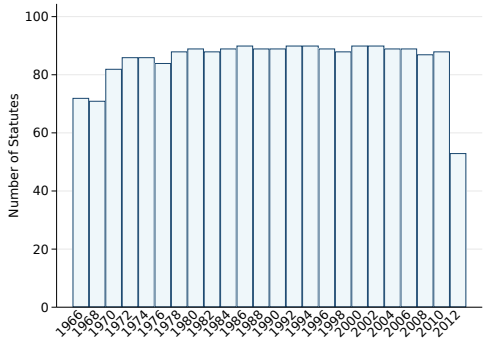
Figure A.6: Time Series: Amend Share, Repeal Share, and OCR Error Rate



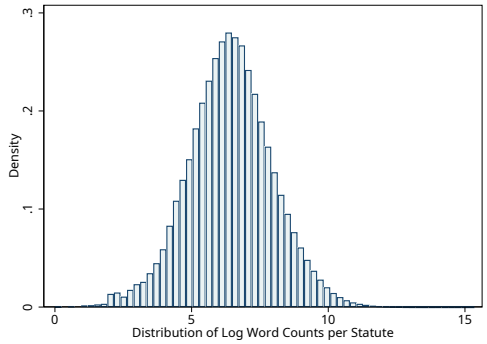
Notes. Time series for the share of amending sentences, share of repeal sentences, and the OCR misspelling rate in the state session laws corpus, over the time period of our analysis. OCR error rate computed as share of common nouns (identified with automated POS tagger) that are not in the open-source dictionary WordNet.

Table A.5: Summary Statistics on Statute Segmentation

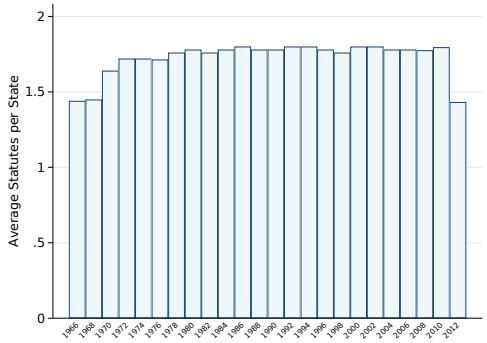
(A) Histogram: Number of Statutes by Biennium



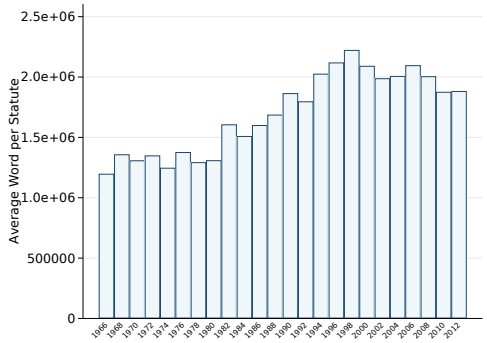
(B) Histogram: Number of Words per Statute



(C) Statutes per State, by Biennium

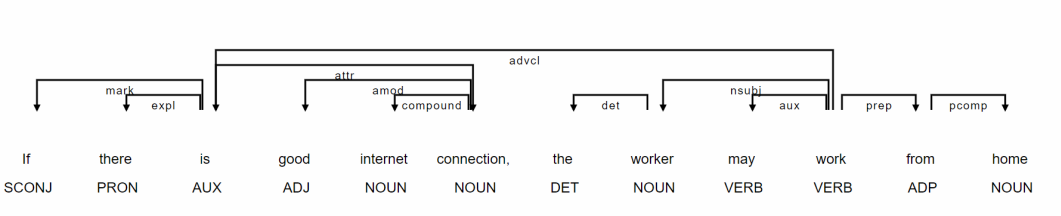


(D) Average Words per Statute, by Biennium



Notes. These figures provide some details on our corpus. The left panels shows the number of statutes by biennium (top) and state (bottom). The left panels show the number of words by statute (top) and the number of words by biennium (bottom).

Figure A.7: Syntactic Parsing for Provision Extraction



Notes. The Figure shows an example of a dependency tree. The letters below the words represent the part of speech (POS) tags. A prerequisite of syntactic dependency parsing, indeed, is POS tagging. The latter assigns labels ('tags') to the tokens in a sentence according to their function, such as noun, verb and adjectives. The arcs above the sentence represent the syntactic relations between words. First of all, the parser identifies the head of the sentence, namely the main verb, in this case 'work'. Then, the parser identifies the subject of the sentence and tells the researcher also that it is a nominal subject, in this case 'worker'. Indeed, in some cases, the subject may by a clause. The subject is then associated to a determiner, 'the'. Then, the parser looks at the other side of the sentence and, in this case, identifies a preposition, namely 'from', and the prepositional complement 'home'. It should be noticed that the verb of the contingent part of the sentence, 'is', is related to the main verb and hence the main sentence with the dependency adverbial clause. The latter is one of the most common syntactic relations that allow identifying a contingency.

B.2 Extracting Legal Provisions

Our information extraction approach relies on two stages: the definition of extraction rules and the syntactic parsing of the text. First, we decide the lexical and syntactic features of the provisions we want to extract. We focus on delegation, constraint, permission, and entitlement. Table A.6 shows the extraction rules, namely the lexical and syntactic rules we expect the main legal provisions above to follow. These are based on large-scale repositories of coded ontologies. These are dictionaries of words and dependencies that have been annotated to serve a theme, such as making a promise. An example of these ontology dictionaries is FrameNet.

Figure A.7 shows the result of the syntactic parser. The dependency parser tells us whether a noun is the subject or the object of the sentence. It tells us rich information about the verb -- whether it is the main verb or just an auxiliary, whether it is active or passive, and so on. These annotations provide the ingredients from which our extraction rules build measures of delegation. Our dependencies are produced using the Python package spaCy (Honnibal et al., 2015). The spaCy parser obtains state-of-the-art performance on the standard computational linguistics metrics. Like most parsers, it is trained on corpora of hand-parsed sentences. A detailed discussion of the process of information extraction can be found in Vannoni et al. (2019).

Table A.6: Types of Legal Provisions, with Extraction Rules

Lexical Units

Strict modals: 'shall','must','will'

Permissive modals: 'may','can'

Delegation verbs: 'require', 'expect', 'compel', 'oblige' , 'obligate', 'have to', 'ought to'

Constraint verbs: 'prohibit', 'forbid', 'ban', 'bar', 'restrict', 'proscribe'

Permission verbs: 'allow', 'permit', 'authorize'

Extraction Rules

Delegation: strict modal + active verb + not negation OR not permissive modal + delegation verb + not negation

Constraint: modal + not delegation verb + negation OR strict modal + constraint verb + not negation OR permission verb + negate

Permission: permission verb + not negation OR permissive modal + not special verb + not negation OR constraint verb + negation

Entitlement: entitlement verb + not negation OR strict modal + passive + not negation OR delegation verb + negation

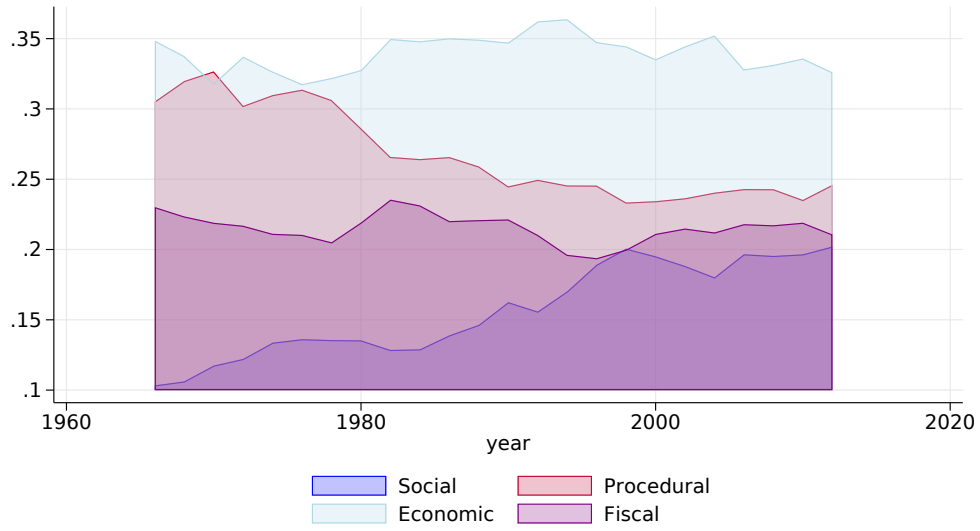
Notes. As enumerated in the table, a delegation is characterized by one of two structures: 1) a non-negated strict modal followed by an active verb ('The worker shall act'), or 2) a non-negated non-permissive modal (either a non-modal or a strict modal) followed by a delegation verb ("The worker is expect to act"). Constraints are characterized by 1) a negated modal ("The worker shall no"), a negated permission verb ("The worker is not allowed), or a non-negated constraint verb ("The worker shall be prohibited from"). Permissions are characterized by a 1) non-negated permission verb ("The worker is allowed to"), 2) a non-negated permissive modal followed by a non-special verb ("The may act"), or a 3) negated constraint verb ("The worker is not prohibited from"). Finally, entitlements are characterized by 1) a non-negated entitlement verb ("The worker retains the power to"), 2) a non-negated strict modal followed by a passive verb ("The worker shall be considered"), or 3) a negated delegation verb ("The worker is not obligated to"). By following these rules, we can see that the sentence in A.7is a permission: "The worker may work".

B.3 Details on Legislative Topics

Table A.7: List of Topics, 42-Topic Specification (with Broader Categories)

Label	Frequency	Category	Most Associated Words
Licensing	0.0318	economic	license fee holder valid such_license card renew proof good age such_person
Energy	0.0267	economic	director control solid gas site air oil coal environment underground tank mine
Partnerships	0.0267	economic	agent partnership foreign partner merger case transact mail demand stock
Payments	0.0258	economic	paid payment pay obligor child_support cost unpaid receipt withheld collect
Credit	0.0241	economic	interest transfer lien instrument issuer debtor seller holder buyer contract
Real Property	0.0227	economic	real loan trust mortgage interest broker common sale lender deed condominium
Traffic	0.0211	economic	motor dealer driver owner plate test vessel trailer weight special accident
Banks	0.0208	economic	institution bank stock deposit higher credit credit_union branch loan account
Insurance	0.0206	economic	life contract small premium carrier surplus risk condition benefit minimum pool
Contracts	0.0205	economic	contract work labor contractor cost repair perform bid job master firm trade
Land	0.0203	economic	land owner park parish port airport forest parcel lot map easement plat portion
Retail	0.0201	economic	sale sold retail sell price distributor fuel product milk liquor aircraft supplier
Torts	0.0201	economic	claim death claimant lieu_thereof loss settlement award case judgment legal
Traffic	0.0182	economic	highway traffic feet railroad state_highway transit load road space front stop
Commodities	0.0176	economic	fish food livestock plant game dog farm seed control sale grain wild owner deer
Land	0.0089	economic	street road feet island run tract river township center_line corner beach
Bonds	0.0336	fiscal	interest bond sale payment sold debt pay cost pledge paid sell interest_thereon
Taxes	0.0294	fiscal	tax gross credit return paid net assessor refund case such_tax homestead state_tax
Budgeting	0.0294	fiscal	budget for_the_fiscal_year so_much_thereof transfer special aid grant biennium
Funding	0.0276	fiscal	fund account money trust_fund transfer special excess deposit state_general_fund
Local Projects	0.0268	fiscal	development project local local_government compact zone urban government cost
Pensions	0.0267	fiscal	age benefit credit paid pension per_cent equal membership death elect final
Taxes	0.0263	fiscal	rate total equal paid calendar_year maximum strikeout subparagraph base excess
Tax Admin	0.0174	fiscal	paid sheriff auditor said_board warrant census audit supervisor cabinet travel
Miscellaneous	0.0202	misc	tile tie sueh lie said_code whieh shal ill supp aid thc tho tle tire aet sha
Courts	0.0390	procedural	court attorney judgment trial case district_court petition circuit_court circuit
Appeals Courts	0.0389	procedural	review appeal final complaint case petition civil receipt mail panel subpoena
Administration	0.0301	procedural	governor chief fire personnel bureau appoint shall_consist volunteer membership
Elections	0.0291	procedural	ballot petition voter township precinct register tenant cast elector referendum
Governance	0.0285	procedural	power invalid control proper event thereon hereof art shall_have_the_power
Policy Research	0.0278	procedural	center data review research staff local access develop implement level task
Elect Districts	0.0217	procedural	district special petition such_district said_district creation portion district_board
Local Govt	0.0207	procedural	council charter mayor special government conflict appoint perform oath organ
Governance	0.0162	procedural	government commonwealth civil attorney_general exempt uniform nonprofit
Local Issues	0.0120	procedural	local local_law new_matter superior such_law event fair race centum thirty-first
Education	0.0291	social	school school_district state_board student teacher pupil school_year tuition
Family Law	0.0275	social	child court parent minor children age guardian placement adult petition youth
Public Health	0.0254	social	health care home health_care social human children medicaid public_health
Healthcare	0.0242	social	treatment physician patient mental drug mental_health dental condition care
Criminal Law	0.0205	social	crime probation fine victim parole jail misdemeanor arrest sex firearm sexual
Water	0.0171	social	town water town_council sewer said_town lake river san town_clerk town_board
Social Issues	0.0087	social	sect team great stricken high_school veteran life honor nation first_paragraph

Figure A.8: Shares across Policy Categories over Time



Notes. This Figure shows the shares of topic groups (social, procedural, economic and fiscal) over time. We can see that economic clauses stay relatively stable over time, whereas social clauses increased drastically. Fiscal and procedural clauses, instead, slowly decreased over time.

Table A.7 shows the words associated with each topic for the 42-topic specification. We also include the assigned policy category for each topic: economic regulation, fiscal policy, procedural law, or social regulation.

B.4 Details on Contingency

An established literature on policy design and legal linguistics has emphasized the special relevance of contingencies in legal texts. In particular, the so-called "institutional grammar" has been used to study how legislation is written. This approach, which builds on the seminal paper by Crawford and Ostrom (1995), extracts relevant semantic features of the language in legislation. One of these semantic features are contingencies (what some of the literature calls "conditions"), which define the scope of application of the provision.

Take as an example the following sentence from U.S. state organic farming legislation (Frantz and Siddiki, 2022) as an example:

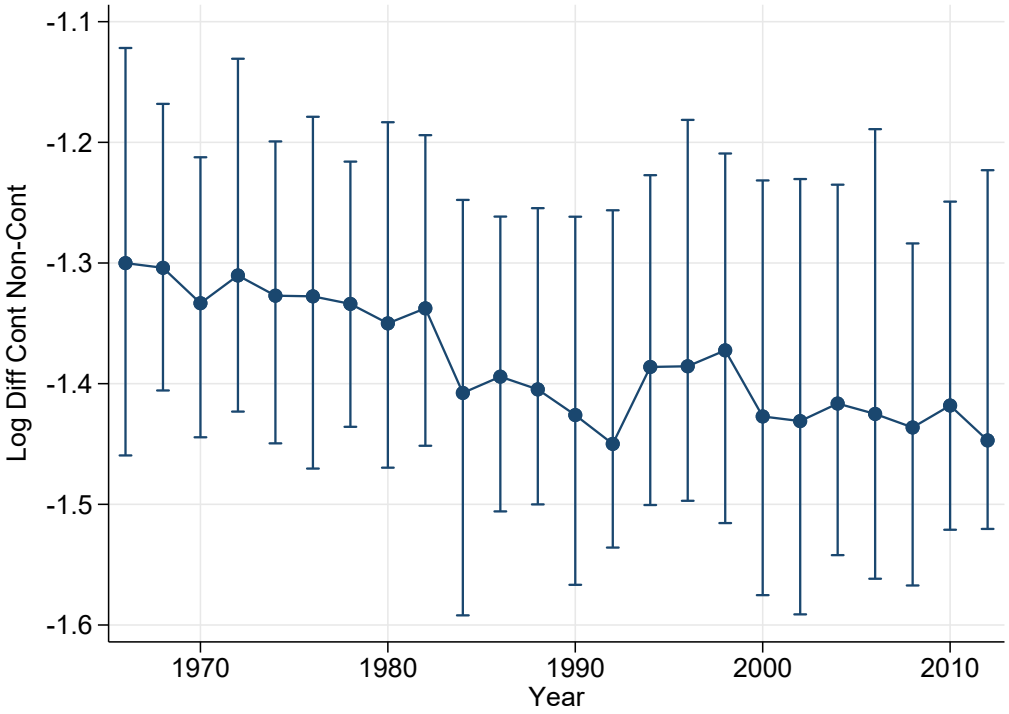
> "A certified operation or a person responsibly connected with an operation whose certification has been revoked will be ineligible to receive certification for a period of 5 years following the date of such revocation, except that Secretary may, when in the best interest of the certification program, reduce or eliminate the period of ineligibility."

In this sentence, there is a contingency that signals an exception: "except that Secretary may. . . ". More specifically, this is a so-called activation condition (Frantz and Siddiki, 2022), namely a condition under which the statement activates – e.g., "if X, then Y". Activation conditions specify the context of the provision, thus making legislation more precise. In the example above, the provision operates in a context with two scenarios: one where the Secretary does not do anything, and another where the Secretary takes action and changes the process of certification. The former would be the only scenario if the condition was not present. In conclusion, the presence of this activation condition (what we call a contingency) provides a more detailed, precise set of rules than a situation where there is no contingency.

As discussed in Section 2 on the corpus, our legislative text output measurement does not distinguish statutes that increase regulations or decrease them (deregulate). An example of a "deregulating" law is Texas Utilities Code Title 2.C Ch. 65, "Deregulation of certain incumbent local exchange company markets", enacted in 2005.[15] That statute has many contingent clauses, reflecting that more contingent clauses does not necessarily mean more intense regulation.

[15] See `https://statutes.capitol.texas.gov/Docs/UT/htm/UT.65.htm`.

Figure A.9: Net Contingencies by Biennium



This figure shows the difference in log contingencies and log non-contingencies over time in our dataset. Error spikes give the 25th and 75th percentiles.

Table A.8: Examples of Contingent Provisions

State	Year	Topic	Provision Text
UT	2009	0	(iv) if liability under the bond filed by the applicant with the division pursuant to Section 40-10-15 shall be for the duration of the underground mining operations and until the requirements of this Subsection (2) and Section 40-10-16 have been fully complied with.
MD	1992	1	Unless authorized by the Board, the consumer member of the Board may not participate in any activity related to examinations under this subtitle.]
TN	2005	2	The prescribing optometrist must sign the handwritten prescription order on the day it is issued unless it is a standing order issued in a hospital, a nursing home or an assisted care living facility as defined in SS68-11-201.
TX	1985	3	The transcription shall be in narrative form unless a party gives written objection to the use of narrative form not later than the fifth day after receiving notice of the request for a statement of facts.
OR	1985	4	Roadside vehicle emergency lighting must be lighted and placed upon the highway where they are clearly visible to the drivers of approaching vehicles for a distance of 500 feet and according to the following.(A)
KS	1987	5	If any provision or clause of this act or application thereof to any person or circumstances is held invalid, such invalidity shall not affect other provisions or applications of the act which can be given effect without the invalid provision or application, and to this end the provisions of this act are declared to be severable.
IL	1979	6	If the taxpayer's average monthly tax liability to the Department under this Act, the "Use Tax Act", the "Service Occupation Tax Act", the "Service Use Tax Act", the "Municipal Retailers' Occupation Tax Act", the "Municipal Service Occupation Tax Act", the "County Retailers' Occupation Tax Act" and the "County Service Occupation Tax Act" was $25,000 or more during the preceding 4 complete calendar quarters or was $10,000 or more if such 4 quarter period ended on or after Mp-ch 31, 1977, he shall file a return with the Depar t-m-1.
CA	2006	7	With respect to each foreign disappearing other business entity previously registered for the transaction of intrastate business in this state, the filing of the agreement of merger pursuant to subdivision (f) automatically has the effect of a cancellation of registration for that foreign other business entity as of the date of filing in this state or, if later, the effective date of the merger, without the necessity of the filing of a certificate of cancellation.
CA	1996	12	The court shall continue the case only if it finds that there is a substantial probability that the minor will be returned to the physical custody of his or her parent or guardian within six months or that reasonable services have not been provided to the parent or guardian.
SD	1994	8	If a draft is payable at a fixed period after sight and the acceptor fails to date the acceptance, the holder may complete the acceptance by supplying a date in good faith.(d)
IN	2010	9	If the electronic mail address or the fax number provided by the voter does not permit the county voter registration office to send the voter an application not later than the end of the first business day after the county voter registration office receives the communication, the county voter registration office shall send the application to the voter by United States mail.
FL	1976	10	(1) DEFINITION.--"Industry trade products" means all food products having any-nenda~ry-pweduet-wh~eh-as the semblance of milk or a milk product defined in this chapter but which does not come within the definition of milk, a milk product, ea filled milk, or filled milk product.(2) LABELING.--Industry trade products shall be labeled with a fanciful name or any other descriptive name that accurately describes the product, but in no case shall an "industry trade roduct" be labeled as an imitation of any product defined in this chapter.

Table A.9: Examples of Non-Contingent Provisions

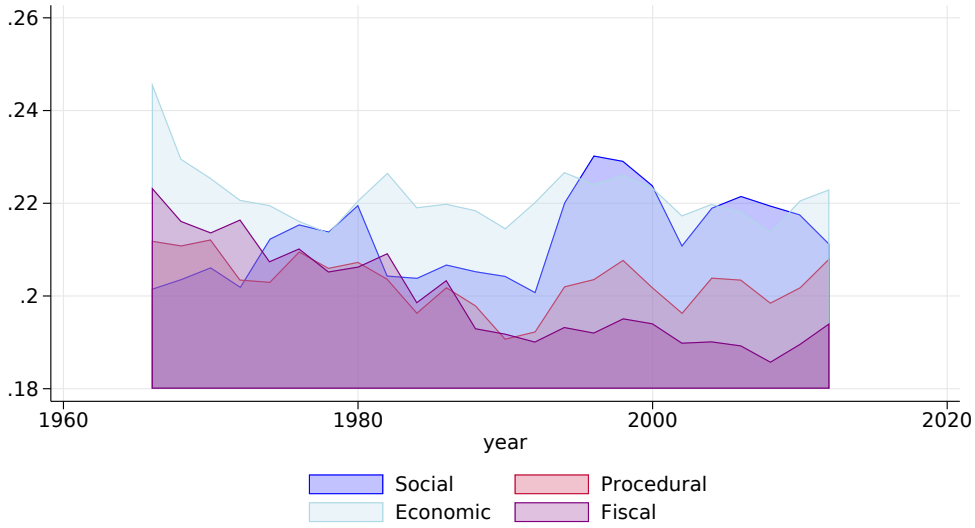
State	Year	Topic	Provision Text
UT	2009	0	these water impoundments will not result in the diminution of the quality or quantity of water utilized by adjacent or surrounding landowners for agricultural, industrial, recreational, or domestic uses.
MD	1992	1	A member may not serve more than [two] 2 consecutive full terms.
TN	2005	2	Nothing in this section shall be construed to prevent a physician assistant from issuing a verbal prescription order.
TX	1985	3	The clerk shall note the payment of the fee on the docket of the court.
OR	1985	4	A rear mounted lighting system shall have a green light, a yellow light and a red light.
IL	1979	6	Such determination shall be subject to review and revision by the Department in the manner hereinafter provided for the correction of returns.
CA	2006	7	The agreement of merger shall be approved on behalf of each other party by those persons authorized or required to approve the merger by the laws under which it is organized.
CA	1996	12	The court shall order that those services be initiated, continued, or terminated.(f)
IN	2010	9	A voter may not submit a registration application by fax or an electronic transmission except.
FL	1976	10	This act shall take effect October 1, 1976.
NY	1992	11	This act shall take effect on the same date as a chapter of the laws of 1992 amending the state law, relating to creating assembly and senate districts, as proposed in legislative bill number S. 7280 - A. 10111 takes effect.
NY	1969	13	Such notes may, among other things, be issued to provide funds t,.
CA	1990	14	The city council shall, within 10 days after the establishment of the district, invite bids for the making of the improvement by ordering a notice of the invitation to be published by two successive insertions in a daily or weekly newspaper published or circulated in the city and designated by the city council for that purpose.
IL	1953	15	Bonds shall be held at their book value.
VA	2002	16	The State Council shall report on the status of the Generalist Initiative to the House Appropriations and Senate Finance Committees at their regularly scheduled meetings in November.2.

Table A.10: Descriptive Statistics on Contingency, by Decade

	Mean	Standard Deviation
1960s		
Log Contingent	7.418	1.0050
Log Non-Contingent	8.779	.8894
Cont-Noncont Diff	-1.361	.274
Share of Contingent	.193	.0406
1970s		
Log Contingent	7.374	.7940
Log Non-Contingent	8.701	.7650
Cont-Noncont Diff	-1.326	.209
Share of Contingent	.1952	.0324
1980s		
Log Contingent	7.490	.8744
Log Non-Contingent	8.8693	.8136
Cont-Noncont Diff	-1.378	.210
Share of Contingent	.189	.0331
1990s		
Log Contingent	7.707	1.087
Log Non-Contingent	9.111	.9601
Cont-Noncont Diff	-1.404	.271
Share of Contingent	.1886	.0397
2000s		
Log Contingent	7.619	1.091
Log Non-Contingent	9.046	.9599
Cont-Noncont Diff	-1.427	.296
Share of Contingent	.18640	.0447

Notes. . **p<.01; *p<.05; +p<.1. This table shows the descriptive statistics for the logged number of contingent and non-contingent provisions by decade.

Figure A.10: Evolution of Contingent Language by Policy Category



Notes. This Figure shows the trends in the shares of contingent clauses by topic category (social, procedural, economic and fiscal) over time.

C Instrument Checks

Table A.11: Descriptive Statistics on Endogenous Regressor and Instrument, by Decade

	Mean	Standard Deviation
1960s		
Shock to Legislation	.447	1.342
Log Provisions	9.08	.8867
1970s		
Shock to Legislation	-.1276	1.094
Log Provisions	9.023	.7584
1980s		
Shock to Legislation	.1469	.875
Log Provisions	9.173	.8073
1990s		
Shock to Legislation	.127	.7515
Log Provisions	9.402	.9554
2000s		
Shock to Legislation	-.4175	.872
Log Provisions	9.334	.9568

Notes. . **p<.01; *p<.05; +p<.1. This table shows the descriptive statistics for the instrument and the endogenous regressor by decade.

Table A.12: First Stage Estimates are Stable Over Time

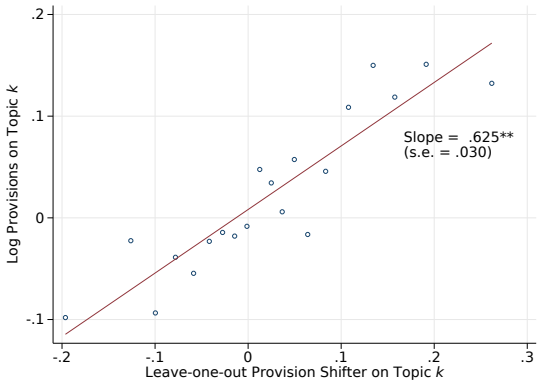
	(1)	(2)	(3)
	Effect on Provisions		
	1960s-1970s	1980s-1990s	2000s
Shock to Legislative Output	-1.153**	-1.292**	-1.591**
	(0.391)	(0.0814)	(0.165)
Observations	348	500	249
State FE	X	X	X
Time FE	X	X	X

Notes. . **p<.01; *p<.05; +p<.1. This table shows the first stage regressions separately by time period.

Figure A.11: Decomposing First Stage Effects of Shift and Share Terms

(A) First-Stage Effect of Shifter

$$\Delta W_{st}^k \sim \sum_{r \neq s} \frac{\Delta \log W_{rt}^k}{49}$$



(B) First-Stage Effect of Share

$$\Delta W_{st}^k \sim \frac{W_{s0}^k}{W_{s0}}$$

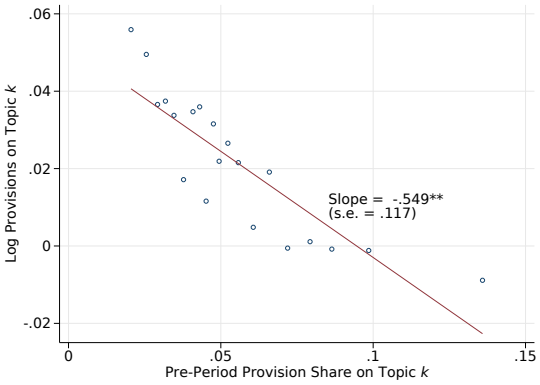
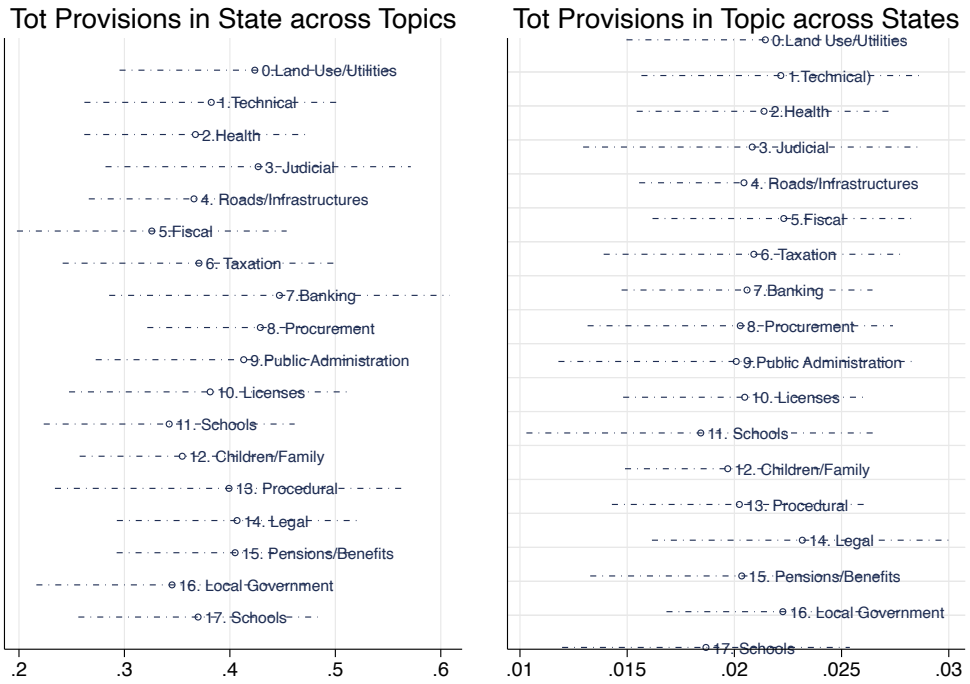
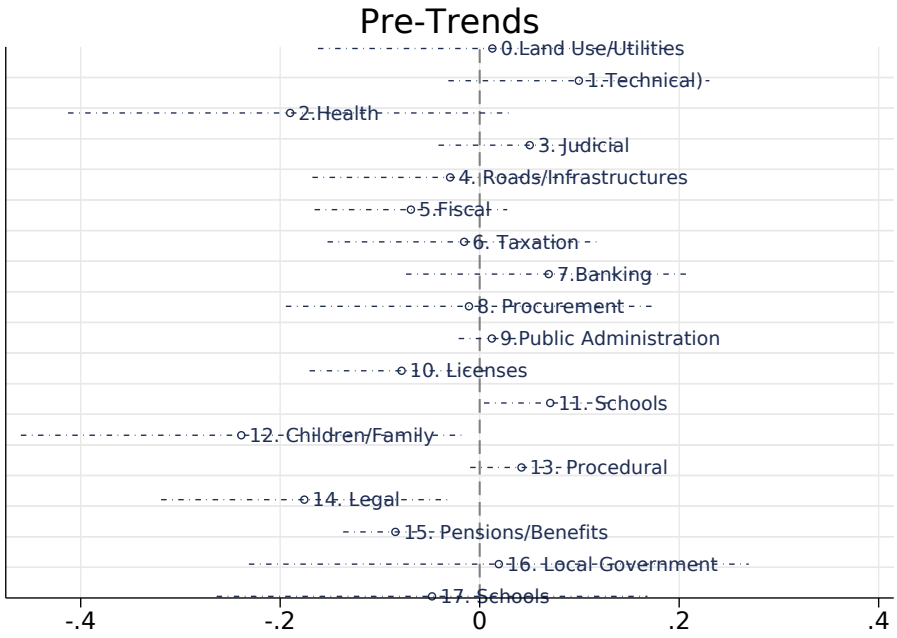


Figure A.12: All Topics Contribute to Instrument



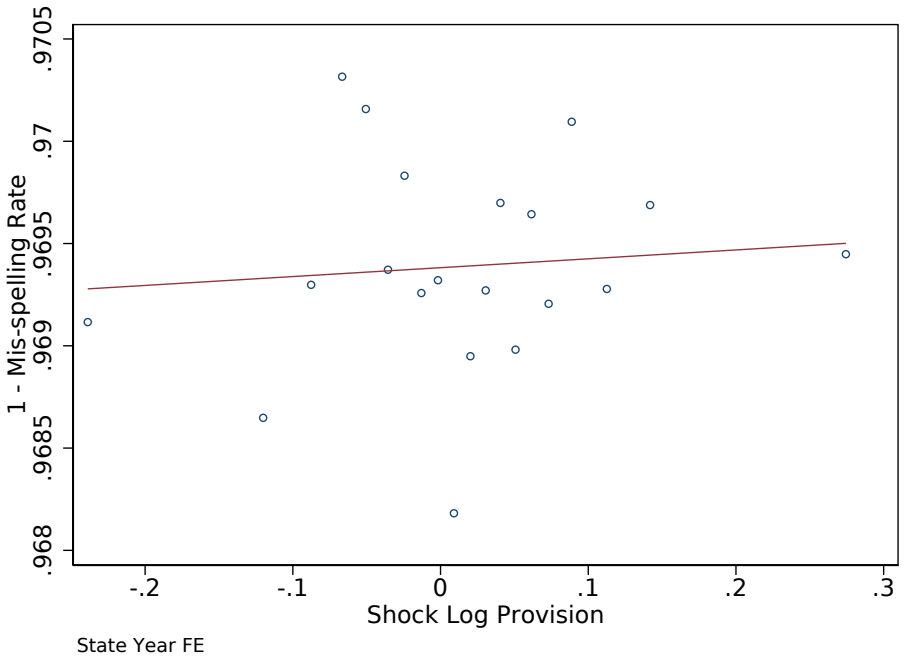
Notes. To check that the relevance of the shift-share instrument is driven by a majority of topics, we regress the increase in provisions related to a topic in a state on the increase in the total provisions related to that topic across states and the increase in the legal provisions in that state, for every topic (including state and year fixed effects and clustering standard errors by state).

Figure A.13: Pre-Treatment Topic Shares do not predict Growth Trends



Notes. We plot the coefficients that show that pre-treatment topic shares are not correlated with growth trends. All specifications include biennium fixed effects and standard errors clustered at the state level.

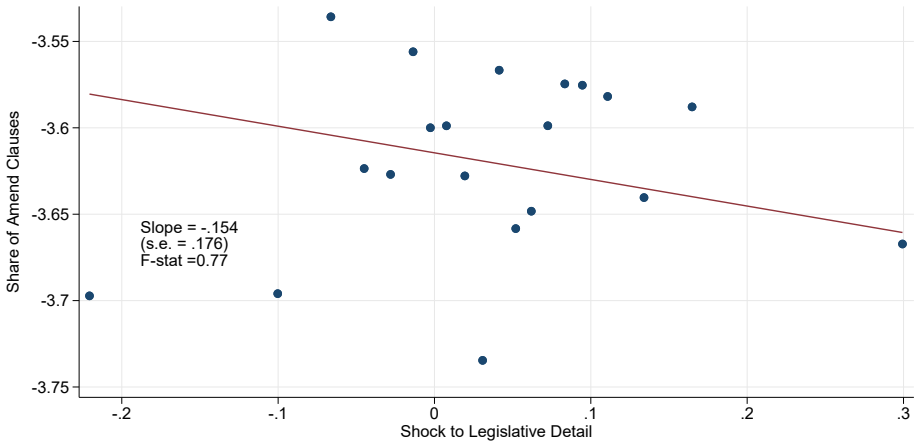
Figure A.14: Instrument is Uncorrelated with OCR Error Rate



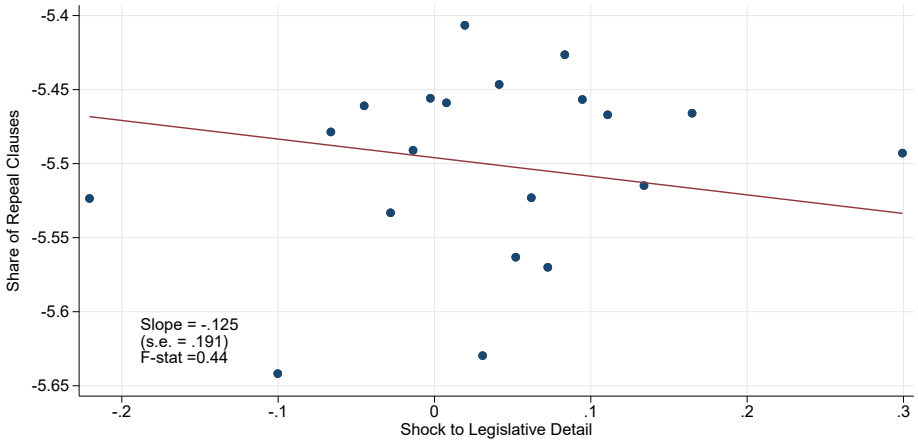
Notes. The Figure shows that the correct spelling rate (computed as the proportion of common nouns that are found in the WordNet dictionary) is not correlated with the instrument.

Figure A.15: Instrument is Uncorrelated with Share of Amending / Repealing Clauses

A. RF Effect of Instrument on Amending-Clause Share



B. RF Effect of Instrument on Repealing-Clause Share



Notes. Binned scatter plot of the relationship between legislative output shock instrument and the share of sentences containing amendment language (Panel A) and the share of sentences containing repealing language (Panel B). These are defined as the presence of the patterns "amend*" and "repeal*", respectively, where * indicates any word suffix.

Table A.13: Placebo Test: No Lead Effect of Legislative Output on Economic Growth

	(1) OLS	(2) RF	(3) 2SLS
Lead Log Provisions	0.00676	0.0174	-0.0135
	(0.00634)	(0.0164)	(0.0124)
Observations	1132	1132	1132
First Stage F-stat	.	.	15.14
State FE	X	X	X
Time FE	X	X	X
State Trends	X	X	X

Notes. Column 1 shows the OLS estimate with state and biennium fixed effect, and controlling for state specific trends, as well as standard errors clustered by state. Column 2 and 3 shows the same but the reduced form and 2SLS estimates. **p<.01; *p<.05; +p<.1.

Table A.14: Instrument Uncorrelated with Initial Characteristics

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Instrument	Instrument	Instrument	Instrument	PCA	PCA
Initial Share of Urban Pop	-0.021	-0.0302				0.688
	(0.0386)	(0.0547)				(3.277)
Initial Share of Foreign Pop	-0.0143	0.126				-12.04
	(0.219)	(0.361)				(18.01)
Initial Log Population	0.079	0.165				2.103
	(0.0753)	(0.123)				(4.902)
Initial Log Population^2	-0.00552	-0.0113				-0.159
	(0.00503)	(0.00824)				(0.316)
Initial Growth per Capita			0.00559	0.0205	0.483	
			(0.0251)	(0.0342)	(1.710)	
Sample (Years)	All	First Ten	All	First Ten	First	First
Observations	1135	526	1183	548	50	48
Time FE	X	X	X	X	X	X

Notes. Columns 1 and 3 show the results for the instrument balance test, using the whole sample. Column 2 and 4 show the results using only the first 10 years. Columns 5 and 6 show the results for the balance test for the first principal component of the pre-treatment topic shares. All specifications are with biennium fixed effects, as well as standard errors clustered by state. **p<.01; *p<.05; +p<.1.

Table A.15: Instrument Balance Checks for Potential Confounders

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Govt Exp	Lagged Govt Exp	Leg Exp	Lagged Leg Exp	Taxes	Lagged Taxes	Party	Lagged Party
Instrument (Z)	0.0523*	0.00789	0.0339	0.0333	0.0318	0.0655	0.0296	-0.0438
	(0.0257)	(0.0326)	(0.0734)	(0.0748)	(0.0535)	(0.0797)	(0.187)	(0.256)
Observations	1,183	1,133	1,183	1,133	1,183	1,133	1,123	1,110
Time FE	X	X	X	X	X	X	X	X

Notes. This table show the results for the balance test, regressing the instrument on the respective variables in each column current general government expenditure, lagged general government expenditure, current legislative expenditure, lagged legislative expenditure, current tax revenue, lagged tax revenue, current Democratic party control of state government, and lagged Democratic party control. Budget variables are in logs. All specifications are with biennium fixed effects, as well as standard errors clustered by state. **p<.01; *p<.05; +p<.1.

D Supporting Results

D.1 Robustness Checks on the Main Results

Table A.16: Effect of Legislative Output on Economic Growth (OLS)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Effect on Growth Rate Per capita						
Legislative Output	0.0146+	0.0152	0.0140*	0.0121*	0.00401+	0.0117+	0.00558*
	(0.00832)	(0.0123)	(0.00608)	(0.00512)	(0.00237)	(0.00680)	(0.00233)
Observations	1,182	1,182	1,182	1,182	1,134	1,074	1,086
R^2	0.431	0.446	0.561	0.746	0.628	0.473	0.862
Time FE	X	X	X	X	X	X	X
State FE	X	X	X	X	X	X	X
State Trends		X					X
Econ Vars × Time			X				X
Sector Shares × Time				X			X
Demog Vars × Time					X		X
Topic Shares						X	X

Table A.17: Effect of Legislative Output on Economic Growth (Reduced Form)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Effect on Growth Rate Per Capita						
Legislative Shock	-0.0200*	-0.0205*	-0.0169*	-0.0150*	-0.0132*	-0.0216*	-0.0118+
	(0.00883)	(0.00940)	(0.00670)	(0.00660)	(0.00626)	(0.00832)	(0.00627)
Observations	1,182	1,182	1,182	1,182	1,134	1,182	1,087
R^2	0.420	0.440	0.552	0.739	0.629	0.430	0.855
Time FE	X	X	X	X	X	X	X
State FE	X	X	X	X	X	X	X
State Trends		X					X
Econ Vars × Time			X				X
Sector Shares × Time				X			X
Demog Vars × Time					X		X
Topic Shares						X	X

Table A.18: Effect of Legislative Output on Economic Growth, Leads and Lags

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Effect on Real GDP Growth Per Capita (2SLS)							Reduced Form
Next Biennium	0.00664	0.00426	0.0009					
(Lead)	(0.00909)	(0.00838)	(0.00525)					
Legislative Output	0.0173*	0.0161+	0.00834+	0.0176*	0.0302	0.0160*	0.0103*	-0.0238*
	(0.00862)	(0.00841)	(0.00436)	(0.00787)	(0.0190)	(0.00720)	(0.00482)	(0.0114)
Last Biennium				0.00453	0.0146	0.00394	0.0033	-0.0123+
(Lag)				(0.00689)	(0.0134)	(0.00641)	(0.00296)	(0.00683)
Two Bienniums Ago					0.0128			
(2nd Lag)					(0.0128)			
First Stage F-stat	8.596	10.17	12.06	9.026	0.962	10.68	19.94	
Observations	1,130	1,130	1,038	1,179	1,176	1,179	1,085	1,179
State FE	X	X	X	X	X	X	X	X
Time FE	X	X	X	X	X	X	X	X
State Trends		X	X			X	X	
Econ Vars × Time			X				X	
Sector Shares × Time			X				X	
Demog Vars × Time			X				X	
Topic Shares			X				X	
Lagged Govt Expend			X				X	

Notes. Columns 1, 2, and 3 show the results with the placebo lead and the contemporaneous effect together. Columns 4, 6, and 7 include together the lag and the contemporaneous effects. Column 5 includes two lag effects and the contemporaneous one together. Column 8 is the reduced form, where the indicated endogenous regressors are replaced with the associated instruments. Specification include state and biennium fixed effects, state specific trends, and additional covariates, as indicated. All specifications have standard errors clustered by state. **p<.01; *p<.05; +p<.1.

Table A.19: Effect of Laws on Growth (2SLS): Alternative Clustering of Standard Errors

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
	Effect on Real GDP Growth Per Capita										
Clustering	*None (Robust SEs)*		*Two-Way: State & Year*		*Initial Topics (12)*		*Initial Topics (16)*		*Initial Topics (20)*		AKM
Legislative Output	0.0182*	0.0168*	0.0182*	0.0168+	0.0182+	0.0168+	0.0182+	0.0168+	0.0182*	0.0168+	0.014**
	(0.00872)	(0.00808)	(0.00854)	(0.00879)	(0.00986)	(0.00921)	(0.00873)	(0.00856)	(0.00835)	(0.00804)	(0.0011)
First Stage F-stat	.	.	.	.	19.06	19.30	20.24	19.81	19.52	19.46	.
Observations	1,182	1,182	1,182	1,182	1,182	1,182	1,182	1,182	1,182	1,182	1,182
Time FE	X	X	X	X	X	X	X	X	X	X	X
State FE	X	X	X	X	X	X	X	X	X	X	
State Trends		X		X		X		X		X	

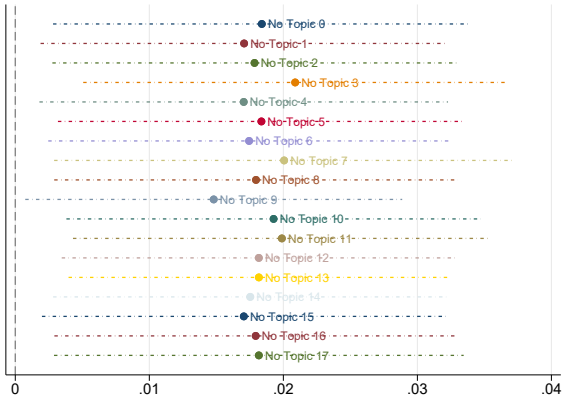
Notes. Columns 1 and 2 report the estimates for the effect of legislative output on growth per capita using robust standard errors. Columns 3 and 4 use standard errors clustered at the state and year level. Columns 5 to 10 use standard errors clustered at the initial topic level, with 12, 16 and 20 topics. All specifications include a first column with time and state fixed effects and a second column with the addition of state specific trends. Column 11 uses standard errors from the ivreg_ss command by Adao et al. (2019), using the default settings and initial share of topic 2 dropped due to collinearity and year fixed effects. For columns 1 to 4 the first stage F-stat is not generated because of the alternative clustering. **p<.01; *p<.05; +p<.1.

Table A.20: Effect of Legislative Output on Economic Growth - Topic Controls

	(1)	(2)	(3)	(4)
	Effect on Growth per Capita			
	2SLS	2SLS	2SLS	2SLS
Legislative Output	0.0182+	0.0168+	0.0182*	0.0168+
	(0.00905)	(0.00864)	(0.00903)	(0.00863)
Observations	1182	1182	1182	1182
First Stage F-stat	22.78	22.11	22.84	22.17
State FE	X	X	X	X
Time FE	X	X	X	X
State Trends		X		X
Frequent Topic Shares	X	X		
PCA			X	X
Growing Topic				

Notes. The table shows the results for baseline 2SLS estimate controlling for the share of the most frequent topics in columns 1 and 2, and for the first principal component in columns 3 and 4. All specifications have state and time fixed effect, and standard errors clustered by state. Column 2, 4 and 6 also control for state trends. **p<.01; *p<.05; +p<.1.

Figure A.16: 2SLS Results are not driven by any single topic



Notes: This is a coefficient plot showing the results of the main 2SLS model, with the instrument constructed by leaving one topic out at a time. State and time FE and clustered SE

Table A.21: Effect of Legislative Output on Economic Growth - Different Number of Topics

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Effect on Real GDP Growth Per Capita						
Topic Number	*6*	*12*	*24*	*30*	*36*	*42*	*48*
Legislative Output	0.0150+	0.0158+	0.0168	0.0146+	0.0142+	0.0139+	0.0132
	(0.00771)	(0.00823)	(0.0101)	(0.00834)	(0.00825)	(0.00760)	(0.00832)
Observations	1182	1182	1182	1182	1182	1182	1182
First Stage F-stat	18.43	18.87	28.63	34.23	36.75	39.2	35.77
State FE	X	X	X	X	X	X	X
Time FE	X	X	X	X	X	X	X

Notes. The table shows the results for baseline 2SLS estimate where the instrument is constructed using different number of topics. All specifications have state and time fixed effects, and standard errors clustered by state. **p<.01; *p<.05; +p<.1.

Table A.22: Effect of Laws on Growth – Adjusting for Words or Pages

	(1)	(2)	(3)	(4)
	Effect on Real GDP Growth Per Capita			
Legislative Output			0.0663	0.0203+
			(0.0632)	(0.0105)
Log Word Count	0.0154	0.0146	-0.0529	
	(0.00933)	(0.00901)	(0.0635)	
Log Page Count				-0.0103
				(0.00768)
First Stage F-stat	12.45	12.35	5.765	20.70
Observations	1,182	1,182	1,182	1,182
State FE	X	X	X	X
Time FE	X	X	X	X
State-Specific Trends		X	X	X

Notes. Columns 1 and 2 report the results for the effect of log of words on growth per capita. Column 3 and 4 report the effect of legislative output on growth per capita controlling for the log of the number of words and pages respectively. All specifications have state and time fixed effect, and standard errors clustered by state. Column 2 to 4 also controls for state trends. **p<.01; *p<.05; +p<.1.

Table A.23: Main Results, Controlling for Amend/Repeal Share

	Effect on Growth Rate Per capita							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Legislative Output	0.0185*	0.0174*			0.0184+	0.0176*		
	(0.00911)	(0.00845)			(0.00920)	(0.00844)		
Cont - Non-Cont			0.0663**	0.0645**			0.0630**	0.0613**
			(0.0216)	(0.0211)			(0.0202)	(0.0202)
First Stage F-stat	23.09	23.97	23.49	33.50	19.91	23.10	25.13	34.97
Observations	1,156	1,156	1,156	1,156	1,156	1,156	1,156	1,156
State FE	X	X	X	X	X	X	X	X
Time FE	X	X	X	X	X	X	X	X
Amend Share	X	X	X	X				
Repeal Share					X	X	X	X
State Trends		X		X		X		X

Notes. This table shows robustness specifications controlling for the current share of amending clauses (Columns 1-4) and share of repealing clauses (Columns 5-8). All specifications are with biennium fixed effects, as well as standard errors clustered by state. **p<.01; *p<.05; +p<.1.

D.2 Relevance of Campaign Finance and Interest Groups

Table A.24: Controlling for Campaign Finance Rules

	(1)	(2)	(3)	(4)
	Effect on Growth Rate Per Capita			
	2SLS	2SLS	2SLS	2SLS
Legislative Output	0.0258*	0.0245*		
	(0.0128)	(0.0104)		
Contingent - Non-Contingent			0.0479*	0.0616**
			(0.0230)	(0.0229)
Observations	870	870	870	870
First Stage F-stat	6.779	11.63	37.02	29.68
State FE	X	X	X	X
Time FE	X	X	X	X
Camp. Fin. Rules X Time FE	X	X	X	X
State-Specific Trends		X		X

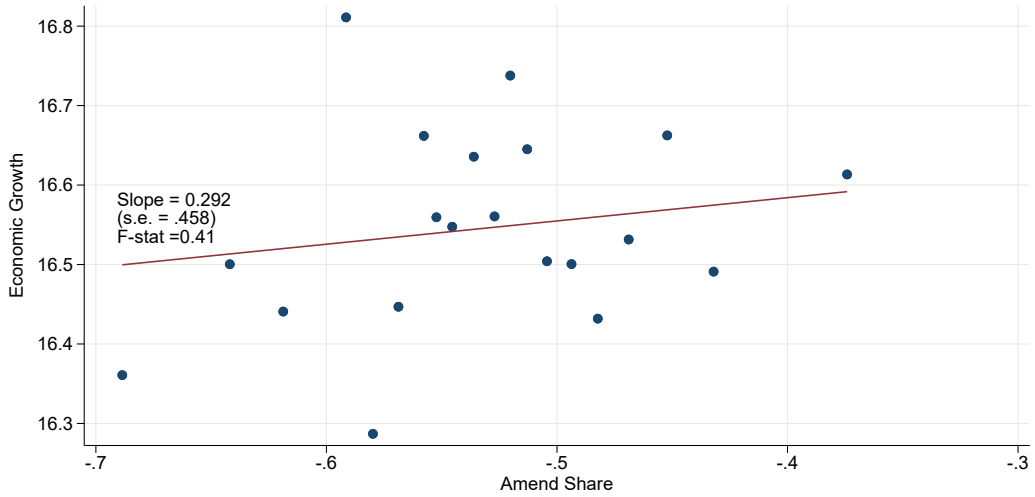
Notes. **p<.01; *p<.05; +p<.1. Regression results controlling for an index for limits on campaign contributions to state races, interacted with time. Smaller sample size is due to some missing years with campaign finance rules data.

Interest groups play a crucial role in proposing and implementing legislation (Bombardini and Trebbi, 2020). If the political system is open to more types of interest groups, then more more diverse proposals will get to legislators. It might also influence the implementation of legislation.

To check whether our results are driven by interest groups, we gathered data on the legislation regulating campaign finance contributions. We inspected the Book of States for the years 1952-2000 and coded whether contributions are restricted/prohibited for everyone, restricted/prohibited for some organizations (e.g., corporations or labor unions), or unlimited for everyone. We encode this as a categorical variable in our dataset.

Table A.24 shows the results for our baseline specification and for contingencies when controlling for the campaign finance index fixed effects, interacted with biennium fixed effects. This specification controls flexibly and allows our effects to be different over time depending on these rules. Results are robust and estimates are similar to

Figure A.17: Legislation Instrument Does Not Affect Campaign Contributions



Notes. Binscatter diagram of campaign donations to state politicians with the legislative-shock shift-share instrument, for the years 2000-2010 (years for which the campaign donations data is available).

those from the main models, suggesting that our results are not driven by lobbying efforts.

As an additional check, we collected information on campaign contributions to candidates for state government offices. This data is available on a set of web pages at `followthemoney.org`, for elections since the year 2000. We built a programmatic web scraper to collect all of these data and summed them by state and biennium. We then linked it to our main dataset for 2000-2010. Figure A.17 shows that our instrument has no linear effect on these contributions, suggesting that they are not an important mediator for the estimated effects on growth.

D.3 Effects of Legislative Output on Other Outcomes

Table A.25: Effect of Legislative Output on Additional Economic Variables

	(1) GDP (Total)	(2) Population	(3) Employment	(4) Profits	(5) Wages	(6) Establishments
Legislative Output	0.0199+	-0.00193	0.00481	0.0486+	0.0106+	-0.00877+
	(0.0102)	(0.00240)	(0.0119)	(0.0244)	(0.00536)	(0.00485)
First Stage F-stat	22.81	22.81	14.84	181.3	22.81	14.84
Observations	1183	1183	821	549	1183	821
State FE	X	X	X	X	X	X
Time FE	X	X	X	X	X	X

Notes. Results for the 2SLS model (Second Stage 1 and First Stage 3) but with different outcome variables. Column 1 explores the effect on state GDP (not per capita). Column 2 shows there is no effect on population. Column 3 uses employment while column 4 looks at firm profits (value added) within the state. Column 5 looks at wages and Column 6 establishment growth. All specifications include state and biennium fixed effects. Standard errors clustered by state. **p<.01; *p<.05; +p<.1.

Table A.26: Effect of Legislative Output on Additional Economic Variables II

	(1) Small Est	(2) Med Est	(3) Large Est	(4) Profit / Worker	(5) Large Est Ratio	(6) Large/Small Est Ratio
Legislative Output	-0.00465	0.00896	0.0172	0.0296	-3.99e-05	-8.69e-05
	(0.00486)	(0.00993)	(0.0315)	(0.0188)	(0.000136)	(0.000425)
First Stage F-stat	14.84	14.84	14.84	181.3	11.04	11.04
Observations	821	821	821	549	798	798
State FE	X	X	X	X	X	X
Time FE	X	X	X	X	X	X
State Trends						

Notes. This table reports 2SLS estimates on a range of additional outcomes, showing that new provisions do not not affect the average firm size or profit per worker. All specifications include state and biennium fixed effect in the first column with the addition of state trends in the second column. Standard errors clustered by state. **p<.01; *p<.05; +p<.1.

Table A.27: Legislative Output Shocks Do Not Affect Spending, Taxes, or Political Control

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Govt Spending		Legis Spending		Taxes		Dem Control	
Model	RF	2SLS	RF	2SLS	RF	2SLS	RF	2SLS
Legislative Output	0.0408	-0.0371	0.0339	-0.0309	0.0272	-0.0248	0.0296	-0.0268
	(0.0354)	(0.0326)	(0.0734)	(0.0620)	(0.0587)	(0.0524)	(0.187)	(0.172)
Observations	1183	1183	1183	1183	1183	1183	1123	1123
First Stage F-Stat	.	22.81	.	22.81	.	22.81	.	21.85
State FE	X	X	X	X	X	X	X	X
Time FE	X	X	X	X	X	X	X	X

Notes. RF and 2SLS effects on other outcomes. There is no effect on government spending, legislative spending, taxes, or political control. The first specification respectively includes state and biennium fixed effects, and the second adds state specific trends. All specifications have standard errors clustered by state. **p<.01; *p<.05; +p<.1.

D.4 Relevance of Regulations and Caselaw

The analysis in this paper has focused on the state session laws – the legislation or statutes enacted by state legislatures to be added to the statutory code. There are two major additional sources of rules for governing the state economy: state regulations and state caselaw.

First, there are the regulations that bureaucratic agencies enact to help implement statutes. These are often much more detailed than statutes. For example, Davis (2017) documents that in the case of the U.S. federal government, regulatory texts dwarf the legislative texts in volume and complexity. McLaughlin and Sherouse (2017) look at the particular case of the Dodd-Frank Act, which by itself resulted in tens of thousands of provisions to be added to the corpus of federal regulations. Federal regulations could be having an important economic impact at the state level, and the states themselves also issue regulations.

Second, the judiciary has an important role in economic governance. First, legislation and regulations have to be interpreted by judges for enforcement. When a regulator challenges a company action, or companies file suit against each other, the state courts are there to adjudicate and also to issue opinions clarifying legal rules. In a common law system like the United States, moreover, judges are often responsible for the rules themselves (e.g. Gennaioli and Shleifer, 2007). For example, Autor et al. (2007) shows that common-law rules on wrongful discharge had effects on employment and firm structure in U.S. states (see also MacLeod and Nakavachara, 2007).

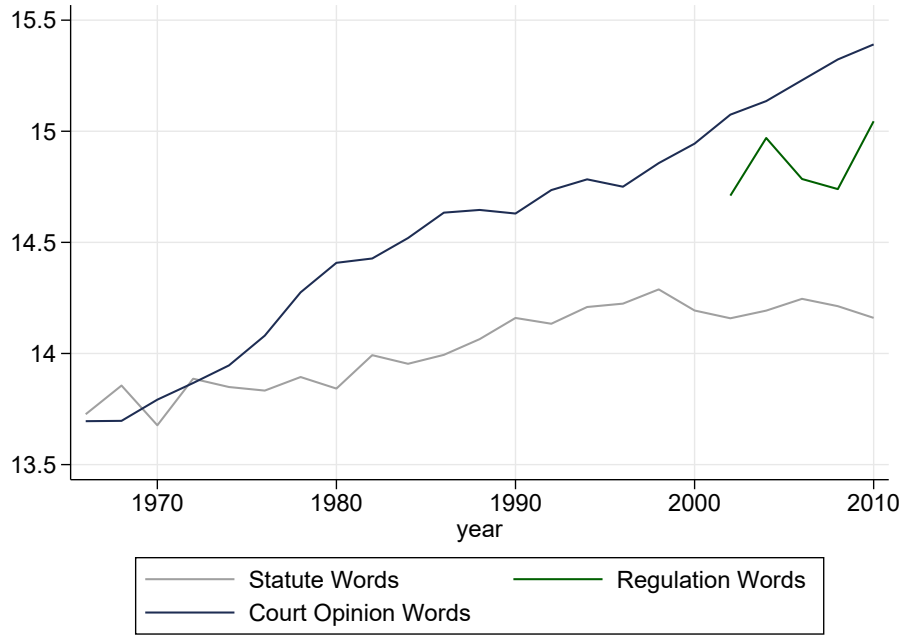
It could be that our main approach, focusing on legislation, leaves out some growth-relevant legal features in these other legal domains. In particular, it could be that our instrument affects growth not just through its effect on legislation, but also through its effect on regulations or caselaw. That would be a violation of the exclusion restriction and call into question the causal interpretation of the 2SLS regressions.[16]

To address these issues, we collected three additional measures of state-level legal output. We built two new corpora on state laws – a corpus of recent state regulations, and a corpus of state appellate court cases, both from Lexis Nexis. Third, we have a measure of the intensity of federal regulations across states from McLaughlin and Sherouse (2016).

State Regulations. We gained access to the proprietary State Net Regulatory

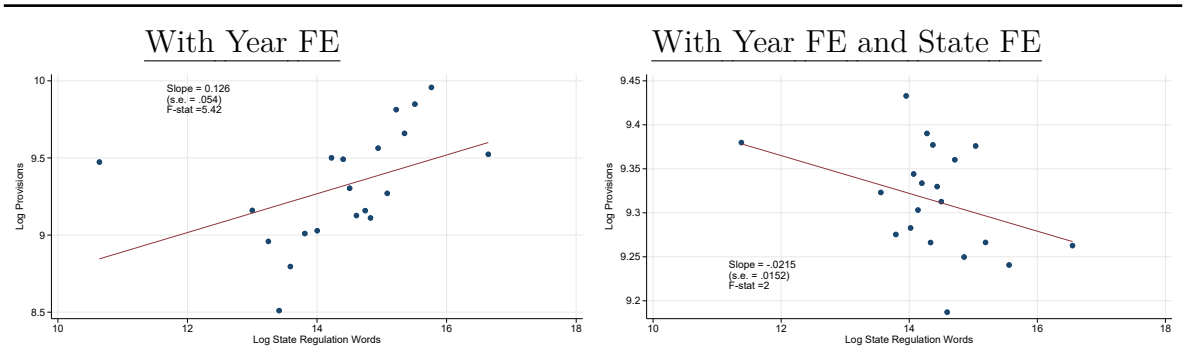
[16]The reduced-form regressions would still be causal, but difficult to interpret since it is a shift-share instrument whose units do not have a clear economic meaning.

Figure A.18: Comparison of Legal Volume in U.S. States: Statutes, Regulations, and Court Opinions



Notes. This figure shows the time series of the log word counts (average across states by biennium) for statutes (gray), regulations (green) and court opinions (black). See text for additional details on data sources.

Figure A.19: State Legislation and State Regulations



Notes. The relationship between logged number of provisions and logged state regulation words, controlling for year FE (left panel) and state and year FE (right panel). SE clustered at state level.

Text corpus from Lexis Nexis. This is a corpus of regulatory activity by state, available for some states starting in 1998 and most states starting in 1999, with complete coverage starting in 2002, up until 2017. It contains a rich collection of records corresponding to regulatory actions, for example new tax rules issued by state tax agencies. The database contains 20GB of XML files, which we processed to extract the regulatory text content. The processed corpus contains 642 thousand documents, adding up to 1.8 billion words. We computed the log word counts in regulations by state-biennium and merged them to our main dataset (through 2010). Figure A.18 plots the average log biennium word counts in regulations across states in the green time series. In the gray time series, we have for comparison the average log biennium word counts in statutes (the state session laws) across states. The regulation word count is about .7 to .9 higher in log scale, reflecting that the volume of regulations is about double the volume of statutes.

Figure A.19 shows binscatter plots relating regulatory output to legislative output. In the cross section (left panel), there is a positive and statistically significant effect. Intuitively, states with more complex legislation also have more complex regulations. When adding in state fixed effects and looking at within-state changes in legislative / regulatory flows, the slope flips sign and is no longer significant.

Table A.28 shows the reduced-form effect of our legislative shock on state regulations. Reassuringly, there is no effect. This means that regulatory detail is likely not a major mediator between our instrument and economic growth. Further, adding recent regulatory detail, averaged by state, and then interacted with time FE, does not change our results (Table A.29).

Table A.28: Reduced Form Effect of Legislation Instrument on State Regulatory Output

	(1)	(2)
	Effect on State Regs	
Shock to Legislation	-0.256	0.116
	(1.053)	(0.949)
Observations	329	329
State FE	X	X
Time FE	X	X
State-Specific Trends		X

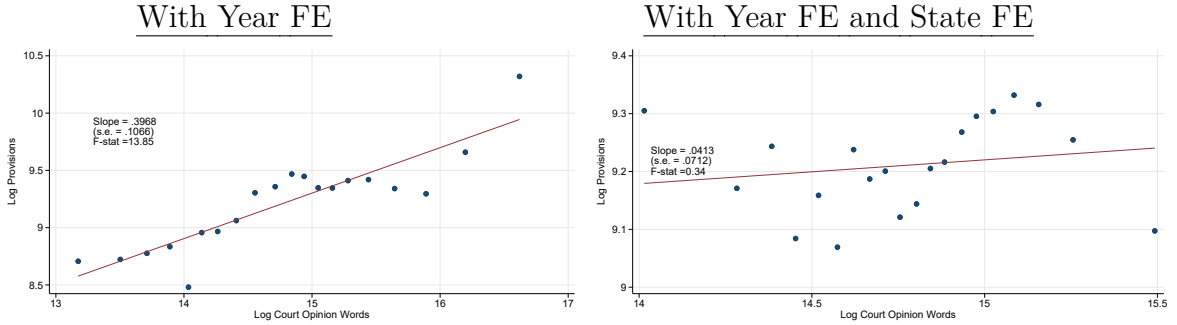
Notes. The table shows the results of regressing the logged state regulation word count on the instrument. **p<.01; *p<.05; +p<.1.

Table A.29: Controlling For State Regulation, Interacted with Time FE

	(1)	(2)
	Effect on Growth per Capita	
	2SLS	2SLS
Legislative Output	0.0190+	
	(0.0103)	
Contingent - Non-Contingent		0.0778**
		(0.0256)
Observations	1,182	1,182
First Stage F-stat	20.06	22.35
State FE	X	X
Time FE	X	X
State Reg x Time	X	X

Notes: Main 2SLS results controlling for a state's state regulation in 2002-2010, interacted with time.

Figure A.20: Log Provisions and Log Court Opinion Words



Notes. The relationship between logged number of provisions and logged state court opinion words, controlling for year FE (left panel) and state and year FE (right panel). SE clustered at state level.

Case Law. The second part of the legal process to explore is the judiciary. We gained access to the corpus of state appellate court opinions from LexisNexis. This corpus contains opinions from state intermediate and high courts from their inception, with the earliest opinion from 1658 and the latest in 2017. These cases are common-law decisions which can interpret statutes/regulations, apply precedents to new cases, or make new precedents. The raw dataset is 230GB of XML files, which we processed to extract the opinion texts. The processed corpus contains 9.7 million written opinions, adding up to 10.5 billion words. We computed the log word counts in cases by state-biennium and merged them to our main dataset. Figure A.18 plots the average log biennium word counts in court opinions across states in the black time series. We can see that at the beginning of the sample, the volume of text was similar between statutes (gray series) and caselaw. But over the last decades, the volume of caselaw has increased more rapidly, such that in recent years the caselaw word count is about 1.3 higher on a log scale.

Figure A.20 shows the descriptive relationship between legislative volume and the volume of laws from courts. As with regulations, in the cross-section there is a positive and statistically significant relationship. This reflects that states with more legislation also have more published court opinions. In the panel, there is still a positive relationship but it is no longer statistically significant. Changes in legislative text flows are not strictly related to changes in judicial text flows.

Table A.30 shows the reduced form effect of the legislative instrument on judicial opinion text volume. Unlike with the regulations, there is a small positive effect of the instrument on court output, but only when including state trends. However, we

Table A.30: Reduced Form Effect of Legislation Instrument on Judicial Opinion Output

	(1)	(2)
	Effect on Growth per Capita	
	State Court Opinion Words	State Court Opinion Words
Shock to Legislative Output	-0.00632	0.0676+
	(0.0807)	(0.0353)
Observations	1,183	1,183
State FE	X	X
Time FE	X	X
State-Specific Trends		X

Notes. The table shows the results of regressing the logged state court case word count on the instrument. **p<.01; *p<.05; +p<.1.

ran our main specifications controlling for the measure of judicial text output (Table A.31) and the results are identical. Further, we found that results are similar when we control for number of cases, number of opinions, and average number of words per opinion. Overall, these results suggest that caselaw is not an important mediator for our economic-growth results.

Finally, we would like to net out any influence of the federal judicial system. For this purpose, we allow for separate trends by federal judicial circuit – groups of 3-7 states that share a federal circuit court (the intermediate court below the U.S. Supreme Court). Table A.32 shows the main results after adding circuit-year interacted fixed effects. Results are robust and estimates are very similar to those from the main models.

FRASE Index for Federal Regulation. We add in information on the FRASE index from McLaughlin and Sherouse (2016). FRASE (Federal Regulation and State Enterprise) measures the impact of federal regulation on the private-sector industries in each state's economy. Cross-state variation is given by the differences in industry composition. Hence, a state's FRASE score represents the degree of impact federal regulations have on a state's economy relative to federal regulations' impact on the national economy.

FRASE is available only for recent years. Figure A.21 shows a positive relationship between state statutory legislation and FRASE in these overlapping years, but that relationship is not statistically significant. Table A.33 shows that the shift-share instrument for legislation has no effect on FRASE, suggesting that federal regulations

Table A.31: Controlling for Legal Detail in State Court Opinions

	(1)	(2)
	Effect on Growth per Capita	
	2SLS	2SLS
Legislative Output	0.0166+	
	(0.00864)	
Contingent - Non-Contingent		0.0695**
		(0.0230)
Observations	1,182	1,182
First Stage F-stat	22.18	36.57
State FE	X	X
Time FE	X	X
Court Opinion	X	X
State Trends	X	X

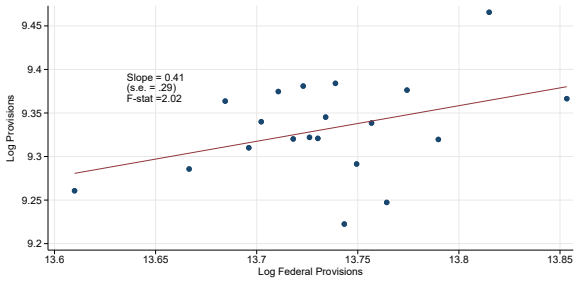
2SLS results when controlling for logged state court opinion words. Notes. . **p<.01; *p<.05; +p<.1.

Table A.32: Federal Judicial Circuits

	(1)	(2)	(3)	(4)
	Effect on Growth per Capita			
	2SLS	2SLS	2SLS	2SLS
Legislative Output	0.0110*	0.0163*		
	(0.00446)	(0.00792)		
Contingent - Non-Contingent			0.0495+	0.0576*
			(0.0274)	(0.0259)
Observations	1,180	1,180	1,180	1,180
First Stage F-stat	17.78	67.95	9.349	14.44
CircuitxTime FE	X	X	X	X
State FE		X		X

Notes. 2SLS results when controlling for circuit-year fixed effects. **p<.01; *p<.05; +p<.1.

Figure A.21: Statutory Legislation and FRASE Federal Regulation Index



Notes. The relationship between logged number of provisions and logged state regulation words, controlling for year FE (left panel) and state and year FE (right panel). SE clustered at state level.

Table A.33: Effect of Legislative Shock on FRASE Federal Regulation Index

	(1)	(2)
	Fed Reg	Fed Reg
Shock to Legislation	-0.0182	0.000935
	(0.0218)	(0.0225)
Observations	385	385
State FE	X	X
Time FE	X	X
State-Specific Trends		X

Notes. The table shows the results of regressing the (logged) FRASE index on the instrument. All models include standard errors clustered at state level. **p<.01; *p<.05; +p<.1.

are not an important mediator for the main effects.

To control for FRASE in the whole sample, we average it by state across years and interact it with time fixed effects. Table A.34 shows the main results of the paper, controlling for FRASE interacted with time effects. Results are robust and the coefficient estimates are very similar to those from the main model.

Table A.34: Adjusting for FRASE Federal Regulation Index

	(1)	(2)
	2SLS	2SLS
Legislative Output	0.0143+	
	(0.00765)	
Contingent Provisions		
Non-Contingent Provisions		
Contingent - Non-Contingent		0.0638**
		(0.0214)
Observations	1,182	1,182
First Stage F-stat	21.67	29.50
State FE	X	X
Time FE	X	X
Fed Reg x Time	X	X

Notes. 2SLS results when controlling for a state's FRASE index interacted with year FE . **p<.01; *p<.05; +p<.1.

D.5 Additional Material for Contingency

Table A.35: Effect of Contingency, Additional Specifications

	(1)	(2)	(3)	(4)	(5)
	Effect on Real GDP Growth Per Capita				
Contingent -	0.0617*	0.0677**	0.120**	0.0637*	0.0642
Non-Contingent	(0.0230)	(0.0215)	(0.0377)	(0.0275)	(0.0390)
First Stage F-stat	31.67	38.33	22.6	33.61	15.24
Observations	1133	1122	1182	1132	1086
Time FE	X	X	X	X	X
State FE	X	X	X	X	X
State Trends	X	X	X	X	X
Lagged Govt Exp	X				X
Democrat Control		X			X
Topic Shares			X		X
Control for Lagged y				X	X
Econ Vars × Time					X
Sector Shares × Time					X
Demog Vars × Time					X

Notes. Effect of the difference in contingent and non-contingent clauses – 2SLS estimates. All specifications include time and state fixed effects, control for state trends and use standard errors clustered at the state level. Column 1 controls for lagged government expenditure. Column 2 controls for democratic control over the state. Column 3 includes the topic shares among the controls. Column 4 includes the lagged dependent variable. Column 5 includes all the aforementioned controls, adding the trends of economics variables, sector shares and demographic variables. **p<.01; *p<.05; +p<.1.

Table A.36: Effect of Contingencies on Additional Economic Variables

	(1) GDP (Total)	(2) Population	(3) Employment	(4) Profits	(5) Wages	(6) Establishments
Contingent	0.0771**	0.0192*	0.00445	0.186	0.0517+	-0.00516
	(0.0240)	(0.00944)	(0.0312)	(0.119)	(0.0271)	(0.0131)
Non-Contingent	-0.0694**	-0.0240+	-0.000140	-0.181	-0.0480+	-0.00201
	(0.0256)	(0.0121)	(0.0315)	(0.137)	(0.0283)	(0.0159)
First Stage F-stat	22.27	22.27	36.52	16.24	22.27	36.52
Observations	1183	1183	821	549	1183	821
State FE	X	X	X	X	X	X
Time FE	X	X	X	X	X	X

Notes. Results for the 2SLS model with contingent and non-contingent clauses but with different outcome variables. Column 1 explores the effect on state GDP (not per capita). Column 2 shows there is no effect on population. Column 3 uses employment while column 4 looks at firm profits (value added) within the state. Column 5 looks at wages and Column 6 establishment growth. All specifications include state and biennium fixed effects. Standard errors clustered by state. **p<.01; *p<.05; +p<.1.

Table A.37: Effect of Contingent and Non-Contingent Clauses by Themselves

	(1)	(2)	(3)	(4)	(5)	(6)
	Effect on Real GDP Growth Per Capita					
Contingent Provisions	0.0199*	0.0185*	0.0443			0.0018
	(0.00797)	(0.00771)	(0.0575)			(0.0970)
Non-Contingent Provisions			-0.0326	0.0166+	0.0153+	0.0133
			(0.0705)	(0.00839)	(0.00794)	(0.110)
First Stage F-stat	43.2	26.34	20.51	22.26	22.12	6.318
Observations	1182	1182	1182	1182	1182	1182
Time FE	X	X	X	X	X	X
State FE	X	X	X	X	X	X
State Trends		X	X		X	X
Contingency Control			X			
Non-Contingency Control						X

Notes. Additional contingency 2SLS specifications. There is an effect for contingent clauses by themselves, and a weaker effect of non-contingent clauses by themselves.

D.6 Additional Material on Concavity and Uncertainty

Table A.38: Concavity: Effect of Provisions on Growth by Recent Detail Level

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Effect on Real GDP Growth Per Capita						
Recent Legal Detail	*Low*			*Medium*		*High*	
Legislative Output	0.0404*	0.0425*		0.00640	0.000205	0.0002	-0.0109
	(0.0167)	(0.0158)		(0.0104)	(0.0107)	(0.00743)	(0.00935)
Contingent - Non-Contingent			0.117**				
			(0.0351)				
First Stage F-stat	66.18	59.26	25.29	48.65	47.87	86.59	67.12
Observations	392	392	392	385	385	382	382
Time FE	X	X	X	X	X	X	X
State FE	X	X	X	X	X	X	X
State Trends		X	X		X		X

Notes. Results for the 2SLS model (Second Stage 1 and First Stage 3), splitting up the data by terciles in recent legislative output (previous five biennia). Columns 1 through 3 report results for states with lower tercile recent legislative output. Columns 4 and 5 report results for those with average recent legislative output and Columns 6 and 7 states with recent legislative output in the higher tercile. All specifications include a first column with time and state fixed effects and a second column with the addition of state specific trends. **p<.01; *p<.05; +p<.1.

Table A.39: Cross-Tabulation: Terciles in Recent Detail and Economic Policy Uncertainty

	Terciles in Economic Uncertainty			
Terciles in Recent Detail	1st	2nd	3rd	Total
1st	83	125	164	372
2nd	107	121	142	370
3rd	179	130	79	388
Total	369	376	385	1130

Notes. This table shows that recent detail (concavity) and economic policy uncertainty recover different dimensions in the dataset.

Table A.40: Concavity Effects, with Residualized Previous Detail Ranking

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Effect on Real GDP Growth Per Capita						
Recent Detail	*Low*			*Medium*		*High*	
Legislative Output	0.0220+	0.018		0.0141	0.0211	0.00703	0.0173
	(0.0121)	(0.0113)		(0.0122)	(0.0168)	(0.0178)	(0.0237)
Contingent - Non-Contingent			0.0889*				
			(0.0403)				
First Stage F-stat	54.34	55.89	12.68	37.54	35.24	77.57	109.2
Observations	389	389	389	389	389	382	382
Time FE	X	X	X	X	X	X	X
State FE	X	X	X	X	X	X	X
State Trends		X	X		X		X

Notes. The main concavity results, but the previous detail variable is residualized on state and year fixed effects before making the ranking. Columns 1, 2 and 3 report results for states with lower tercile recent legislative output. Columns 4 and 5 report results for those with average recent legislative output and Columns 6 and 7 states with recent legislative output in the higher tercile. All specifications include a first column with time and state fixed effects and a second column with the addition of state specific trends. **p<.01; *p<.05; +p<.1.

Table A.41: Uncertainty Effects, with Residualized Uncertainty Ranking

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Effect on Real GDP Growth Per Capita									
Economic Uncertainty	Low		Medium		High					
Legislative Output	0.0201		0.0146		0.0191	0.0199+				
	(0.0152)		(0.0211)		(0.0132)	(0.0102)				
Contingent Provisions							0.0624	0.0637		
							(0.0561)	(0.0759)		
Non-Contingent Provisions							-0.0508	-0.0522		
							(0.0676)	(0.0857)		
Contingent -		0.0275		0.0613					0.0847+	0.113+
Non-Contingent		(0.0613)		(0.0547)					(0.0470)	(0.0588)
First Stage F-stat	39.95	3.998	2.42	9.512	6.721	257.2	3.578	4.873	6.87	9.285
Observations	362	362	381	381	355	355	355	355	355	355
Time FE	X	X	X	X	X	X	X	X	X	X
State FE	X	X	X	X	X	X	X	X	X	X
State Trends						X		X		X

Notes. The main economic policy uncertainty results, but the uncertainty variable is residualized on state and year fixed effects before making the ranking. Columns 1-2 show results for states with lowest tercile uncertainty. Columns 3-4 report results for those with median uncertainty while Columns 5-10 states with uncertainty in the higher tercile. All specifications include state and biennium fixed effects, while for High Uncertainty states, results controlling for state specific trends are also included (as indicated). **p<.01; *p<.05; +p<.1.

Table A.42: Uncertainty Effects, with Lagged Economic Growth Control

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Effect on Real GDP Growth Per Capita						
Economic Uncertainty	Low	Medium	High				
Legislative Output	-0.0136	0.00796	0.0445*				
	(0.0115)	(0.0112)	(0.0169)				
Contingent Provisions				0.142*	0.151*		
				(0.0567)	(0.0703)		
Non-Contingent Provisions				-0.126+	-0.131		
				(0.0641)	(0.0804)		
Contingent - Non-Contingent						0.185**	0.202**
						(0.0463)	(0.0601)
Lagged Growth P.C.	0.454**	0.443**	0.186**	0.190**	0.172**	0.206**	0.172**
	(0.0840)	(0.0616)	(0.0616)	(0.0511)	(0.0588)	(0.0588)	(0.0625)
First Stage F-stat	48.45	3.599	24.33	8.971	8.198	9.488	9.484
Observations	335	348	363	363	363	363	363
Time FE	X	X	X	X	X	X	X
State FE	X	X	X	X	X	X	X
State Trends					X		X

Notes. The main uncertainty results, but adding lagged growth per capita as a control. Columns 1-2 show results for states with lowest tercile uncertainty. Columns 3-4 report results for those with median uncertainty while Columns 5-10 states with uncertainty in the higher tercile. All specifications include state and biennium fixed effects, while for High Uncertainty states, results controlling for state specific trends are also included (as indicated). **p<.01; *p<.05; +p<.1.

Table A.43: Effects when Splitting by Terciles in Recent Growth

	(1)	(2)	(3)	(4)	(5)	(6)
	Effect on Real GDP Growth Per Capita					
Recent Growth in State	Low		Medium		High	
Legislative Output	0.0409*		-0.00129		0.00924	
	(0.0157)		(0.00964)		(0.00878)	
Contingent -		0.113*		0.00672		0.0812+
Non-Contingent		(0.0456)		(0.0391)		(0.0431)
First Stage F-stat	30.39	7.111	6.480	11.43	3.658	8.572
Observations	347	347	370	370	408	408
Time FE	X	X	X	X	X	X
State FE	X	X	X	X	X	X

Notes. 2SLS estimates separating out by recent growth. Columns 1-2 show results for states with lowest tercile of recent growth. Columns 3-4 report results for those with median growth while Columns 5-6 states with recent growth in the higher tercile. All specifications include state and biennium fixed effects and use standard errors clustered by state. **p<.01; *p<.05; +p<.1.

E Two Models of Legislative Detail and Growth

This appendix outlines two formal models for how legislative detail could influence economic growth. The main result of the paper is that increasing legislative detail in U.S. states has been good for growth. And that is particularly strong when new laws and rules contain relatively more legislative contingencies, in topics and industries with a relatively small stock of existing laws and rules (concavity), and in periods and states with a relatively higher level of economic policy uncertainty. In the main text we focused on the standard hold-up model, where more complete legislation increases specific investments by reducing the threat of ex post hold-up. In this appendix we offer two formal models to undergird and complement the hold-up model, and in particular to derive predictions for the mechanisms analysis in Section 6. First, in Section E.1 we present a model based on writing costs. Second, Section E.2 uses a concise decision theory framework for understanding the legislator's choice when to legislate. Both models provide microfoundations for the broader view that greater detail enhances relationship-specific investments by producing a more complete – and thus more enforceable – contract.

E.1 Legislating as Incomplete Contracting

Here we describe how the writing costs approach of Battigalli and Maggi (2002, 2008) can be adapted to the legislative process in order to derive a set of hypotheses on the causes and consequences of legislative output. We start by using the logic of Battigalli and Maggi (2002) to describe the law as an incomplete contract; we will then describe how some insights can also be derived from Battigalli and Maggi (2008), where the focus is not on the degree of completeness of a law but on the type of clauses (contingent or spot) and on their evolution over time. We will derive from this framework a set of hypotheses that motivate the mechanisms analysis from Section 6.

E.1.1 What can we learn from the writing costs approach?

A law can be viewed as an incomplete contract between the legislator (the principal) and the citizens (agents), with an efficiency objective. Incompleteness can take the form of rigidity (non-contingent clauses) or discretion (empty clauses). The chosen degree of incompleteness depends on writing costs, which could include for example the cost of figuring out the relevant legal requirements, the cost of thinking how to describe them, and the cost of time needed to draft and enact the laws. These are all costs related to

the details and precision of the language of the law.

The language of the law consists of primitive sentences that describe (1) elementary events and (2) elementary actions, plus logical connectives (e.g., "not," "and," "or"). This language can be used to describe state- dependent constraints on behavior, or in other words, a correspondence from states to allowable behaviors. Each primitive sentence has a cost and the total cost of writing the contract is a function of the costs of its primitive sentences about events and actions. It follows naturally that contingent clauses are more costly than non-contingent clauses.

A contingency is a formula about the environment. It could include different events with different logical connectives, so one contingency might be conditioned on event 1 or event 2, while another contingency could be conditioned on event 1 and event 3. An instruction is a formula for behavior – i.e., a set of actions with some logical connectives, such as action 1 and/or action 2. Omitting from the text of a law an elementary sentence about the possible events or situations that could occur saves on the cost of describing contingencies, but makes the contract rigid. Omitting from the contract an elementary action saves on the cost of describing behavior, but gives discretion to the agent.

We can adapt the main characterization of Battigalli and Maggi (2002) to our context. Informally, we restate their Proposition 1 as saying that optimal legislation should have contingent clauses for the most important decisions regulated by the law, while less important decisions can be regulated by rigid or non-contingent clauses, and the least important decisions can be left to discretion.

According to Proposition 2(II) in Battigalli and Maggi (2002), in more uncertain environments, the optimal law contains more contingent clauses and fewer rigid clauses, and it leaves more discretion to the agent. When uncertainty is higher the efficiency cost of ignoring low-probability events and writing rigid clauses is higher.

In Battigalli and Maggi (2008), agent discretion is regulated by informal contracts or spot clauses, which become possible with repeated play. Further, when the cost of describing contingencies is low relative to the cost of re-negotiation after unregulated events, then contingent clauses are optimal to begin with. A spot approach is optimal when this relative cost is high. Under intermediate values, an enrichment approach – where when a new unregulated contingency occurs it induces an enrichment of the contingent clauses in the law – may be optimal.

E.1.2 Deriving testable predictions

Now we use these ideas to derive a set of testable predictions.

Completeness. The first aggregate prediction coming out of the optimal contract framework is that if more legislation is added by a benevolent principal, it must be because the clauses are beneficial – that is, specifying more enforceable rules rather than giving hand-outs to lobbyists. Then more complete rules increase investment due to reduction in hold-up opportunities. In other words:

H0: If legislators are benevolent, then the greater the completeness of law, the better the economic outcomes to be expected.

Contingency. The second prediction that we can derive from the Battigalli and Maggi (2002, 2008) framework is related to contingent clauses. Suppose that for each issue or topic there are plenty of contingencies that one could potentially differentiate, but each contingency requires a constant marginal writing cost. Even if the marginal writing cost of an extra contingency is constant, the marginal benefit depends on many things that could vary a lot from place to place and from year to year, as well as some common component that relates to technological changes or other exogenous transformations of the topic to be regulated.[17] As a result, given a fixed marginal writing cost but wide variation in the benefit function, the state legislators choose different levels of contingent laws across states.

The optimal level of completeness of contracts is increasing in the marginal benefit of adding contingencies. Hence we should expect the relation between contingent clauses and growth to be stronger than for other clauses. That is, clauses along these lines: "if a worker has such characteristics... then a firm with such other characteristics could employ him or her with a special tax treatment, transfer, labor law relaxation, etc..." should be expected to have a positive effect given that it more precisely describes the economic environment, leaving less for subsequent renegotiation. These clauses are more costly to write and hence a rational legislator who has decided to introduce it must have anticipated a higher marginal benefit from it. The testable hypothesis that

[17] For example, in a state where all employees are in one or two sectors without many differentiations of skills, the marginal benefit from new contingent statements related to different sectors, seniority, education or other observables would be low. Hence, that state might have relatively simple labor laws and tax laws with non-contingent statements. On the other hand, in a state where skill differentiation matters, there is a higher marginal benefit from more clauses as, for example, the planner might find it important to give incentives to workers to switch from one sector to another.

corresponds to this reasoning is:

H1: The changes in legislative output that most contribute to growth prospects are contingent, rather than non-contingent, clauses.

Say that at some point a shock arises, such as the advent of the internet, where new elementary events and actions arise. The existing legislation is not optimal, so legislators should write more clauses, and more specifically write more contingent clauses based on the new set of events. Now clauses like: "if there is a good internet connection, the worker has the right to work from home" could be added and support more economic activity. A side prediction would be that contingent clauses would be even more beneficial in states with greater economic complexity – more sectors, more levels, more segmentation, more strategic incentives to be given, etc.

Concavity. We now turn to a third implication of the Battigalli and Maggi (2002) framework. Assume for simplicity that each contingent clause has the same cost c. Thus, a law that includes l contingent clauses has cost cl. The state j's marginal benefit from adding a contingent clause is a function $B(l, t, w_j)$, where $t \in R_+$ is a parameter capturing a common factor (like technological change) and $w_j \in R_+$ is a state specific parameter capturing the degree of complexity of the economy to be regulated in state j. Let $\frac{\partial B}{\partial t} > 0$, $\frac{\partial B}{\partial w_j} > 0$, and $\frac{\partial B}{\partial l} < 0$ (the latter capturing a concavity assumption).

In state j with a low w_j, prescribing a rigid clause to always be at the office from 9 to 5 could be optimal. In state k with $w_k > w_j$, however, there may already be a contingent clause that working from home is possible when some condition on traffic or weather is met. In other words, state k with high w_k may optimally have $l_k^* > l_j^*$. Suppose that this is the case at time 0 with common technology t_0. Consider an exogenous shock at time 1 determining $t_1 > t_0$ (like the invention of internet), such that $l_k^*(t_1) = l_k^*(t_0) + 1$ and $l_j^*(t_1) = l_j^*(t_0) + 1$. It follows naturally, given the concavity assumption, that the effects must be bigger in state j. When a change in t makes it convenient for both states to add a contingent clause like "if there is good internet connection, the worker shall work from home" then this addition benefits relatively more state j.

H2: An exogenous increase in legislative completeness will have a greater growth differential in the states with lower initial level of legislative stock.

Uncertainty. The fourth implication of the Battigalli and Maggi (2002) framework concerns the role of uncertainty. That is, it is plausible that the marginal benefit of contingent clauses is higher in states that are exposed to greater uncertainty. Under low uncertainty, a rigid clause that follows the likely state works best. The more

uncertain are the relevant situations, the more valuable will it be to account for different possible contingencies and state a context-dependent action. The functional form for the marginal benefit of an additional contingent clause could be enriched by adding an additional parameter $u_j \in R_+$ capturing the degree of uncertainty in state j. The simple hypothesis to be tested is that indeed the marginal benefit of more contingent clauses is higher when u_j is higher.

H3: The greater or the more frequent the sources of **uncertainty** in a state, the greater will be the growth benefit from higher legislative completeness, and especially from more contingent clauses.

E.2 A model of optimal legislating under uncertainty

Now we present an alternative model based on legislative decision-making. A decision maker (DM henceforth) has to decide every period whether to adopt a legislative reform regulating an industry. The proposal arriving to the DM's desk every period can be complex or simple, y_c or y_s, and let's assume that the arrival of one or the other is governed by an exogenous process. A proposal y_c involves a large number of contingent clauses and/or exemptions, while y_s has no clauses or exemptions to check, hence has a lower implementation cost: $c(y_c) > c(y_s)$. These different costs for the bureaucracy also have cost implications for the DM, since a reform that takes longer to be fully implemented yields reputation benefits for the DM with lower probability by re-election time. Given that we do not want to introduce the bureaucracy as an explicit additional player, let's consider that $c(y)$ is directly the cost for the DM.

The implementation costs of the reform are known at the time of the decision, but the benefits are uncertain. If the reform passed at time t is good, it adds $b_t = 1$ to growth, while if the reform is bad it subtracts 1 ($b_t = -1$). The growth potential metric can be any increasing function of the likelihood of good reforms being passed. In our overall intuition for state economic legislation, we understand good legislation as increasing the completeness of the legislative contract and thereby increasing investment. Bad legislation is just adding costly boilerplate, or else making suboptimal rules.

The DM receives a signal (from experts, from her understanding, or from the relevant staff) on the potential benefits of the reform, and decides whether to adopt it or not. The DM's calculus requires the expected benefits of the reform to outweigh her cost, even though only the expected benefit matters for growth potential.

Formally, there are two states of the economy, one where the reform considered, y,

is *good* (θ^G) and one where it is *bad* (θ^B). The state θ is unknown to the DM, who has to appraise the effect of the reform on the basis of her prior $\kappa \in (0,1)$ that the reform is good and on the basis of her signal s, which can be either good (s^G) or bad (s^B). The signal identifies the true state with precision $1-z$, where $z \in [0, \frac{1}{2}]$ captures the difficulty of appraising the benefits of the reform on the basis of the signal.[18] The likelihood of the good state is updated by Bayes' rule as follows:

$$\kappa(s^G) = \frac{(1-z)\kappa}{(1-z)\kappa + z(1-\kappa)}, \qquad \kappa(s^B) = \frac{z\kappa}{z\kappa + (1-z)(1-\kappa)}.$$

If the DM rejects the reform, the status quo is maintained, and his payoff is normalized to zero. If she adopts the reform, the realized benefits and the cost $c > 0$ of its implementation determine a DM utility depending on the state. The utility is $u(y|\theta^B) = -1 - c$ if it is a bad reform, while $u(y|\theta^G) = 1 - c > 0$ if it is a good reform.

Obviously the DM adopts the reform if she expects positive effects on the economy net of her costs relative to the status quo, that is, if and only if

$$\mathbb{E}[u|s] = \kappa(s)u(y|\theta^G) + (1-\kappa(s))u(y|\theta^B) > 0 \quad \Leftrightarrow \quad \kappa(s) > \frac{1+c}{2}.$$

We can thus characterize the decision of the DM as a function of her signal, as follows:

Lemma 1: *The reform y is adopted under s^G when*

$$\mathbb{E}[u|q^G] > 0 \quad \Leftrightarrow \quad (1-z)\kappa v - z(1-\kappa)\ell > c$$

Note that the noise z, which correlates with the complexity of the stock of legislation, depresses the expected benefits of reforms adopted on the basis of the signal q^G. In words, when signals are noisy, the DM puts a lower weight on positive signals received in her updating.

Lemma 2: *Under the bad signal q^B, the reform is adopted only when*

$$\mathbb{E}[u|q^B] > 0 \quad \Leftrightarrow \quad z\kappa v - (1-z)(1-\kappa)\ell > c.$$

Together, the two lemmas imply that if the proposal is a simple one with $c(y_s) = 0$ it passes the scrutiny of the DM for any signal she might get when $\kappa > z > (1-\kappa)$.

[18] When $z = \frac{1}{2}$, the signal is uninformative, while when $z = 0$, the signal is perfect.

Thus, assuming $\kappa \geq 1/2$ throughout, only the second inequality matters, $z > (1-\kappa)$. Otherwise, for $z < (1-\kappa)$, a simple proposal passes only after a good signal.

When the cost is positive, the condition for passing a proposal after a bad signal becomes

$$z > \hat{z} \equiv \frac{(1-\kappa)(c+\ell)}{\kappa(v-c)+(1-\kappa)(c+\ell)} \text{ and } \hat{z} < 1/2. \tag{12}$$

$\hat{z} < 1/2$ requires that $(1-\kappa)(l+c) < \kappa(v-c)$, i.e.,

$$c < \kappa v - (1-\kappa)l. \tag{13}$$

In contrast, the reform is accepted under signal q^G if and only if

$$z < z^* \equiv \frac{\kappa(v-c)}{\kappa(v-c)+(1-\kappa)(c+\ell)}. \tag{14}$$

Thus:

Proposition 1:

1. *If (13) is violated (high enough c), then the reform is adopted if and only if the DM receives a good signal and $z < z^*$.*
2. *If (13) holds, then: (a) the DM adopts the reform under both signals for $z \in (\hat{z}, 1/2]$; and (b) the DM adopts the reform only under a good signal when $z \leq \hat{z}$.*

For a given implementation cost c, the prediction is that when κ is sufficiently high then there exists a range of z under which the signals are noisy enough that the DM makes the reform regardless of the signal. On the other hand, if z is low enough, then the DM approves only after a good signal, and it is in such cases that the expected benefit of the reform is higher conditional on the reform passing. In other words, the cases in which the DM follows her signal and hence carefully selects proposals are when c is high, κ is low (closer to 1/2, hence high uncertainty), and z is low (low stock of existing regulations). These are exactly the three conditions highlighted respectively as contingency, uncertainty, and concavity.

Proposition 2:

1. **Contingencies prediction:** When the reform considered is more complex, it is more likely that (13) does not hold, and hence that the DM adopts the reform only after a good signal. In other words, conditional on being adopted, a reform with more contingent clauses yields higher benefits.

2. **Concavity prediction:** When a topic or an industry is not heavily regulated yet and hence the evaluation of the merits of a reform is easier (low z), then again the DM approves only in case of a good signal, and hence conditional on passing, the reforms have a higher chance to be beneficial.

3. **Uncertainty prediction:** In contexts with high uncertainty about economic policies, κ must be tending to $1/2$, and in such cases, like for the above contingencies prediction, (13) is unlikely to hold, and hence once again the DM applies the good signal filter, implying that conditional on passing, the reform is likely to be beneficial.

Proposition 2 highlights how the three parameters c, κ, z capture in the most concise manner the three predictions we made and tested about heterogeneity in the results. This parsimony is obtained at the cost of strong assumptions. One implicit assumption is that a law or rule with more contingencies and exemptions has higher implementation costs, but (1) such higher costs do not matter for growth, and (2) the difficulty of evaluation for the DM, captured by z, is not impacted. We instead consider it intuitive that z depends on the existing stock of laws and rules already in place. It is ultimately an empirical question whether a law or rule with lots of contingencies and exemptions is harder or easier to evaluate for the DM, and it could even be the case that evaluation is actually easier with more contingencies. So both the above strong assumptions seem reasonable. A third strong assumption of this simple model is that simple or complex proposals arrive to the DM as an exogenous process, unmodeled. However, Foarta and Morelli (2022) show that when the choice to go for a simple or complex proposal is endogenized, it remains true that complex proposals are more likely in a world with greater uncertainty, and that produces good outcomes when the relevant DM is benevolent. In our settings, benevolence comes from the context of state legislators, who work in a competitive legal environment with potential learning from other states. Hence the simplification made in this simple model does not seem too problematic.

The value of the simple decision theory rationalization model is, in any case, the robustness and generality that it could give to our findings. The fact that our findings are consistent with both a standard hold-up reduction model, the Battigalli-Maggi writing costs framework, and this simple decision theory model gives us confidence that these uncovered empirical relationships are not coincidental. These empirically validated theoretical results could guide future research, for example on E.U. allocation of legislative and regulatory jurisdictions.