

# Persistence in Policy: Evidence from Close Votes

Zach Freitas-Groff\*

December 17, 2023

[\(Click for Current Version\)](#)

## Abstract

Policy choices sometimes appear stubbornly persistent, even when they become politically unpopular or economically damaging. This paper offers the first systematic empirical investigation of how persistent policy choices are, defined as whether an electorate's or legislature's decisions affect whether a policy is in place decades later. I create a new dataset that tracks the historical record of more than 800 policies that were the subjects of close U.S. state referendums since 1900. In a regression discontinuity design, I estimate that passing a referendum increases the chance a corresponding policy is operative 20, 40, or even 100 years later by over 40 percentage points. I collect additional data on U.S. Congressional legislation and international referendums and use existing data on state legislation to document similar policy persistence for a range of institutional environments, cultures, and topics. I develop a theoretical model to distinguish between possible causes of persistence, and I present evidence that persistence arises because policies' salience systematically declines over time. Calibrating my model suggests that many policies remain in place—or not—regardless of popular support.

KEYWORDS: Political Economy, Democracy, Persistence, Regression Discontinuity.

JEL codes: D72, D78, H11, K15, K16, K40, N4, P48.

---

\*Stanford University, Department of Economics, 579 Serra Mall, Stanford, CA, 94305, USA, [zgroff@stanford.edu](mailto:zgroff@stanford.edu). I am indebted to B. Douglas Bernheim, Matthew Gentzkow, Caroline Hoxby, and Ran Abramitzky for invaluable mentorship and advice. I thank Matt Brown, Mark Duggan, Liran Einav, Lucas Freitas-Groff, Mike Geruso, Ben Grodeck, Andy Hall, Nathan Hendren, Saumitra Jha, Zane Kashner, Greg Martin, Neil Malhotra, Rossa O'Keefe-O'Donovan, Petra Persson, Keller Scholl, Dean Spears, Alessandra Voena, Mark Walsh, and seminar participants at Stanford University and the Global Priorities Institute for comments and suggestions. I am grateful to Milena Baltic, Bogdan Ilic, Petra Jelesijevic, Lav Velimirovic, and Keith Wynroe for excellent research assistance. This material is based upon work supported by the National Science Foundation Graduate Research Fellowship Program under Grant No. (DGE - 1656518). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. I gratefully acknowledge the support of the Institute for Research in the Social Sciences at Stanford University, the Leonard W. Ely and Shirley R. Ely Graduate Student Fellowship, and the Hawley-Shoven Fellowship, the latter two through a grant to the Stanford Institute for Economic Policy Research.

# 1 Introduction

If voters and their representatives obtain their desired policies in the long run, then we might expect that past policy choices will generally cease to matter over time. Decisions to alter tax rates, adopt new public programs, or grant new civil rights may matter in the short run, but eventually policies will end up in place only to the extent that they have public support. Policies that pass and prove unpopular will get repealed, and policies that fail to pass but prove popular will pass later.

Theoretical and empirical work, however, raise the possibility that policy choices might persist for puzzlingly long periods of time. A small theoretical literature explains how policy choices can persist, typically focusing on endogenous responses to the prevailing policy (Fernandez and Rodrik, 1991; Coate and Morris, 1999; Gieczewski, 2021; Acemoglu et al., 2021). Economic history offers examples of institutional persistence including extractive colonial institutions (Acemoglu et al., 2001, 2002), common versus civil law systems (Porta et al., 1998), public good provision (Dell, 2010), and liberal democracy (Woodberry, 2012). Yet as Kelly (2019) and Voth (2021) note, we only study the cases where there is persistence: few would write detailed papers documenting political decisions hundreds of years ago that have no effects today. While policy persistence has deep implications for the functioning of democracy, there is no empirical estimate of how long policy choices typically persist.

In this paper, I collect detailed statutory histories of close U.S. state-level referendums to show that policy choices are remarkably, differentially persistent. Mere approval increases the chance that a law is operative decades or even a century after the referendum by around 40 percentage points. The question of causal persistence depends on how often approved policies get repealed and how often rejected policies pass later. In keeping with the theoretical literature (Coate and Morris, 1999), I define “persistence” as the difference between the shares of (otherwise identical) passed and rejected proposals that are operative  $t$  years later. Persistence captures a causal phenomenon where today’s policy determines tomorrow’s. To measure persistence, I estimate the effect of referendum passage on the status of the relevant law over time in a regression discontinuity panel design. The results show that the pivotal voter’s decision in a single referendum exerts a lasting influence. I then replicate this result with international referendums and legislatively-passed policies.

The histories of the income tax in Oregon and Washington illustrate why persistence is puzzling and why it matters. Between 1912 and 1940, many U.S. states adopted income taxes for the first time, with Oregon and Washington each having several referendums on the topic (National Conference of State Legislatures, 2023; Ballotpedia, 2023a). In Oregon, 52.49% of voters approved an income tax in 1930, and it remains in place today. In Washington, on the other hand, voters rejected an income tax in 1938 and in 1942 after courts threw out an initially-approved tax. The

contrast is vexing; the two states were and remain politically, economically, and demographically alike, yet they have fundamentally different tax structures, apparently because of the persistence of close decisions made nearly a century ago. Understanding how common this apparent policy persistence is requires finding and comparing more examples of policies that pass or fail for quasi-random reasons.

My data collection draws on the rich historical documentation of U.S. referendums, which are well suited to study political decision-making because of their number, diversity of topics, and varied institutional features. Voters in every U.S. state except Delaware face referendums, with over 8,000 occurring since 1900. Referendums have shaped state laws on topics including property and income taxation, school funding, gambling, labor unions and drug policy. Referendums can be proposed either by legislators or by citizen petition depending on the state. State referendum processes differ in various ways, including the difficulty of proposing a referendum and whether approval requires more than a fifty-percent majority. The referendum process is an especially interesting context in which to study policy persistence because, since citizens can initiate referendums in many states, it might seem like a process in which policy should reliably update over time based on voters' evolving attitudes.

To study the persistence of referendum decisions, my novel dataset tracks whether a proposed law is operative from its referendum through 2022. My dataset builds on existing data on referendum votes, topics, and features from the National Conference of State Legislatures and Ballotpedia. Working with a team of Upwork assistants, I manually collect the full history of adoption, modification, and repeal for a subset of the wider universe of referendums given the intensive data collection requirements. I focus on laws that pass or fail narrowly in order to compare laws that are statistically similar. Specifically, I collect outcomes for 825 referendums that pass or fail by fewer than 2.5 percentage points, which I select for being well within the typical bandwidth in electoral regression discontinuity designs and validate with a battery of empirical tests (Song, 2018). Bandwidth selection necessarily occurs before data collection, but the resulting subset of referendums retains the diversity of topics and institutional backgrounds of the wider dataset.

My analysis employs a causal approach that is different from prior historical work on persistence. Previous research on this topic has taken geographic areas as the units of observation and local adoption of a specific institution as the source of variation (Voth, 2021). My work complements and extends this literature by taking laws as the units of observation and exploiting the quasi-random variation in which policies pass.<sup>1</sup> Using a local linear regression, I estimate persistence as the effect of a referendum's passage on whether a substantively similar policy is operative later. The flexibility of the design allows me to investigate possible mechanisms of persistence by estimating the effect

---

<sup>1</sup>The variation is quasi-random in the sense that whether a given policy passes depends on the decision of a small group of voters. Any last-minute news, weather, or other event uncorrelated with unobservables can tip the balance.

of passage on intermediate outcomes and interacting passage with features of state institutions.

The effect of passage on whether a policy is operative decreases from one hundred percentage points immediately after a referendum to forty percentage points after a couple of decades, where it plateaus. I distinguish between two reasons this effect can decline: policies that pass get repealed, and policies that fail pass later. In fact, just over 20% of passed referendums are later repealed, while just under 40% of failed referendums pass later. The overall effect is robust across data specifications, constant over time, uncorrelated with measures of policy importance, and unaffected by excluding referendums that would later become obsolete. I also find similar persistence across topics including social policy, criminal justice, taxes, spending, and the structure of state government.

I interpret the reduced-form results with a model of the policymaking process, which shows that long-term persistence contradicts a simple dynamic extension of a canonical (Downsian) approach to political economy, where policies arise through efficient political deliberation. Two propositions demonstrate that persistence must be short-lived in a simple democracy where decisionmakers have stable policy preferences. The first proposition is that when proposing is costless, the unique Nash equilibrium exhibits no persistence. That is, after a single election cycle, the policy converges to whatever voters prefer. The second proposition is that, even when there are proposal costs, persistence steadily declines in any Markov perfect equilibrium because the likelihood of someone successfully changing the policy is constant over time. The two results, calibrated to the data, show that there should be no persistence after a few decades in an archetypal democracy, capturing a world in which issues work themselves out over time (Liebowitz and Margolis, 1995; Downs, 1972).

After presenting the model, I tweak its assumptions to illustrate two possible (non-exhaustive) mechanisms that can generate long-term persistence: endogenous political responses and declining issue salience (or “punctuated equilibrium”). For the first mechanism, I show that if either intrinsic motivation or external incentives make decisionmakers more supportive of policies the longer they are in place, there can be arbitrarily long-term persistence, which is in line with the economics literature on the topic (Fernandez and Rodrik, 1991; Coate and Morris, 1999; Acemoglu et al., 2021). The second mechanism is that if policies’ salience varies over time, policy change will tend to occur during narrow windows of heightened salience, after which the policy is unlikely to change for a long time. The mechanism is similar to “punctuated equilibrium” in political science but, but I show that the phenomenon is broader and relies on simpler foundations than has previously been appreciated (Pierson, 2004; Baumgartner and Jones, 2010). The two mechanisms have different implications for the welfare implications and prevalence of persistence, with the latter mechanism generally implying fairly uniform persistence across topics and a weak relationship between policy and voters’ preferences.

My empirical tests tend to favor the issue salience mechanism over the endogenous political responses mechanism. Data on referendum repeat attempts and news data from the Library of



Congress are consistent with the theory that referendum topics' salience is abnormally high at the time of a typical referendum. Further underlining the importance of issue salience, policies are less persistent when more people directly observe their effects. For instance, broad-based tax is likely to be more persistent than a narrow-based one. At odds with a world in which persistence arises from endogenous reactions to policies, I find no evidence that repeals slow down over time, that people tire of repeat attempts, that voters become more sympathetic to passed policies, or that policy choices are more persistent for policies for which there are theoretical reasons to expect greater path dependence.

To illustrate the implications of these results, I calibrate a special case of my theoretical model, which suggests that voters' preferences do not match the statuses of 45% of all policies proposed by petition in the previous 100 years. In other words, 45% of those policies are either operative but lack voter support, or not operative but have voter support. The mismatch between voter preferences and policies' status occurs despite voters' ability to revisit these referendums via petition. I conduct a counterfactual exercise to evaluate institutional reforms aimed at reducing this mismatch. The exercise shows that periodic votes to re-approve previously passed policies can increase the alignment between policy and voters' preferences. However, making policy changes easier does not help because it allows unpopular policies to pass during unusual election years.

The pervasiveness of persistence offers credibility to long-term policy event studies and implies that a large share of policies' impact occurs over the long term. Methodologically, the quasirandom, long-term policy variation documented in this paper supports an assumption behind long-term event-study-type designs: two identical states can end up with different policies for long stretches of time. At a normative level, political persistence implies that the impacts of coin-flip policy choices may be sizable for longer than a century. I quantify the distribution of policies' impacts over time with a duration index, which shows that the first few decades account for less than half of the average policy reform's effective lifetime. Among other implications, this result may substantially affect the generational incidence of transfer programs or pollution control.

To better understand how widespread policy persistence is, I replicate the main result with another novel dataset on legislative bills and non-U.S. national referendums. My approach to studying non-U.S. national referendums largely follows my approach to studying state referendums, and it obtains similar estimates. To estimate the persistence of state policy changes, I use data from Caughey and Warshaw (2016) and Grossmann et al. (2021) to identify over a thousand state policy changes and to match each state that adopts a policy with politically similar states. I validate my matching method against regression discontinuity estimates for those policy changes attributable to statewide referendums. To estimate the persistence of congressional legislation, I draw on contemporary congressional histories to identify legislative proposals whose narrative is consistent with narrow passage or failure, inspired by the narrative approach pioneered by Romer

and Romer (1989, 1994) for monetary policy. In all three settings, I find that the level of persistence is remarkably similar to the level of persistence of U.S. state-level referendums.

Relative to empirical literature, this paper offers the first quasi-experimental estimate of long-term policy persistence. No prior work has used a regression discontinuity design to study persistence or, to my knowledge, compiled data on both passed and failed policies' subsequent histories. In some ways, this paper resembles work on the legislative incumbency advantage, in which regression discontinuity designs have proven credible (Lee, 2008; Eggers et al., 2015; Hainmueller et al., 2015; Fowler and Hall, 2017). Other research documents the repeal and modification of passed federal policies but lacks a comparison group of similar rejected policies (Bickers, 1991; Lewis, 2002; Carpenter and Lewis, 2004; Corder, 2004; Post and Pierson, 2005; Maltzman and Shipan, 2008; Berry et al., 2010; Ragusa, 2010; Thrower, 2017). Previous research analyzes referendums' political and economic dynamics but not persistence (Matsusaka, 1992, 1995; Besley and Coate, 2000; Cellini et al., 2010; Burszтын et al., 2023). Finally, recent work studies other ways in which policy can be path dependent, such as diffusion and social backlash (Collins, 2003; Wheaton, 2020; Wang and Yang, 2021; Carollo et al., 2022; DellaVigna and Kim, 2022; Shigeoka and Watanabe, 2023).

The historical persistence literature studies the related question of whether specific past upheavals have economic effects today. A substantial literature identifies cases of institutional persistence but does not measure the persistence of public policies or how common persistence is (Porta et al., 1998; Acemoglu et al., 2005; Botticini et al., 2006; Dell, 2010; Woodberry, 2012; Nunn, 2014; Alesina and Giuliano, 2015; Guiso et al., 2016; Alesina et al., 2017; Dell et al., 2018; Persson and Tabellini, 2020; Acemoglu et al., 2021; Voth, 2021; Cirone and Pepinsky, 2022). Moreover, the institutional persistence literature typically shows the effect of a distant historical event on a current outcome, while this paper estimates effects over the intervening time.

The data in this paper allow for the first empirical tests of the small theoretical literature on policy persistence and policy dynamics more broadly. The rich dataset and diversity of policies I study allow me to evaluate previously untested explanations for path dependence (Fernandez and Rodrik, 1991; Coate and Morris, 1999; Pierson, 2004; Baumgartner and Jones, 2010; Roberts, 2015; Gieczewski, 2021; Acemoglu et al., 2021). Moreover, the small but growing literature on optimal policy with political dynamics has until recently lacked an empirical measure of how policy choices generally affect future policy (Feldstein, 1976; Acemoglu et al., 2010; Scheuer and Wolitzky, 2016).

The structure of the paper is as follows. Section 2 describes the background and empirical strategy, and Section 3 presents the main results. Section 4 then introduces a game-theoretic model of the referendum process and investigates the possible mechanisms of persistence based on the model and reduced-form facts. Section 5 discusses what the results mean for constitutional design, policymaking, and research. Section 6 replicates these results for other policymaking processes, and Section 7 concludes.

## 2 U.S. State Referendums: Empirical Strategy

I measure persistence use a regression discontinuity panel design on a novel historical dataset. For close referendums in a subset of U.S. states since 1900, the dataset records the history of the proposed policy from the referendum to the present day. The focus on close (i.e., narrowly passed or failed) referendums maximizes causal identification given the cost of data collection, and I show that the resulting subsample is broadly representative of the universe of referendums. To understand what drives persistence, I compile data additional outcome data ad collect the timing and outcomes of attempts to revisit a referendum. In addition to the main regression discontinuity estimates, additional analyses measure heterogeneity, slippery slope and backlash effects, and the effect of successful referendum campaigns.

### 2.1 Setting and Context of U.S. Referendums

The referendums I study represent a set of political decisions and directly related laws across a range of topics and institutional settings. Referendums offer a rich and appropriate setting to help us understand political dynamics generally. They represent a major source of changes in U.S. state laws, and they are typically the way U.S. states ratify and amend constitutions.

Referendums have played a broad and significant role in seemingly every area of state policy. Every state except Delaware requires voter approval for constitutional amendments, so every topic of state constitutions has faced voters at the ballot box (Ballotpedia, 2023b). Before 1900, referendums largely came about via legislative referral. The Progressive Era changed this with a growth in citizens' ability to propose referendums by petition. The Prohibition movement participated in this push, which offered one of the main ways in which Prohibition proponents and opponents pursued their goals, including prohibition of alcohol, increased worker protections, and expansion of the right to vote (Ballotpedia, 2023b). In recent decades, referendums have been a key battleground for taxpayer revolts to limit state and local spending, same-sex marriage, marijuana legalization, sports betting, tobacco taxes, capital punishment, abortion, and animal welfare.

One attractive feature of studying referendums is that I can observe variation along multiple dimensions in who proposes referendums and how difficult they are to pass. Referendums typically concern state constitutional amendments or statutes and can be proposed by the legislature (“legislative referrals”) or petition (“initiatives”).<sup>2</sup> State requirements for legislative referrals differ in terms of whether they require a supermajority vote from the legislature, mandate votes by consecutive legislatures, impose deadlines, or specify criteria for referendum content. (For instance, some

---

<sup>2</sup>A small number of referendums get proposed via other processes, such as a constitutional convention, an appointed commission, or a “veto referendum,” where a petition seeks to override recently approved legislation (Ballotpedia, 2023b).

states require that each referendum only concern a single subject.) State requirements for initiatives vary in terms of the number of petition signatures required, the submission deadline, criteria for referendum content, and filing costs (National Conference of State Legislatures, 2022). For all referendums, states vary in the size of a majority required for passage and whether the legislature can alter an approved referendum. These varied costs enable inference about the distribution of willingness to pay for a referendum and whether reducing political frictions reduces persistence.

A final feature of referendums worth noting is that historical documentation of proposed referendums is typically much richer and more systematic than that for proposed laws. Because referendums go before the voters, secretaries of state typically produce pamphlets describing what each proposed law would do, often including the exact text. Moreover, states typically retain the election results from past referendums more consistently than they do legislative votes. This makes referendums a useful arena for the study of policy persistence.

## 2.2 State-Level Referendum Data

The data collection begins from a dataset of referendums compiled by the National Conference of State Legislatures, supplemented by data from Ballotpedia. It includes the topic, the referendum outcome, and details on the type of referendum. To understand policies' evolution, I work with a team of Upwork assistants to manually collect statutory histories for a subset of proposed laws from the year of each law's corresponding referendum through 2022.

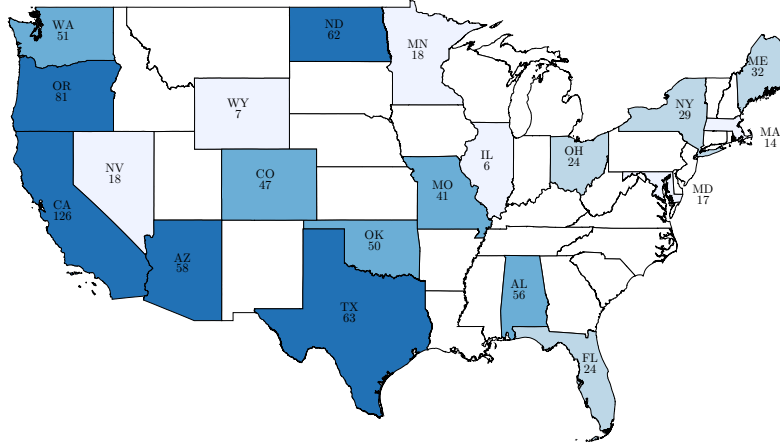
### 2.2.1 Full Universe of Referendums

To arrive at the dataset of close referendums required for causal inference, I begin with a larger dataset of all state-level referendums.

The National Conference of State Legislatures (NCSL) dataset provides the primary covariates of interest for most measures. The NCSL is a nonpartisan, quasigovernmental organization that collects and provides information about U.S. state legislatures. The NCSL has compiled a dataset of 7,772 referendums from states' websites that includes each referendum's name, year, brief description, election type (general, primary, or special), proposal type (legislative referral, initiative, veto referendum, or other), whether the referendum passed, and the vote share up to two decimal places. The vote percentage and indicator for passage serve as my independent variables in the main specification. NCSL also includes a set of 39 non-exclusive indicator variables for the referendum's topic, such as local government or taxation, which I use to understand the heterogeneity of persistence. NCSL's dataset covers most but not all referendums, since some referendums are not available in the state archives NCSL consults but are available in newspaper archives.

To cover referendums not in the NCSL dataset, I scrape the nonprofit website Ballotpedia,

Figure 1: Distribution of Sampled Referendums by State



*Notes:* The figure shows the number of referendums from each state in the final sample, with states selected for availability of documentation and institutional variation. For each included state, the sample covers all referendums since 1900 within a 2.5% bandwidth that have sufficient documentation to observe policies' evolution over time. Darker shades indicate states with more referendums in the sample. The sample covers each of the U.S.'s main regions.

which contains most of the same information. Ballotpedia is a civic education organization that offers resources on elections, policies, and how to vote. Ballotpedia increases the total number of non-duplicate referendums in our dataset by about 7.5%.<sup>3</sup> Ballotpedia does not record the topic variables that are present in the NCSL dataset, so I record these by hand for the referendums I use in the analysis. Ballotpedia also serves as an important source of historical information for the full sample.

### 2.2.2 Subsample for Further Data Collection

From the initial sample, I select a subsample for manual data collection that is narrow enough to enable causal identification but broad enough to capture a wide array of referendums. I select 20 states for consideration based on the volume of referendums, availability of historical documentation (regarding both referendums and subsequent law), and institutional variation (e.g., supermajority requirements and existence of initiative processes). The numbers of referendums from each state appear in Figure 1. I select 1900 as a cut-off date because many states' referendum processes only begin around the turn of the century, and the quality and completeness of the historical records

---

<sup>3</sup>To identify measures that are already in the NCSL data, I drop any measure on Ballotpedia that matches a NCSL measure by vote share and year. My data checks find that this does not falsely identify any duplicates. I then prune any remaining duplicates (since there are occasional small errors in reported vote share in either source) during data collection.

deteriorate as one goes back in time.

### 2.2.3 Bandwidth Selection

To focus on collecting data on close referendums, I choose a bandwidth to define the sample before data collection. I select a very conservative bandwidth of 2.5 percentage points and perform a variety of empirical tests to confirm the validity of estimates with this bandwidth.

From the selected states, I select all those referendums between 1900 and 2020 where the share of votes in favor differed from the threshold by no more than a 2.5 percentage-point bandwidth (e.g., support between 47.5% and 52.5% for a 50% threshold). Selecting a bandwidth for a regression discontinuity (RD) design typically involves a bias-variance tradeoff, where wider bandwidths reduce variance by allowing in more observations. My constraints allow me to collect outcome data from approximately 800 referendums, so even a 1.5 percentage-point bandwidth would be wide enough to contain my desired sample size. Given the overhead cost of understanding each state’s referendum process and records, I opt for the slightly wider 2.5 percentage-point bandwidth to increase the number of observations per state. This bandwidth is quite conservative, around one quarter of the typical bandwidths used in electoral RD designs (Song, 2018). Appendix Section A.1 details empirical tests which confirm that this bandwidth yields unbiased measures of persistence.

### 2.2.4 Data Collection

Measuring persistence requires original data collection to track the histories of policy changes proposed in close referendums. To understand how persistence varies with policies’ characteristics, I gather original and pre-existing data on covariates of practical and theoretical interest.

The primary outcomes of interest require manually tracing state law pertaining to a given referendum, which a team of history graduate students on Upwork carries out for each state using one of three alternative methods. The first method is a review of each historical amendment to a given article or section of a state’s constitution. It is only possible for those states that list all such amendments. The second method, “manual binary search”, applies to states that maintain an online record of the relevant historical statutes or constitution. The first step is to make a binary comparison between the current law on the referendum’s topic and the original law prior to or modified by the referendum in question. If these two differ, we check the historical statutes and constitution in the intervening years to find each point at which the law changed.<sup>4</sup> The third

---

<sup>4</sup>In many cases, the law changes multiple times. We continue the search process until we identify a series of changes that appears continuous. To verify that we do not omit any changes, we search secondary sources and reference materials for any time period in which a law or amendment could have changed without appearing in the records we have.

method, which we use when the other methods are not possible, is to simply review secondary sources and reference materials. Appendix A.2 provides more details on each method included in the data entry instructions for the Upwork assistants.

For each referendum, we record whether the proposed law is operative and in what form using a structured process. For each referendum and each year from the election through 2022, we construct a vector of outcomes that indicate whether the proposed law is adopted and in what form. We note, first, whether the exact text of the law is on the books (i.e., “operative”) as formulated. However, this outcome is overly literal: it ignores substantively similar laws that pass later with different wording and overestimates the frequency of repeal. For that reason, we construct indicators for each of the following outcomes:

- A) Each component of the proposed law is fully operative.
- B) At least one component of the proposed law is operative, possibly in a weaker form.
- C) At least one component of the proposed law is operative in a strictly stronger form.
- D) For at least one component of the proposed law, an opposite law is operative.

Note that B, C, and D can all hold simultaneously, as can A and B. Though opposite laws are rare, partial and stronger versions of a proposed law are common, and any measure of whether a policy is operative will need to allow for these possibilities if it is not overly literal.

To illustrate the different outcomes, consider a referendum to establish a personal income tax at a rate of 1.25%. Outcome A will be equal to one if and only if there is a personal income tax at a rate of at least 1.25% that is in no way legally weaker than what the referendum proposed. Outcome B will be equal to one if and only if there is a personal income tax but at a rate of less than 1.25% or with some restrictions that were not in the initial referendum. Outcome C will be equal to one if there is a personal income tax rate above 1.25% or the tax has been expanded, for instance to cover corporate income. Finally, the fourth outcome will be equal to one if there has been a constitutional amendment to explicitly prohibit an income tax or there is an income credit against other taxes.

One challenge in constructing these outcomes is that there may be ambiguity about what should qualify as operative, weaker, stronger, or opposite. For example, a proposed income tax of 1.25% that remains in place twenty years later, but with a slightly different enforcement mechanism, might be stronger in one sense but weaker in another. For this reason, we classify laws as weaker, stronger, or opposite for particular reasons; there are numeric reasons (e.g., a tax rate is even higher than the proposed one), substantive reasons (e.g., enforcement is stricter), or legal reasons (e.g., the law is

harder to repeal).<sup>5</sup> In addition, we record one version of each outcome that is “narrow” in the sense that we only consider the particular part of the law or constitution that the referendum sought to modify and another version that is “holistic” in that it considers any related law in that state. Related to this last distinction, we also record a version that considers superseding federal laws. I adopt the holistic version as the main specification and employ the others as robustness checks.

I validate the outcome data by contracting an independent audit of the outcome data and, separately, by comparing my outcome data with similar existing data from Caughey and Warshaw (2016). For the independent audit, an Upwork assistant not previously involved in this project’s data entry records each outcome for a random 10% subsample for the year 2022. The auditor’s entry matches my data for the four outcomes listed above (using the holistic version of each outcome) at rates of 87%, 86%, 74%, and 79%, respectively. In addition, around twenty referendums in my dataset affect policies whose enactment Caughey and Warshaw (2016) track in a dataset on state-level policy variation.<sup>6</sup> Appendix Figure E2 shows the match rate, which averages around a stable 80%, with most mismatches resulting from differing definitions in the two datasets.<sup>7</sup>

To understand which sorts of policies are most persistent and why, I collect additional data on referendums’ baseline characteristics, the history of repeat and repeat attempts, and news coverage. The Upwork assistants construct, and I manually review, subjective indicators including whether a policy is left- or right-leaning, concerns a numeric parameter (e.g., a tax rate), concerns a nominal dollar amount, concerns an inflation-adjusted dollar amount, would eventually become obsolete, or would be costly to reverse. I construct subjective categorical variables for how large the impact of the policy seems, what share of the population the policy would directly affect, and what share the policy would indirectly affect. Sample instructions for entering these variables are available in Appendix A.3. I obtain existing data on historical voter turnout in primary and general elections (McDonald, 2002; Clubb et al., 2005; Leip, 2002), campaign spending (Open Secrets, 2022), and news coverage from the API for the Library of Congress’s “Chronicling America” database of pre-1963 historical newspapers.<sup>8</sup> Finally, Upwork assistants identify earlier and later referendums on

---

<sup>5</sup>Where relevant, we also record whether in our subjective judgment an outcome is only true as a matter of technicality (e.g., the only component of the law that is missing is unimportant).

<sup>6</sup>Importantly, the variables in Caughey and Warshaw (2016) are not directly comparable because Caughey and Warshaw (2016) observe whether a single binary policy is in place, while my data record more detail on the structure and level of the policy. For example, a referendum to enact a property tax ceiling will be coded as operative in my data only if the ceiling is in place at the exact level proposed but operative in Caughey and Warshaw (2016) if it is operative at any level. For this reason, my measure of whether a weaker version of the proposed policy is in place should also be comparable, imperfectly, to Caughey and Warshaw (2016).

<sup>7</sup>Appendix Figure E3 shows estimates of persistence using the Caughey and Warshaw (2016) data. The estimates are noisy given the small sample but consistent with the estimates presented in Section 3.

<sup>8</sup>The Library of Congress database covers 296 of the referendums in my sample. From each referendum’s description, I select the two words that are most distinctive from that referendum’s description using the commonly-used



the same policy in order to understand referendum timing and the evolution of voter preferences.<sup>9</sup>

## 2.3 Measuring Persistence for State-Level Referendums

I use a regression discontinuity panel design to measure the effect of narrowly passing a law on whether it is operative at a later date, defined above as persistence. Intuitively, estimating the discontinuity at the threshold for passage captures the way a pivotal voter’s yes-or-no choice affects the policy over time, and additional analysis tests various mechanisms and implications of the observed dynamics.

My estimating equation takes the form of a local linear regression:

$$Y_{rt} = \beta_{0t} + \beta_{1t}\mathbb{1}\{v_r > \tau\} + \beta_{2t}(v_r - \tau) + \beta_{3t}(v_r - \tau) \times \mathbb{1}\{v_r > \tau\} + \varepsilon_{rt} \quad (1)$$

where  $Y_{rt}$  is an outcome  $t$  years after referendum  $r$ , typically whether the law, or a version of it, is operative (as defined in Section 2.2);  $v_r$  is the vote share the referendum received; and  $\tau$  is the share of votes required for passage.  $\beta_{1t}$  is the coefficient of interest and captures the effect of passage on the relevant outcome  $t$  years later. That is,  $\beta_{1t}$  measures persistence. Intuitively,  $\beta_{1t}$  captures the effect of passage at the point that  $v_r = \tau$ , estimated by the jump in  $Y_{rt}$  when we go from just below to just over the threshold. The linear vote share terms, which capture any first-order correlation between a law’s baseline popularity and later history, minimizes bias and overfitting relative to the commonly used higher-order polynomials (Imbens and Lemieux, 2008; Gelman and Imbens, 2019).

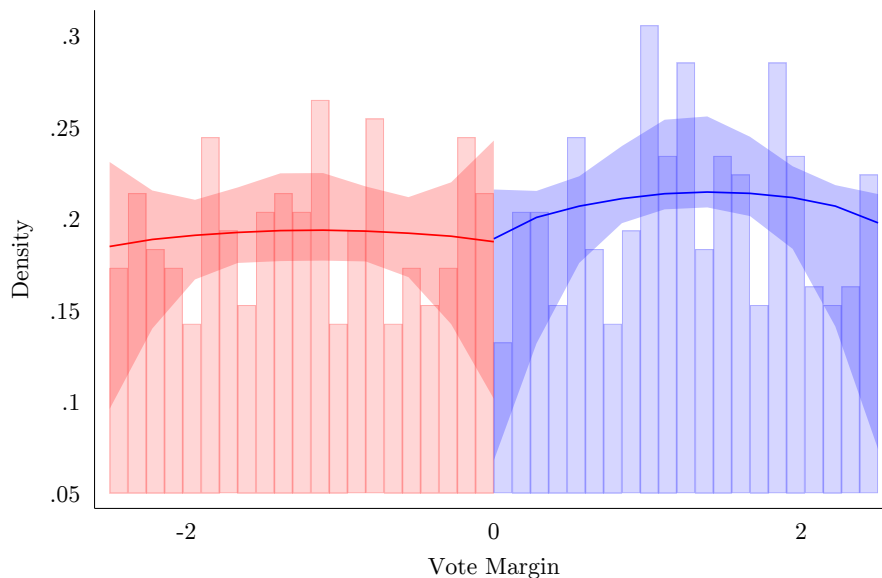
Since I observe outcomes for most laws through the present day, I face tradeoffs associated with a panel dataset. My main specification is an unbalanced panel where, for each  $t$  years, I include all those referendums that happened  $t$  years ago. Data are missing for fewer than 5% of the total sample. I include balanced panels in Appendix E for comparison. I also offer separate plots of persistence for older and for more recent referendums in Appendix E. In all cases, I cluster standard errors at the level of the state and topic because multiple referendums on the same topic share a common outcome.

---

term frequency-inverse document frequency measure (Kelly et al., 2021; Koffi, 2021). I then search the database for the number of newspaper articles including those terms and the name of the state in a given time period relative to the referendum in question, and I linearly detrend the number of articles using a linear time trend because the number of articles is strongly decreasing over time. See Appendix Figure E4 for a plot of the strongly linear time trend.

<sup>9</sup>An important challenge in identifying attempts to revisit a referendum is that if we use our outcome data to identify such attempts (e.g., by finding each referendum that modified the policy), we will identify only successful attempts. This would bias our measure of voters’ preferences and would generally prevent us from observing cases that had repeated failed proposals. Instead, we search for the full NCSL and Ballotpedia universe for the words in each of our sampled referendums’ descriptions, dropping articles and generic terms like ”amendment.” This method is almost certainly incomplete, but it should accurately characterize the timing of attempts to revisit the same topic and the votes on such attempts.

Figure 2: Vote Share is Smoothly Distributed around the Threshold for Approval



*Notes:* The figure shows the distribution of vote shares relative to the threshold for referendums in the sample as well as a polynomial fit to the distribution. Each bar shows the density (i.e., percentage of observations divided by the width) for an equally-sized bin (around 0.10 percentage points). A local quadratic polynomial fits the density on either side. If the identifying assumption is violated, the fitted polynomials should differ at zero, which they do not.

The identifying assumption is that referendums that pass and fail are identical in the limit as we approach the threshold for passage. I assume that any difference between the predicted likelihood that a policy is operative if it passes or if it fails exactly at the threshold is attributable to its having passed. Conceptually, the most obvious way for the identifying assumption not to hold would be if proponents fine-tune the referendum texts months in advance or their election-cycle advertisements to secure a bare majority. For this to be plausible, proponents would need to be able to distinguish minute differences in support, but even much-better-pollled presidential elections have average polling errors of about 2–6 percentage points (Jennings and Wlezien, 2018). One previously posited confounder in the electoral-RD literature is that a state’s dominant party may skew close House elections, but Eggers et al. (2015) show empirically and conceptually that this is likely due to chance. Empirically, a violation of the identifying assumption would imply bunching just above the threshold, which Figure 3 rules out, and imbalance across baseline covariates, which Appendix Table D1 rules out.

Two alternative ways to study persistence are measuring the hazard rate and measuring the effect of proposing a referendum under a set of additional assumptions. The hazard rate is intuitively

interesting because in a world where some share of policy changes are effectively permanent, the hazard rate should approach zero over time. I measure a hazard rate by building on equation (1) and taking the rate of change of the coefficient of interest  $t$  years after the referendum:

$$\log\left(\frac{\beta_{1t+1}}{\beta_{1t}}\right) \quad (2)$$

I calculate standard errors using the delta method. To measure the effect of proposing a referendum on later policy, I describe a method in Appendix Section A.3.1 that uses a recursive algorithm developed by Cellini et al. (2010). This strategy estimates the effect of the choice to campaign for a policy change under the assumption that the effect of passing a policy change is constant over time.

### 2.3.1 Heterogeneity

To better understand what drives persistence, I also look at the heterogeneity of effects by topics, institutional features, and relevant aspects of the election by estimating equations of the following form:

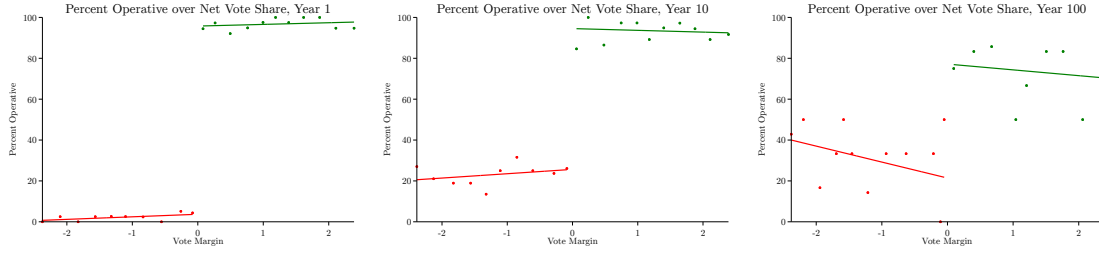
$$Y_{rt} = \beta_{0,t} + \beta_{1t} \mathbb{1}\{v_r > \tau\} + \theta_{0t} X_{rst} + \theta_{1t} X_{st} \times \mathbb{1}\{v_r > \tau\} + W_{s0} + \varepsilon_{rt} \quad (3)$$

where  $Y_{rt}$  are the outcomes described in Section 2.2;  $X_{rst}$  are covariates for the relevant referendum  $r$  in state  $s$  and year  $t$ , such as the number of voters, the topic of the referendum, or whether the legislature can alter the approved law; and  $W_{rs0}$  are baseline control variables. Here the coefficient of interest is the possibly vector-valued  $\theta_{1t}$ , which measures the difference in the effect of passage when we change the value of covariates. An important difference from equation (1) is that here I drop the running variable  $v_r - \tau$  to preserve statistical power, which requires the stronger identifying assumption that passed and failed referendums are identical within the bandwidth even away from the threshold. As we will see in the next section, the running variable has no significant effect on the outcomes.<sup>10</sup> Dropping it in the main specification narrows the error bars while leaving the point estimates essentially unchanged. When interacting passage with covariates, dropping the running variable avoids overfitting.

---

<sup>10</sup>Appendix Table D2 shows covariate balance between passed and failed referendums without the vote share term controlled for in Appendix Table D1. Two variables, whether a referendum is proposed by a legislature and whether it is proposed at a general election, do differ between passed and failed referendums when we do not control for vote share. The difference is small in practical terms and relative to measures of persistence. Appendix Figure E5 shows that there is no discontinuity in these variables at the threshold. I include these variables as controls when analyzing heterogeneity.

Figure 3: Discontinuity at the Threshold—Narrowly Approved Laws Are More Likely to Be in Place



*Notes:* Each panel shows a binscatter and lines of best fit for the share of referendums operative over the vote share in favor minus the threshold (typically 50%) within the 2.5 percentage-point bandwidth. The lines have slopes indistinguishable from zero and indicate no relationship between vote share and the likelihood a law is in place above and below the threshold. The effect of passage is estimated by the jump between the lines at zero, which is sizable even 100 years later. At 100 years post-referendum, the sample size is small, with each dot representing only four observations.

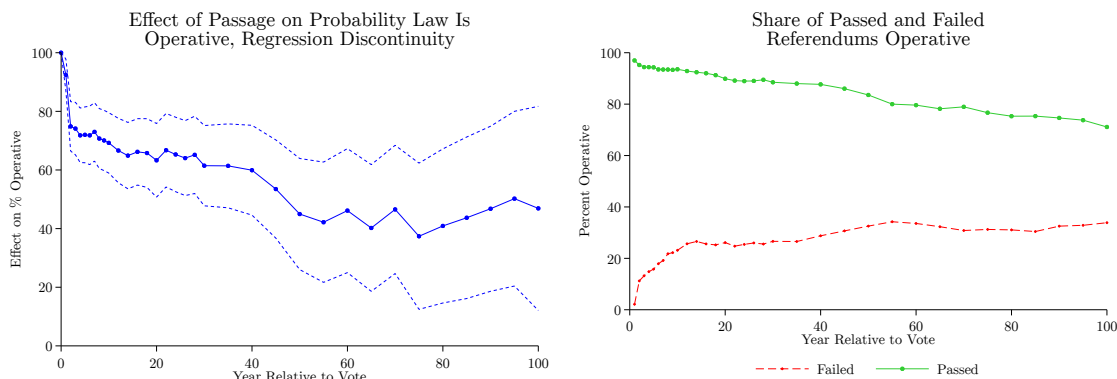
### 3 U.S. State Referendums: Main Results

Applying my estimation approach to the statutory histories, I find that approving a referendum increases the chance the policy is operative 10, 20, 50, and even 100 years later by over forty percentage points. The effect is robust to different data collection procedures and similar across topics and political backgrounds. Intuitively, this persistence indicates a substantial and lasting mismatch between policy and voters’ preferences if there is even a moderate degree of drift in voters’ attitudes over time.

Persistence, measured as the effect of passage on whether a law is operative later, declines rapidly in the first few years following a referendum before nearing a steady forty-percentage point level. Figure 3 plots the share of referendums that are in place as a function of the vote share at three points in time. The gap between the failed and passed referendums is sizable 5, 10, and even 100 years later, with the relationship between the vote share and the shares of referendums that are operative essentially flat. Figure 4 and Appendix Table D5 present the primary measure of persistence over time, capturing the difference between these two lines at zero for the holistic outcome described in Section 2.2. Many policy choices do not persist for even the first few years after the referendum, but after the first few years, persistence nearly flattens. Appendix Figure E6 shows the declining rate of change of the curve in Figure 4. Appendix Figure E20 estimates the persistent effect of a successful policy campaign following a strategy in Appendix Section A.3.1, which assumes that the same set of referendum proponents would never again pass the policy except for the current referendum.

The estimated effects are robust to the way I define the outcome. In Appendix Figure E7, I present versions of Figure 4 using four other definitions of whether a law is operative (described in

Figure 4: Persistence Over Time—Approval Has a Long, Stable Effect on Whether a Law Is in Place



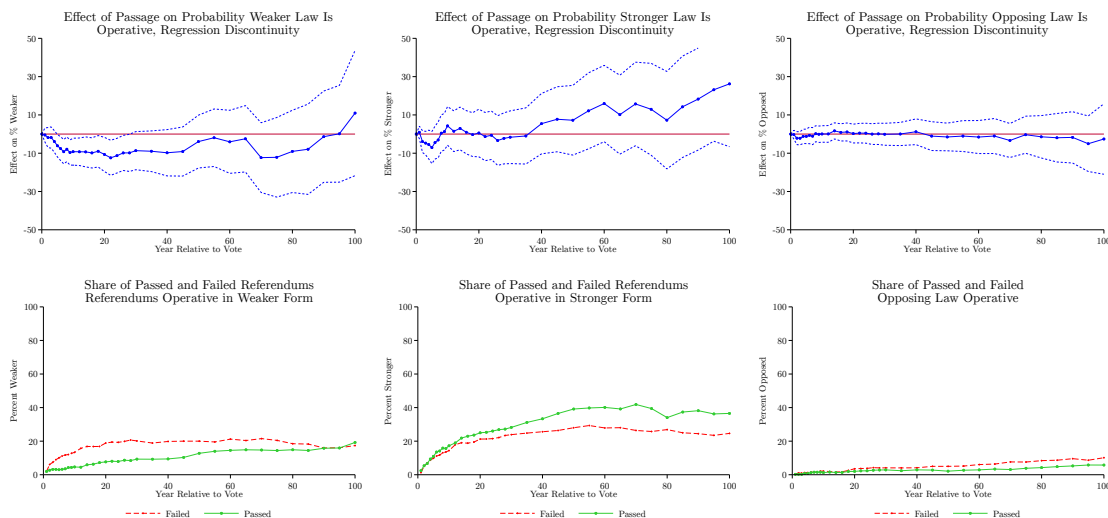
*Notes:* This plot shows the effect of passing a policy or law on whether it is operative later on. The left panel shows the effect of passage on whether a policy is operative later in percentage points, estimated by equation (1) in an unbalanced panel. Dotted lines signify the 95% confidence interval with standard errors at the state-topic level. For purposes of exposition, the graph shows coefficients every year for the first ten, every two years for the next twenty, and then every five years. The right panel shows the percent of narrowly passed and failed referendums that are operative over time. Many failed policies pass in the first few decades, after which persistence plateaus. See Appendix Table D5 for point estimates. These figures use the holistic classification of the law described in Section 2.2, considering any numeric, substantive, or legal changes that weaken the law.

Section 2.2). The only version that yields a clearly different pattern is when I define “operative” in a strictly literal sense. Under that approach, far fewer failed referendums ever pass, and successful referendums get much more steadily repealed. This is to be expected given that this outcome treats even slightly modified laws as no longer operative. This pattern confirms that the data collection process is regularly picking up modifications of the law throughout the entire sample frame.

The pattern of persistence remains the same when I use a balanced panel or restrict the time period I consider. In a balanced panel (Appendix Figure E8, top left), I find nearly the exact same pattern as in Figure 4. I also examine whether persistence has changed over time by comparing older and more recent referendums (Appendix Figure E8, top right). The effects nearly coincide, indicating that the mechanisms driving persistence have not fundamentally changed. The histories of repeals and repeat attempts look nearly identical over time (Appendix Figure E8, bottom panels). Dropping referendums that would plausibly become obsolete or that involve a nominal dollar value leaves persistence unchanged or slightly greater (Appendix Figure E9).

Passing a referendum has suggestive intensive-margin effects: it appears to increase the strength of the law that is ultimately put in place. Figure 5 shows the estimated effect of passage from (1) where the outcome is that a weaker (left), stronger (middle), or opposite (right) version of the proposed law is operative. In the left column, we see that proponents often propose weaker versions

Figure 5: Intensive Margin—Initial Approval Has a Small Effect on the Law’s Ultimate Strength



*Notes:* The top panel in each column shows, from left to right, the effect of passage on whether a weaker version of a law is in place, whether a stronger version is in place, and whether an opposing law is in place. The estimated effects are equal to the coefficient on passage in equation (1) with outcome indicators described in Section 2.2. The bottom panels show the percent of passed and failed referendums for which the outcome is equal to one. Appendix Figure E13 presents alternative versions of the lefthand panels estimated in a simple difference equation. Passing a referendum makes it less likely that a weaker version of that law is adopted in the following few decades, and in the simple difference specification, passage increases the probability of having a stronger version later on.

of failed referendums in the following few years, but this does not weaken the law in the long run. The middle column shows that over the longer term, around 35% of passed laws are reenacted in a stronger form, compared to only 25% of initially rejected laws.<sup>11</sup> In the third column, we see no indication that passing a law either triggers or prevents a policy backlash, despite evidence that cultural backlash to social policies is quite common (Wheaton, 2020). However, for all of these three outcomes, we observe only changes in state law on the specific topic of the referendum, so there could very plausibly be additional effects on policies not observed.

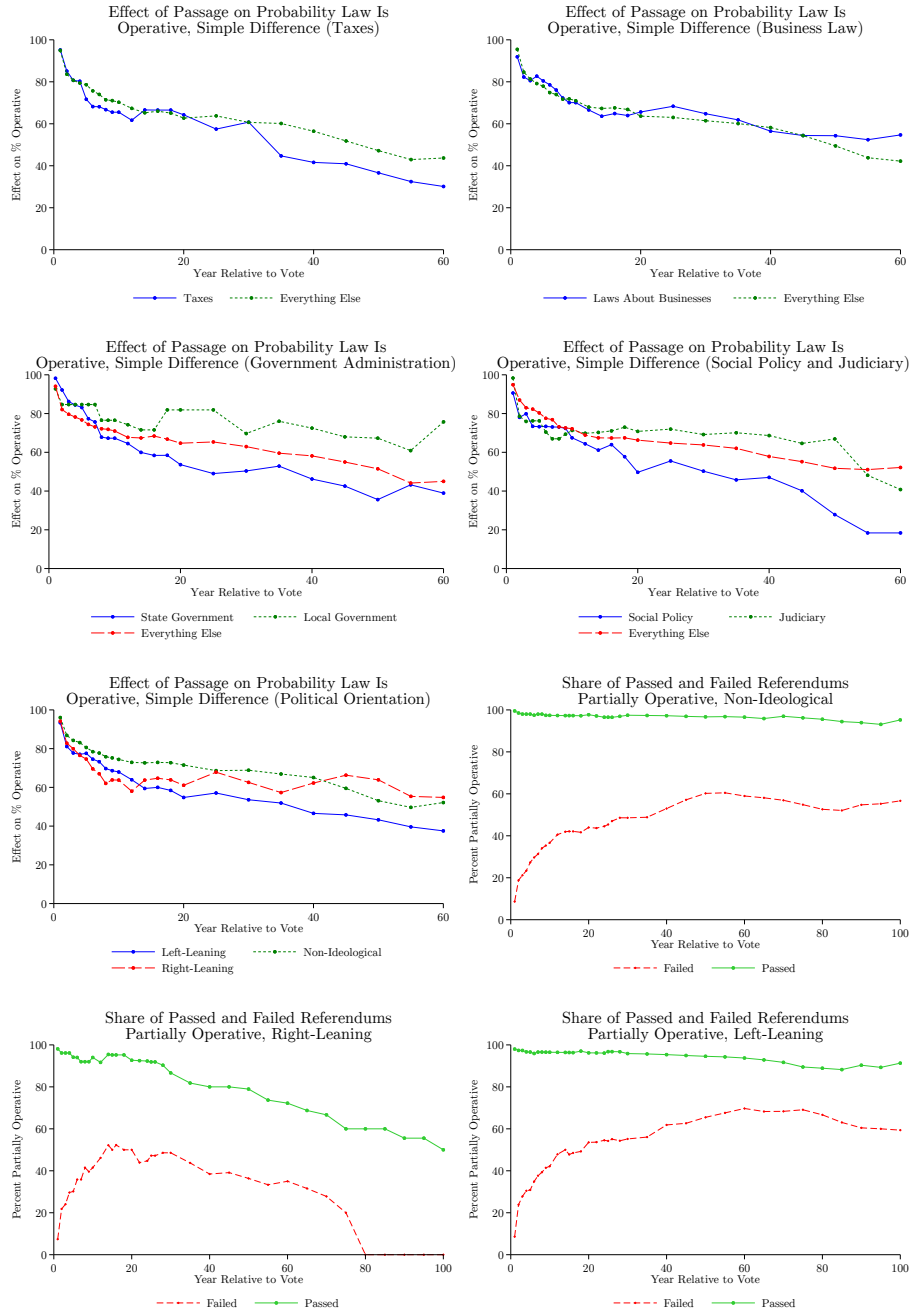
Figure 6 shows that persistence is similar across a diverse set of policy areas.<sup>12</sup> Though there are slight quantitative differences, few are significant, and the shape of each curve is broadly similar. Some curves that do differ significantly are social policy, which is somewhat less persistent, and local government, which is somewhat more persistent than the rest.<sup>13</sup> The lack of a heterogeneous

<sup>11</sup>The coefficients are not significant in the regression discontinuity design but are significant for the middle of the sample period in the simple difference design shown in Appendix Figure E13.

<sup>12</sup>See Appendix Figure E10 for figures with confidence intervals.

<sup>13</sup>Appendix Figure E12 shows heterogeneity for tax policy. Rejected taxes often pass in a weaker form later, leading the effect of passage on whether some version of a tax is in place to approach zero.

Figure 6: Heterogeneity—Persistence Is Similar across Topics and Political Orientations



*Notes:* Each curve plots the persistence of close referendums by baseline characteristics, estimated by equation (3). Appendix Figure E10 shows standard errors, Appendix Figure E11 shows the coefficients when we include the running variable (vote share) in the regression, and Appendix Figure E12 shows the effect on whether any version of the policy is operative, including a weaker version. There is little heterogeneity across most dimensions.

effect between right-leaning, left-leaning, and non-ideological referendums is noteworthy because the political dynamics for left- and right-leaning referendums appear to differ substantially (Figure 6, bottom row). Fewer successful right-leaning referendums remain in place after 100 years than left-leaning ones, yet this is exactly compensated by fewer unsuccessful ones passing later.<sup>14</sup> This shows that persistence does not appear to depend on the specific ideological content of the referendum and raises the question of what generates persistence.

## 4 Theoretical Framework and Mechanisms

To delineate some potential causes and initial implications of persistence, I propose a game in which advocates can pay costs to propose laws to voters. Varying my assumptions about how advocates' and voters' payoffs and beliefs evolve reveals a key fact and two possible mechanisms of persistence. The key fact is that in a world where current voters' preferences determine policy, modeled as the simplest dynamic extension of the canonical (i.e., Downsian) model of democracy, persistence is either nonexistent or follows a specific empirical pattern inconsistent with the results in Section 3. The first possible mechanism relaxes this simple model by allowing advocates' or voters' policy views to be path dependent. The second possible mechanism adds a feature to the model that allows the salience of a policy to decline after it is voted on because of a regression-to-the-mean dynamic. Note that both mechanisms could contribute to persistence, as could others not considered here. Empirical patterns support the second mechanism, which implies that many policies are not what current voters or policy advocates would choose.

### 4.1 Setup

I model the time path of policies, referendum proposals, and votes as a repeated game between a proponent ( $a_1$ ) and opponent ( $a_0$ ) of some policy  $p_t \in \{0, 1\}$  where  $t$  denotes the years post-referendum. After describing the structure of the game, I formally define persistence as the difference in the likelihood that a policy is operative after a given number of years as a function of whether it is operative after an initial referendum.

The proponent and opponent represent the collection of actors who propose a referendum, and I refer to them as “advocates.” In the case of referendums by petition, an advocate is typically an interest group. In the case of legislative referrals, the proponent is the legislature or a faction

---

<sup>14</sup>Another dimension of heterogeneity that might be of interest is supermajority requirements. Though few states require a supermajority to pass referendums, Appendix Figure E15 shows that the pattern is similar for referendums subject to a supermajority requirement. There may be somewhat less persistence in the first few years, seemingly because failed referendums' proponents try again more in supermajority states.



within the legislature, and the opponent is an opposing interest group or faction in the legislature. As in these examples, each advocate may represent a collection of individuals. The defining feature of  $a_1$  and  $a_0$  is that  $a_1$  prefers  $p_t = 1$ , and  $a_0$  prefers  $p_t = 0$ . In each period  $t$ , either advocate  $a_x$  may pay a given cost  $c_{xt}$  to propose a referendum that would change the policy from  $p_{t-1}$  to  $1 - p_{t-1}$ . Let  $\theta_{xt} \in \{0, 1\}$  be equal to 1 if  $a_x$  proposes a referendum in period  $t$ .

Voting occurs in the following way. There is a unit mass of voters with single-peaked preferences. For each period, a share of voters  $v_t$  prefers policy  $p_t = 1$ , with  $v_t$  common knowledge. In each referendum, there is an independent and identically distributed horizontal shift to voters' preferences of  $\varepsilon_t \sim F_\varepsilon$  (e.g., last-minute news, weather, or unknown aspects of voters' preferences), yielding a final vote share in favor of  $p_t = 1$  of  $v_t + \varepsilon_t$ . If the vote share in favor exceeds a threshold  $\tau_t \in [0, 1]$  (e.g.,  $\tau = 0.5$  for a simple majority), the referendum passes, and  $p_t = 1 - p_{t-1}$ ; otherwise  $p_t = p_{t-1}$ .  $v_t$  and  $\tau_t$  are random variables drawn at the start of period  $t$  with possibly degenerate distributions and the only requirement that  $\mathbb{E}_{v_t}[F_\varepsilon(\tau_t - v_t)]$  and  $\mathbb{E}_{v_t}[F_\varepsilon(v_t - \tau_t)]$  are well-defined.

Advocates receive benefits from their preferred policies. Advocate  $a_x$  receives a payoff  $b_{xt}$ , which is a random variable but will generally be taken to have a degenerate distribution (i.e., to be constant). The CDF of  $b_{xt}$  is strictly decreasing in  $p_t$  for  $x = 1$  and increasing for  $x = 0$  so that the advocate benefits from their preferred policy.<sup>15</sup> Let the game be infinitely repeated, and advocates discount their payoffs in period  $t$  by some factor  $\delta$ . The advocates' payoffs capture the welfare impacts of the policy on those who participate in the policy process, while the vote share  $v_t + \varepsilon_t$  captures the share of the population that benefits from the policy according to their own preferences.

The timing of the game from period 0 on is as follows:

1. The game begins with some policy  $p_{-1} \in P$ .
2.  $b_{0t}$ ,  $b_{1t}$ , and  $v_t$  are drawn from some distribution.
3. Both advocates choose whether to propose a referendum and the policy to propose.
4. If either advocate proposes a referendum, there is an election. Voters rank the policies.
5. If there is an election and  $v_t + \varepsilon_t > \tau_t$ ,  $p_t = 1 - p_{t-1}$ ; otherwise  $p_t = p_{t-1}$ .
6. The game repeats ad infinitum in the next period starting from step 2.
7. Advocate  $a_x$  gets the total payoff  $\sum_{t=0}^{\infty} \delta^t (b_{xt} \mathbb{1}\{p_t = x\} - c_{xt} \theta_{xt})$ .

---

<sup>15</sup>For simplicity, I do not consider proposal costs as a random variable. Any dynamics that could arise from random variation in costs can instead be captured by assuming random variation in  $b_{xt}$  in the proposal period.

Finally, define persistence of a policy  $\hat{p}$  in period  $t$  as follows, letting  $H_t$  denote the full history of state variables and actions prior through period  $t$ :

$$P_t = \mathbb{P}[p_t = 1 | p_0 = 1, H_t] - \mathbb{P}[p_t = 1 | p_0 = 0, H_t]$$

Persistence, which I define consistently with prior literature (Coate and Morris, 1999), is the difference in the likelihood that a policy is operative in a given period if it was operative after the first period compared to if it was not, holding all other context and policy features constant. Persistence is counterfactual: it is a comparison of how long the policy is operative in two different worlds, one where it initially passes and one where it does not. It is perfectly possible to have a world with no persistence but where most policies last for a very long time. In such a world, the lack of persistence would imply that whether a policy is operative is deterministic rather than contingent.

## 4.2 Persistence in an Idealized Democracy

Persistence is generally either nonexistent or, if the distribution of advocate and voter beliefs and preferences is stationary, short-lived. Intuitively, the stationary-preference world captures the simplest extension of the benchmark model of democracy from Downs (1957), which assumes no systematic trend in advocates' or voters' policy interest following a policy change.

The starkest result we obtain is that, when proposing a referendum is costless, and the required threshold is a simple majority, there is no persistence. This holds because if anybody wants to change the law and has even the slightest possibility of winning, they will propose to change it.

PROPOSITION 1 (frictionless democracy): *Suppose the following conditions hold for every  $t, t' < t, x$ :*

(1)  $v_t, \varepsilon_t \perp p_{t'}, v_{t'}, \varepsilon_{t'}, \theta_{xt'}$

(2)  $\tau_t = 0.5$

(3)  $c_{xt} = 0$

(4)  $b_{xt} = b_x$  is constant over time.

(5) All parameters except  $\varepsilon_t$  are common knowledge at the time of referendum proposals.

Then  $P_t = 0$  in any Nash equilibrium.

All proofs are in the appendix, but I offer an intuitive explanation here. The first and second assumptions guarantee that a referendum outcome is independent of the history to date. The third, fourth, and fifth assumptions imply that whenever a referendum could pass, someone will propose a referendum.<sup>16</sup> The inevitability of a referendum means that the policy operative at the

---

<sup>16</sup>Intuitively, the fourth condition is necessary to avoid a world where the two advocates coordinate so that they

end of a period must always be the policy that would win if there were a referendum, which is independent of the history and, therefore, the initial policy.

Introducing frictions can allow for persistence, but the next results show that persistence should decline steadily in a basic Downsian model. Obviously, proposing a referendum is not entirely costless. Either a legislature needs to coordinate a vote or an interest group needs to secure signatures on a petition, with any number of potential additional restrictions on legislative referrals and petition signatures. For this reason, the next proposition introduces frictions but keeps strategies and the relevant variables history independent. We define a new object,  $-\frac{1}{t} \log P_t$ , which is the hazard rate of persistence, or the rate at which the relevance of past policy choices declines. The Markov perfect equilibrium concept limits the set of equilibria to the set of subgame perfect equilibria where strategies depend only on the current state and not the prior history.

**PROPOSITION 2 (stationary democracy):** *Suppose the following conditions hold for every  $t, t' < t, x$ :*

(1) *All variables except  $\varepsilon_t$  are constant over time:  $(v_t, \tau_t, b_{xt}, c_{xt}) = (v, \tau, b_x, c_x)$ .*

(2) *All variables except  $\varepsilon_t$  are common knowledge at the time of referendum proposals.*

*Then  $\exists q \in [0, 1] : P_t = q^t$  in any Markov perfect equilibrium. In any equilibrium where a referendum occurs in the first period,  $q < 1$ .*

Proposition 2 states that persistence declines at a constant hazard rate in a standard model, which follows almost immediately from the definition of a Markov strategy. Persistence in period  $t$  is a function of how likely it is that the policy changes between period 0 and period  $t$ . In a Markov perfect equilibrium, the likelihood of an advocate proposing depends only on the state variables. If the state variables are constant, the likelihood of there being a referendum will therefore be constant.<sup>17</sup> If the threshold and vote share are also constant, it follows that the likelihood of a policy change in a given period will be constant. The average likelihood of a policy change in a

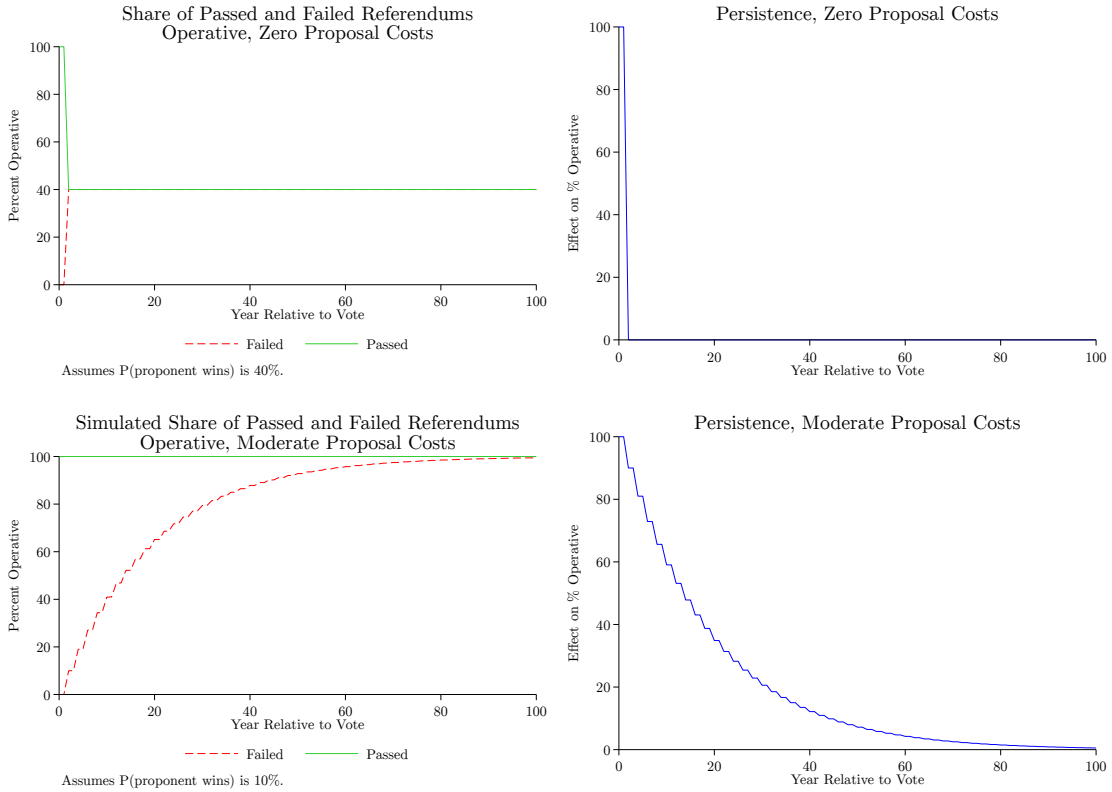
---

each get their preferred policy in some periods where their relative payoff from it is especially high. A way this would be falsified is if two conflicting interest groups coordinated so that one group had their preferred policy in years when it was especially important to them. This condition should be fairly uncontroversial; even if the payoffs did vary somewhat over time, such a degree of coordination seems unlikely.

Note that we might expect policy preferences captured by  $b_{xt}$  to vary somewhat over time, as when legislatures change hands. The result from Proposition 1 can also be derived from the slightly weaker condition that the distribution of  $b_{xt}$  is stationary if we adopt Markov perfect equilibrium as our equilibrium concept. The equilibrium restriction in this case is again necessary to avoid coordination between opposing advocates to get their preferred policies when their preferences are strongest.

<sup>17</sup>Note that the condition Proposition 2 relies on is actually weaker than the first condition here, similarly to Proposition 1. The distribution of the state variables need only be stationary, not constant, for the proof to work. I use the stronger assumption here for simplicity.

Figure 7: Simulated Policy Histories—Little or No Persistence in an Idealized Democracy



*Notes:* This figure simulates policy histories in a simple model of democracy with a two-year election cycle. The upper-left panel shows the likelihood of a policy being operative over time if it initially passes or fails, assuming the proponent has 40% odds of winning a referendum in each period. The upper-right panel displays the measure of persistence, which is the difference between the two lines on the left. The lower-left and right panels are similar but assume that costs are high enough to deter an opponent, and the odds of a successful repeat are lower. There is no persistence in the upper case and a constant hazard rate in the lower case.

given period is therefore constant, generating the result.

A Markov perfect equilibrium is intuitively likely since a non-Markov equilibrium would require a coordinated dance between the two opposing advocates. Specifically, a non-Markov equilibrium would require the advocates to have coordinated expectations concerning play that vary based on the prior history of the game even when the state variables are the same. For example, in a non-Markov equilibrium, a referendum proponent or opponent might propose a referendum because the other side had proposed one previously even if the state of the world is the same in every payoff-relevant way. Opponents of tax limits, for example, would need to elect not to attempt to repeal a tax limit because of the amount of time proponents had waited between successive

attempts. The sort of scenario necessary for a non-Markov equilibrium is implausible. The scenario is especially implausible because, as we saw in Section 3, attempts to repeal previously approved policies are rare.

Propositions 1 and 2 provide empirical signatures of a world where the policy steadily converges toward the policy preferred by voters or advocates. An influential line of argument claims that where there is any degree of rational optimizing, there should be little path dependence because any deviation from the optimum should be remedied quickly (Pierson, 2004; David, 2007; Liebowitz and Margolis, 2014). Relatedly, some work argues that short-term changes in voters' or advocates' interests will exert little lasting influence (Downs, 1972; Edelman, 1985). Proposition 1 formalizes this convergence argument without frictions, while Proposition 1 offers a weaker version when there are frictions.

The results in Section 3 are at odds with Propositions 1 and 2, suggesting an empirical puzzle and a substantial mismatch between policy and voters' preferences.<sup>18</sup> Figure 7 illustrates the dynamic characterized by Propositions 1 and 2 in a simulation of the average time path of passed and failed referendums within the model in this section, assuming a probability of later passage similar to what the data show.<sup>19</sup> The picture is strikingly different from Figure 4, and Appendix Figure E6 rejects the constant hazard rate implied by Proposition 2. A key mystery is why, while few successful referendums ever get repealed, so many attempted referendums never end up in place. Moreover, that policy persistence is so much greater than the democratic benchmark would predict, even after 100 years, indicates that our political process results in policies that are far from what voters want.

---

<sup>18</sup>It is important to rule out a mechanical issue when examining persistence across a collection of policy changes. Specifically, persistence may decline heterogeneously for different policy changes (i.e., the quantity  $q$  in Proposition 2 may vary across policy changes). This will mean that persistence does not decline exponentially on average even if persistence declines exponentially for each policy change individually (Follmann and Goldberg, 1988; Weitzman, 1998). To see this, consider two policies for which Proposition 2 holds with respective values for  $q$  of  $q_1$  and  $q_2$  where  $q_1 > q_2$ . Persistence in period  $t$  will equal  $(q_1^t + q_2^t)/2$ . Observe that  $((q_1^t + q_2^t)/2)^{1/t}$  is strictly increasing in  $t$  and approaches  $q_1$  in the limit. This can be accounted for with empirical analyses of the frequency of attempts to revisit policy changes conditional on their not having been successfully revisited.

Appendix Figure E19 rules out this mechanical issue. In Appendix Figure E19, I plot the distribution of repeat attempts over time, i.e., the main driver of the right panel of Figure 4. I restrict the figure only to those referendums that do not pass within 25 years and show that even when conditional on not passing, any attempts to revisit a referendum are concentrated in the immediate aftermath.

<sup>19</sup>Appendix Figure E16 depicts the time path for a policy where advocates' payoffs are stochastic but stationary, and Appendix Figure E17 shows results for additional repetition frequencies.

### 4.3 Persistence and Endogenous Policy Views

Prominent explanations for policy persistence in economic theory involve endogenous responses of advocates or voters to the prior history of policies and referendums. After formally describing this mechanism and its empirical implications, I conduct a series of empirical tests that suggest endogenous responses do not easily explain which policies are most persistent.

#### 4.3.1 How Endogenous Policy Views Can Generate Persistence

Theories of endogenous policy propose that people become increasingly supportive of successful policies the longer they are operative or increasingly wary of failed policies with each attempt to revisit a referendum. This assumption can explain arbitrarily high long-term persistence and generates testable implications.

Theoretical models for why advocates and voters may be counterfactually more supportive of policies once they are operative differ in their details but share a high-level structure. Possible models include endogenous cultivation of a preference for the status quo (Bernheim et al., 2021), fixed adaptation costs to new policies (Coate and Morris, 1999), and endogenous changes in the distribution of political power (Acemoglu et al., 2021).<sup>20</sup> Formally, we can consider a setting where the cumulative distribution function (CDF) of the proponent’s policy payoff  $b_{1t}$  is decreasing, and the opponent’s payoff  $b_{0t}$  increasing, in  $\iota$  periods’ prior policies. That is,  $\forall b_{1t}, F_{b_{1t}}(b_{1t}|p_{t-1}, p_{t-2}, \dots, p_{t-\iota})$  is decreasing in each of  $p_{t-1}, p_{t-2}, \dots, p_{t-\iota}$ , and  $b_{0t}$  similarly. Since an advocate will not propose when their payoff is sufficiently low, and will propose when their payoff is high, sufficiently elastic payoffs with response to prior policy can create indefinite persistence even following an initial referendum. There can also be indefinite persistence if voters’ attitudes are endogenous, with  $v_t(p_{t-1}, p_{t-2}, \dots, p_{t-\iota})$  a decreasing function of each of  $p_{t-1}, p_{t-2}, \dots, p_{t-\iota}$ , and for some intermediate range of vote shares, proponents and opponents do not attempt referendums.

As an initial observation, theories of endogenous policy views do not obviously fit the rate of repeal and repeat attempts over time. Section 3 documents a slowing down of repeat attempts over time, but repeal attempts happen at a steady pace. Endogenous policy views generally do not predict a slowdown in attempts to repeat a referendum by proponents because a failed referendum leaves the policy history unchanged from the prior period.<sup>21</sup> If  $p_{t-1} = \dots = p_{t-\iota} = 0$ , then when

---

<sup>20</sup>Another theoretical possibility is migration. That is, voters move to jurisdictions where they agree with the policies, resulting in a shift in voter preferences toward the status quo. I do not discuss this in depth because (i) a migration response large enough to generate persistence across such a broad range of policies seems implausible, and (ii) a migration mechanism would typically result in similar empirical features to a preference for the status quo.

<sup>21</sup>Theories in Coate and Morris (1999) and Bernheim et al. (2021) support this implication. A generalization of the theory in Fernandez and Rodrik (1991), which relates persistence to uncertainty about policies’ impacts, would also yield a similar implication as voters and advocates gradually lose their memory of the former policy’s impacts.

$p_t = 0$ , the policy history for the previous  $t$  periods is unchanged in period  $t+1$ .<sup>22</sup> While endogenous policy views are consistent with this empirical pattern, they do not easily explain it.

Specific theories of endogenous policy views yield empirically testable hypotheses. Theories that relate persistence to specific types of policy, such as economic policy (Coate and Morris, 1999) or political reforms (Acemoglu et al., 2021) would predict greater persistence for these areas, which is the first hypothesis I test below:

HYPOTHESIS 1: There will be greater persistence for economic policy or for political reforms than for other policies.

Theories that predict endogenous policy views based on the difficulty of reversing policy predict another form of heterogeneity (Coate and Morris, 1999):

HYPOTHESIS 2: There will be greater persistence when policy reversal is more costly.

If persistence is a product of endogenous changes in voters' attitudes, voters will become more sympathetic to policies once they are in place:

HYPOTHESIS 3: Voters will become more sympathetic to policies after they pass, and less sympathetic after they fail.

Finally, policy views could be endogenous not to the policy history but rather to the political history, with voters or advocates increasingly averse to proposing a policy the more times it has been proposed before:

HYPOTHESIS 4: There will be fewer attempts to repeat a referendum after observing more recent votes on the topic.

Endogenous policy views are practically important because they add complexity to the welfare implications of persistence, though they do not necessarily change the conclusion that persistence implies a substantial mismatch between policy and voters' preferences. Specifically, if voters' preferences change to favor the policy that is in place, then there are unresolved normative issues since welfare analysis typically assumes a fixed set of preferences (Bernheim and Taubinsky, 2018; Bernheim et al., 2021). If policy views change in response to economic incentives or altered political structures, it may be that, while the prevailing policy is not what voters would choose *ex ante*, it is

---

<sup>22</sup>Two cases in which this might not hold are when a policy is responding to a recent change or when the goal of a policy is to prevent future changes (e.g., preventing the expansion of government). Despite this distinction, there is no more persistence in the case of policies that respond to novel crises (Appendix Figure E23) or that are conservative (Figure 6).

not worth it to switch. Persistence mediated by endogenous responses to prior policies would imply that which policy is optimal could depend, at least in part, on which policy is already in place, which raises novel normative questions.

### 4.3.2 Empirical Evidence on Endogenous Policy Views

The empirical data do not support the above four hypotheses, indicating that endogenous policy views do not predict which policies are most persistent and do not readily fit the data.

There is little heterogeneity across the axes predicted in Hypotheses 1 and 2. In Figure 6, the only policy area where we see greater persistence to a statistically and practically significant degree consistent with Hypothesis 1 is local government but not state government, the judiciary, taxes, or business law. Appendix Figure E23 shows that, contrary to Hypothesis 2, policies we classify as costly to reverse are no more persistent than others.<sup>23</sup> In the data, the ability of policies to trigger endogenous responses does not seem to predict persistence.

The data do not imply an effect of approving a referendum on subsequent voter support for that policy (Hypothesis 3), but I am unable to rule out some effect. The right panel of Figure 8 shows the average vote share in favor of a policy in each decade in which we identify a vote on that policy as described in Section 2.2.4. For successful referendums, this is generally the vote share against attempted repeals. For failed ones, it is the vote share in favor of subsequent repeat attempts. There is no significant difference between the two lines, though the confidence interval is large. This comes with an important caveat: the average observed vote share years after failed and passed referendums should be selected upward and downward, respectively, relative to the average vote share across all referendums. This is because repeat and repeal attempts are respectively proposed by proponents and opponents of the policy in question.<sup>24</sup> Given this caveat, it is not possible to reject some positive effect of passing a referendum on subsequent voter support despite the lack of empirical confirmation.

I test Hypothesis 4 by exploiting the feature that some states prohibit nearly identical referendums within a certain window (typically between one and five years). Under Hypothesis 4, we should see an uptick in attempts to repeat a referendum after this window passes, and the states should slowly but eventually catch up to the other states in the share of rejected referendums that later pass. Figure 8 displays the share of rejected referendums that pass over time in states with and without repeat limits.<sup>25</sup> While there is significant noise from the small number of observations,

---

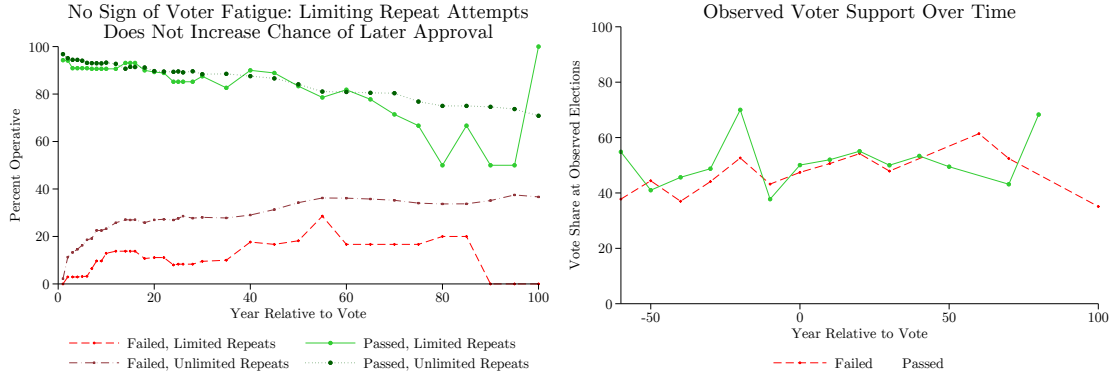
<sup>23</sup>See Section 2.2.4 and Appendix A.3 for details on how we construct this variable.

<sup>24</sup>This might seem like a natural setting to correct for selection via the Heckman correction (Heckman, 1979). In particular, this selection issue should differ with costs across different states, which can be used to fit a model of selection. This method simply confirms there is too little to say here: using the Heckman correction, the estimated effect of passage on later vote share spans dozens of percentage points.

<sup>25</sup>Appendix Figure E24 gives more detail on this heterogeneity. Comparing the hazard rates immediately after



Figure 8: No Evidence of Endogenous Responses to Policy or Politics



*Notes:* The left panel shows the share of passed and failed referendums that are operative over time for states with and without limits on the ability to repeat a referendum one to five years later, analogously to the right panel of Figure 4. If, per Hypothesis 4, voters or policy proposers tire of attempts to revisit a referendum, states with such limits should eventually catch up, with the two dashed red lines eventually converging. The right panel shows the average vote for referendums in each decade on repeal or repeat attempts. The difference between the two lines offers a lower bound on the effect of passing a policy on voter support. There is no sign of an effect, but since the data are only a lower bound, I cannot rule one out. The vote share is roughly flat over time.

the data suggest that these states do not catch up and that reactions to votes are not the primary driver of persistence.

While the data do not offer evidence of endogenous responses to prior policy, it is worth noting the limitations of what this data can tell us. The patterns documented thus far are informative about what sorts of policies tend to persist at a high level, but given the broad diversity of topics, there is insufficient statistical power to interrogate the patterns for specific policy topics (e.g., pensions or tax breaks to industries) in depth. As such, there may be important endogenous dynamics at play for specific topics that do not determine the high-level pattern. There also could be endogenous responses to prior policy that are homogenous across issue areas and not related to reversal costs. The data do indicate, however, that in thinking about which policies are persistent at a high level, stories about endogenous responses to policy are not the best explanation.

#### 4.4 Persistence and Issue Salience

An alternative explanation for policy persistence is that an issue’s salience tends to decline after that issue is voted on for exogenous reasons, a dynamic characterized in less general terms by Baumgartner and Jones (2010) as “punctuated equilibrium,” a notion borrowed from evolutionary

---

the limit on repeats is lifted in states with repeats and those without reveals no “catch up” effect. There is some evidence of substitution from exact repeat attempts toward passing weaker versions of the policy.

biology (Gould and Gould, 2009).<sup>26</sup> I propose a model, with support in the data, in which interest in a policy follows an autoregressive process, and the cost of proposing a policy change selects for policy proposals to occur when policies' salience is abnormally high.

#### 4.4.1 How Variation in Issue Salience Can Generate Persistence

Policy changes can be arbitrarily persistent if brief periods of elevated political interest punctuate long periods of time with little change, to paraphrase Baumgartner and Jones (2010). In my proposed model, political interest in a policy follows an autoregressive process, and the cost of proposing a policy change selects for policy proposals with abnormally high interest that fades away with time.

The issue salience explanation for persistence that I propose requires two simple but crucial assumptions. The first is that policy interest follows an autoregressive process. Consider a setting, for example, where  $b_{1t} = \phi b_{1,t-1} + (1 - \phi)\bar{b}_1 + \eta_t$  for some long-term average  $a_1$  policy payoff  $\bar{b}_1$  and proposal cost  $c_1$ .<sup>27</sup> Let  $\frac{1}{1-\phi}\bar{b}_1 < c_1$ ,  $\phi < 1$  and  $\eta_t \sim F_\eta$  be iid with  $\mathbb{E}[\eta_t] = 0$ . If  $\bar{b}_1$  is lower than  $c_1$ , then for any  $b_{1t} > c_1$ , the likelihood that  $b_{1t'} > c_1$  will decline over time for  $t' > t$ . This happens without any endogenous response simply because the proposal occurs when the advocate is especially motivated. Motivations might fluctuate over time, for instance, because of shocks to economic conditions, social movements, or the media landscape.

The second crucial assumption for this issue-salience explanation of persistence is that proposing a policy must always be sufficiently costly to deter most referendums people would want to propose, an assumption that can also explain the low rate of repeals. This assumption follows from the observations that elections and legal changes are expensive, and there is a large space of possible policies anyone could propose (e.g., laws restricting the consumption of raw onions by customer service agents). For this reason, costs must be high enough to deter the vast majority of possible policy changes. States erect such barriers by requiring a minimum number of signatures and additional geographic, temporal, financial, and legal requirements for petitions. For legislative referrals, some states have supermajority requirements or require passage at multiple legislatures. In addition to generating a specific dynamic pattern I describe below, this assumption can endogenously generate the relatively low rate of repeals we observe in the data, as shown in Appendix B.1.

---

<sup>26</sup>Mokyr (1990) argues that technological change also exhibits a pattern of punctuated equilibrium.

<sup>27</sup>For simplicity, I focus on the case where  $b_{1t}$  is autoregressive and keep  $b_{0t}$  constant given that the low rate of repeals indicates that movement in  $b_{0t}$  is relatively unimportant. It is of course possible to obtain the same dynamic in this section with some evolution in  $b_{0t}$ , via an autoregressive or other process, but the movement in  $b_{0t}$  must be somewhat restricted. Because a higher value of  $b_{0t}$  will increase the threat of repeal,  $b_{0t}$ , policy proposals by  $a_1$  select for a high  $b_{1t}$  and a low  $b_{0t}$ . If  $b_{0t}$  fluctuates enough in absolute terms, this can lead to the opposite dynamic from what we observe in the data: an increasing hazard rate rather than a declining one. The low rate of repeals suggests that  $b_{0t}$  should not play a major role in proponents' calculus.

Fluctuating issue salience will yield arbitrarily greater persistence than in the model that assumes the distribution of advocates' interest in a topic is stationary. Proposal costs select for proposed policies whose popularity and interest are abnormally high, so mere mean regression implies fewer repeat attempts as time passes. The next result formalizes this effect by assuming  $b_{1t}$  evolves according to a vector autoregression. I take  $b_{1t}$  to capture overall interest in a policy, consisting of both the policy's impact and its salience. I note, however, that both the punctuated equilibrium literature and research on voters' policy attitudes indicate that salience matters more than (potentially indirect) impact (Baumgartner and Jones, 2010; Dal Bó et al., 2018).

**PROPOSITION 3** (punctuated equilibrium): *Suppose the following conditions hold for every  $t, t' < t, x$ :*

- (1)  $(v_t, \tau_t, c_{xt}, b_{0t}) = (v, \tau, c_x, b_0)$  are constant over time.
- (2)  $b_{1t} = \phi b_{1,t-1} + (1 - \phi)b + \eta_t$  where  $\frac{1}{1-\delta}b < c$ ,  $\phi < 1$  and  $\eta_t \sim F_\eta$  is iid with  $\mathbb{E}[\eta_t] = 0$ .
- (3)  $c_0 > \sum_{t=0}^{\infty} \delta^t b_0$
- (4)  $a_1$  proposes a referendum in period 0 ( $\theta_{10} = 1$ ).

Then  $\forall t > 1, P_t > P_1^t$ . In addition, the following hold:

- (A)  $\mathbb{E}[b_{1,t}] < b_{1,0}$ .
- (B) Let  $b_{1,-t}$  denote  $a_1$ 's payoff from  $p_{-t}$  in the  $t^{\text{th}}$  period before period 0. If  $F_\eta$  is in the exponential family of distributions with support on an open subset of  $\mathbb{R}$ , then  $\forall t, \mathbb{E}[b_{1,-t}] < b_{1,0}$
- (C) For a given  $v, \tau, c_x, \phi, b_0, b_{10}$ , persistence  $P_t$  is increasing in  $\bar{b}_1$  for all  $t$ .

The key insight of Proposition 3 is that with fluctuating political interest, the hazard rate need not steadily decline to explain the data. The crucial assumption in Proposition 3 is the second one, which tells us that policy preferences vary over time but regress to the mean. The second assumption also requires that, in general, proposing is not worthwhile, with the stricter third assumption motivated by the above argument that the opponents do not propose.<sup>28</sup> Together, these conditions imply that proponents will attempt a referendum less frequently with time, and for a suitable choice of parameters, policy choices can persist for arbitrarily long periods after the initial period of elevated interest. It is worth noting that a similar dynamic can occur if voter preferences  $v_t$  follow an autoregressive process, with  $b_{0t}$  and  $b_{1t}$  fixed, and  $v_t$  having mean  $\hat{v}$ ,  $(1 - F_\varepsilon(\tau - \hat{v}))\frac{1}{1-\delta}\bar{b}_1 < c_1$ , and  $F_\varepsilon(\tau - \hat{v})\frac{1}{1-\delta}b_0 < c_0$ .<sup>29</sup> Intuitively, if advocates only propose when they have sufficient voter sympathy, then we might see proposals only at periods of unusual voter

<sup>28</sup>I make the latter assumption because is necessary to avoid a scenario where referendum proposals select for times of especially low interest on the other side, leading repeal attempts increase over time. A similar result would follow if we assumed that attempts to repeal a referendum correlate with attempts to repeat one.

<sup>29</sup>Specifically, in order to ensure this result, it must be that the advocate who is not the initial proposer is not much more likely than the initial proposer to have a successful referendum. If they were, then repeal attempts would

enthusiasm, and whatever happens in these periods of unusual interest sticks.

Proposition 3 also yields two distinct empirical predictions. The first conclusions,  $A$  and  $B$ , say that there should be a hump shape in the level of interest in a particular policy around a referendum:<sup>30</sup>

HYPOTHESIS 5: Interest in a given policy will peak around attempts to change the policy.

My chief measure of interest over time is news reports, which intuitively capture a composite of a policy’s impact and its salience. The second prediction ( $B$ ) is that when the average level of interest in a policy is higher, there should be less persistence, which occurs because attempts to revisit the policy are more sustained.

HYPOTHESIS 6: Policy choices will be less persistent for topics with more sustained interest.

In line with my discussion of Hypothesis 5, I consider Hypothesis 6 to have support if either more important or more salient policies are less persistent than others.

The basic dynamic proposed here fits with the notion of punctuated equilibrium proposed by Baumgartner and Jones (2010) but is more formal and general. Baumgartner and Jones (2010) show that policymaking in the U.S. and other countries resembles a pattern of long periods of incremental change punctuated by occasional major spurts of activity (Baumgartner et al., 2006, 2009). Baumgartner and Jones (2010) characterize punctuated equilibrium as arising out of the division of responsibility in policymaking (among “policy subsystems”) with varying understandings of the policy (“policy images”). Proposition 3 shows that a version of punctuated equilibrium should arise under much more minimal conditions and for a far broader set of policy reforms than previously recognized, though temporal variation in political interest remains important in both models.

#### 4.4.2 Issue Salience: A Punctuated Equilibrium

Empirical patterns broadly support the connection between issue salience and persistence. In particular, news reports, policies’ salience, and advocates’ idiosyncratic preferences appear to fluctuate over time in a way that causes a single policy choice to matter for decades.

---

increase over time as voter support approached its new normal, similar to the issue Proposition 3 rules out with assumption three.

<sup>30</sup>The assumption on the distribution of  $F_\eta$  is a weak one necessary to rule out an unusual distribution where observing a positive shock implies a negative update on the prior value of  $b_{1,-t}$ . The exponential family includes virtually any plausible distribution on  $F_\eta$ .

News coverage of the topic of a referendum exhibits a peak over time, consistent with Hypothesis 5. In the left panel of Figure 9, we see that the volume of news articles related to a referendum description’s keywords peaks in the years before a referendum and then declines.<sup>31</sup> Appendix Figure E19 shows that attempts to change a policy are heavily concentrated in the ten years immediately before and after a referendum, consistent with a sharp spike in political interest.<sup>32</sup> There is no noteworthy evolution in voters’ support over time, indicating that it is primarily advocates, not voters, who drive this punctuated equilibrium dynamic.

There is some evidence that policies with greater salience are less persistent, as Hypothesis 6 predicts. The right panel of Figure 9 shows persistence by the share of the population that would observe the policy in question being carried out. Specifically, the lower line shows estimated persistence for policies we determine affect a population at least as large as an entire industry, demographic group, region, or city.<sup>33</sup> For this subset of referendums, persistence declines steadily toward a 100-year effect of 20 percentage points without a clear plateau. Appendix Figure E21 shows a similar pattern for referendums whose topics were the subject of greater news coverage prior to the referendum campaign.<sup>34</sup> At the same time, two measures of interest that relate less to a policy’s salience—the share of the population indirectly affected and the share of voters at an election who vote on the referendum—do not predict persistence (Appendix Figure E21).<sup>35</sup> It appears to be salience, and not impact, that predicts persistence, as neither estimates of fiscal impacts nor subjective indicators of a policy’s impact predict persistence (Appendix Figure E22).

Washington and Oregon’s divergence on income taxation described in the introduction illustrates the punctuated equilibrium mechanism. In the early twentieth century, Washington state had three income tax referendums in a ten-year window—1932, 1938, and 1942—and Oregon had ten referendums in a ten-year window: 1922, 1923, 1924, 1926, 1926, 1926, 1927, 1928, 1930, 1932 (Ballotpedia, 2023b). After these respective periods, neither state had referendums at anything approaching this frequency (Ballotpedia, 2023b). These similar timings were not coincidental: the

---

<sup>31</sup>Appendix Figure E19 shows that the evolution is similar for passed and failed referendums, confirming that this peak is exogenous to the policy. Both figures resemble the media patterns Baumgartner and Jones (2010) document for a handful of policy issues, with news typically peaking shortly before a policy change.

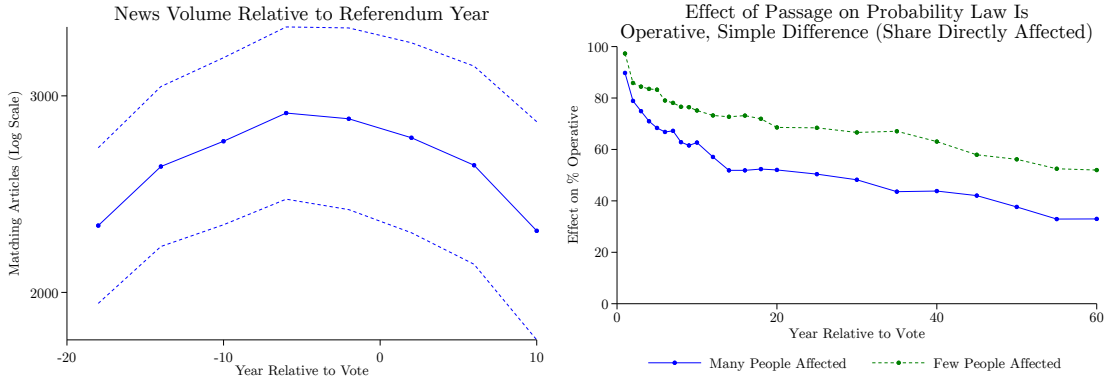
<sup>32</sup>Baumgartner et al. (2018) characterize a punctuated equilibrium as a dynamic where the distribution of policy changes has a far heavier tail than the normal distribution, with a high kurtosis value. The distribution of repeat attempts over time in Appendix Figure E19 has a kurtosis of 9.51, nearly three times that of a normal distribution.

<sup>33</sup>See Section 2.2.4 and Appendix A.3 for details on how we construct this variable.

<sup>34</sup>Specifically, motivated by Proposition 3 (B), I consider the difference between the logarithm of the number of related articles in years 20-1 before the vote and the number in years 0-3.

<sup>35</sup>Theoretical, observational, and experimental evidence indicate that a lack of information is the primary driver of decisions to abstain from a ballot item (Feddersen and Pesendorfer, 1996; Wattenberg et al., 2000; Battaglini et al., 2008). As such, the share of voters who vote on a referendum is a measure of how many voters heard about the referendum and how many voters sought out information about it.

Figure 9: Variation in Issue Salience Can Explain Persistence



*Notes:* The left panel displays the logarithm of the number of news articles in the Library of Congress’s “Chronicling America” historical news database that match selected keywords for four-year bins before and after the referendum year, as described in Section 2.2.4. A rise and fall in news around a referendum is consistent with policy changes happening when issue salience is especially high. The right panel shows persistence for policies with above- and below-median shares of the population affected by them. Policies that are more salient, measured by the number of people who directly observe them, are more persistent.

United States had recently adopted a federal income tax, and the Progressive Era and New Deal added pressure for new and more progressive forms of taxation (National Archives, 2022). The status of each state’s income tax at the end of this punctuated period remained the same until today because interest in the income tax did not return to its earlier peak.

## 5 Implications

The widespread persistence of policy choices bears important implications both for policy and for econometric methods. Using both a calibrated model and an intuitive argument, I show that persistence implies a substantial and lasting mismatch between prevailing policy and popular preferences. In addition, persistence lends support to long-term policy event study designs.

### 5.1 Implications for Policy with a Calibrated Model

Persistence implies that there is a mismatch between optimal and actual policy. I demonstrate this under weak assumptions via an intuitive argument. I then calibrate a version of my game-theoretic model with data on initiatives to demonstrate the scale of this divergence as well as the effects of potential institutional reforms.

Intuitively, persistence implies a mismatch between actual and optimal policy because a policy

may remain in place—while another, equally or more desirable policy is not—for long stretches of time. For simplicity, I assume that current voters’ attitudes and social welfare are independent of the policy history, consistent with the evidence presented above.<sup>36</sup> Consider any criterion for selecting the optimal policy from a binary policy space which is independent of prior election outcomes. The requirement that the criterion be independent of prior election outcomes is more general than it might seem: any criterion that depends only on the present state, including majority rule, egalitarianism, or utilitarianism, would satisfy it. For a referendum with roughly a 50% chance of passing, one-half of persistence ( $P_t$  as defined in Section 4.1) is a lower bound on the probability that the policy’s status is not optimal after  $t$  years.<sup>37</sup> Given the results in Section 3, for at least 20% of marginal referendums, the policy does not match the optimum for any history-independent criterion.

To measure the scale of this policy mismatch, I calibrate my model of the policymaking process from Section 4.1 using data on initiatives (i.e., referendums proposed by petition). I focus on initiatives because petition signature requirements offer an observable source of variation in the cost of proposing a referendum. I add to the model from Section 4.1 the following assumptions:

1. In each state, there are  $N$  binary policies.
2. Policymakers are perfectly myopic:  $\delta = 0$
3.  $\log b_{1t} = \phi_b \log b_{1,t-1} + (1 - \phi_b) \log \bar{b}_1 + \eta_t$ , where  $\eta_t \sim \mathcal{N}(0, \sigma_\eta^2)$  is iid., and  $\phi_b \in [0, 1]$
4.  $\log \bar{b}_1 \sim \mathcal{N}(\mu_b, \sigma_b^2)$
5.  $\log b_{0t} = \log b_0 \sim \mathcal{N}(\mu_b, \sigma_b^2 + \sigma_\eta^2)$ , with  $\text{Corr}(\log b_0, \log \bar{b}_1) = \rho$
6.  $v_t = \max(\min(\phi_v \log v_{t-1} + (1 - \phi_v)\bar{v} + \gamma_t, 100), 0)$ , where  $\gamma_t \sim \mathcal{N}(0, \sigma_\gamma^2)$  is iid., and  $\phi_v \in [0, 1]$
7.  $\bar{v} \sim \mathcal{N}(\mu_v, \sigma_v^2)$ .

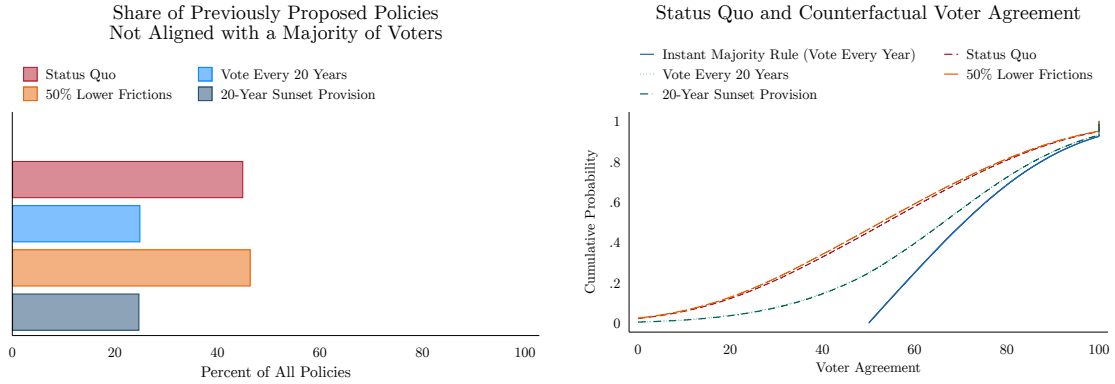
The above yields eleven parameters to estimate:  $N, \phi_b, \sigma_\eta, \mu_b, \sigma_b, \rho, \phi_v, \mu_v, \sigma_v, \sigma_\gamma$ , and  $\sigma_\varepsilon$ . I estimate the parameters via maximum likelihood and a grid search using the number of initiatives (i.e., referendums proposed by petition) and the frequency of repeat and repeal attempts in each state and year within a given vote share range (e.g., 45%–55%). With these assumptions and the model from Section 4.1, the likelihood of observing a given number of initiatives is the expectation of a

---

<sup>36</sup>A version of the argument in this section may very well go through without this assumption. The normative issues involved with endogenous preferences are currently unresolved (Bernheim et al., 2021). If persistence is a result of switching costs, the arguments in this section still imply that the policies in place frequently do not match voters’ *ex ante* preferences, though it may not be worth switching depending on the weight one gives to future periods.

<sup>37</sup>See Appendix B.3 for a formal proof.

Figure 10: Around Half of Policies Do Not Match Voter Preferences



*Notes:* The figure depicts the level of voter agreement with previously proposed policies after 100 years of policymaking in a calibrated model. The left panel shows the share of policies whose status matches voters’ preferences, meaning the share of policies that are currently operative and lack voter support or are not operative and have voter support. The right panel shows the cumulative distribution of voter agreement with policies. Both panels show the status quo and three possible institutional reforms. Requiring votes every 20 years or 20-year sunset provisions increase policy alignment with voters’ preferences, while reducing frictions does not.

binomial  $\text{Bin}(N, p)$  density where  $p$  is the likelihood of a policy change being proposed with observed voter support  $v_t$  in the given vote range. Appendix Section C.1 gives more details on the estimation and shows that the calibration fits key features of the data.

Considering the set of all policies proposed by initiative in a 100-year period, the calibration implies that 45% of the policies’ statuses differ from what voters would choose today. To measure voter agreement with a policy’s status, I take (i) the share of voters who support a policy if the policy is operative or (ii) the share who oppose a policy if the policy is not operative, in the 100th year of the simulation. The left panel of Figure 10 shows the share of policies whose statuses match the views of a majority of voters under the status quo, along with three possible institutional reforms to make policies more voter-responsive. Under the status quo, 47% of policies are not what voters would currently choose. The level of disagreement is not trivial: the average level of voter support for current policies is 54%, only slightly higher than if policies were passed at random (50%) and significantly lower than the average level of agreement if policies were perfectly in line with majority rule (73%). Figure 10 shows the full cumulative distribution function of voter agreement, confirming that the status quo is far from the theoretical maximum.

Resolving the mismatch between policies and voter preferences requires making subtle tradeoffs, and one of the most obvious solutions has only a limited effect. Figure 10 shows that making it easier to change policies does not improve policy mismatch. This is because, while making it easier to change policies increases the likelihood that a given policy gets revisited, it also expands the



playing field and leads to a proliferation of policies that have more marginal levels of support.<sup>38</sup> On the other hand, requiring a vote on a policy or sunseting (automatically repealing) a policy every 20 years *can* reduce the policy mismatch. Appendix Figure C15 captures one benefit of requiring a vote at some frequency rather than adopting sunset provisions: sunset provisions tend to create more instability. Both sunset provisions and, to a lesser degree, periodic votes have the effect of increasing the number of policy changes and leaving fewer policies in place than voters would generally prefer.<sup>39</sup>

The mismatch I observe between policy and voter preferences is attributable largely to the drift in voters' preferences over time for passed policies. In other words, the main driver of the mismatch is that many policies pass during brief moments of popularity but become unpopular later on. Appendix Figure C16 presents the extent of mismatch for policies that are and are not operative, respectively. The share of operative policies that lack voter support is well over 50%, while the share of policies that are not operative but have voter support is below 20%. The institutional reforms that reduce this mismatch do so largely by repealing unpopular policies. In alternative parameterizations in Appendix Figure C18, I show that varying the estimated mean vote share in favor of an initiative ( $\mu_v$ ) or the degree of noise ( $\sigma_\varepsilon$ ) does not qualitatively change the results. Assuming the distribution of voter attitudes is constant over time does reduce the degree of mismatch, implying that persistence combined with drift in voters' preferences over time drives the mismatch in the calibrated model.

## 5.2 Implications for Cost-Benefit Analyses

Policy persistence implies that a large share of a policy's impacts occur after its first decade, which informs how we should evaluate second-best policies. I study the implications of persistence for policies' impacts from the perspective of the pivotal voter, but a similar logic should apply to an advocate using the method described in Appendix Section A.3.1.

Persistence is a key parameter for the long-term impact of a voter's policy choice. Suppose a voter considers whether to support a policy that will produce some output worth  $\delta_v^t o_t \in \mathbb{R}$  for each period  $t$  it is in place. The voter's payoff to the policy passing in period 0 (i.e., the payoff if they are pivotal) is as follows:

---

<sup>38</sup>Consistent with the model, Appendix Figure E28 shows that there is no less persistence in states where policies are harder to revisit or when we lower policy frictions even more.

<sup>39</sup>In line with this feature of sunseting provisions, alternative calibrations of the model in Appendix Figure C18 show that sunseting can actually increase policy mismatch under some circumstances. In contrast, the reduction in mismatch from periodic votes is robust across alternative parameterizations.

$$\mathbb{E}\left[\sum_{t=0}^{\infty} \delta_v^t o_t p_t \mid p_0 = 1\right] - \mathbb{E}\left[\sum_{t=0}^{\infty} \delta_v^t o_t p_t \mid p_0 = 0\right] = \sum_{t=0}^{\infty} \delta_v^t o_t P_t \quad (4)$$

One immediate implication of this framework is that, to the extent that some policies are more persistent than others, voters should be more concerned with the more persistent policies, all else equal. While heterogeneity is limited in the data, the results described in Sections 3 and 4 suggest that choices about reforms to less-salient policies are somewhat more influential than choices about other equally important policies.<sup>40</sup>

The expression in (4) lends itself to thinking about various tradeoffs a voter may face. Consider, for example, a policy that is good in the short term but bad in the long term. Formally, suppose the policy yields a constant outcome  $o$  up until period  $\bar{t} - 1$  but  $-o$  starting in  $\bar{t}$ . The voter will want to support the policy if

$$\sum_{t=0}^{t=\bar{t}-1} \delta_v^t P_t > \sum_{t=\bar{t}}^{t=\infty} \delta_v^t P_t \quad (5)$$

The inequality implies that the more persistent the policy choice, the less the voter will want to support the policy.

I empirically address this normative question of the relative importance of a policy's long-term impacts by constructing a policy duration index in Figure 11. To construct the index, I assume a policy yields a fixed benefit (or cost) in each year, and I estimate the benefits that accrue each year following the passage of the policy discounted at a rate  $\delta$ . Appendix Figure E31 illustrates the index's construction. As Figure 11 shows, the first ten years—a common timeframe for policy analysis—cover less than a third of the policy's total impact under all but the most aggressive discount rate.

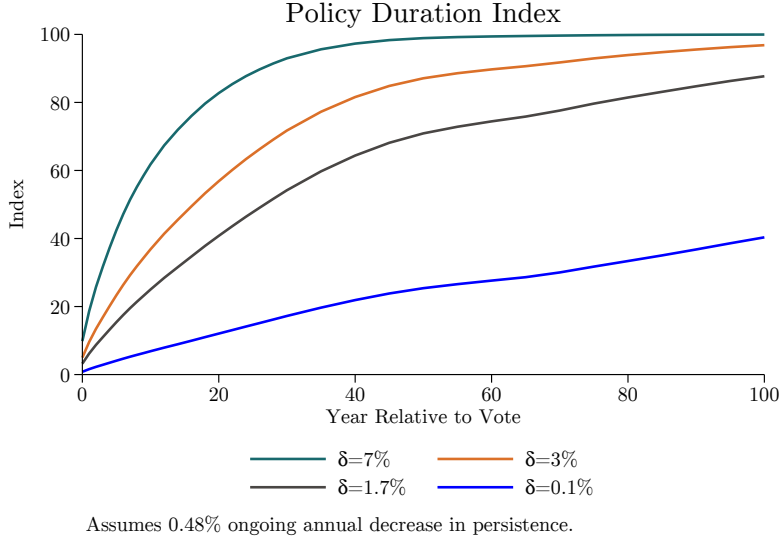
Persistence can inform our understanding of which imperfect policies are welfare improving given political economy constraints. Consider a voter faced with an imperfect version of a policy, such as a carbon tax they support but with designated spending they oppose, or a drug legalization law they support with an objectionable regulatory structure.<sup>41</sup> Formally, suppose a proposed policy yields some constant outcome  $o$ , but the voter believes that, should another version pass later, the

---

<sup>40</sup>The somewhat greater persistence of reforms to local governments and judiciaries is consistent with the long literature in political economy and economic history on the importance of institutions (North, 1991; Acemoglu et al., 2005), though my analysis of the mechanisms underlying persistence suggests that institutions might be persistent because of their low salience rather than their effects on incentives.

<sup>41</sup>Both examples resemble choices voters faced recently. For the first, Washington state had referendums in 2016 and 2018 on carbon taxes, each with different and controversial rebate mechanisms. For the second, the state of Ohio had a cannabis legalization referendum opposed by some pro-legalization groups because of the monopoly it would have granted to certain businesses.

Figure 11: Policy Duration Index—Persistence Implies Policy Choices Have a Long Lifetime



*Notes:* Each line indicates the share of a policy choice’s total absolute impact that accrues by a given year for a policy that has a fixed impact per year. I discount this constant impact by the rates in the legend, which include standard consumption discount rates of 1.7% and 3% (Rennert et al., 2022), an extreme discount rate of 7%, and a “pure rate of time preference” of 0.1% (Nordhaus, 2007). Appendix Figure E31 illustrates the index’s construction. Policy persistence implies that the first 20 years account for only about half of a policy’s effective lifetime under the moderate discount rates.

later version would yield some greater outcome  $o'$ . The voter will support the policy based on whether the following modified version of (4) is positive:

$$\mathbb{E}\left[\sum_{t=0}^{\infty} \delta_v^t o p_t \mid p_0 = 1\right] - \mathbb{E}\left[\sum_{t=0}^{\infty} \delta_v^t o' p_t \mid p_0 = 0\right] = \sum_{t=0}^{\infty} \delta_v^t o P_t - \sum_{t=0}^{\infty} \delta_v^t (o' - o) \mathbb{E}[p_t \mid p_0 = 0] \quad (6)$$

The voter will be more inclined to accept the imperfect policy whenever  $\frac{o' - o}{o} > \frac{\sum_{t=0}^{\infty} \delta_v^t P_t}{\sum_{t=0}^{\infty} \delta_v^t \mathbb{E}[p_t \mid p_0 = 0]}$ . Intuitively, the more persistent the policy choice is, the more the voter will want to take the available option, because saying no could be long-lasting.

The degree of persistence in my sample of close referendums implies that the pivotal voter should generally want to accept an imperfect policy unless they expect a later alternative policy to be twice as valuable. Specifically, we can estimate the ratio  $\frac{\sum_{t=0}^{\infty} \delta_v^t P_t}{\sum_{t=0}^{\infty} \delta_v^t \mathbb{E}[p_t \mid p_0 = 0]}$  under a variety of discount rates. I illustrate this calculation in Appendix Figure E32. The results appear in Appendix Figure D6, as well as the results of the same calculation using theoretical benchmarks in the simple Downsian framework of Section 4.2. At any of the higher discount rates, the voter needs to prefer a 50/50 gamble with the alternative policy in order for it to be in their interest to wait. The high

degree of persistence implies that, at first approximation, it is better to accept an improvement than to wait for a marginally better policy.

### 5.3 Implications for Econometric Methodology

A final implication of persistence is that it supports empirical work that compares geographic units with different policies. Specifically, persistence supports the identification of both cross-sectional analysis and long-term event studies.

The empirical finding of widespread persistence validates a range of designs by demonstrating that quasirandom variation in an initial policy in an important setting is long-lived. For instance, the critical assumption for the classic Abadie et al. (2010) study of the effect of a cigarette tax passed by referendum in California is that California is similar to a weighted collection of states even though the other states differ with regard to their tobacco policy. As time passes, this assumption becomes less credible if we expect a state to eventually converge to the policy that reflects its political fundamentals. Long-term persistence implies that a state policy on nearly any topic can vary for nearly a century in a quasirandom way. This lends support to the assumptions behind difference-in-difference, synthetic controls, and other event study designs on long time horizons.

There are two empirical designs in which close dependence of the present on the past is useful. One is event study designs where observations are matched in an early period, and we want to assume that they remain similar later on. When persistence is low, long-term policy divergence might be a reason to think that two states are less similar than they initially appeared. The second empirical design is cross-sectional designs, i.e., comparisons across states. In this setting, we might worry that policy differences across states have to do with current voters' and politicians' preferences. If we learn that current policy is a function of past policy, then current policy may be partly a function of past tastes and less correlated with present tastes.

To understand these implications, consider the effort to identify the effect of a policy on some outcome across two states. Let there be two states  $i$  and  $j$  with outcomes  $Z_{it}, Z_{jt}$  and policies  $p_{it} = 1$  and  $p_{jt} = 0$  in period  $t$ . Suppose we want to estimate the difference between the potential outcome  $Z(p)$  as a function of the policy  $p$  for  $p = 1$  and  $p = 0$ . A standard approach might simply take the difference across the two states:  $Z_{it} - Z_{jt}$ . In Appendix Section B.3, I show that the difference across the two states  $Z_{it} - Z_{jt}$  will be an unbiased estimate of the difference between potential outcomes when there is complete persistence if the two states are similar at baseline. More generally, when persistence is high, the difference in outcomes across two states conditional on present policies approaches the difference conditional on historical policies. This finding is helpful because it is often more plausible to assume that potential outcomes are independent of historical policies than current policies for the reasons reviewed above.

Table 1: Frequency of Referendums by Country

|                           | Frequency |
|---------------------------|-----------|
| Denmark                   | 10        |
| Estonia                   | 1         |
| France                    | 3         |
| Ireland                   | 4         |
| Italy                     | 11        |
| Madagascar                | 1         |
| Palau                     | 1         |
| Peru                      | 1         |
| Switzerland               | 7         |
| Taiwan, Province of China | 5         |
| Turkey                    | 1         |
| United Kingdom            | 2         |
| Uruguay                   | 6         |
| Venezuela                 | 2         |
| Total                     | 55        |

## 6 Other Policy Processes: Strategy and Results

The persistence of non-U.S. national referendums, U.S. state legislation, and U.S. federal legislation follow a similar pattern to U.S. state referendums. The latter two domains, which involve different identification strategies, further underline the pervasiveness of policy persistence.

### 6.1 National Referendums

Persistence of non-U.S. national referendums is both qualitatively and quantitatively similar to that of U.S. state-level referendums.

I draw on two main data sources for baseline information on national referendums, which I supplement in a similar manner to U.S. state-level referendums using secondary sources. Elkins and

Hudson (2019) offers a dataset of constitutional referendums around the world, drawing on a dataset from the Centre for Research on Direct Democracy and several academic sources (Nohlen et al., 1999, 2001; Qvortrup and Qvortrup, 2005; Nohlen, 2005; Nohlen and Stöver, 2010). I supplement this by collecting non-constitutional referendums from the Center for Research on Direct Democracy (2023) for countries with frequent referendums, focusing on those outside of Europe. Table 1 lists the frequency of referendums for each country. Inclusion criteria are the same as for state-level referendums, except I widen the bandwidth to 10% since there are less available data. I also require each country to have been listed as at least “partly free” at the time of the referendum according to Freedom House (2023). I construct outcomes exactly as described in Section 2.2, except I rely more heavily on secondary sources such as news reports.

National referendums show the same pattern of a rapid initial decline in persistence followed by a plateau. The top-left plot in Figure 12 shows the effect of passage of a national referendum on whether the proposed policy is in place over time.<sup>42</sup> The estimates do not rule out a degree of persistence similar to that for U.S. states though the point estimates for national referendums are somewhat higher. Appendix Figure E29 shows that the intensive-margin effects resemble the U.S. state-level effects. Together, these figures broadly suggest that referendums in other countries are similarly persistent.

## 6.2 U.S. State Legislation

Policy adoption choices by state legislatures appear to be as persistent as referendum outcomes. The identifying assumptions for a regression discontinuity design do not apply in a legislative context, so I use a matching method for state policies, which I validate against my main results.

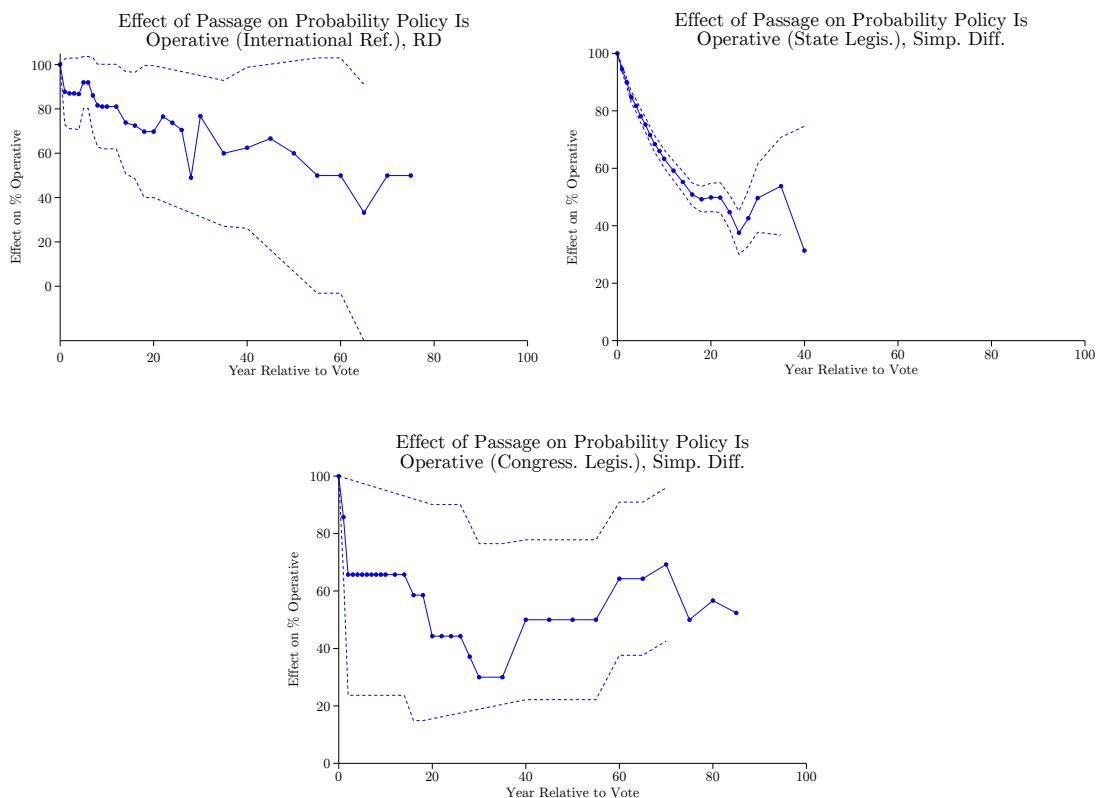
To study the persistence of state policies, I draw on existing data that records state policies and relevant covariates over time. Caughey and Warshaw (2016) collect variables for a wide range of state policies from 1934 through 2014, including both binary indicators (e.g., “Does the state ban discrimination in private housing?”) and continuous variables (e.g., the income tax rate). I use these as my measures of whether a policy is in place over time and to identify when states adopt or repeal policies. The data typically only cover a few decades for each policy, and generally only during periods in which the policy was changing in many states, which could plausibly bias estimates of persistence downward. To match states, I take covariates from Grossmann et al. (2021), who compile a broad set of state policy correlates that include the state’s demographics, other measures of state policy adoption, and the state’s partisan alignment.

To estimate the persistence of state policies, I compare policy histories for each time a state changes a policy with a control group of state policy histories. I adopt this strategy because the size

---

<sup>42</sup>The graph ends at year 80 when we no longer have any passed referendums in the sample.

Figure 12: Policy Persistence Is Similar for Legislation and International Referendums



*Notes:* Each panel displays the persistence of policies, analogous to the left panel of Figure 4, but for international referendums, state legislation, and Congressional legislation. The top-left panel shows the effect of passage on whether a proposed policy is in place later on for non-U.S. national-level referendums, estimated by equation (1). The top-right panel shows the effect of a state’s adopting a policy on whether that policy is operative later on as compared to matched comparison states, using data from Caughey and Warshaw (2016) and Grossmann et al. (2021). The third panel shows the effect of passing congressional legislation on whether the proposed policy is operative later on, selected using a narrative approach similar to Romer and Romer (1989, 1994) to identify narrowly passed and failed bills in the U.S. Congress.

of state legislatures and consequent ability of parties to coordinate both votes and the legislation that reaches the floor violates the identifying assumption of a regression discontinuity design.<sup>43</sup> Similar to my analysis of referendums, a treatment observation here is a state, policy, and year

<sup>43</sup>Instead of studying close legislation, one possibility would be to use closely-divided state legislatures’ party control as an instrument for policy change. Unfortunately, the rate of policy change in state legislatures and the share of legislatures that are closely divided are too small for there to be a strong first stage. Moreover, the weak instruments problem (Hahn and Hausman, 2003) is especially severe when studying persistence because the correlation between whether a policy is in place initially and later on is extraordinarily high.

when that policy is adopted or repealed (e.g., California, mandatory car insurance policy, 1975). For each state-policy history, I select a set of five states that resemble the treated state. If the control states capture what *would* have happened had the treated state *not* adopted the policy, I can identify the effect of adopting the policy on whether it is operative later by taking the difference between the two groups' averages. The same policy can be reused multiple times, which I account for by clustering standard errors at the state-policy and state-policy-year levels.

I select control groups using the Mahalanobis distance, which measures how far the vector of covariates for each possible control observation is from the treatment observation, weighted by the inverse of the covariance matrix. This gives the most weight to the most informative dimensions of similarity. For each state that adopts or repeals a policy, I construct a Mahalanobis distance between that state and all other states on the states' political liberalism according to the Caughey and Warshaw (2016) index, upper chamber Democratic share, lower chamber Democratic share, and gubernatorial party as well as the 4- and 8-year lags of these variables. I restrict my sample to those states that had the same value of the outcome variable for the previous four years, and I select the five closest matches for each treatment observation. This yields a final sample of 1,561 state-policy time series post-adoption, and 7,805 state-policy time series for control states. In Appendix Figure E30, I output effect estimates for a small sample of referendums that I can match to policies in the Caughey and Warshaw (2016) data and confirm that the estimates closely resemble the regression discontinuity estimates in Figure 4.

Having validated a matching method for state-level policies, we can now examine the persistence of state legislative policy changes. The upper-right panel of Figure 12 shows the difference in the share of policies that are in place after adoption compared to matched states that did not initially adopt them. Given the limitations of the dataset, the effect can only be estimated for forty years. In the first forty years, we see the same pattern as for state-level referendums: a sharp, initial decline followed by a plateau at just more than forty percentage points.

### 6.3 U.S. Federal Legislation

For a final investigation of the generality of persistence, I use a narrative approach to identify U.S. congressional legislation that nearly passed or nearly failed, similar to the approach used by Romer and Romer (1989, 1994) to identify exogenous monetary policy shocks.

To identify congressional legislation that narrowly passed or failed, I review descriptions of U.S. Congress activity from 1925 through 1956. As in the case of state legislatures, the identifying assumption of a regression discontinuity design is not likely to be satisfied in the U.S. Congress. Vote whipping and the decision to hold back doomed bills create systematic differences between laws approved and laws rejected by small margins. Instead, I review contemporary narratives of



the activity of each two-year Congress from Editorial Research Reports (1956) to identify proposed legislation that came close to fruition and cases where bills appeared to squeak through. To do this, I feed each two-year narrative record of Congress into the large-language model Claude and ask it to produce a table listing narrowly passed and failed legislation.<sup>44</sup> I then review the history of each of these candidates in the congressional record to determine which were genuinely close. Out of a total of 75 candidates, I select 19 pieces of proposed legislation.

An illustrative piece of legislation is a 1940 bill to systematize and review federal regulations. A dramatic expansion in the scope of the federal government under President Franklin Roosevelt had left it unclear when and how a regulation could be contested, in contrast with more established procedures for congressionally-approved laws (Elias, 2016). The 1940 bill proposed establishing a court specifically to review agency rulings. The bill passed both houses of Congress by large majorities, but President Roosevelt vetoed the bill. Congress then sustained the veto by a significant margin, preventing the law from passing. Six years later, the Administrative Procedure Act achieved some of what the 1940 bill sought to do, but without establishing a court and with considerably less independent oversight than proposed in the 1940 bill (Elias, 2016). I record the bill as being partially operative after 1946 given that there remain views that modern administrative law falls significantly short of the 1940 bill’s proposed restrictions, with the issue a matter of renewed legal debate (Rabkin, 2020).

The bottom panel of Figure 12 shows the status of these narrowly approved and rejected policies over time. Given the small sample, the results are noisy, but the point estimates are strikingly similar to those for state-level referendums. All is not exactly the same, however: Appendix Figure E29 shows a different pattern when we look at the share of such policies that are at least partially operative over time. In contrast to the state referendum case, I find no effect after three decades of passing congressional legislation on whether there is some, possibly weaker, version of the policy in place. While noisy, the picture suggests that the primary effect of Congress’s policy choices over long time horizons is to determine the nature and strength of a policy. The ability to bargain may mean that policies that reach the point of serious consideration eventually get passed in a form that is sufficiently acceptable at the right moment.

## 7 Conclusion

The statutory histories of state-level referendums and other close votes across a variety of settings and topics show that political decisions persist for decades, often beyond a century. A theoretically-guided analysis shows that this implies a substantial mismatch between policy and voter preferences,

---

<sup>44</sup>See Appendix A.4 for an example of this conversation with Claude.

or even social welfare. These findings also carry implications for cost-benefit analysis and policy event-study analyses.

Policy persistence is widespread, suggesting that the policy processes studied here are far from a world where contemporary voters' attitudes determine policy. While many failed referendums pass in the first few subsequent decades, there is a lasting 40-percentage point gap in the eventual adoption of narrowly passed and failed referendums. The mechanisms driving the general effect do not appear to be idiosyncratic to particular policy areas, such as specific economic incentives or cultural changes. Rather, the data are consistent with the explanation that issue salience drives persistence. Referendums tend to occur during brief, selected periods in which there is unusual interest in a given policy. The data indicate that, for a given policy, there are long periods of time when few take enough of an interest to revisit it for reasons largely independent of the policy's impacts.

Additional data and a calibration exercise show that there is a substantial and widespread mismatch between existing policy and voters' preferred policy because of how slowly policies update. A model of the initiative process indicates that many policies are as good as random given the extent of policy persistence and changes in voters' preferences over time. Evidence from international referendums and state and congressional legislation confirms that institutional persistence is broader than simply state referendums. Any obvious way to reduce policy mismatch faces tradeoffs between keeping policies up to date and allowing a proliferation of policy changes. Yet policy persistence also tells us that long-term changes to policy are possible, as choices made today can matter for a very long time and are amenable to empirical study given the degree of long-term, random variation.

This paper is the first attempt to quantify persistence at the policy level across a broad set of policies, leaving many open questions for future research. One important avenue for research is to identify other sources of exogenous variation in policy adoption to better understand other contexts. Another related question is whether culture is similarly persistent as recent work in economic history argues (Alesina and Giuliano, 2015). The final line of inquiry that I will mention here concerns the implications of persistence. The body of literature studying the political constraints for optimal policy and the effects of small perturbations is small (Feldstein, 1976; Scheuer and Wolitzky, 2016), and there is little work precisely analyzing this issue. The methods in this paper lend themselves to further empirical analysis of the political impacts of policy changes, which could serve as a starting point for an analysis of optimal policy with empirically-calibrated political dynamics.

## References

- Abadie, A., A. Diamond, and J. Hainmueller (2010). Synthetic control methods for comparative case studies: Estimating the effect of california’s tobacco control program. *Journal of the American statistical Association* 105(490), 493–505.
- Acemoglu, D., G. Egorov, and K. Sonin (2021). Institutional change and institutional persistence. In *The Handbook of Historical Economics*, pp. 365–389. Elsevier.
- Acemoglu, D., M. Golosov, and A. Tsyvinski (2010). Dynamic mirrlees taxation under political economy constraints. *The Review of Economic Studies* 77(3), 841–881.
- Acemoglu, D., S. Johnson, and J. A. Robinson (2001). The colonial origins of comparative development: An empirical investigation. *American economic review* 91(5), 1369–1401.
- Acemoglu, D., S. Johnson, and J. A. Robinson (2002). Reversal of fortune: Geography and institutions in the making of the modern world income distribution. *The Quarterly journal of economics* 117(4), 1231–1294.
- Acemoglu, D., S. Johnson, and J. A. Robinson (2005). Institutions as a fundamental cause of long-run growth. *Handbook of economic growth* 1, 385–472.
- Alesina, A. and P. Giuliano (2015). Culture and institutions. *Journal of economic literature* 53(4), 898–944.
- Alesina, A., G. Tabellini, and F. Trebbi (2017). Is europe an optimal political area? Technical report, National Bureau of Economic Research.
- Ballotpedia (2023a). *Ballot Measures overview*. Ballotpedia.
- Ballotpedia (2023b). *History of initiative and referendum in the U.S. History of initiative and referendum in the U.S. History of initiative and referendum in the U.S.* Ballotpedia.
- Bar-Lev, S. K., P. Enis, and G. Letac (1994). Sampling models which admit a given general exponential family as a conjugate family of priors. *The Annals of Statistics*, 1555–1586.
- Battaglini, M., R. B. Morton, and T. R. Palfrey (2008). Information aggregation and strategic abstention in large laboratory elections. *American Economic Review* 98(2), 194–200.
- Baumgartner, F. R., C. Breunig, C. Green-Pedersen, B. D. Jones, P. B. Mortensen, M. Nuytemans, and S. Walgrave (2009). Punctuated equilibrium in comparative perspective. *American Journal of Political Science* 53(3), 603–620.

- Baumgartner, F. R., M. Foucault, and A. François (2006). Punctuated equilibrium in french budgeting processes. *Journal of European Public Policy* 13(7), 1086–1103.
- Baumgartner, F. R. and B. D. Jones (2010). *Agendas and instability in American politics*. University of Chicago Press.
- Baumgartner, F. R., B. D. Jones, and P. B. Mortensen (2018). Punctuated equilibrium theory: Explaining stability and change in public policymaking. *Theories of the policy process*, 55–101.
- Bernheim, B. D., L. Braghieri, A. Martínez-Marquina, and D. Zuckerman (2021). A theory of chosen preferences. *American Economic Review* 111(2), 720–54.
- Bernheim, B. D. and D. Taubinsky (2018). Behavioral public economics. *Handbook of behavioral economics: Applications and Foundations* 1 1, 381–516.
- Berry, C. R., B. C. Burden, and W. G. Howell (2010). After enactment: The lives and deaths of federal programs. *American Journal of Political Science* 54(1), 1–17.
- Besley, T. J. and S. Coate (2000). Issue unbundling via citizens’ initiatives.
- Bickers, K. N. (1991). The programmatic expansion of the us government. *Western Political Quarterly* 44(4), 891–914.
- Botticini, M., Z. Eckstein, et al. (2006). *Path dependence and occupations*, Volume 3. Boston Univ., Department of Economics, Inst. for Economic Development.
- Bursztyn, L., D. Cantoni, P. Funk, F. Schönenberger, and N. Yuchtman (2023). Identifying the effect of election closeness on voter turnout: Evidence from swiss referenda. *Journal of the European Economic Association*, jvad038.
- Calonico, S., M. D. Cattaneo, and M. H. Farrell (2020). Optimal bandwidth choice for robust bias-corrected inference in regression discontinuity designs. *The Econometrics Journal* 23(2), 192–210.
- Carollo, N. A., J. F. Hicks, A. Karch, and M. M. Kleiner (2022). The origins and evolution of occupational licensing in the united states.
- Carpenter, D. P. and D. E. Lewis (2004). Political learning from rare events: Poisson inference, fiscal constraints, and the lifetime of bureaus. *Political Analysis* 12(3), 201–232.
- Caughey, D. and C. Warshaw (2016). The dynamics of state policy liberalism, 1936–2014. *American Journal of Political Science* 60(4), 899–913.

- Cellini, S. R., F. Ferreira, and J. Rothstein (2010). The value of school facility investments: Evidence from a dynamic regression discontinuity design. *The Quarterly Journal of Economics* 125(1), 215–261.
- Center for Research on Direct Democracy (2023). *C2D*. Center for Research on Direct Democracy.
- Cirone, A. and T. B. Pepinsky (2022). Historical persistence. *Annual Review of Political Science* 25, 241–259.
- Clubb, J. M., W. H. Flanigan, and N. H. Zingale (2005). Electoral data for counties in the united states, 1840-1972. *Ann Arbor, MI: ICPSR*.
- Coate, S. and S. Morris (1999). Policy persistence. *American Economic Review* 89(5), 1327–1336.
- Collins, W. J. (2003). The political economy of state-level fair employment laws, 1940–1964. *Explorations in Economic History* 40(1), 24–51.
- Corder, J. K. (2004). Are federal programs immortal? estimating the hazard of program termination. *American Politics Research* 32(1), 3–25.
- Dal Bó, E., P. Dal Bó, and E. Eyster (2018). The demand for bad policy when voters underappreciate equilibrium effects. *The Review of Economic Studies* 85(2), 964–998.
- David, P. A. (2007). Path dependence, its critics and the quest for ‘historical economics’. *The evolution of economic institutions: A critical reader*, 120–142.
- Dell, M. (2010). The persistent effects of peru’s mining mita. *Econometrica* 78(6), 1863–1903.
- Dell, M., N. Lane, and P. Querubin (2018). The historical state, local collective action, and economic development in vietnam. *Econometrica* 86(6), 2083–2121.
- DellaVigna, S. and W. Kim (2022). Policy diffusion and polarization across us states. Technical report, National Bureau of Economic Research.
- Diaconis, P. and D. Ylvisaker (1979). Conjugate priors for exponential families. *The Annals of statistics*, 269–281.
- Downs, A. (1957). An economic theory of political action in a democracy. *Journal of political economy* 65(2), 135–150.
- Downs, A. (1972). Up and down with ecology-the issue-attention cycle. *The Public Interest* 28, 38.
- Edelman, M. J. (1985). *The symbolic uses of politics*. University of Illinois Press.

- Editorial Research Reports (1925-1956). *Congressional Records and Roll Calls*. Editorial Research Reports.
- Eggers, A. C., A. Fowler, J. Hainmueller, A. B. Hall, and J. M. Snyder Jr (2015). On the validity of the regression discontinuity design for estimating electoral effects: New evidence from over 40,000 close races. *American Journal of Political Science* 59(1), 259–274.
- Elias, R. A. (2016). The legislative history of the administrative procedure act. *Fordham Environmental Law Review* 27(2), 207–224.
- Elkins, Z. and A. Hudson (2019). The constitutional referendum in historical perspective. *Comparative Constitution Making* 142, 152.
- Feddersen, T. J. and W. Pesendorfer (1996). The swing voter’s curse. *The American economic review*, 408–424.
- Feldstein, M. (1976). On the theory of tax reform. *Journal of public economics* 6(1-2), 77–104.
- Fernandez, R. and D. Rodrik (1991). Resistance to reform: Status quo bias in the presence of individual-specific uncertainty. *The American economic review*, 1146–1155.
- Follmann, D. A. and M. S. Goldberg (1988). Distinguishing heterogeneity from decreasing hazard rates. *Technometrics* 30(4), 389–396.
- Fowler, A. and A. B. Hall (2017). Long-term consequences of election results. *British Journal of Political Science* 47(2), 351–372.
- Freedom House (2023). *Freedom in the World*. Freedom House.
- Gelman, A. and G. Imbens (2019). Why high-order polynomials should not be used in regression discontinuity designs. *Journal of Business & Economic Statistics* 37(3), 447–456.
- Gieczewski, G. (2021). Policy persistence and drift in organizations. *Econometrica* 89(1), 251–279.
- Gould, S. J. and S. J. Gould (2009). *Punctuated equilibrium*. Harvard University Press.
- Grossmann, M., M. P. Jordan, and J. McCrain (2021). The correlates of state policy and the structure of state panel data. *State Politics & Policy Quarterly* 21(4), 430–450.
- Guiso, L., P. Sapienza, and L. Zingales (2016). Long-term persistence. *Journal of the European Economic Association* 14(6), 1401–1436.
- Hahn, J. and J. Hausman (2003). Weak instruments: Diagnosis and cures in empirical econometrics. *American Economic Review* 93(2), 118–125.

- Hainmueller, J., A. B. Hall, and J. M. Snyder Jr (2015). Assessing the external validity of election rd estimates: An investigation of the incumbency advantage. *The Journal of Politics* 77(3), 707–720.
- Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica: Journal of the econometric society*, 153–161.
- Imbens, G. and K. Kalyanaraman (2012). Optimal bandwidth choice for the regression discontinuity estimator. *The Review of economic studies* 79(3), 933–959.
- Imbens, G. W. and T. Lemieux (2008). Regression discontinuity designs: A guide to practice. *Journal of econometrics* 142(2), 615–635.
- Jennings, W. and C. Wlezien (2018). Election polling errors across time and space. *Nature Human Behaviour* 2(4), 276–283.
- Kelly, B., D. Papanikolaou, A. Seru, and M. Taddy (2021). Measuring technological innovation over the long run. *American Economic Review: Insights* 3(3), 303–320.
- Kelly, M. (2019). The standard errors of persistence. Technical report, UCD Centre for Economic Research Working Paper Series.
- Koffi, M. (2021). Innovative ideas and gender inequality. Technical report, Working Paper Series.
- Lee, D. S. (2008). Randomized experiments from non-random selection in us house elections. *Journal of Econometrics* 142(2), 675–697.
- Leip, D. (2002). *Dave Leip’s Atlas of U.S. Presidential Elections*.
- Lewis, D. E. (2002). The politics of agency termination: Confronting the myth of agency immortality. *The Journal of Politics* 64(1), 89–107.
- Liebowitz, S. J. and S. E. Margolis (1995). Path dependence, lock-in, and history. *The Journal of Law, Economics, and Organization* 11(1), 205–226.
- Liebowitz, S. J. and S. E. Margolis (2014). *Path dependence and lock-in*. Edward Elgar Publishing Limited.
- Maltzman, F. and C. R. Shipan (2008). Change, continuity, and the evolution of the law. *American Journal of Political Science* 52(2), 252–267.
- Matsusaka, J. G. (1992). Economics of direct legislation. *The Quarterly Journal of Economics* 107(2), 541–571.

- Matsusaka, J. G. (1995). Fiscal effects of the voter initiative: Evidence from the last 30 years. *Journal of political Economy* 103(3), 587–623.
- McDonald, M. P. (2002). The turnout rate among eligible voters in the states, 1980–2000. *State Politics & Policy Quarterly* 2(2), 199–212.
- Mokyr, J. (1990). Punctuated equilibria and technological progress. *The American Economic Review* 80(2), 350–354.
- National Archives (2022). *16th Amendment to the U.S. Constitution: Federal Income Tax (1913)*. National Archives.
- National Conference of State Legislatures (2022). *Initiative and Referendum Processes Initiative and Referendum Processes*. National Conference of State Legislatures.
- National Conference of State Legislatures (2023). *Statewide Ballot Measures Database*. National Conference of State Legislatures.
- Nohlen, D. (2005). *Elections in the Americas A Data Handbook Volume 1: North America, Central America, and the Caribbean*. OUP Oxford.
- Nohlen, D., F. Grotz, and C. Hartmann (2001). *Elections in Asia and the Pacific: A Data Handbook: Volume I: Middle East, Central Asia, and South Asia*. OUP Oxford.
- Nohlen, D. and P. Stöver (2010). *Elections in Europe*. Nomos Verlagsgesellschaft mbH & Co. KG.
- Nohlen, D., B. Thibaut, and M. Krennerich (1999). *Elections in Africa: A data handbook*. OUP Oxford.
- Nordhaus, W. D. (2007). A review of the stern review on the economics of climate change. *Journal of economic literature* 45(3), 686–702.
- North, D. C. (1991). Institutions. *Journal of economic perspectives* 5(1), 97–112.
- Nunn, N. (2014). Historical development. *Handbook of economic growth* 2, 347–402.
- Open Secrets (2022). *Open Secrets: Ballot Measures*. Open Secrets.
- Persson, T. and G. Tabellini (2020). Culture, institutions and policy. *Available at SSRN 3680457*.
- Pierson, P. (2004). *Politics in time: History, institutions, and social analysis*. Princeton University Press.



- Porta, R. L., F. Lopez-de Silanes, A. Shleifer, and R. W. Vishny (1998). Law and finance. *Journal of political economy* 106(6), 1113–1155.
- Post, A. E. and P. Pierson (2005). How a law stays a law: The durability of us tax breaks, 1967-2003. In *annual meeting of the American Political Science Association, Washington*.
- Qvortrup, M. and M. Qvortrup (2005). *A comparative study of referendums: Government by the people*. Manchester University Press.
- Rabkin, J. (2020). The origins of the apa: Misremembered and forgotten views. *Geo. Mason L. Rev.* 28, 547.
- Ragusa, J. M. (2010). The lifecycle of public policy: An event history analysis of repeals to landmark legislative enactments, 1951-2006. *American Politics Research* 38(6), 1015–1051.
- Rennert, K., F. Errickson, B. C. Prest, L. Rennels, R. G. Newell, W. Pizer, C. Kingdon, J. Wingenroth, R. Cooke, B. Parthum, et al. (2022). Comprehensive evidence implies a higher social cost of co2. *Nature* 610(7933), 687–692.
- Roberts, K. (2015). Dynamic voting in clubs. *Research in Economics* 69(3), 320–335.
- Romer, C. D. and D. H. Romer (1989). Does monetary policy matter? a new test in the spirit of friedman and schwartz. *NBER macroeconomics annual* 4, 121–170.
- Romer, C. D. and D. H. Romer (1994). Monetary policy matters. *Journal of Monetary Economics* 34(1), 75–88.
- Scheuer, F. and A. Wolitzky (2016). Capital taxation under political constraints. *American Economic Review* 106(8), 2304–2328.
- Shigeoka, H. and Y. Watanabe (2023). Policy diffusion through elections. Technical report, National Bureau of Economic Research.
- Song, B. (2018). Estimating incumbency effects using regression discontinuity design. *Research & Politics* 5(4), 2053168018813446.
- Thrower, S. (2017). To revoke or not revoke? the political determinants of executive order longevity. *American Journal of Political Science* 61(3), 642–656.
- Voth, H.-J. (2021). Persistence—myth and mystery. In *The handbook of historical economics*, pp. 243–267. Elsevier.

- Wang, S. and D. Y. Yang (2021). Policy experimentation in china: The political economy of policy learning. Technical report, National Bureau of Economic Research.
- Wattenberg, M. P., I. McAllister, and A. Salvanto (2000). How voting is like taking an sat test: An analysis of american voter rolloff. *American Politics Quarterly* 28(2), 234–250.
- Weitzman, M. L. (1998). Why the far-distant future should be discounted at its lowest possible rate. *Journal of environmental economics and management* 36(3), 201–208.
- Wheaton, B. (2020). Laws, beliefs, and backlash. *Unpublished Working Paper*.
- Woodberry, R. D. (2012). The missionary roots of liberal democracy. *American political science review* 106(2), 244–274.

## A Further Details on Empirical Strategy

### A.1 Robustness to Bandwidth Selection

While I choose the bandwidth for my dataset before collecting data, multiple empirical tests confirm that the selected bandwidth yields unbiased empirical estimates.

I validate the validity of my bandwidth by linking referendum data with an existing dataset of state-level policy indicators in order to perform standard calculations of the optimal bandwidth for an RD design. Caughey and Warshaw (2016) collect variables for a wide range of state policies from 1934 through 2014, including both binary indicators (e.g., “Does the state ban discrimination in private housing?”) and continuous variables (e.g., the income tax rate). Only a small minority of referendums involve the direct adoption or repeal of a policy in the data, but we link those we can identify to obtain a sample of around 205 referendums in both datasets. Since the optimal bandwidth depends on the sample size, I augment the dataset such that the number of observations between -2.5 and 2.5 is the same as the number in my main sample.<sup>45</sup> I assign each synthetic observation the outcome of the nearest neighbor by vote share. This gives me a synthetic dataset of 4,784 referendums with vote shares ranging from 10% to 87% and a matched outcome variable.

In my synthetic dataset, I find that 2.5% is well within the optimal bandwidth according to the benchmark method. Calonico et al. (2020) offer an algorithm for suggesting the optimal, bias-corrected bandwidth in a regression discontinuity design which yields an optimal bandwidth of 5.6%.<sup>46</sup> Though my manual outcome data do not cover this wider 5.6% bandwidth, I perform robustness checks to examine whether my results are sensitive to the bandwidth. As checks, I compare my empirical estimates with estimates using a wider bandwidth and the Caughey and Warshaw (2016) outcomes for the smaller, overlapping sample for the shorter time series in the Caughey and Warshaw (2016) data. I also produce estimates for a 1% bandwidth in Appendix Figure E1.

### A.2 Sample Instructions for Outcome Data Collection

The next pages contain example instructions to an Upwork assistant for recording referendums’ statutory histories.

---

<sup>45</sup>I assume vote share follows a truncated normal distribution, which I confirm fits the data.

<sup>46</sup>Imbens and Kalyanaraman (2012) offer a different algorithm that minimizes mean squared error but which Calonico et al. (2020) show yields suboptimal confidence intervals. The Imbens and Kalyanaraman (2012) method yields the much wider bandwidth of 12.9%.

## **Ballot Measure Data Input Instructions: Minnesota**

### **Overall Instructions:**

1. **Please ignore the red columns for now.** These are rows Zach is working on, so please do not touch these.
2. **Removing Duplicates:** We've collected data from two sources (Ballotpedia (BP) and NCSL). Given this, there might be a few measures that appear twice, and we want to check this. On the **Flag\_Duplicates** sheet, there will be a list of measures which may be duplicates. The **bp** and **nsl** columns are marked with a 1 to indicate the measure came from that data source. The process for removing duplicates is:
  - a. Read through the list of measures on this sheet
  - b. If two measures have the same year, have vote totals within 0.1% of each other, **and** come from different data sources (e.g. one from BP, one from NCSL), then:
  - c. Compare their **ballot\_id** and **ballotdescrip** fields. If **in your judgement** these two fields are similar enough (in terms of keywords etc.) to almost certainly be referring to the same measure, then:
  - d. Look them up on Google, and mark **only the measure with a 1 in the BP columns as duplicate**. To do this, add some text to the **DUPLICATE** column and it'll be automatically copied to the relevant measure in the main sheet, so if there is anything we should know from the duplicates process, write that here.
3. **Checking Overall Variables:** The main fields to be filled in for each measure are on the **Checks** sheet. If you cannot figure one of these out quickly, just flag it and write that in notes. Ensure each of the following variables are entered and correct:
  - a. Skip any row you've marked as a duplicate (which should automatically show up in the "DUPLICATE" column).
  - b. **flag** and **notes:** Set **flag** to 1 if there is anything you are unsure about or that could be an issue in interpreting the variables for a given measure. Enter any important things to know under **notes**, regardless of whether **flag** is 0 or 1.
  - c. **link1:** Include a link to a primary source for the text of the proposed measure (e.g. the text according to the state government's or a library's website). You can find the link in the following way:
    - i. Using this link, click on the link in the "Ballot Language" column. This should offer the text of the proposed law. It will look like this:

*An act proposing an amendment to the Constitution of the State of Minnesota, Article IV, Section 5, pertaining to the distribution of the excise tax on petroleum products.*

Be it enacted by the Legislature of the State of Minnesota :

Section 1. The following amendments to the Constitution of the State of Minnesota, Article IV, Section 5, is hereby proposed to the people of the state for their approval or rejection, which section, when amended, shall read as follows:

Sec. 5. Proposed amendment to Minnesota Constitution, Article IX, Section 5. For the purpose of defraying extraordinary expenditures, the state may contract public debts, but such debts shall never, in the aggregate, exceed two hundred and fifty thousand dollars; every such debt shall be authorized by law, for some single object, to be distinctly specified therein; and no such law shall take effect until it shall have been passed by the vote of two-thirds of the members of each branch of the legislature, to be recorded by yeas and nays on the journals of each house respectively; and every such law shall levy a tax annually sufficient to pay the annual interest of such debt, and also a tax sufficient to pay the principal of such debt within ten years from the final passage of such law, and shall specially appropriate the proceeds of such taxes to the payment of such principal and interest; and such appropriation and taxes shall not be repealed, postponed, or diminished, until the principal and interest of such debt shall have been wholly paid. The state shall never contract any debts for works of internal improvements, or be a party in carrying on such works, except as authorized by Section 16 of Article 9, and by Article 16 of this Constitution, but it may levy an excise tax upon any substance, material, fluid, force or other means or instrumentality, or the business of dealing in, selling or producing any or all thereof, used or useful, in producing or generating power for propelling motor or other vehicles used on the public highways of this state, and shall place one-half of the proceeds of such tax in the Trunk Highway Fund provided for in Section 2 of said Article 16, apportion for road and street purposes on the basis of population 6% thereof to cities of over 100,000 population,

- d. **link2:** If **link1** does not include the vote count, please include a link to the vote count. This will pretty much always be [this spreadsheet with information on all measures](#).
- e. You may include additional informative links as **link2** (if not filled in), **link3**, and **link4**.
- f. **disqualified:** This indicates that a measure should not count in the sample. Note that if this column is anything other than a “0”, it means you can skip the measure, do nothing else, and move on to the next row. To fill in this column, look the measure up on Google. This column should be equal to
  - i. 1 if the ballot measure is a bond measure proposing the short-term borrowing and spending of a fixed amount of money
  - ii. 2 if it is an advisory measure that does not actually change the law
  - iii. 3 if it is a measure that changes something only temporarily, such as a one-time expense or temporary tax
  - iv. 4 if it is an amendment to the U.S. constitution
  - v. 5 if it should for some other reason not be considered to propose a permanent legal change.
  - vi. 0 otherwise
- g. **ballotname:** This is the proposition/amendment/initiative number (e.g. “Proposition 8”).
- h. **ballotdescrip:** This is the title of the ballot measure. Leave unchanged if filled in unless it includes, e.g. “Amendment 9” or ‘Proposition 3” in which case that remove that, and include it in **ballotname**, not **ballotdescrip**.
- i. **type:** This should be “Initiative”, “Legislative Referendum”, “Popular Referendum,” or “Other.” *For Minnesota, it should be “Legislative Referendum” unless otherwise noted.*
- j. **typeOther:** If **type** is equal to “Other,” say what type of measure it is (e.g. a commission referral).

- k. **electiontype:** This should be equal to “General,” “Primary,” “Special,” or “Recall.” *In Minnesota, this is probably always “General”, but the text of the measure should say something like “This proposed amendment shall be submitted to the people of this state for their approval or rejection at the general election”*
- l. **threshold:** This is the percent of votes needed to pass, 50% times the total votes at the election divided by the number voting on the measure.
- m. **pctyesvotes:** This is the vote share in favor of passage, which you can verify as the “Yes” votes divided by the total of the “Yes” and “No” votes [at this link](#).
- n. **totalVotesMeasure:** This is the total number of votes on the measure in question, i.e. the number of people voting yes or no on it. You can find this by adding the “Yes” and “No” columns [at this link](#).
- o. **totalVotesElection:** This is the total number of votes at the election. If it is available, it will be available in the “Total” column [on this page](#).
- p. **totalVotesOther:** Only fill this in if you could not fill in totalVotesElection. This is the largest total number of votes across the state on something else voted on at the election. You can use [this link to congressional votes](#) and add the number of votes for senate and add all the votes on House representatives; choose the greater number between senate and House representatives.
  - i. **totalVotesOtherDesc** describes what you put in totalVotesOther. It should usually be “Total Senate votes”, “Total House votes”, or something similar.
- q. **passed:** This is 1 if the measure passed, and 0 if it did not, as indicated by an “A” in the Action column [at this link](#).
- r. **constitutional:** This is 1 if the measure is an amendment to the state constitution and 0 if not, as indicated by the “Type of Action” column [at this link](#).
- a. **politicalLean:** This is generally 1 if the measure proposed made government bigger or more democratic, bestowed more rights on women or minority groups, made drug use easier, or weakened religion, and 0 if it did the opposite. Specifically, it should be a 1 if the measure did any of the following, in order, and 0 if the opposite:
  - i. Higher taxes
  - ii. Higher spending (including public salaries)
  - iii. Higher debt (including bonds) or lower savings
  - iv. Making passing taxes, spending, or debt easier
  - v. Greater economic regulations (e.g. rules about pollution or how a business would function)
  - vi. Civil rights, equal rights for women, or rights for a minority group
  - vii. Greater permissiveness for drug use, substance use, or sexual acts
  - viii. Expanding who can vote
  - ix. Lower state support for religion/laws that are less in line with prominent religious doctrines
  - x. Tighter ethics rules or campaign spending regulations
  - xi. Greater direct democracy (except when it conflicts with iv above, such as a measure requiring voter approval of new taxes)
  - xii. Greater power for the legislature, including longer terms and higher pay (except when it conflicts with x or xi)

- xiii. Greater leniency for those accused or convicted of crimes and fewer protections for crime victims
- s. **clearAlignment:** This is equal to 3 if it is obvious what **politicalLean** should be equal to, 1 if it is completely unclear, and 2 if it is somewhat unclear.
- t. **proponent:** This is only relevant if at the link included in link1, you found arguments in support of the measure in addition to the measure itself. In that case, list the parties who write arguments in support of the ballot measure or are accused by opponents of backing the measure (e.g. “League of Women Voters” or “Hospital Workers Union”).
- u. **opponent:** This is only relevant if at the link included in link1, you found arguments in support of the measure in addition to the measure itself. List the parties who write arguments against the ballot measure or are accused by proponents of opposing the measure (e.g. “League of Women Voters” or “Hospital Workers Union”).
- v. **implementationDate:** This is the date at which the ballot measure goes into effect. Unless otherwise noted, it is the day after the general election.
- w. **abort-termlimits:** Please quickly set the variable to 1 for any topic the measure applies if it’s obvious, but don’t spend too much time on this.
4. For every event in which the proposed law changed (including it going into effect for the first time), there are a set of columns to be filled out at the end of the **Checks** sheet. Fill out only the columns that have the matching **Transition No.** Instructions for a recommended process are on the second-to-last page of this document, but first is a description of the columns.

*The rules for completing these fields is as follows:*

- a) **transitionDate** - When did the change in question first go into effect?*
- b) **literalStatus** - Is the exact proposed law in effect as it was originally worded?*

*For now, please ignore the **administrative** and **enforcementNews** variables.*

*The next set of questions need to be answered three times, with different scopes. The scopes are:*

- ***Narrow Scope:** When answering these questions, restrict your attention to the phrases modified by the measure itself either in the exact same location or in a substitute location (e.g. if the constitution was reorganized or the law failed and later passed with a different section number).*
- ***Broad Scope:** When answering these questions, search the state constitution and code and count anything that seems like it could possibly qualify. The constitution and statute links at the end of this document are searchable; please search for keywords (e.g. “water development bonds”) and see if any “broad” category could be true.*
- ***Federal Scope:** When answering these questions, consider whether the provisions of the proposed law have been adopted into Federal Law*

*With these scopes in mind, there are 4 questions to answer:*

- **InPlace:** Is the reform in effect; i.e. are there laws on the books with the same meaning as the ballot measure proposed and no further changes?
- **Stronger:** For at least one provision of the original bill, is there now a stronger version of that provisions in place/has the law moved even further in the direction supported by the original proposal? e.g the original bill contained a provision for a 5% tax on X, now there is a 10% tax on X
- **Partial:** For at least one provision of the original bill, is a partial or weaker version of that provision in-place?
- **Opposite:** For at least one provision of the original bill, is there now a provision on the books which has moved the law in the opposite direction to the original proposal e.g. the measure proposed introducing a 10% tax on X, now there is a 5% subsidy

These questions need to be answered for each Scope - e.g. **PartialBroad** is asking “is there a weaker version of some of the proposed measures now in effect on a Federal level?”.

**How to fill in these cells:**

Only fill in the “Broad” variables if one of them is true and different from the same “Narrow” variable. If not, you can ignore them.

**InPlaceNarrow** should be 0 or 1 depending on whether every provision of the measure seems to be in place.

**Stronger, Partial, or Opposite** should be blank or zero unless one of them is true. If one of them is true, the code depends on why it is true.. Use the following codes:

- **Enter a 1 in the cell** if, for some provisions, the answer to the question is yes because of a **numeric fact** e.g. if the question is **PartialBroad** and you believe this category applies because the measure proposed a 20% tax on X, and there is now a 10% state tax on X, then you would enter a 1 into the **PartialBroad** column for the measure.
- **Enter a 2 in the cell** if, for some provisions, the answer to the question is yes because of the **qualitative substance of the law** e.g. if the original measure was to legalize marijuana, and there is a state law to decriminalize (but not legalize) it, then you would enter a 2 into the **PartialBroad** column.
- **Enter a 3 in the cell** if, for some provisions, the answer to the question is yes because of a **change to the legal structure of the law**, e.g. if the original law proposed a tax on X, and now that tax is harder to repeal (e.g. requires a supermajority), then you would enter a 3 into the **PartialBroad** column
- **Enter a 4 in the cell** if the answer to the question is yes for every provision of the proposed measure; for example, if there is a weaker version of each provision of the measure, you would enter a 4 into the **PartialBroad** column.
- **Enter a 5 in the cell** if the only reason a status holds is that there are new provisions not in the original law (in the case of “Stronger”) or some provisions not in place (in the case of “Partial”).



- **Enter an 8 in the cell** if, for some provisions, the answer to the question is yes because of a technicality, but the question is not really a “yes” in spirit. For example, a measure might propose a tax increase while adding a minor provision to ensure taxes are well-spent. If this minor provision is repealed, this would technically mean a yes answer to “PartialNarrow,” but in practice the repealed provision does not actually weaken the reform.
- **Enter a 9 in the cell** if you’ve judged the measure fits the question for reasons other than the above possibilities. Please outline in the **notes** column at the top of the sheet what your criterion was instead

**Please note: More than one of the above can be true simultaneously:** e.g. if a tax was increased by 5% **AND** was made harder to repeal, then you would enter a **1 and a 3** into **StrongerNarrow** row. It’s fine to enter these as one number e.g. as 13 or 31. However please make sure your input contains only the 4 digits above and no duplicates.

Filling in all changes:

To complete the above questions, please use the following strategy.

1. If the measure passed and went into effect, record the implementation date as **transitionDate1**. Record **literalStatus** and **narrowInPlace** as 1. If the measure failed and did not go into effect immediately, record the implementation date as **transitionDate1**, and record **literalStatus** and **narrowInPlace** as 0.
  - a. Include **link1** and **link2** as the source.
2. Find the text of the measure.
  - a. [Using this link](#), click on the link in the “Ballot Language” column. This should offer the text of the proposed law. It will look like this:

747] OF MINNESOTA FOR 1949 1497  
 CHAPTER 747—S. F. No. 91  
 [Not Coded]

*An act proposing an amendment to the Constitution of the State of Minnesota, Article IV, Section 5, pertaining to the distribution of the excise tax on petroleum products.*

Be it enacted by the Legislature of the State of Minnesota:

Section 1. The following amendments to the Constitution of the State of Minnesota, Article IV, Section 5, is hereby proposed to the people of the state for their approval or rejection, which section, when amended, shall read as follows:

Sec. 5. Proposed amendment to Minnesota Constitution, Article IX, Section 5. For the purpose of defraying extraordinary expenditures, the state may contract public debts, but such debts shall never, in the aggregate, exceed two hundred and fifty thousand dollars; every such debt shall be authorized by law, for some single object, to be distinctly specified therein; and no such law shall take effect until it shall have been passed by the vote of two-thirds of the members of each branch of the legislature, to be recorded by yeas and nays on the journals of each house respectively; and every such law shall levy a tax annually sufficient to pay the annual interest of such debt, and also a tax sufficient to pay the principal of such debt within ten years from the final passage of such law, and shall specially appropriate the proceeds of such taxes to the payment of such principal and interest; and such appropriation and taxes shall not be repealed, postponed, or diminished, until the principal and interest of such debt shall have been wholly paid. The state shall never contract any debts for works of internal improvements, or be a party in carrying on such works, except as authorized by Section 16 of Article 9, and by Article 16 of this Constitution, but it may levy an excise tax upon any substance, material, fluid, force or other means or instrumentality, or the business of dealing in, selling or producing any or all thereof, used or useful, in producing or generating power for propelling motor or other vehicles used on the public highways of this state, and shall place one-half of the proceeds of such tax in the Trunk Highway Fund provided for in Section 2 of said Article 16, apportion for road and street purposes on the basis of population 6% thereof to cities of over 100,000 population,

3. See whether that measure is part of today’s constitution.
  - a. Go to the relevant section on [this page](#).

- b. Note that for pre-1974 measures, it will often be the case that the article number has changed. You should look at the [1974 constitutional revisions here](#) to find what the status was prior to 1974 and where the section moved to, if anywhere.
  4. If today's status is the same as the outcome of the referendum, skip to step 6.
  5. If today's status is not the same as the outcome of the election, observe when that section of the constitution was amended. Starting from the earliest amendment after the measure in question, do the following:
    - a. Find the amendment using [this page](#).
    - b. Note whether the amendment repealed, passed, or modified the original measure you are working on (i.e. the one in the **ballotname** column). If so:
      - i. Record the date in the next **transitionDate** column.
      - ii. Fill in **literalStatus**, **InPlaceNarrow**, and, if relevant, any of **StrongerNarrow** through **OppositeFederal** as applicable.
      - iii. Include the link in **source** (and any other links in the following source columns).
      - iv. If this is the last transition you find, include a link to the current law in the next **source** column.
    - c. If the change in question indicates that the prior law did not match what you had written in the previous set of columns, flag it, and leave a note saying this.
  6. For any period of 20 years or more in which the status did not change, do the following:
    - a. Search the ballot measures spreadsheet for other measures on the same topic passed after the initial one, and ensure none of the affect its status.
    - b. You do not need to do this, but I will search newspapers.com and proquest.com separately for the state and each of (a) the name and year of the measure (e.g. North Dakota "Proposition 9" 1986), (b) the topic of the measure plus the phrase "ballot measure" (e.g. North Dakota abortion "ballot measure"), and (c) the topic of the measure plus the word "referendum" (e.g. California abortion referendum). I will record each change to the status of the referendum you find in the news sources.

5. I expect this will generally not apply to Minnesota. If you find arguments for or against the measure, we want to copy and paste this into the Arguments sheet. On the **Arguments** sheet you'll find another list of the ballot\_ids like in the Checks sheet. For each measure, simply paste pro- and con- arguments you find into the respective columns. Copying the text as you find it is fine; please don't spend time editing or summarizing these arguments.

#### Sources:

- Ballot measure details, including links to text:
  - <https://www.lrl.mn.gov/mngov/constitutionalamendments>
- Minnesota constitution:
  - <https://www.revisor.mn.gov/constitution/>
- 1974 Minnesota constitutional revision/reorganization:
  - <https://www.revisor.mn.gov/laws/1974/0/Session+Law/Chapter/409/pdf/>

### **A.3 Instructions for Additional Variables**

The next pages contain instructions to an Upwork assistant for recording additional baseline covariates. Note that the variable referred to as “compelled” in these instructions is what the paper describes, for purposes of clarity, as the share of a population “directly affected” by a policy, and the variable referred to as “exposed” is what I refer to as the share “indirectly affected.”

### Ballot Measure Data Input Instructions: International

I want some additional information on each of the referendums I've collected data on so far. Please go through the "Checks" sheet for the following states in the batches folder:

- Texas
- Oklahoma
- North Dakota
- Ohio
- Missouri
- California
- Oregon
- Arizona
- Alabama
- Colorado
- Illinois
- Minnesota
- Nevada
- Utah
- Washington
- Wyoming
- New York
- Massachusetts
- Maryland
- Maine

I have done an example of the work for Florida if you look at that sheet. You can insert the columns where they are on the "Checks" sheet for other states.

For each one, please add the following variables. You may want to do one or a few variables at a time for the entire sheet rather than going measure by measure. Do what is most efficient.

|             |   |  |
|-------------|---|--|
| <b>need</b> | Is this in response to a pressing need or crisis? | <u>1</u> for any referendum that seems likely to be caused by a perceived pressing need or crisis (e.g., growing government deficit, administrative overload, moral panic)<br><u>2</u> if it is clearly in response to a crisis (a war, financial crisis, government insolvency)<br><u>0</u> otherwise |
|-------------|---|--|

|                           |   |   |
|---------------------------|---|---|
| <b>monetary</b>           | Is this about a dollar amount?                        | <u>1</u> if the referendum involves a dollar amount (e.g. a tax rate, a number of legislatures, a percent of voters)<br><u>2</u> if the main point of the referendum is the dollar amount (e.g., all the referendum does is modify a number). Note that percentages do NOT count—all that counts is something in actual dollars or cents.<br><u>0</u> otherwise<br>Add 0.5 if the dollar figure is (or was, in the case of deletion of a dollar figure) inflation-adjusted.   |
| <b>numeric</b>            | Is this about changing, adding, or deleting a number? | <u>1</u> if the referendum involves a number (e.g. a tax rate, a number of legislatures, a percent of voters)<br><u>2</u> if the main point of the referendum is the number (e.g., all the referendum does is modify a number). Anything that gets rid of a number and replaces it with, e.g, letting the legislature choose should be a “1”. Anything requiring a vote is numeric since it could be a 50% requirement or a supermajority; beware of other cases like this where there is a number that is not obvious.<br><u>0</u> otherwise |
| <b>material</b>           | Does this have a direct, material effect?             | <u>1</u> if the referendum plausibly has some direct, material impact on someone (e.g., gives someone money, affects someone’s job, imprisons someone)<br><u>2</u> if the referendum definitely has a material impact<br><u>0</u> if not (e.g., changes the structure of government without an obvious material impact)   |
| <b>reversalAdjustment</b> | Would reversing this require adjustment costs?        | <u>2</u> if reversing the referendum would likely impose economic   |

|                      |  |   |
|----------------------|--|---|
|                      |  | <p>costs because of adjustments people make. For example, a property tax cut or exemption would lead people to purchase more of that sort of property so that undoing the tax cut or exemption is now costly to them.</p> <p><u>1</u> if reversing the referendum would impose administrative costs because of adjustments people make. For example, establishing a new city or government agency might lead people to depend on that city/government agency.</p> <p><u>0</u> otherwise</p>           |
| <b>compelled</b>     | What share of the population does this compel or allow to take a certain action? | <p><u>1</u> = specific, identified offices (e.g. treasurer, legislator)</p> <p><u>2</u> = people in unusual situations; one or two small towns</p> <p><u>3</u> = an entire industry; a demographic group; a large region or city; multiple small towns; everyone who interacts with the court system or other non-routine government contact</p> <p><u>4</u> = all property taxpayers; all income taxpayers; all people of a given sex; all parents</p> <p><u>5</u> = all residents or all voters</p> |
| <b>exposed</b>       | What share of the population would witness this being implemented?               | same as for compelled   |
| <b>impactAverage</b> | How large an observable impact would this have on people on average?             | <p><u>1</u> no clear benefit</p> <p><u>2</u> incidental benefits, less than next category</p> <p><u>3</u> modification of rights and legal activities, or change to government transfers</p> <p><u>4</u> granting or revocation of right or majority of government transfers received</p> <p><u>5</u> life or livelihood</p>  |

|                            |  |   |
|----------------------------|--|---|
|                            |  |   |
| <b>impactExtreme</b>       | How large an observable impact would this have on those most affected by it? | same as for impactAverage   |
| <b>obsolete</b>            | Would this measure likely or definitely become obsolete with time?           | 0 = not likely<br>1 = likely<br>2 = definitely<br><br>Note that anything that includes a dollar figure that is <b>not</b> inflation-adjusted would qualify here.  |
| <b>obsoletePossibility</b> | If this measure became obsolete, how could it do so?                         | 0 = nothing<br>1 = inflation<br>2 = economic or technological change (e.g., the technology becomes obsolete)<br>3 = changes to the structure of government (e.g., a referendum concerns the salary for a particular office, which could disappear)<br>4 = changes to another law it depends on (e.g., a referendum concerns an exemption from a tax, and the tax could be eliminated) |

### A.3.1 The Persistence of Referendum Campaigns: Treatment on the Treated

While equation 1 can estimate the effect of passing a policy proposal on whether it is operative years later, a related question concerns the effect of launching a successful policy campaign. Borrowing a strategy from the study of school bond referendums in Cellini et al. (2010), I estimate the effect of a successful campaign using a recursive algorithm.

My measure of the effect of a successful referendum campaign estimates the effect of passage in a world where the same set of proponents would never again pass the policy except for the current referendum. To do this, I identify, for each referendum, each subsequent year in which the exact same policy passes as part of the same campaign using historical records. Specifically, I construct an indicator  $C_{rt}$  that indicates that the campaign behind the original referendum successfully passes the policy. I then estimate equation 1 with this as the outcome, and I refer to the coefficient from this regression,  $\beta_{1t}$ , as the probability of a “defier”,  $\Pi_t$ , in line with the program evaluation literature. Intuitively,  $\Pi_t$  is a measure of how many campaigns succeed in year  $t$ . I refer to my estimates of  $\beta_{1t}$  from equation 1 with my primary outcome (i.e., whether a policy is operative) as my “intent to treat” estimate,  $ITT_t$ .

With the estimates of defiers together with my main effects, I apply a recursive method to estimate the effect of a successful campaign on policy over time. Taking my estimate of “defiers”  $\Pi_t$  and “intent to treat”  $ITT_t$ , I estimate a “treatment on the treated” (TOT) effect following Cellini et al. (2010):

$$\begin{aligned}
 TOT_0 &= 1 \\
 TOT_1 &= ITT_1 - TOT_0\Pi_1 \\
 TOT_2 &= ITT_2 - TOT_0\Pi_2 - TOT_1\Pi_1 \\
 &\dots \\
 TOT_N &= ITT_N - \sum_{i=0}^{N-1} TOT_i\Pi_{N-i}
 \end{aligned} \tag{7}$$

This method captures the effect of passing a referendum on whether the policy is operative decades later were the original proponents never to attempt it again. Cellini et al. (2010) employ this method to estimate the effect of approving a school bond in California on property values and use the method to remove the effect on property values via causing or preventing further bond measures. In a similar sense, my TOT estimates remove the effect of the same proponents later on and allow us to understand the persistence of a policy campaign, all considered.

The key assumption behind the TOT approach is that the effect of a campaign succeeding on the first try is the same as the effect of that campaign succeeding  $t$  years later except for the effect



on whether that campaign itself succeeds. That is, there must be a consistent TOT effect over time. This would be falsified if, for example, proponents are more likely to give up as time goes on. In practice, this is a fairly strong assumption. If in fact proponents are more likely to give up as time goes on, this would bias the TOT estimates downward. Nevertheless, this procedure allows us to get a sense of how persistent proponents' choices to pursue a policy are.

#### A.4 Sample Dialogue with Claude about Federal Legislation

**1943 Congressional Records on Contentious Legislation** ▾

PDF
Compiled\_Pa...  
6.20 MB

Could you please go through this series of congressional records and create a table that lists each piece of legislation that either narrowly passed or just barely failed?

The table should have six columns: the name of the legislation, the year, an indicator (1 or 0) for whether it passed, the outcome, vote details, and an explanation of why you think it was close.

Be sure to cover the entire time period in the document and only that, i.e. 1943-1947. Please also note that whether something is close is a matter of context, such as how contentious the debate was, and is not simply a matter of how close the vote was.

Here is a table listing pieces of close legislation from the congressional records from 1943-1947:

| Legislation                    | Year | Passed (1/0) | Outcome                       | Vote Details                          | Explanation  |
|--------------------------------|------|--------------|-------------------------------|---------------------------------------|--|
| Connally-Smith Anti-Strike Act | 1943 | 1            | Passed over presidential veto | House: 244 to 108<br>Senate: 56 to 25 | The presidential veto and contentious debate indicate this was a close and controversial bill. |

ZF

|                               |      |   |  |   |  |
|-------------------------------|------|---|--|---|--|
| Current Tax Payment Act       | 1943 | 1 | Passed after disagreement between House and Senate | Final House vote: 256 to 114<br>Final Senate vote: 62 to 19 | Multiple votes and revisions during passage through Congress suggest this was a close call.    |
| Revenue Act of 1943           | 1943 | 1 | Passed with large bipartisan majority              | House: voice vote   | Large bipartisan support but contentious tax policy debates indicate this was not a slam dunk. |
| Philippine Rehabilitation Act | 1946 | 1 | Passed unanimously                                 | Senate: voice vote<br>House: voice vote                     | Uncontroversial passage but large price tag suggests some backroom dealmaking was required.    |

|                                      |      |   |                                    |   |  |
|--------------------------------------|------|---|------------------------------------|---|--|
| School Lunch Act                     | 1946 | 1 | Passed with overwhelming support   | House: 276 to 101<br>Senate: voice vote         | Strong bipartisan support but inclusion of controversial non-discrimination provision brought this close to failing. |
| Federal Airport Act                  | 1946 | 1 | Passed after conference compromise | House: voice vote<br>Senate: 49 to 32           | Divisions between House and Senate versions required compromise to pass.   |
| Fair Employment Practices Commission | 1946 | 0 | Filibustered in Senate             | Senate vote to end debate: 48 to 36 (needed 67) | Contentious issue that sparked a lengthy filibuster indicates an extremely close and controversial bill.             |

## B Additional Theoretical Framework and Mechanisms

### B.1 Explaining the Infrequency of Repeal Attempts

High proposal costs predict far fewer attempts to repeal referendums than to repeat them under reasonable assumptions about proponents' and opponents' willingness to propose. To illustrate, consider a proposed referendum, and suppose that  $b_{1t}$  and  $b_{0t}$  are drawn according to some symmetric, joint distribution  $F_b$ . For a policy that is unopposed, supporters will pay to pass the policy whenever the lifetime payoff  $\sum_{t=0}^{\infty} \delta^t b_{1t}$ , times the chance of winning, is sufficiently large relative to the proposal cost. When there are dedicated opponents, the calculus is less favorable: given the same payoff to having the policy operative, the value of proposing decreases since the passed policy may be repealed. Opponents also face a less favorable calculus than unopposed supporters because supporters will attempt to undo any successful repeal of the policy. Even if supporters' and opponents' benefits were perfectly correlated, a random draw will be more likely to yield a proposed referendum without an active opposition.

To illustrate, Appendix Figure E18 draws values of  $b_{1t}$  and  $b_{0t}$  for various distributions and shows, under plausible assumptions on the discount factor and cost of proposing, that there is rarely a repeal attempt even conditional on there being an initial proposal. In all plots, I assume the costs of proposing are sufficient to deter 99% of proponents, and I assume that proponents' and opponents' willingness to pay is highly correlated. In the lognormal plots, I draw a distribution of benefits to match a desired correlation. In the Pareto distribution plots, I assume an exponent of one, so the correlation is not well-defined. To draw the willingness to pay from Pareto distributions, I draw variables from a multivariate normal distribution with a desired correlation and transform the variables such that the Pareto CDF of the transformed variable is the same as the normal CDF of the original variable. In the lognormal plots, it is virtually impossible for there to be a repeal attempt because of how rarely both sides are willing to attempt a referendum with the threat of repeal. In the Pareto plots, we observe repeal attempts around a quarter of the time, but this is under a fairly extreme assumption. These distributions are quite simplified, as they assume common knowledge of each side's willingness to pay, but they show that with plausible distributions and even an extremely high correlation between the two sides' willingness to pay, there is rarely an opponent dedicated enough to attempt repeal.

Two empirical patterns support this selection-based explanation for the low rate of repeals. First, data from the nonprofit Open Secrets show that the pro-referendum side does in fact typically spend substantially more, with 25% of "no" campaigns spending nothing compared to only 10% of "yes" campaigns. See Appendix Figure E14 for the full distribution, with proponents also spending at higher levels conditional on spending. Second, if the selection effect operates over time, we should

expect the rate of repeals to be fairly constant over time, in contrast to repeat attempts. This is exactly what Figure 4 shows.

## B.2 Persistence and Political Learning

### B.2.1 How Learning from Political Defeat Can Generate Persistence

Referendum outcomes can also have a lasting effect on long-term policy if narrow failure makes proponents durably pessimistic. Under this story, some policies are operative for long stretches of time simply by virtue of advocates' limited information about voter preferences.

Suppose that advocates are uncertain about voters' attitudes. Formally, suppose voter preferences are unknown but drawn from some known distribution at the start of the game. This can generate a dynamic where obtaining a vote share near the threshold for passage is a negative update on the odds of success, as the next result shows.

PROPOSITION A1 (Persistence from pessimistic updates): *Suppose the following hold:*

- (1)  $(v_t, \tau_t, c_{xt}, b_{xt}) = (v, \tau, c_x, b_x)$  are constant over time.
- (2)  $v$  is drawn from a normal distribution with mean  $\mu_v > \tau$  and standard deviation  $\sigma_v$  censored at 0 and 1.
- (3)  $\varepsilon_t$  is drawn from a normal distribution with mean 0 and standard deviation  $\sigma_\varepsilon$  censored at  $-v$  and  $1 - v$ .
- (3) All parameters except  $\varepsilon_t$  and  $v$  are common knowledge at the time of referendum proposals.
- (4)  $c_0 > \sum_{t=0}^{\infty} \delta^t b_0$

There exist values  $\mu_v, \sigma_v, \sigma_\varepsilon, \bar{b}_1, c_1$  such that  $\lim_{t \rightarrow \infty} \mathbb{E}[P_t | \theta_{10} = 1, v_0 = \tau] > 0$  in any Nash equilibrium. In addition, the following hold:

- (A)  $\lim_{t \rightarrow \infty} \mathbb{E}[P_t | \forall t \in T, \theta_{1t} = 1, v_t = \tau]$  is increasing in  $N$ .
- (B) Holding all other parameters constant,  $P_t$  is decreasing in  $\sigma_\varepsilon$  for all  $t$ .

Observing a close referendum tells a proponent that their public support is weaker than they had thought, which can cause them to not propose again even if their prior was that proposing was worth doing. A nuance of this dynamic is that those who succeed narrowly would also receive a negative signal from the narrow outcome. They have no interest in repealing their preferred policy once it is operative, however, so this does not affect subsequent events. If repeal attempts are rare enough, this generates a lasting wedge between the share of successful and failed referendums that are operative.

For empirical data, Proposition A1 predicts that there should be greater persistence following more informative votes and, all else equal, more votes. Proposition A1 (A) predicts that there will

be more persistence as the number of observed failed votes rises, which coincides with Hypothesis 9. Proposition A1 (B) predicts that when observed votes are more informative about public sentiment, there will be more persistence.<sup>47</sup> We can test this by analyzing heterogeneity with regard to low compared to moderate voter turnout, as low voter turnout implies both a smaller and a more unusual share of the population is voting.<sup>48</sup>

### B.2.2 An Intensive-Margin Effect of Political Learning

The political learning mechanism from Appendix Section B.2.1 yielded two hypotheses predicting that there will be persistence after proponents of a referendum receive more negative news. In the data, I find little support for these hypotheses with my primary measure of persistence, but I do find support for the hypotheses when I broaden my measure of persistence to include weaker versions of the proposed policy.

Political learning predicts that there should be more repeat attempts after low-turnout elections (Hypothesis 11) and fewer repeat attempts as the number of recent votes on the topic rises (Hypothesis 9). I discussed and offered evidence against the latter in the previous section, but Appendix Figure E24 gives evidence that in states where proponents cannot repeat a referendum exactly, they substitute toward weaker versions of the same policy. Hypothesis 11 former hypothesis I study in Appendix Figure E25, which shows heterogeneity of persistence by turnout at the election. Turnout should generally be exogenous to details of the referendum itself since the big-ticket races are rarely referendums. When using my primary measure of persistence, I reject Hypothesis 11. When we look at whether a weaker version of a policy is in place, however, we see that proponents are far more likely to attempt a weaker version of the proposed policy after low-turnout elections and less likely after high-turnout elections, suggesting there is some learning from election results (Appendix Figure E27).

## B.3 Proofs

**Proof of Proposition 1.** Assume 1–4 are satisfied.

Let it be the case that  $\forall t, c_{0t} = c_{1t} = 0, \tau_t = 0.5$ . Define the function  $\Omega_x(v_t, F_{\varepsilon t})$  to be the likelihood that a referendum passes when the initial policy is  $1 - x$ , i.e., the likelihood that a referendum to change the policy to  $x$  passes:

---

<sup>47</sup>Certain signals may not only be noisier signal but also be distorted, i.e.,  $\varepsilon_t$  may not have mean 0. However, a rational advocate can correct for the distortion.

<sup>48</sup>The effect of especially high turnout is unclear because extremely high turnout is abnormal but also may capture the full population's preferences more accurately.

$$\Omega_x(v, \tau, F_\varepsilon) = \begin{cases} F_\varepsilon(0.5 - v) & \text{if } x = 0 \\ 1 - F_\varepsilon(0.5 - v) & \text{if } x = 1 \end{cases}$$

Note that  $\Omega_x(v_t, F_{\varepsilon t}) = 1 - \Omega_{1-x}(v_t, F_{\varepsilon t})$ .

Consider any subgame starting from some period  $\underline{t}$ . We can estimate a lower bound on advocate  $a_x$ 's payoff in the following way. Suppose  $a_x$  proposes a referendum in any period  $t \geq \underline{t}$  where they do not like the policy, i.e., where  $p_{t-1} \neq x$ . In any period where  $p_{t-1} = x$ , the advocate gets an expected payoff of at least  $(1 - \Omega_{1-x}(v_t, F_{\varepsilon t}))b_{xt} = \Omega_x(v_t, F_{\varepsilon t})b_{xt}$ , since even if the opponent proposes, they have a probability  $\Omega_x(v_t, F_{\varepsilon t})$  of the referendum going their way; in any period where  $p_{t-1} = 1 - x$ , the advocate also gets an expected payoff of  $\Omega_x(v_t, F_{\varepsilon t})b_{xt}$ . Hence in the overall game advocate  $a_x$  can always get a payoff of at least  $\sum_{t=\underline{t}}^{\infty} \delta^t \Omega_x(v_t, F_{\varepsilon t})b_{xt}$ .

Finally, let  $\pi_{xt}$  be the expected payoff to advocate  $x$  in period  $t$ . Note that  $\forall t, \pi_{xt} + \pi_{1-x,t} \frac{b_{xt}}{b_{1-x,t}} = b_{xt}$ . In a Nash equilibrium of the subgame, the following follows:

$$\begin{aligned} \sum_{t=\underline{t}}^{\infty} \delta^t \pi_{xt} &= \sum_{t=\underline{t}}^{\infty} \delta^t (b_{xt} - \pi_{1-x,t} \frac{b_{xt}}{b_{1-x,t}}) = \sum_{t=\underline{t}}^{\infty} \delta^t (1 - \frac{\pi_{1-x,t}}{b_{1-x,t}}) b_{xt} \\ &\leq \sum_{t=\underline{t}}^{\infty} \delta^t (1 - \frac{\Omega_x(v_t, F_{\varepsilon t}) b_{1-x,t}}{b_{1-x,t}}) b_{xt} = \sum_{t=\underline{t}}^{\infty} \delta^t \Omega_x(v_t, F_{\varepsilon t}) b_{xt} \end{aligned}$$

Therefore,  $a_x$ 's total payoff in a Nash equilibrium of the subgame must be exactly  $\sum_{t=\underline{t}}^{\infty} \delta^t \Omega_x(v_t, F_{\varepsilon t}) b_{xt}$ .

Now suppose in some period  $p_{\underline{t}-2} = 1 - x$ , and in period  $\underline{t} - 1$ ,  $a_x$  does not propose. By the above argument,  $a_x$ 's payoff in the subgame starting from period  $\underline{t} - 1$  is  $\sum_{t=\underline{t}}^{\infty} \delta^t \Omega_x(v_t, F_{\varepsilon t}) b_{xt}$ . If they do propose,  $a_x$ 's payoff is  $\Omega_x(v_t, F_{\varepsilon_{\underline{t}-1}}) b_{x,\underline{t}-1} + \sum_{t=\underline{t}}^{\infty} \delta^t \Omega_x(v_t, F_{\varepsilon t}) b_{xt}$

This implies that if  $a_x$  does not propose in period  $\underline{t}$ , their payoff is  $\sum_{t=\underline{t}}^{\infty} \delta^t \Omega_x(v_t, F_{\varepsilon t}) b_{xt}$ . If they propose, their payoff is  $\sum_{t=\underline{t}-1}^{\infty} \delta^t \Omega_x(v_t, F_{\varepsilon t}) b_{xt}$ , which is greater whenever  $\Omega_x(v_{\underline{t}-1}, F_{\varepsilon_{\underline{t}-1}}) > 0$ . Therefore, in a Nash equilibrium each advocate must propose whenever they have a chance of winning.

It follows from the above that

$$\begin{aligned} \mathbb{P}[p_t = 1 | p_{t-1} = 1] &= (1 - \Omega_0(v_t, F_{\varepsilon t})) \max(\theta_{0t}, \theta_{1t}) \\ &= (1 - \Omega_0(v_t, F_{\varepsilon t})) \max(\mathbb{1}\{\Omega_0(v_t, F_{\varepsilon t}) > 0\}, 0) \\ &= (1 - \Omega_0(v_t, F_{\varepsilon t})) \end{aligned}$$

But also,

$$\mathbb{P}[p_t = 1 | p_{t-1} = 0] = \Omega_1(v_t, F_{\varepsilon t}) \max(\theta_{0t}, \theta_{1t})$$

$$\begin{aligned}
&= \Omega_1(v_t, F_{\varepsilon t}) \max(0, \mathbb{1}\{\Omega_1(v_t, F_{\varepsilon t}) > 0\}) \\
&= \Omega_1(v_t, F_{\varepsilon t}) = (1 - \Omega_0(v_t, F_{\varepsilon t})) = \mathbb{P}[p_t = 1 | p_{t-1} = 1]
\end{aligned}$$

So  $\mathbb{P}[p_t = 1 | p_{t-1} = 1] = \mathbb{P}[p_t = 1 | p_{t-1} = 0]$ . Note also that this holds regardless of  $p_\tau$  for  $\tau < t - 1$ . We therefore obtain the following:

$$\begin{aligned}
P_t &= \mathbb{P}[p_t = 1 | p_0 = 1] - \mathbb{P}[p_t = 1 | p_0 = 0] \\
&= \mathbb{P}[p_t = 1 | p_1 = 1, p_0 = 1] \mathbb{P}[p_1 = 1 | p_0 = 1] + \mathbb{P}[p_t = 1 | p_1 = 0, p_0 = 1] \mathbb{P}[p_1 = 0 | p_0 = 1] \\
&\quad - (\mathbb{P}[p_t = 1 | p_1 = 1, p_0 = 0] \mathbb{P}[p_1 = 1 | p_0 = 0] + \mathbb{P}[p_t = 1 | p_1 = 0, p_0 = 0] \mathbb{P}[p_1 = 0 | p_0 = 0]) \\
&= \mathbb{P}[p_t = 1 | p_1 = 1] \mathbb{P}[p_1 = 1 | p_0 = 1] + \mathbb{P}[p_t = 1 | p_1 = 0] \mathbb{P}[p_1 = 0 | p_0 = 1] \\
&\quad - (\mathbb{P}[p_t = 1 | p_1 = 1] \mathbb{P}[p_1 = 1 | p_0 = 0] + \mathbb{P}[p_t = 1 | p_1 = 0] \mathbb{P}[p_1 = 0 | p_0 = 0]) \\
&= \mathbb{P}[p_t = 1 | p_1 = 1] (\mathbb{P}[p_1 = 1 | p_0 = 1] - \mathbb{P}[p_1 = 1 | p_0 = 0]) + \mathbb{P}[p_t = 1 | p_1 = 0] (\mathbb{P}[p_1 = 0 | p_0 = 1] - \mathbb{P}[p_1 = 0 | p_0 = 0]) \\
&= \mathbb{P}[p_t = 1 | p_1 = 1] (0) + \mathbb{P}[p_t = 1 | p_1 = 0] (0) = 0
\end{aligned}$$

**Proof of Proposition 2.** Assume 1–2 hold.  $a_x$ 's payoff in the subgame starting from period  $t$  can be given by the following:

$$\begin{aligned}
V_t(p_{t-1} = x, \theta_{xt}) &= [1 - \max(\theta_{xt}, \theta_{1-x,t}) + \max(\theta_{xt}, \theta_{1-x,t}) \Omega_x(v, \tau, F_\varepsilon)] \times [b_x + \delta \mathbb{E}_{\theta_{x,t+1}} [V_{t+1}(p_t = x, \theta_{x,t+1})]] \\
&\quad - \theta_{xt} c_x + \max(\theta_{xt}, \theta_{1-x,t}) (1 - \Omega_x(v, \tau, F_\varepsilon)) \times \delta \mathbb{E}_{\theta_{x,t+1}} [V_{t+1}(p_t = 1 - x, \theta_{x,t+1})]
\end{aligned}$$

Note that no variable from earlier periods is relevant to the payoff. The same holds when  $p_{t-1} = 1 - x$ . This implies that  $\theta_{xt}$  will be independent of all variables from earlier periods except  $p_{t-1}$  in any Markov strategy. The expression for  $a_{1-x}$  is symmetric, implying  $\theta_{1-x,t}$  is also independent of all variables from earlier periods except  $p_{t-1}$ . Note therefore that

$$\begin{aligned}
&\mathbb{P}[p_t = x | p_{t-1} = x] - \mathbb{P}[p_t = x | p_{t-1} = 1 - x] \\
&= \mathbb{E}[\max(\theta_{xt}, \theta_{1-x,t}) (1 - \Omega_x(v, \tau, F_\varepsilon)) | p_{t-1} = x] - \mathbb{E}[\max(\theta_{xt}, \theta_{1-x,t}) \Omega_x(v, \tau, F_\varepsilon) | p_{t-1} = 1 - x]
\end{aligned}$$

which is constant. Define  $q$  to be equal to this quantity. Observe that  $P_1 = \mathbb{P}[p_1 = 1 | p_0 = 1] - \mathbb{P}[p_1 = 1 | p_0 = 0] = q$ , and also:

$$P_t = \mathbb{P}[p_t = 1 | p_0 = 1] - \mathbb{P}[p_t = 1 | p_0 = 0]$$



$$\begin{aligned}
&= \mathbb{P}[p_t = 1 | p_{t-1} = 1, p_0 = 1] \mathbb{P}[p_{t-1} = 1 | p_0 = 1] + \mathbb{P}[p_t = 1 | p_{t-1} = 0, p_0 = 1] \mathbb{P}[p_{t-1} = 0 | p_0 = 1] \\
&- (\mathbb{P}[p_t = 1 | p_{t-1} = 1, p_0 = 0] \mathbb{P}[p_{t-1} = 1 | p_0 = 0] + \mathbb{P}[p_t = 1 | p_{t-1} = 0, p_0 = 0] \mathbb{P}[p_{t-1} = 0 | p_0 = 0]) \\
&= \mathbb{P}[p_t = 1 | p_{t-1} = 1] (\mathbb{P}[p_{t-1} = 1 | p_0 = 1] - \mathbb{P}[p_{t-1} = 1 | p_0 = 0]) \\
&+ \mathbb{P}[p_t = 1 | p_{t-1} = 0] (1 - \mathbb{P}[p_{t-1} = 1 | p_0 = 1] - (1 - \mathbb{P}[p_{t-1} = 1 | p_0 = 0])) \\
&= (\mathbb{P}[p_t = 1 | p_{t-1} = 1] - \mathbb{P}[p_t = 1 | p_{t-1} = 0]) (\mathbb{P}[p_{t-1} = 1 | p_0 = 1] - \mathbb{P}[p_{t-1} = 1 | p_0 = 0]) \\
&= q (\mathbb{P}[p_{t-1} = 1 | p_0 = 1] - \mathbb{P}[p_{t-1} = 1 | p_0 = 0])
\end{aligned}$$

Hence, by induction,  $P_t = q^t$ .

**Proof of Proposition 3.** Assume 1–4 hold. As shown in the proof of Proposition 2, the likelihood of a policy being in place in a given period is independent of the prior policy history. Since  $a_0$  never proposes,  $P_t$  is equal to one minus the probability that  $a_1$  ever successfully proposes. By Boole's inequality, this is at least as great as one minus the sum of the probabilities that  $a_1$  ever successfully proposes. Therefore

$$P_t \geq 1 - \mathbb{E}_{\eta_1, \eta_2, \dots, \eta_t} \left[ \prod_{i=1}^t [1 - (1 - F_\varepsilon(\tau - v)) \times \mathbb{P}((1 - F_\varepsilon(\tau - v))(b_{1i} + \mathbb{E}[\sum_{j=i+1}^{\infty} \delta^{j-2} b_{1j} | b_{1i}]) > c_1 | (1 - F_\varepsilon(\tau - v))(b_{10} + \mathbb{E}[\sum_{j=1}^{\infty} \delta^{j-1} b_{1j} | b_{10}]) > c_1))] \right]$$

Note first that  $b_{1i} + \mathbb{E}[\sum_{j=i+1}^{\infty} \delta^{j-i} b_{1j} | b_{1i}] = b_{1i} + \sum_{j=i+1}^{\infty} \delta^{j-i} (\phi^{j-i} b_{1i} + (1 - \phi^{j-i}) \bar{b}_1 + \sum_{k=i+1}^j \phi^{k-i-1} \mathbb{E}[\eta_k]) = \frac{b_{1i} - \delta \phi \bar{b}_1}{1 - \delta \phi} + \frac{\delta \bar{b}_1}{1 - \delta}$ . Let  $\bar{\eta}_i = \sum_{j=i}^{\infty} \phi^{i-j} \eta_j$ , and let  $\bar{\eta}_{i,k} = \sum_{j=i}^k \phi^{i-j} \eta_j$ . Then  $b_{1i} + \mathbb{E}[\sum_{j=i+1}^{\infty} \delta^{j-i} b_{1j} | b_{1i}] = \frac{\bar{b}_1}{1 - \delta} + \frac{\bar{\eta}_i}{1 - \delta \phi} = \frac{\bar{b}_1}{1 - \delta} + \frac{\bar{\eta}_{i,k}}{1 - \delta \phi} + \frac{\phi^{i-k} \bar{\eta}_{k-1}}{1 - \delta \phi}$  for any  $k$ . Hence we can rewrite the above as

$$\begin{aligned}
P_t &= \mathbb{E}_{\eta_1, \eta_2, \dots, \eta_t} \left[ \prod_{i=1}^t [1 - (1 - F_\varepsilon(\tau - v)) \times \right. \\
&\left. \mathbb{P}\left(\left(\frac{\bar{\eta}_{i,1}}{1 - \delta \phi} > \frac{c_1}{1 - F_\varepsilon(\tau - v)} - \frac{\bar{b}_1}{1 - \delta} - \frac{\phi^{i-1} \bar{\eta}_0}{1 - \delta \phi} \mid \frac{\bar{\eta}_0}{1 - \delta \phi} > \frac{c_1}{1 - F_\varepsilon(\tau - v)} - \frac{\bar{b}_1}{1 - \delta}\right)\right)] \right]
\end{aligned}$$

Next, by the law of iterated expectations, this is equal to

$$P_t = \mathbb{E}_{\eta_1, \eta_2, \dots, \eta_t} \left[ \mathbb{E} \left[ \prod_{i=1}^t [1 - (1 - F_\varepsilon(\tau - v)) \times \mathbb{P}\left(\left(\frac{\bar{\eta}_{i,1}}{1 - \delta \phi} > \frac{c_1}{1 - F_\varepsilon(\tau - v)} - \frac{\bar{b}_1}{1 - \delta} - \frac{\phi^{i-1} \bar{\eta}_0}{1 - \delta \phi} \mid \frac{\bar{\eta}_0}{1 - \delta \phi} > \frac{c_1}{1 - F_\varepsilon(\tau - v)} - \frac{\bar{b}_1}{1 - \delta}\right)\right)] \mid \eta_1, \dots, \eta_{t-1} \right] \right]$$

Note that the conditional expectation inside the unconditional expectation is zero if, for any  $j \in [1, t-1]$ ,  $\frac{\bar{\eta}_{j,1}}{1 - \delta \phi} > \frac{c_1}{1 - F_\varepsilon(\tau - v)} - \frac{\bar{b}_1}{1 - \delta} - \frac{\phi \bar{\eta}_0}{1 - \delta \phi}$ ; otherwise, it is equal to

$$\begin{aligned}
& 1 - (1 - F_\varepsilon(\tau - v)) \times \\
\mathbb{P}\left(\left(\frac{\eta_t}{1 - \delta\phi} > \frac{c_1}{1 - F_\varepsilon(\tau - v)} - \frac{\bar{b}_1}{1 - \delta} - \frac{\phi^{t-1}\bar{\eta}_0 + \bar{\eta}_{t-1,1}}{1 - \delta\phi} \mid \frac{\bar{\eta}_0}{1 - \delta\phi} > \frac{c_1}{1 - F_\varepsilon(\tau - v)} - \frac{\bar{b}_1}{1 - \delta}, \right.\right. \\
& \left.\left. \forall j \in [1, t], \frac{\bar{\eta}_j}{1 - \delta\phi} < \frac{c_1}{1 - F_\varepsilon(\tau - v)} - \frac{\bar{b}_1}{1 - \delta}\right)
\end{aligned}$$

Compare this term to  $P_1$ :

$$\begin{aligned}
& P_1 = 1 - (1 - F_\varepsilon(\tau - v)) \times \\
\mathbb{P}\left(\left(\frac{\eta_1}{1 - \delta\phi} > \frac{c_1}{1 - F_\varepsilon(\tau - v)} - \frac{\bar{b}_1}{1 - \delta} - \frac{\bar{\eta}_0}{1 - \delta\phi} \mid \frac{\bar{\eta}_0}{1 - \delta\phi} > \frac{c_1}{1 - F_\varepsilon(\tau - v)} - \frac{\bar{b}_1}{1 - \delta}\right)\right)
\end{aligned}$$

Since  $\eta_1$  and  $\eta_t$  are identically distributed, these two terms are the same except the expression on the right-hand side of the probability differs by

$$-\frac{\phi^{i-1}\bar{\eta}_0 + \bar{\eta}_{t-1,1}}{1 - \delta\phi} - \left(-\frac{\bar{\eta}_0}{1 - \delta\phi}\right) = \frac{(1 - \phi^{t-1})\bar{\eta}_0 - \bar{\eta}_{t-1,1}}{1 - \delta\phi}$$

We know that  $\frac{\bar{\eta}_0}{1 - \delta\phi} > \frac{c_1}{1 - F_\varepsilon(\tau - v)} - \frac{\bar{b}_1}{1 - \delta} > 0$ . Given  $\frac{\bar{\eta}_j}{1 - \delta\phi} < \frac{c_1}{1 - F_\varepsilon(\tau - v)} - \frac{\bar{b}_1}{1 - \delta}$  and  $\frac{\bar{\eta}_0}{1 - \delta\phi} > \frac{c_1}{1 - F_\varepsilon(\tau - v)} - \frac{\bar{b}_1}{1 - \delta}$ , it follows that  $\bar{\eta}_{t-1,1} < 0$ . Hence,  $P_1 < \mathbb{E}[P_t \mid \forall j \in [1, t], \frac{\bar{\eta}_j}{1 - \delta\phi} < \frac{c_1}{1 - F_\varepsilon(\tau - v)} - \frac{\bar{b}_1}{1 - \delta}]$ . This implies that

$$\begin{aligned}
& P_t > \mathbb{E}_{\eta_1, \eta_2, \dots, \eta_t} \left[ P_1 \prod_{i=1}^t \left[ 1 - (1 - F_\varepsilon(\tau - v)) \times \right. \right. \\
& \left. \left. \mathbb{P}\left(\left(\frac{\bar{\eta}_{i,1}}{1 - \delta\phi} > \frac{c_1}{1 - F_\varepsilon(\tau - v)} - \frac{\bar{b}_1}{1 - \delta} - \frac{\phi^{i-1}\bar{\eta}_0}{1 - \delta\phi} \mid \frac{\bar{\eta}_0}{1 - \delta\phi} > \frac{c_1}{1 - F_\varepsilon(\tau - v)} - \frac{\bar{b}_1}{1 - \delta}\right)\right) \right] \right] \\
& = \mathbb{E}_{\eta_1, \eta_2, \dots, \eta_t} [P_1 P_{t-1}]
\end{aligned}$$

Applying induction, we have that  $P_t > \mathbb{E}[P_1^t] \geq (\mathbb{E}[P_1])^t = P_1^t$ , where the latter follows from Jensen's inequality.

To prove the first part of (A), observe that (1)-(4) imply the following for  $b_{1,t}$ :

$$\mathbb{E}[\bar{b}_1, t \mid b_{1,0}] = \mathbb{E}[\phi^t b_{1,0} + (1 - \phi^t)\bar{b}_1 + \sum_{i=1}^t \phi^{t-i} \eta_i] = \phi^t b_{1,0} + (1 - \phi^t)\bar{b}_1$$

From the above, we know that  $b_{1,0} > \bar{b}_1$ , implying  $\mathbb{E}[\bar{b}_1, t \mid b_{1,0}] = \phi^t b_{1,0} + (1 - \phi^t)\bar{b}_1 < \bar{b}_1$ . To prove the second part of (A), we can take the stationary distribution of  $b_{1k}$  for random  $k$  as our prior for the value of  $b_{1,-t}$ . Since the distribution of  $\phi b_{1k} + (1 - \phi)\bar{b}_1 + \eta_k$  must be the same as the distribution of  $b_{1k}$ , the distribution of  $b_{1k}$  must form a conjugate prior for  $F_\eta$ . Given the conditions in Proposition 3 (A),  $b_{1k}$  must be in the exponential family (Diaconis and Ylvisaker, 1979). Furthermore, we know

that  $\phi^i b_{1k} + (1 - \phi^i) \bar{b}_1 + \sum_{j=0}^i \phi^j \eta_{k_j}$  must also have the same distribution as  $b_{1k}$  for all  $i$ , implying that  $\sum_{j=0}^i \phi^j \eta_{k_j}$  must be a conjugate sampling distribution to  $b_{1k}$ . This implies that  $\sum_{j=0}^i \phi^j \eta_{k_j}$  must be in the exponential family of distributions (Bar-Lev et al., 1994). This in turn implies that, for some  $c \in (0, 1)$ ,

$$\mathbb{E}[\bar{b}_1, -t | b_{1,0}] = c \mathbb{E}[\bar{b}_1, -t] + (1 - c) b_{1,0} = c \bar{b}_1 + (1 - c) b_{1,0} < b_{1,0} \text{ for } b_{1,0} > \bar{b}_1$$

Hence,  $\mathbb{E}[b_{1,-t} | a_1 \text{ proposes in period } 0] = \mathbb{E}[\mathbb{E}[b_{1,-t} | b_{1,0}] | a_1 \text{ proposes in period } 0] < \mathbb{E}[b_{1,0} | a_1 \text{ proposes in period } 0]$

To prove (B), inspect the second expression for  $P_t$  above. Holding  $b_{10}$  and  $\phi$  fixed implies that  $\frac{\bar{\eta}_0}{1 - \delta \phi}$  is fixed. It therefore follows that any increase in  $\bar{b}_1$  increases the probability  $a_1$  successfully proposes a later referendum, lowering  $P_t$ .

**Proof of Proposition A1.** First, let us observe that  $a_x$  learns from the first referendum. Specifically, each decisionmaker's subjective prior distribution over  $v + \varepsilon_t$  is  $\mathcal{N}(\mu, \sigma_\varepsilon^2 + \sigma_v^2)$ . After observing  $v + \varepsilon_0 = \tau$ , each decisionmaker's posterior is

$$\mathcal{N}\left(\frac{\frac{\mu}{\sigma_v^2} + \frac{\tau}{\sigma_\varepsilon^2}}{\frac{1}{\sigma_v^2} + \frac{1}{\sigma_\varepsilon^2}}, \frac{\frac{\sigma_\varepsilon^2}{\sigma_v^2} + 2}{\frac{1}{\sigma_v^2} + \frac{1}{\sigma_\varepsilon^2}}\right)$$

Therefore, the subjective prior over the likelihood that  $v_t \leq \tau$  is  $\Phi\left(\frac{\tau - \mu}{\sigma_\varepsilon^2 + \sigma_v^2}\right)$ , while the posterior is

$$\Phi\left(\frac{\frac{\tau}{\sigma_v^2} + \frac{\tau}{\sigma_\varepsilon^2} - \left(\frac{\mu}{\sigma_v^2} + \frac{\tau}{\sigma_\varepsilon^2}\right)}{\frac{\sigma_\varepsilon^2}{\sigma_v^2} + 2}\right) = \Phi\left(\frac{\frac{\tau}{\sigma_v^2} - \frac{\mu}{\sigma_v^2}}{\frac{\sigma_\varepsilon^2}{\sigma_v^2} + 2}\right) = \Phi\left(\frac{\tau - \mu}{\sigma_\varepsilon^2 + 2\sigma_v^2}\right) > \Phi\left(\frac{\tau - \mu}{\sigma_v^2 + \sigma_\varepsilon^2}\right)$$

Suppose  $\mathbb{P}[(1 - \Phi(\frac{\tau - \mu}{\sigma_v^2 + \sigma_\varepsilon^2})) \sum_{t=0}^{\infty} \delta^t \bar{b}_1 > c_{1t}] > 0$  but  $\mathbb{P}[(1 - \Phi(\frac{\tau - \mu}{\sigma_\varepsilon^2 + 2\sigma_v^2})) \sum_{t=0}^{\infty} \delta^t \bar{b}_1 > c_{1t}] = 0$ . Given  $\sum_{t=0}^{\infty} \delta^t b_0 < c_{0t}$ ,  $a_0$  will never propose. This implies  $a_1$  will propose with positive probability in period 0 but will never propose after observing a vote share of  $\tau$  in period 0. As a result,  $P_t = 1$ .

To prove (A), suppose the decisionmaker observes  $N$  votes with vote totals  $v_{t_1}, v_{t_2}, \dots, v_{t_N} = \tau$ . Now, their posterior is:

$$\mathcal{N}\left(\frac{\frac{\mu}{\sigma_v^2} + \frac{N\tau}{\sigma_\varepsilon^2}}{\frac{1}{\sigma_v^2} + \frac{N}{\sigma_\varepsilon^2}}, \frac{\frac{\sigma_\varepsilon^2}{\sigma_v^2} + N + 1}{\frac{1}{\sigma_v^2} + \frac{N}{\sigma_\varepsilon^2}}\right)$$

Accordingly, the posterior over the likelihood that  $v_t \leq \tau$  is

$$\Phi\left(\frac{\frac{\tau}{\sigma_v^2} + \frac{N\tau}{\sigma_\varepsilon^2} - \left(\frac{\mu}{\sigma_v^2} + \frac{N\tau}{\sigma_\varepsilon^2}\right)}{\frac{\sigma_\varepsilon^2}{\sigma_v^2} + N + 1}\right) = \Phi\left(\frac{\frac{\tau}{\sigma_v^2} - \frac{\mu}{\sigma_v^2}}{\frac{\sigma_\varepsilon^2}{\sigma_v^2} + N + 1}\right) = \Phi\left(\frac{\tau - \mu}{\sigma_\varepsilon^2 + (N + 1)\sigma_v^2}\right)$$

which is increasing in  $N$ .

To prove (B), note that as  $\sigma_\varepsilon$  decreases, the posterior probability of winning  $\Phi\left(\frac{\tau-\mu}{\sigma_\varepsilon^2+2\sigma_v^2}\right)$  grows.

**Proof that persistence bounds divergence from optimal policy.** Consider any function  $S$  that is independent of prior election outcomes and maps the binary policy space considered in Section 4.1 to a single optimal policy  $p_t^*$  in any period  $t$ . Assume that  $S$  is independent of past election outcomes.

Allowing for the possibility that there is some unknown information about  $p^*$ , let  $s = \mathbb{P}[p^* = 1]$ ; that is,  $s$  denotes the likelihood that the optimal policy is  $p^* = 1$ . Unknown information might include, for example, future economic or cultural changes.

For referendums that fail in period 0, the likelihood that  $p_t \neq p_t^*$  must be at least  $|\mathbb{P}[p_t = 1|p_0 = 1] - s|$ ; for those that fail,  $p_t \neq p_t^*$  must be at least as great as  $|s - \mathbb{P}[p_t = 1|p_0 = 0]|$ . For a referendum with even odds of passing, then,  $\mathbb{P}[p_t \neq p_t^*] \geq 0.5|\mathbb{P}[p_t = 1|p_0 = 1] - s| + 0.5|s - \mathbb{P}[p_t = 1|p_0 = 0]| \geq 0.5P_t$ .

More generally, for a referendum that has a chance of passing  $q$ ,  $\mathbb{P}[p_t \neq p_t^*] > q(\mathbb{P}[p_t = 1|p_0 = 1] - s) + (1 - q)(s - \mathbb{P}[p_t = 1|p_0 = 0])$ , which is greater than  $0.5P_t$  when  $q$  is high and  $s$  is low and less than  $0.5P_t$  when  $q$  is high and  $s$  is high.

**Proof that complete persistence implies unbiased event study estimate given baseline similarity.** Specifically, assume that the two states' policies are independent of each other, and  $Z$  is independent of past policies conditional on the current policy.<sup>49</sup> We then obtain the following:

$$\begin{aligned} & \mathbb{E}[Z_{it} - Z_{jt}|p_{it} = 1, p_{jt} = 0] \\ = & \frac{\mathbb{E}[Z_{it} - Z_{jt}|p_{i0} = 1, p_{j0} = 0] - \mathbb{E}[Z_{it} - Z_{jt}|p_{it} = 0, p_{jt} = 1]\mathbb{P}[p_{it} = 0|p_{i0} = 1]\mathbb{P}[p_{jt} = 1|p_{j0} = 0]}{\mathbb{P}[p_{it} = 1|p_{i0} = 1]\mathbb{P}[p_{jt} = 0|p_{j0} = 0]} \end{aligned}$$

Notice that when persistence  $P_t = 1$ , the latter expression simply becomes  $\mathbb{E}[Z_{it} - Z_{jt}|p_{i0} = 1, p_{j0} = 0]$ . That is, when persistence is high, the expectation conditional on present policies will approach the expectation conditional on historical policies. This is helpful, because it is often more plausible to assume that potential outcomes are independent of historical policies than current policies for the reasons reviewed above.

---

<sup>49</sup>Alternatively, we could allow  $Z$  to be a function of past policies but be interested in the effect of the past policy history, i.e., the effect of having the policy in place from period 0 through period  $t$ .

## C Additional Details on the Calibrated Model

### C.1 Estimating the Calibrated Model

Using the modeling assumptions described in Section 5.1, I identify parameter estimates that simulate the degree of persistence and the distribution of votes on initiatives I observe. I base my modeling assumptions on the reduced-form data and modified versions of the model in Section 5.1.

I estimate the baseline distributions of  $b_{10}$ ,  $v_0$ , and  $N$  by maximizing the likelihood function and the remaining parameters by manual inspection. As described in Section 5.1, the likelihood for a given observation is the expectation of a binomial density function. Specifically, for a given vote-share range  $[\underline{v}, \bar{v}]$  and proposal cost  $c$ , the chance of observing  $k$  initiatives is given by the following:

$$\int_{\underline{v}}^{\bar{v}} \int_{-\infty}^{\infty} f(k|N, v_0, \tau, \sigma_\varepsilon) f_{\mathcal{N}}\left(\frac{\hat{v} - v_0}{\sigma_\varepsilon}\right) f_{\mathcal{N}}\left(\frac{v_0 - \mu_v}{\sigma_v}\right) dv d\hat{v} \quad (8)$$

where

$$f(k|N, v, \tau, \sigma_\varepsilon) = \text{Bin}\left(N, \Phi\left(\frac{1}{\sigma_b^2 + \frac{\sigma_\eta^2}{1 - \phi_b^2}} (\log c - \log \Phi\left(\frac{\tau - v}{\sigma_\varepsilon}\right) - \mu_b)\right)\right) \quad (9)$$

The estimates of  $N$ ,  $\mu_b$ , and  $\sigma_b$  are not sensitive to starting values. I do a grid search over possible starting values of  $\sigma_\varepsilon$ ,  $\sigma_v$ , and  $\mu_v$ .

To estimate the time-series parameters  $\phi_b$ ,  $\sigma_\eta$ ,  $\rho$ ,  $\phi_v$ , and  $\sigma_\gamma$ , I start with a grid search and then adjust the parameters  $\phi_b$ ,  $\sigma_\eta$ , and  $\rho$  by manual inspection to fit the patterns in the data. The chance of observing  $k'$  repeats from  $N'$  baseline initiatives follows a similar form, but the expectation is also over  $b_{10}$ , with  $\mu_b$  and  $\sigma_b$  replaced by the conditional mean and conditional variance of  $b_{1t}$  given  $b_{10}$ . The chance of observing  $k'$  repeats from  $N'$  baseline initiatives is similar but with the conditional mean and variance of  $b_0$ . To make the integrals computationally tractable, I discretize the vote shares by transforming each vote share to the nearest multiple of 10 from 10 to 90. The likelihood function is flat in this region, so I adjust the parameters  $\phi_b$ ,  $\sigma_\eta$ , and  $\rho$  to fit the pattern of persistence in the data, yielding the parameters in Appendix Table C2. I also output the results of the calibration exercise with alternative parameters to understand how the estimates vary with alternative choices.

The assumptions of my calibrated model assume, based on the reduced form data and the estimates of alternative models, that vote share is uncorrelated with advocates' willingness to pay, and opponents' payoffs are constant. To test whether there is any substantial correlation between voter preferences and advocates' willingness to pay, I maximize a modified version of the likelihood function that allows for correlation. I estimate a correlation of -0.0009 between vote share and the

logarithm of the proponents' willingness to pay in this model. This assumption is also consistent with the observation that advocates willing to pay to propose a referendum are in the tail of the distribution of benefits and likely unrepresentative of voters. Finally, I assume opponents' willingness-to-pay is constant for computational simplicity and given the roughly constant rate of repeals throughout my sample period.

The parameters yield a pattern of persistence and distribution of vote share similar to what we see in the data. Appendix Table C2 provides the parameter estimates. The minimum willingness to pay for a proponent to attempt a referendum in any state in the data is about -5, so advocates willing to attempt a referendum are in the tails of the distribution. Around half the variance in vote share is accounted for by unanticipated noise, and half is accounted for by the known component. The parameters imply a small positive correlation between proponents' and opponents' willingness to pay, a high degree of autocorrelation in advocates' preferences, and a low degree of autocorrelation in voters' preferences.

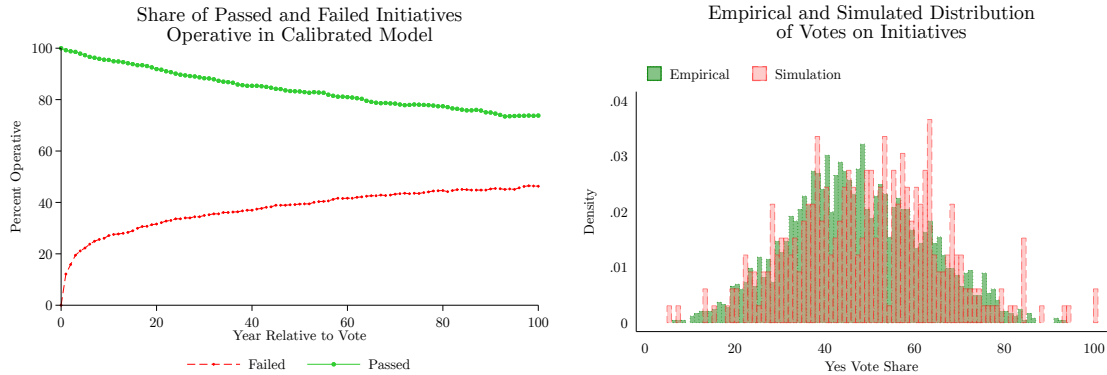
The simulated data with adjusted parameters fit the pattern of persistence I observe and the baseline vote share distribution. Appendix Table C3 presents the initial parameters output by the maximum likelihood estimation before manual adjustment. The initial parameters predict more repeat and repeal attempts than actually occur, as shown in Appendix Figure C14. With adjustment, the fit improves considerably. The left panel of Appendix Figure C13 shows that the predicted share of marginally passed and failed initiatives in place over time with the adjusted parameters is similar to the share in Figure 4. The right panel of Appendix Figure C13 shows that the predicted distribution of vote share is similar to the empirical distribution. There is somewhat of a jump in predicted vote share at 50% because the model assumes advocates are able to anticipate when they will win and are more likely to propose in these cases.

Appendix Table C2: Model Parameters

| $\mu_b$ | $\sigma_b$ | $\sigma_\varepsilon$ | $\mu_v$ | $\sigma_v$ | $N$     | $\sigma_\eta$ | $\phi_b$ | $\rho$ | $\sigma_\gamma$ | $\phi_v$ |
|---------|------------|----------------------|---------|------------|---------|---------------|----------|--------|-----------------|----------|
| -34.46  | 7.64       | 8.78                 | 28.94   | 12.51      | 2055.30 | 3.02          | 0.75     | 0.18   | 15.81           | 0.11     |

*Notes:* This table shows the parameters used for the counterfactual exercise. The parameters result from a maximum likelihood estimation and grid search, after which I adjust  $\phi_b$ ,  $\sigma_\eta$ , and  $\rho$  to fit the pattern of persistence in the data. Appendix Figure C13 shows the model fit with these parameters. Appendix Table C3 provides the values of the parameters before manual adjustment, and Appendix Figure C14 shows the model fit with the unadjusted parameters.

Appendix Figure C13: A Calibrated Model Fits Key Empirical Facts



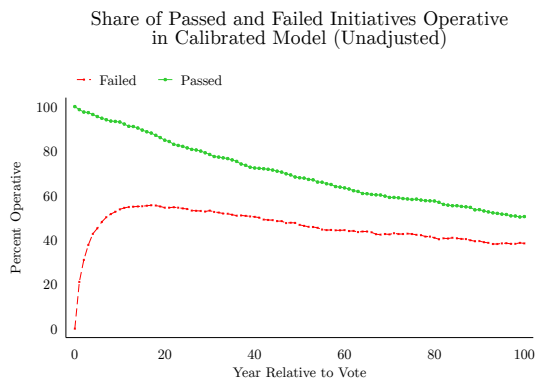
Notes: The two panels illustrate model fit under the calibrated model, using the parameters in Appendix Table C2. The left panel shows the simulated history of passed and failed initiatives that pass or fail by less than 2.5 percentage points for comparison with Figure 4. The right panel shows the empirical and simulated distributions of the vote share in favor of initiatives.

Appendix Table C3: Unadjusted Model Parameters

| $\mu_b$ | $\sigma_b$ | $\sigma_\varepsilon$ | $\mu_v$ | $\sigma_v$ | $N$     | $\sigma_\eta$ | $\phi_b$    | $\rho$      | $\sigma_\gamma$ | $\phi_v$ |
|---------|------------|----------------------|---------|------------|---------|---------------|-------------|-------------|-----------------|----------|
| -34.46  | 7.64       | 8.78                 | 28.94   | 12.51      | 2055.30 | <b>0.58</b>   | <b>0.96</b> | <b>0.29</b> | 15.81           | 0.11     |

Notes: This table shows the parameters of the model estimated by maximum likelihood and a grid search. Appendix Figure C14 depicts the model fit for the degree of persistence. The estimates predict too many repeat and repeal attempts, so I adjust  $\phi_B$  and  $\rho$  downwards, and  $\sigma_\eta$  accordingly.

Appendix Figure C14: Unadjusted Parameter Estimates Underestimate Persistence

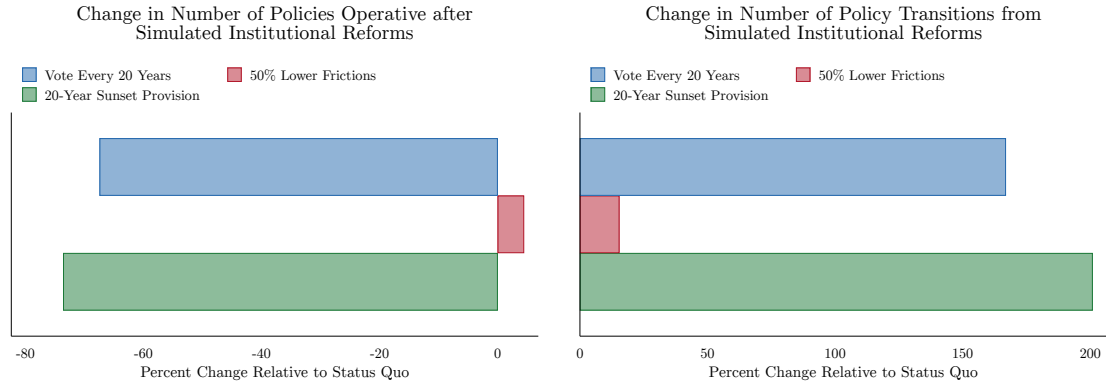


*Notes:* The left panel shows the simulated history of passed and failed initiatives that pass or fail by less than 2.5 percentage points using the initial parameters shown in Appendix Table C3, produced by maximum likelihood estimation and a grid search. These parameters produce too many repeat and repeal attempts. The alternative parameterization presented in Appendix Table C2 produces a better fit to the data, as can be seen by comparing Appendix Figure C13 with Figure 4.



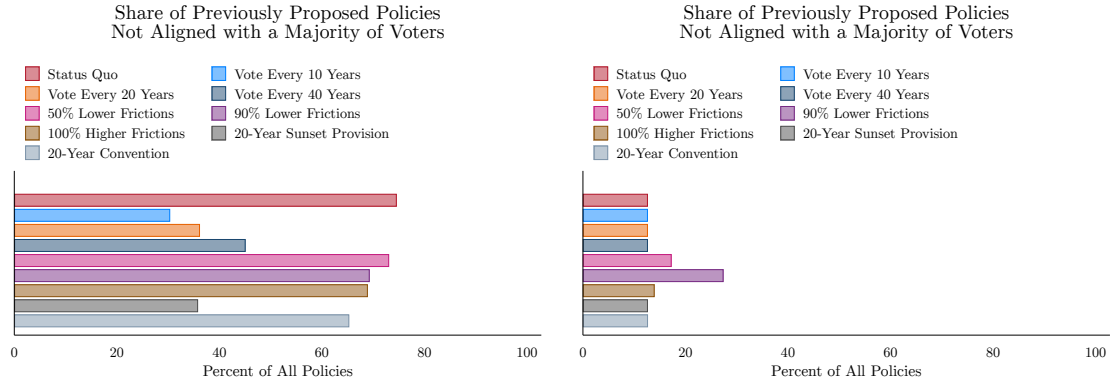
## C.2 Additional Results from the Calibrated Model

Appendix Figure C15: Tradeoffs—Institutional Reforms Increase Policy Instability, Repeals



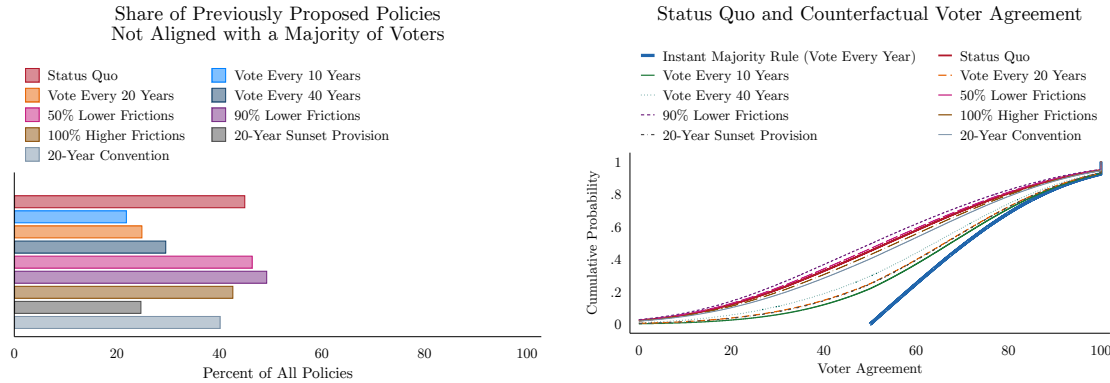
*Notes:* The left panel shows the percentage change in the number of previously proposed policies that are operative after 100 years of a simulated policymaking process for various institutional reforms and an idealized, instantaneous democracy compared to the status quo. The right panel shows the percentage change in the number of policy transitions relative to the status quo. An instantaneous democracy would involve 40 times as many policy transitions as occur now. Requiring votes every 20 years and sunset provisions both decrease the share of policies that are operative relative to the status quo more than the idealized democracy would select.

Appendix Figure C16: Policy Mismatch Happens Primarily by Passing Policies that Become Unpopular



Notes: The left and right panel show the share of policies whose status does not match voters' preferences, i.e., right panel of Figure 10, for policies that are operative and not operative, respectively. Policy mismatch is much greater for operative policies, and institutional reforms cause fewer unpopular policies to be in place.

Appendix Figure C17: Policy Mismatch with Additional Institutional Reforms



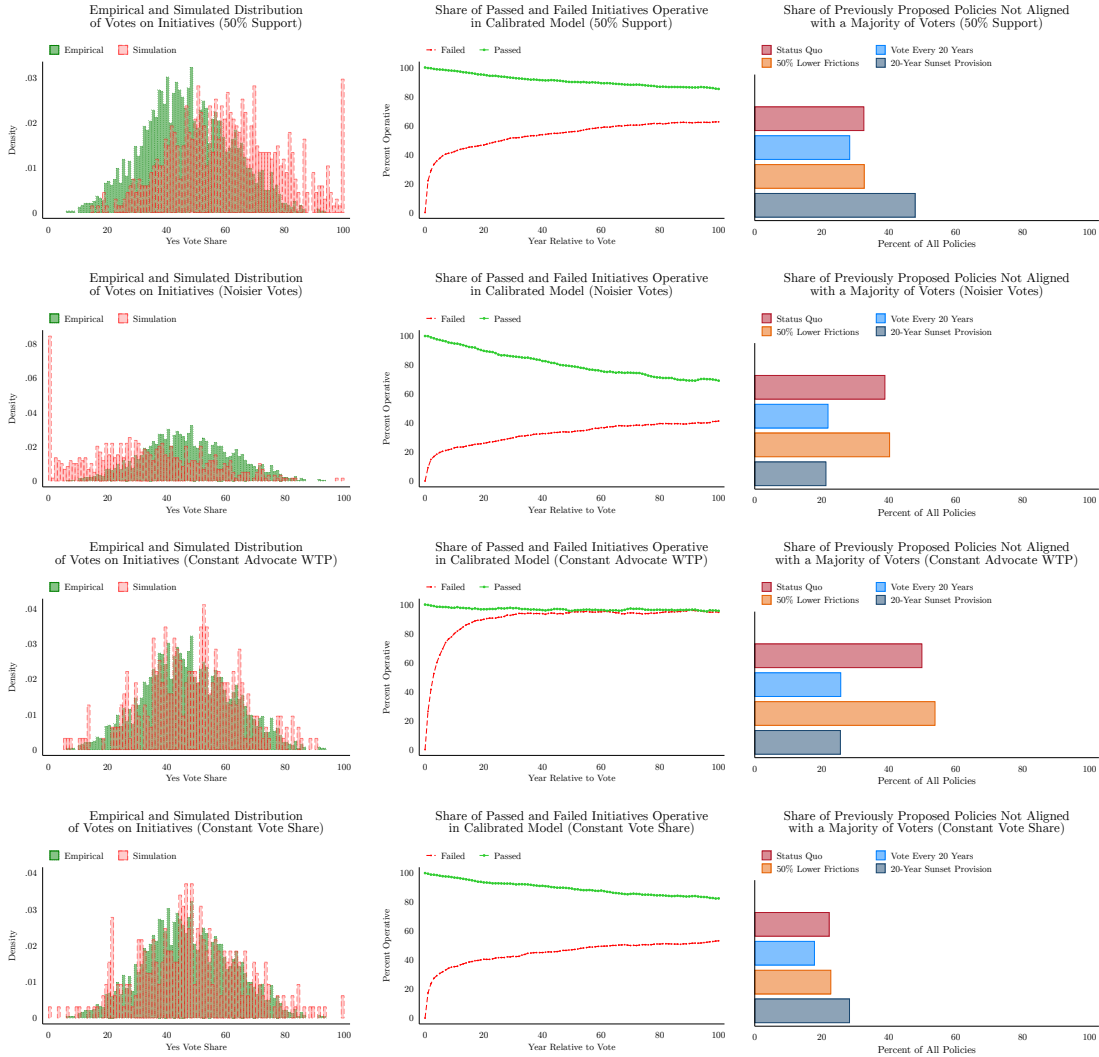
Notes: This figure shows analogous results to Figure 10 but for a wider set of institutional reforms. The figure depicts the level of voter agreement with previously proposed policies after simulating a calibrated model of 100 years of policymaking. The left panel shows the share of policies whose status matches voters' preferences, meaning the share of policies that are currently operative and have voter support or are not operative and lack voter support. The right panel shows the cumulative distribution of voter agreement with policies. Both panels show the status quo and three possible institutional reforms. Requiring votes periodically or 20-year sunset provisions increase policy alignment with voters' preferences, while reducing frictions does not.

### C.3 Alternative Parametrizations

Appendix Table C4: Alternative Model Parameters

| Parameterization      | $\mu_b$ | $\sigma_b$ | $\sigma_\varepsilon$ | $\mu_v$   | $\sigma_v$  | $N$    | $\sigma_\eta$ | $\phi_b$ | $\rho$ | $\sigma_\gamma$ | $\phi_v$ |
|-----------------------|---------|------------|----------------------|-----------|-------------|--------|---------------|----------|--------|-----------------|----------|
| Baseline              | -34.46  | 4.6306     | 8.78                 | 28.94     | 6.092       | 2055.3 | 2.4761        | 0.6      | 0.22   | 7.2545          | 0.7      |
| 50% Support           | -34.46  | 4.6306     | 8.78                 | <b>50</b> | 6.092       | 2055.3 | 2.4761        | 0.6      | 0.22   | 7.2545          | 0.7      |
| Noisier Votes         | -34.46  | 4.6306     | <b>6.092</b>         | 28.94     | <b>8.78</b> | 2055.3 | 2.4761        | 0.6      | 0.22   | 7.2545          | 0.7      |
| Constant Advocate WTP | -34.46  | 10.3       | 8.78                 | 28.94     | 6.092       | 2055.3 | <b>0</b>      | <b>0</b> | 0.22   | 7.2545          | 0.7      |
| Baseline              | -34.46  | 4.6306     | 8.78                 | 28.94     | 20.3        | 2055.3 | 2.4761        | 0.6      | 0.22   | <b>0</b>        | <b>0</b> |

Appendix Figure C18: Calibration and Policy Mismatch for Alternative Parameters



Notes: This figure shows model fit and results for alternative parameterizations described in Appendix Table C4. The first two columns show the same graphs as in Appendix Figure C13 for each parameterization, and the final column shows the right panel of Figure 10 for the alternative parameterization.

## D Additional Tables

Appendix Table D1: Balance—Sample Characteristics and Predictions at Threshold

|                | NCSL   | Within BW | Failed | Passed | RD P-Value |
|----------------|--------|-----------|--------|--------|------------|
| <i>Overall</i> |        |           |        |        |            |
| Passed         | 0.580  | 0.513     | 0.000  | 1.000  | .          |
| Vote Share     | 53.284 | 50.580    | 50.411 | 50.740 | 0.480      |
| General        | 0.849  | 0.723     | 0.744  | 0.703  | 0.527      |
| Primary        | 0.093  | 0.093     | 0.079  | 0.106  | 0.552      |
| Legislative    | 0.619  | 0.747     | 0.729  | 0.763  | 0.602      |
| Initiative     | 0.321  | 0.203     | 0.198  | 0.208  | 0.882      |
| <i>Topics</i>  |        |           |        |        |            |
| Taxation       | 0.192  | 0.186     | 0.113  | 0.256  | 0.013      |
| Budgets        | 0.067  | 0.076     | 0.078  | 0.073  | 0.910      |
| Business       | 0.068  | 0.092     | 0.116  | 0.069  | 0.242      |
| State Govt.    | 0.131  | 0.121     | 0.131  | 0.112  | 0.685      |
| Local Govt.    | 0.112  | 0.157     | 0.127  | 0.185  | 0.267      |
| Supermajority  |        | 0.065     | 0.051  | 0.077  | 0.436      |
| Constitutional |        | 0.774     | 0.766  | 0.782  | 0.790      |
| Observations   | 7772   | 824       | 401    | 423    | 824        |

*Notes:* The first column shows the values of baseline variables for referendums in the NCSL data (supermajority and constitutional variables are not in the NCSL data and require manual entry). The second column gives characteristics for those referendums in our sample of states and bandwidth. The third and fourth columns present the predicted value of each variable for passed and failed referendums at the threshold, and the fifth column presents the p-value of the coefficient on passage from estimating equation (1) with the respective variable as the outcome. The only significant imbalance is with regard to whether a referendum concerns taxation. Appendix Table D3 shows that imbalance by topic is approximately that which we would expect by random chance.

Appendix Table D2: Balance—Sample Characteristics for Passed and Failed

|                | Failed | Passed | T-Test P-Value |
|----------------|--------|--------|----------------|
| <i>Overall</i> |        |        |                |
| Passed         | 0.000  | 1.000  | .              |
| Vote Share     | 49.335 | 52.078 | 0.000          |
| Supermajority  | 0.052  | 0.064  | 0.483          |
| Constitutional | 0.742  | 0.796  | 0.065          |
| General        | 0.736  | 0.678  | 0.069          |
| Primary        | 0.104  | 0.108  | 0.850          |
| Legislative    | 0.668  | 0.755  | 0.006          |
| Initiative     | 0.264  | 0.206  | 0.051          |
| <i>Topics</i>  |        |        |                |
| Taxation       | 0.160  | 0.265  | 0.000          |
| Budgets        | 0.085  | 0.102  | 0.406          |
| Business       | 0.100  | 0.073  | 0.176          |
| State Govt.    | 0.135  | 0.109  | 0.255          |
| Local Govt.    | 0.150  | 0.175  | 0.326          |
| Observations   | 401    | 423    | 1              |

*Notes:* The columns show the values of baseline variables for passed and failed referendums in the sample.

Appendix Table D3: Balance—Sample Topics

|                                  | Failed | Passed | RD P-Value |
|----------------------------------|--------|--------|------------|
| Abortion                         | -0.002 | 0.020  | 0.068      |
| Agriculture                      | 0.028  | 0.035  | 0.741      |
| Animals/Hunting/Fishing          | 0.007  | 0.010  | 0.913      |
| Arts/Culture                     | 0.007  | 0.010  | 0.818      |
| Banks/Finance                    | 0.040  | 0.010  | 0.110      |
| Bonds                            | 0.037  | 0.004  | 0.066      |
| Government Budgets               | 0.078  | 0.073  | 0.910      |
| Business/Commerce                | 0.116  | 0.069  | 0.242      |
| Civil/Const. Law                 | 0.117  | 0.072  | 0.237      |
| Crime                            | 0.022  | 0.039  | 0.547      |
| Drugs/Alcohol                    | 0.047  | 0.062  | 0.630      |
| Economic Development             | 0.008  | 0.020  | 0.626      |
| Higher Education                 | 0.058  | 0.030  | 0.310      |
| Schools                          | 0.061  | 0.092  | 0.396      |
| Elections                        | 0.087  | 0.103  | 0.716      |
| Initiatives/Referendum           | 0.014  | 0.025  | 0.617      |
| Energy/Electricity               | 0.022  | 0.005  | 0.223      |
| Environment                      | 0.023  | 0.036  | 0.549      |
| Ethics/Lobbying/Campaign Finance | 0.007  | 0.010  | 0.844      |
| Observations                     | 401    | 423    | 825        |

Appendix Table D4: Balance—Sample Topics (CONT'D)

|                       | Failed | Passed | RD P-Value |
|-----------------------|--------|--------|------------|
| Federal Gov.          | 0.001  | 0.021  | 0.097      |
| Gambling Lottery      | 0.022  | 0.037  | 0.452      |
| Health                | 0.053  | 0.066  | 0.656      |
| Human Services        | 0.042  | 0.011  | 0.122      |
| Insurance             | 0.004  | -0.000 | 0.623      |
| Judiciary             | 0.107  | 0.039  | 0.056      |
| Juvenile Justice      | -0.002 | 0.004  | 0.467      |
| Labor/Employment      | 0.006  | 0.057  | 0.116      |
| Land Use/Prop. Rights | 0.044  | 0.030  | 0.570      |
| Legislature           | 0.073  | 0.058  | 0.733      |
| Local Gov.            | 0.127  | 0.185  | 0.267      |
| Military Veteran      | 0.004  | 0.004  | 0.998      |
| Natural Resource      | 0.029  | 0.037  | 0.771      |
| Redistricting         | 0.018  | 0.002  | 0.346      |
| State Gov.            | 0.131  | 0.112  | 0.685      |
| Native Americans      | -0.003 | 0.005  | 0.209      |
| Taxes/Revenue         | 0.113  | 0.256  | 0.013      |
| Telecommunications    | -0.001 | 0.000  | 0.781      |
| Term Limits           | 0.001  | 0.002  | 0.938      |
| Transportation        | 0.052  | 0.039  | 0.680      |
| Observations          | 401    | 423    | 825        |



Appendix Table D5: Persistence Over Time—Point Estimates for Figure 4

|                      | Year 5              | Year 10             | Year 40             | Year 100           |
|----------------------|---------------------|---------------------|---------------------|--------------------|
| Passed               | 0.720***<br>(0.049) | 0.692***<br>(0.053) | 0.599***<br>(0.078) | 0.469**<br>(0.178) |
| Dist. from Threshold | 0.045<br>(0.027)    | 0.022<br>(0.030)    | -0.019<br>(0.046)   | -0.061<br>(0.084)  |
| Passed x Dist.       | -0.039<br>(0.032)   | -0.035<br>(0.036)   | 0.030<br>(0.055)    | 0.048<br>(0.130)   |
| Observations         | 754                 | 722                 | 472                 | 107                |

\*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ . Standard errors in parentheses.

*Notes:* The table presents the coefficients from estimating equation (1) for 5, 10, 40, and 100-year outcomes, the same estimates shown graphically in Figure 4. The number of observations declines over time as the number of referendums long enough ago to have outcome data goes down.

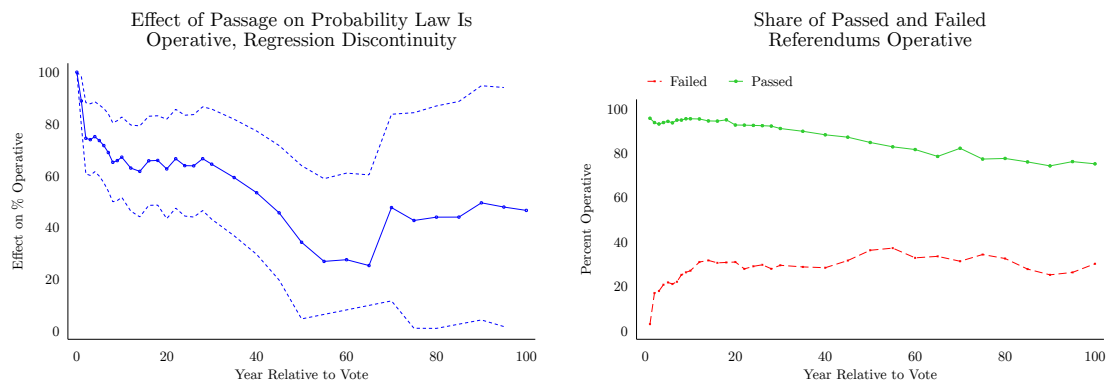
Appendix Table D6: Approval is Equivalent to a 200% Increase in the Amount of Time a Policy Is Operative

|   | $\delta = 7\%$ | $\delta = 3\%$ | $\delta = 1.7\%$ | $\delta = 0.1\%$ |
|---|----------------|----------------|------------------|------------------|
| Empirical Estimate                              | 355            | 236            | 189              | 42               |
| Theoretical Benchmark, $c=0$ , $p= 40\%$        | 19             | 8              | 4                | 0                |
| Theoretical Benchmark, Moderate $c$ , $p= 40\%$ | 17             | 7              | 4                | 0                |
| Theoretical Benchmark, $c=0$ , $p= 20\%$        | 38             | 15             | 9                | 1                |
| Theoretical Benchmark, Moderate $c$ , $p= 20\%$ | 34             | 15             | 8                | 0                |

*Notes:* Each column uses persistence and the likelihood of failed policies passing later to calculate  $\frac{\sum_{t=0}^{\infty} \delta^t P_t}{\sum_{t=0}^{\infty} \delta^t \mathbb{E}[p_t | p_0=0]}$  under a different discount rate  $\delta$ . The first row uses empirical estimates. The lower three rows offer theoretical benchmarks as depicted in Figure 7 in a simple Downsian setting where costs are zero, so opponents attempt to repeal the law (rows two and four) or where costs are moderate, and opponents never attempt repeals (rows three and five). The second and third rows assume the likelihood of a referendum going in favor of the initially proposed policy change is 40%, and the second two rows 20%.

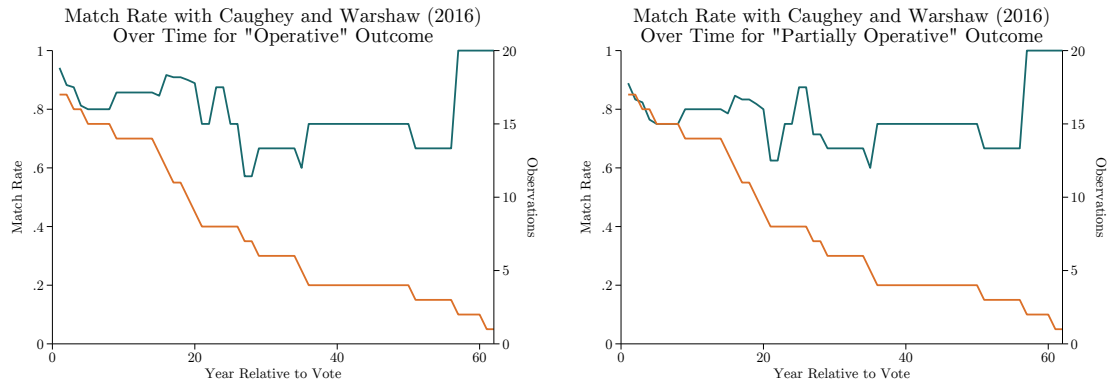
## E Additional Figures

Appendix Figure E1: Persistence Over Time—Estimates are Robust to a 1% Bandwidth



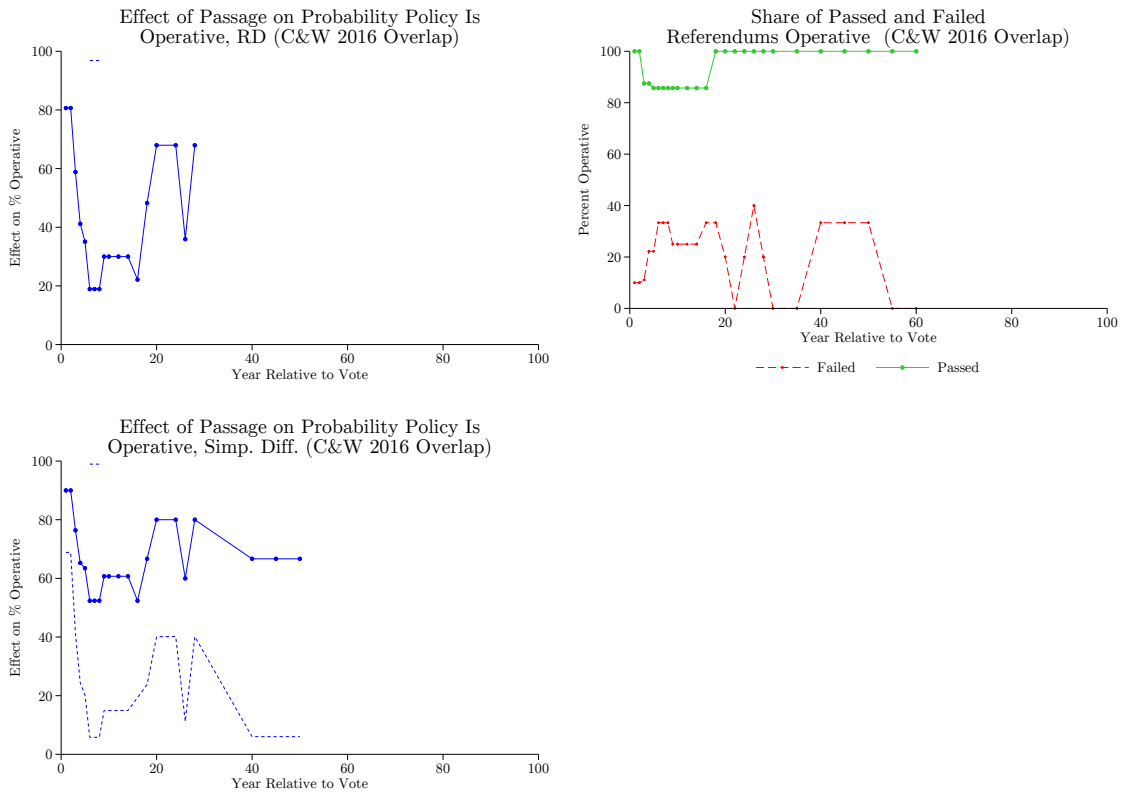
*Notes:* This figure shows analogous estimates to those in Figure 4 but only with observations within a 1% bandwidth. Each panel describes the evolution of the shares of narrowly passed and failed laws that are operative, using a holistic definition of the law in question and accounting for any numeric, substantive, or legal changes that weakened the law as described in Section 2.2. The left panel shows the coefficient on passage in percentage points, estimated with equation (1) in an unbalanced panel. Dotted lines signify the 95% confidence interval with standard errors at the state-topic level. For purposes of exposition, the graph shows coefficients every year for the first ten, every two years for the next twenty, and then every five years. The right panel shows the percent of narrowly passed and failed referendums that are operative over time. Many failed policies pass in the first few decades, after which persistence reaches an approximate plateau.

Appendix Figure E2: Match Rate with Caughey and Warshaw (2016) Indicators



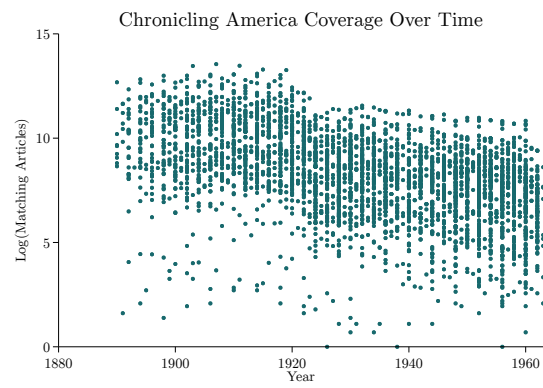
Notes: Each figure shows the share of referendums in a given year for which our data match Caughey and Warshaw (2016) and, on the right-*y*-axis, the number of observations.

Appendix Figure E3: Estimating Persistence using Caughey and Warshaw (2016) Indicators



Notes: Upper left panel shows persistence estimated in equation (1) where the outcome is a policy indicator from Caughey and Warshaw (2016). Upper right shows the share of passed and failed referendums that are in place using Caughey and Warshaw (2016) indicators. The lower panel shows the estimates in a simple difference specification.

Appendix Figure E4: Chronicing America Coverage Over Time



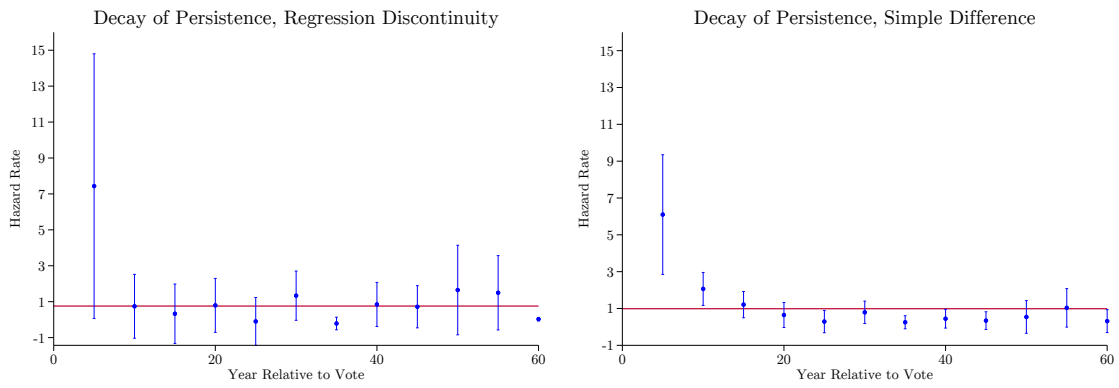
*Notes:* Each point represents the natural logarithm of the number of news articles matching a given referendum's keywords in the "Chronicing America" database in a four-year bin starting in that year.

Appendix Figure E5: Manipulation Testing for Variables that Correlate with Vote Share



*Notes:* The upper panels plot the share of observations that are proposed by the legislature (left) or at a general election (right) over the vote share. The lower panels plot the share as a histogram with a polynomial fit to the distribution for referendums proposed by the legislature (left) or at a general election (right).

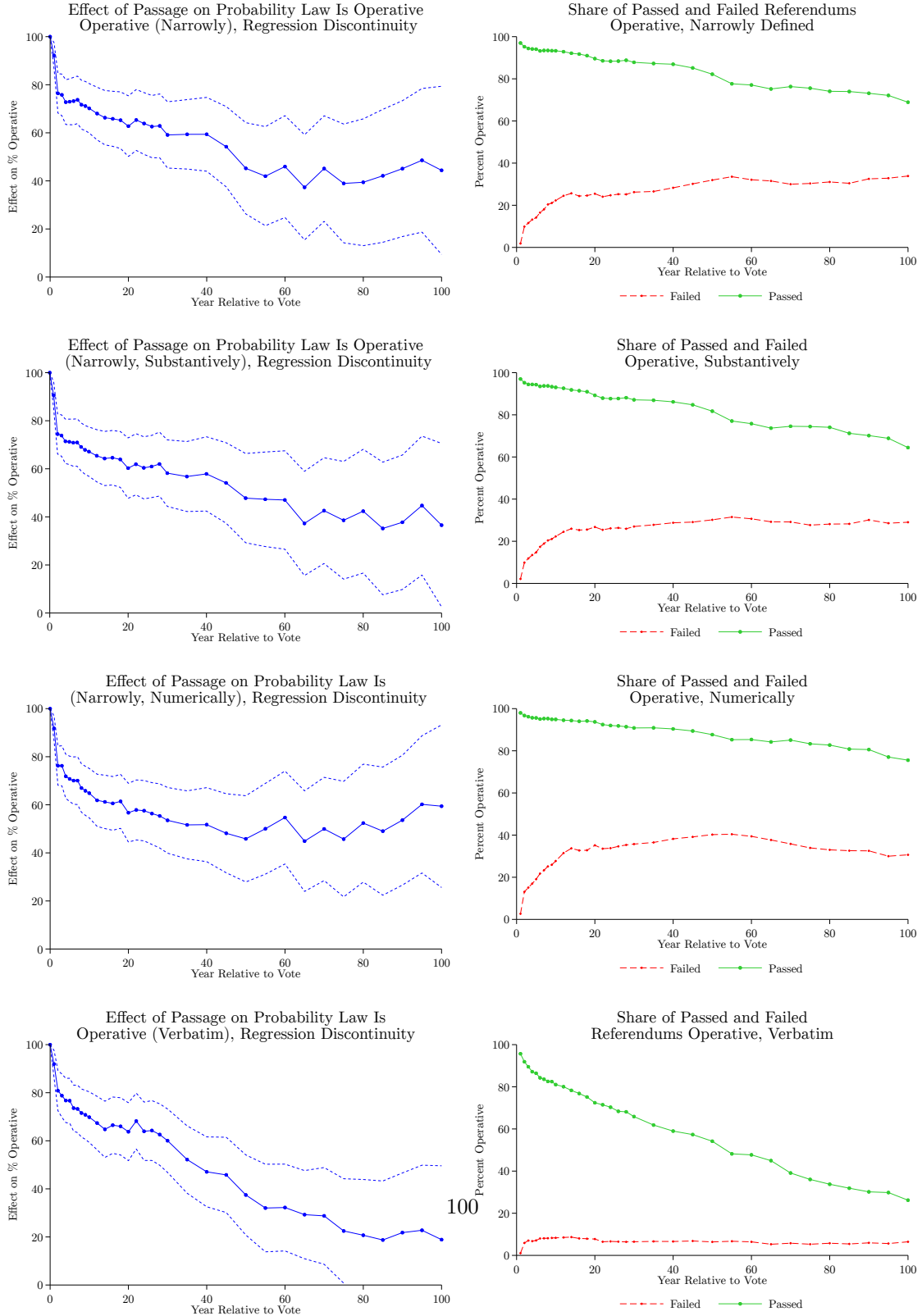
Appendix Figure E6: Persistence Hazard Rate



*Notes:* The graphs show the decay of persistence over time. The left panel shows the five-year average hazard rate of persistence estimated using equation (2) off of the regression discontinuity estimates, and the lower-right shows the five-year average hazard rate, again via equation (2), but for a simple-difference specification, i.e., without the vote share terms in equation (1). The dashed red lines in the bottom two panels indicate the average rate of decay over the 60-year period.

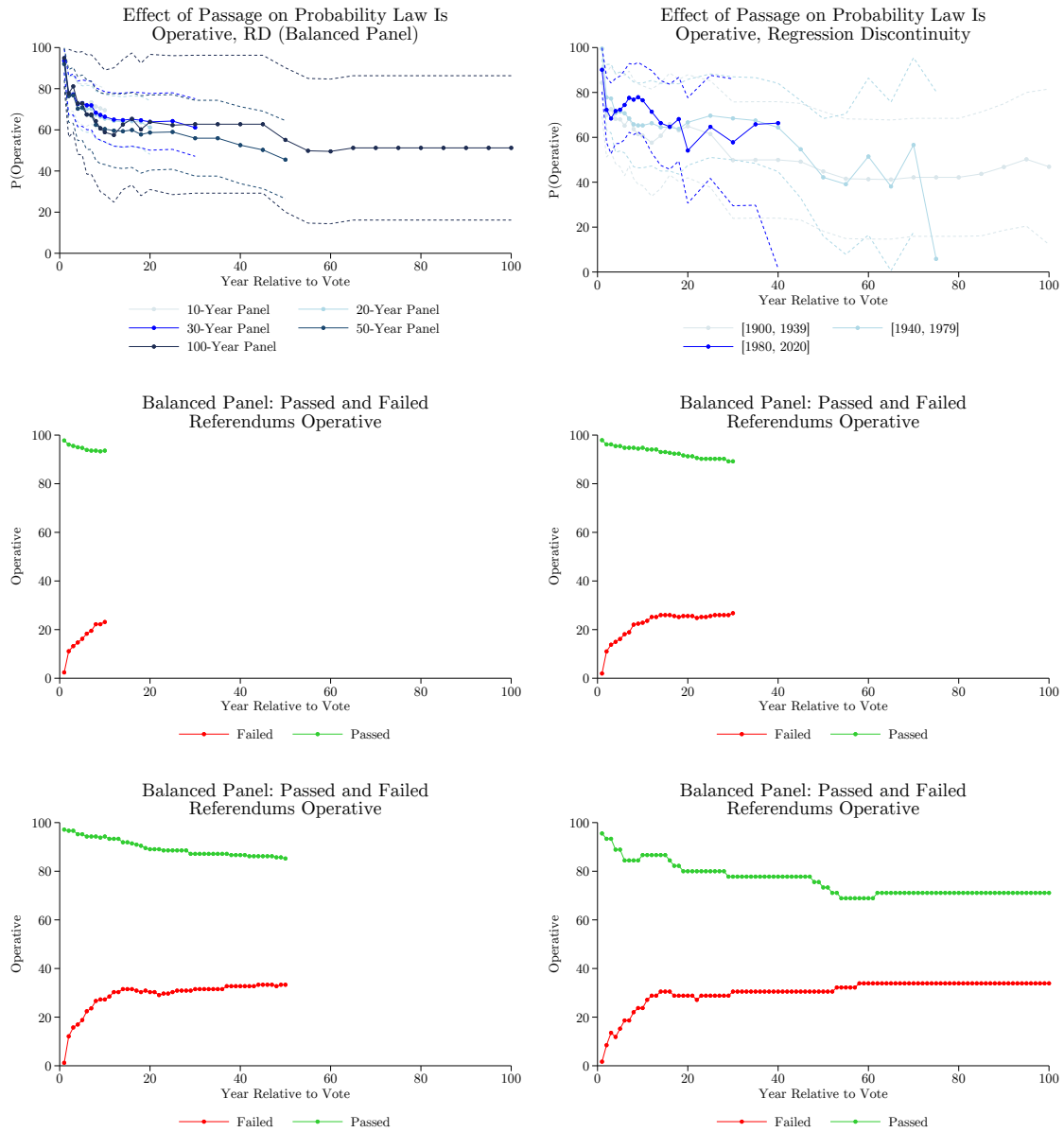


Appendix Figure E7: Robustness of Average Persistence to Outcome Definitions



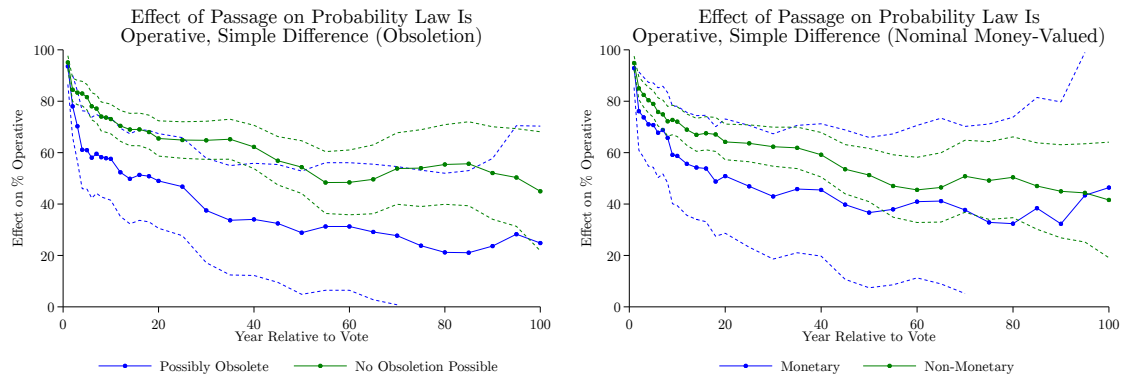
Notes: Each row shows an alternative definition of whether a law is operative. The panels on the left show persistence, and panels on the right show the shares of laws that are in place over time. The top row ignores any areas of state law and constitutions, executive orders, news reports or federal law not directly addressed by the proposed referendum. The second row does the same and counts a law as operative even if its legal structure is weaker (hence a somewhat lower bar for being considered operative). The third row does the same and counts a law as operative even if it is substantially or legally weaker (hence an even lower bar for being considered operative). The bottom row considers the strictest definition of

Appendix Figure E8: Persistence in Balanced Panel and Over Time



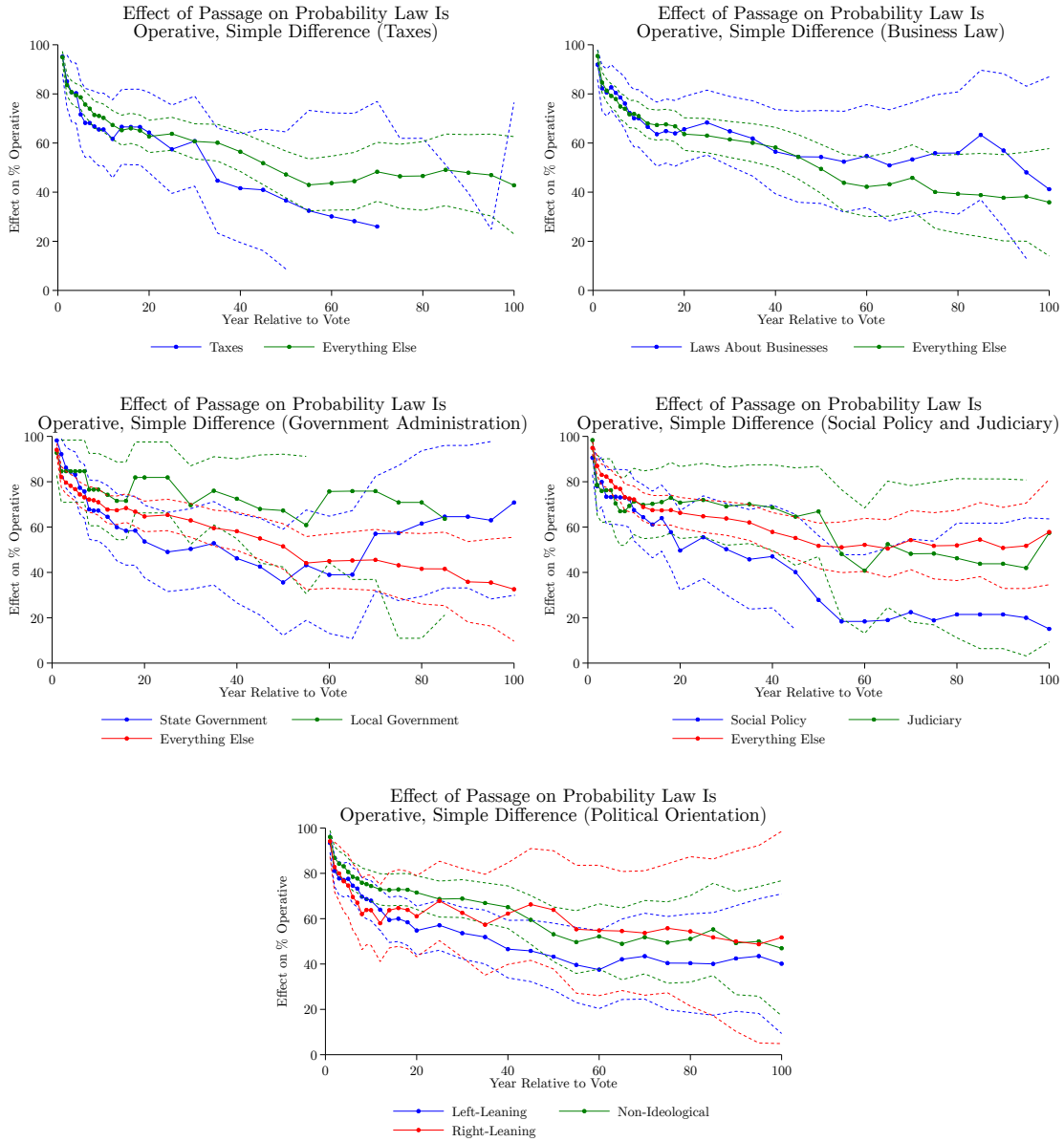
Notes: Top left panel shows the effect of passage on whether a proposed law is operative from estimating equation (1) on balanced panels consisting of all observations with a sufficient number of years of outcome data. Top right panel shows the effect of passage (in an unbalanced panel) separately for referendums passed in three different 40-year periods. The lower two rows show the shares of passed and failed referendums that are operative over time for the same balanced panels in the top-left.

Appendix Figure E9: Persistence Is at Least as High after Dropping Obsolete Referendums



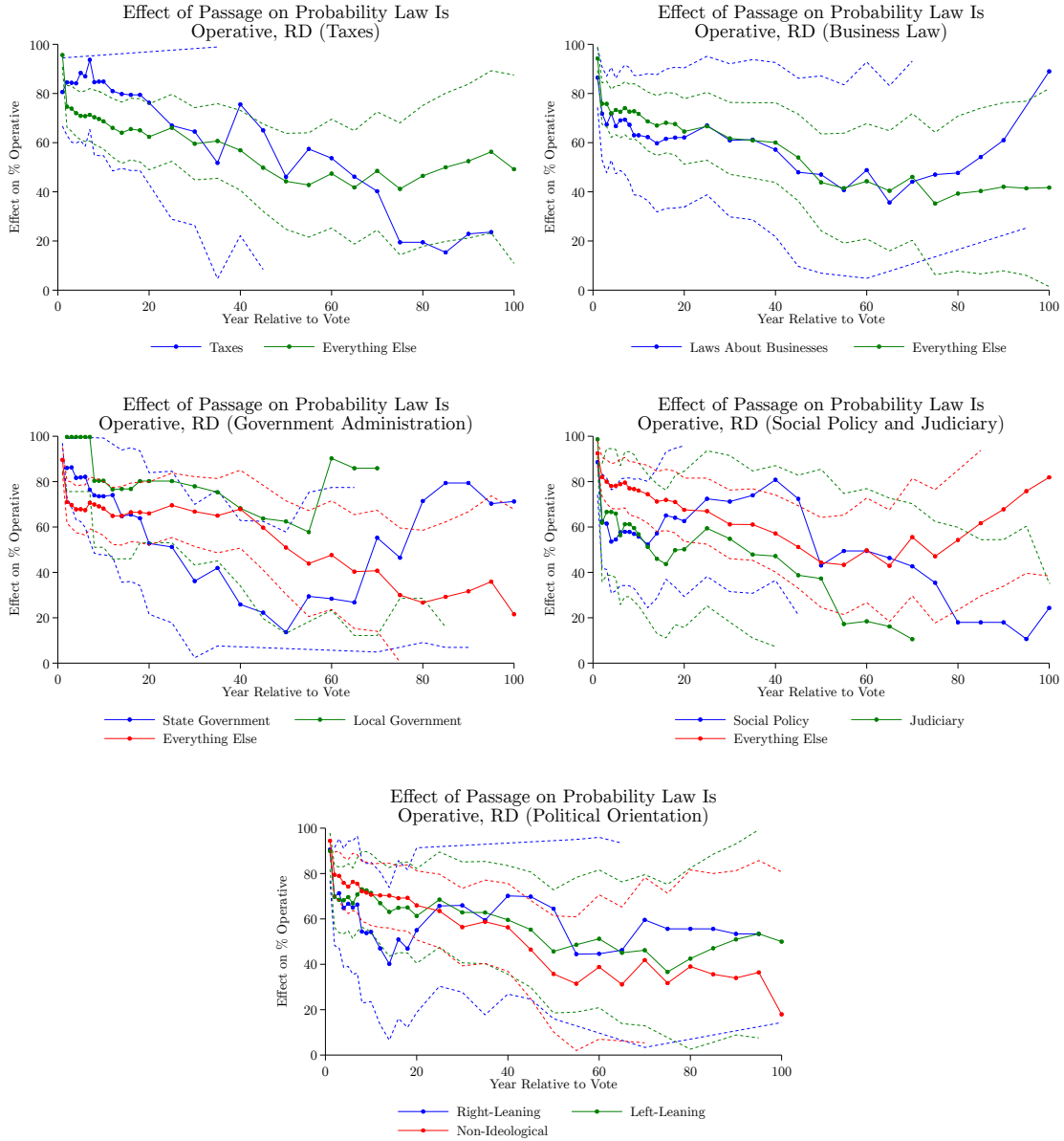
Notes: Each panel shows heterogeneity of persistence by a measure of whether a proposed policy would likely or inevitably go obsolete. The left panel shows persistence for policies that we subjectively determine would likely go obsolete. The right panel shows persistence for policies that involve a change to a dollar value in nominal terms.

Appendix Figure E10: Heterogeneity by Topic and Political Orientation, CIs



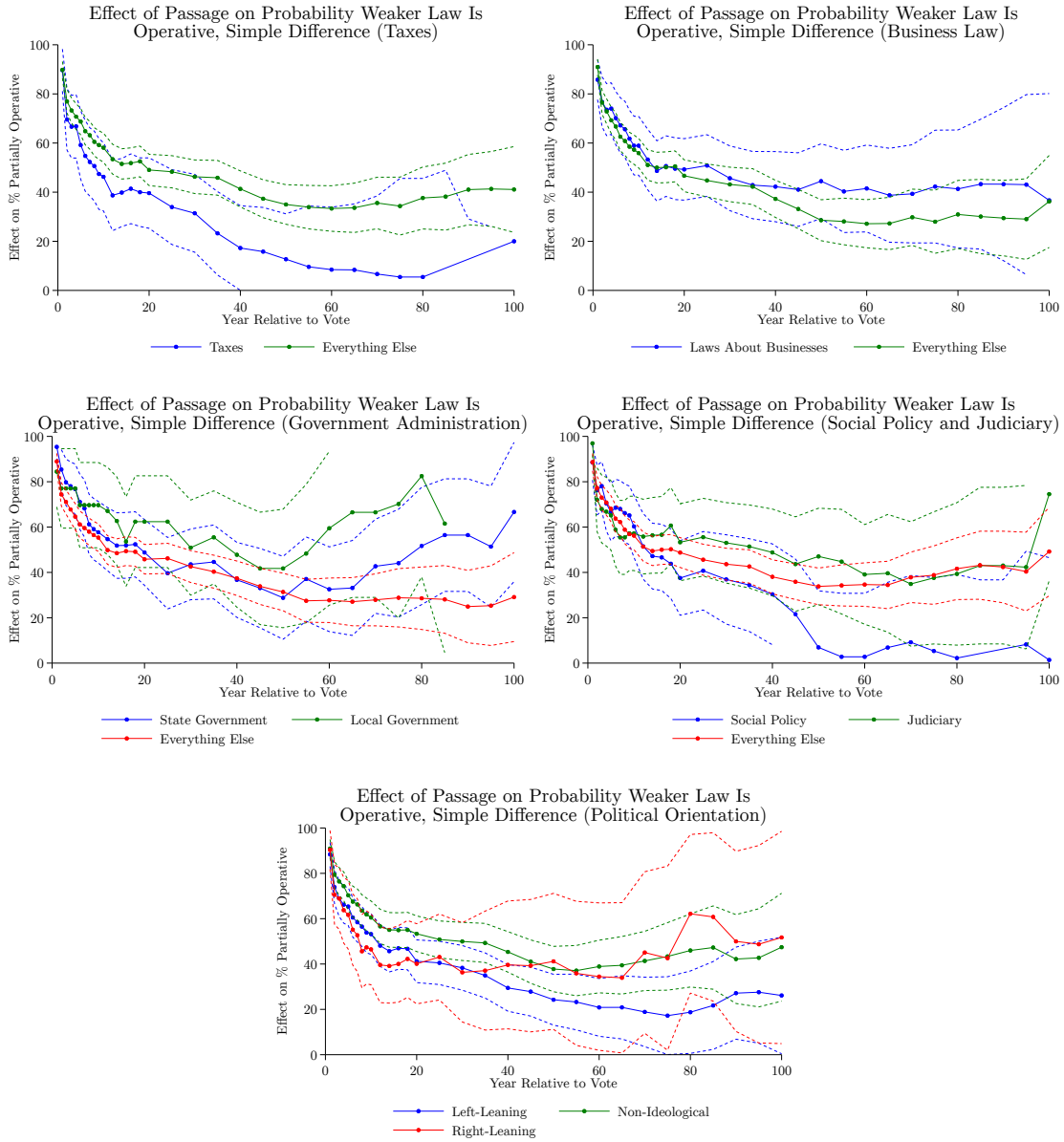
Notes: Each curve plots the persistence of close referendums by policy topic estimated by equation (3). The plots display the same coefficients as in Figure 6 but with confidence intervals.

Appendix Figure E11: Heterogeneity by Topic and Political Orientation, Regression Discontinuity



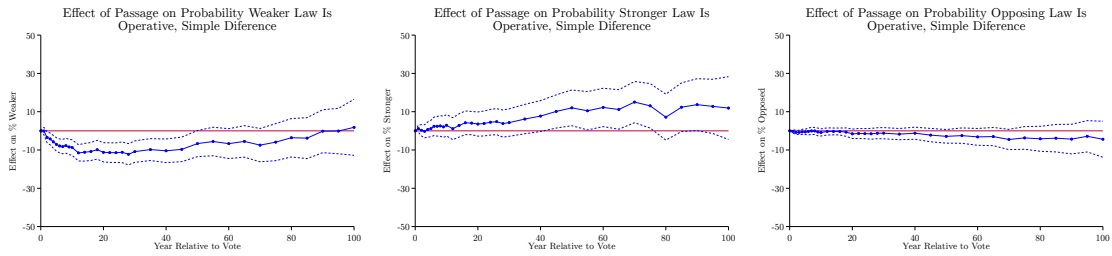
Notes: Each curve plots the persistence of close referendums by policy topic estimated by a modification of equation (3) to include the running variable terms from equation (1).

Appendix Figure E12: Heterogeneity by Topic and Political Orientation, Including Weaker Version



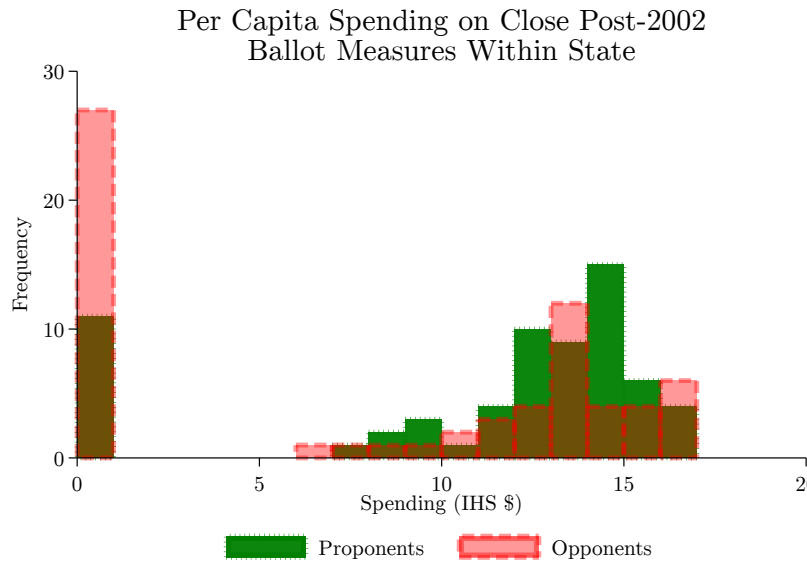
Notes: Each curve plots the persistence of close referendums by policy topic estimated by equation (3) but where the outcome is equal to one so long as at least a weaker version of the policy is in place.

Appendix Figure E13: Intensive Margin Effects, Simple Difference



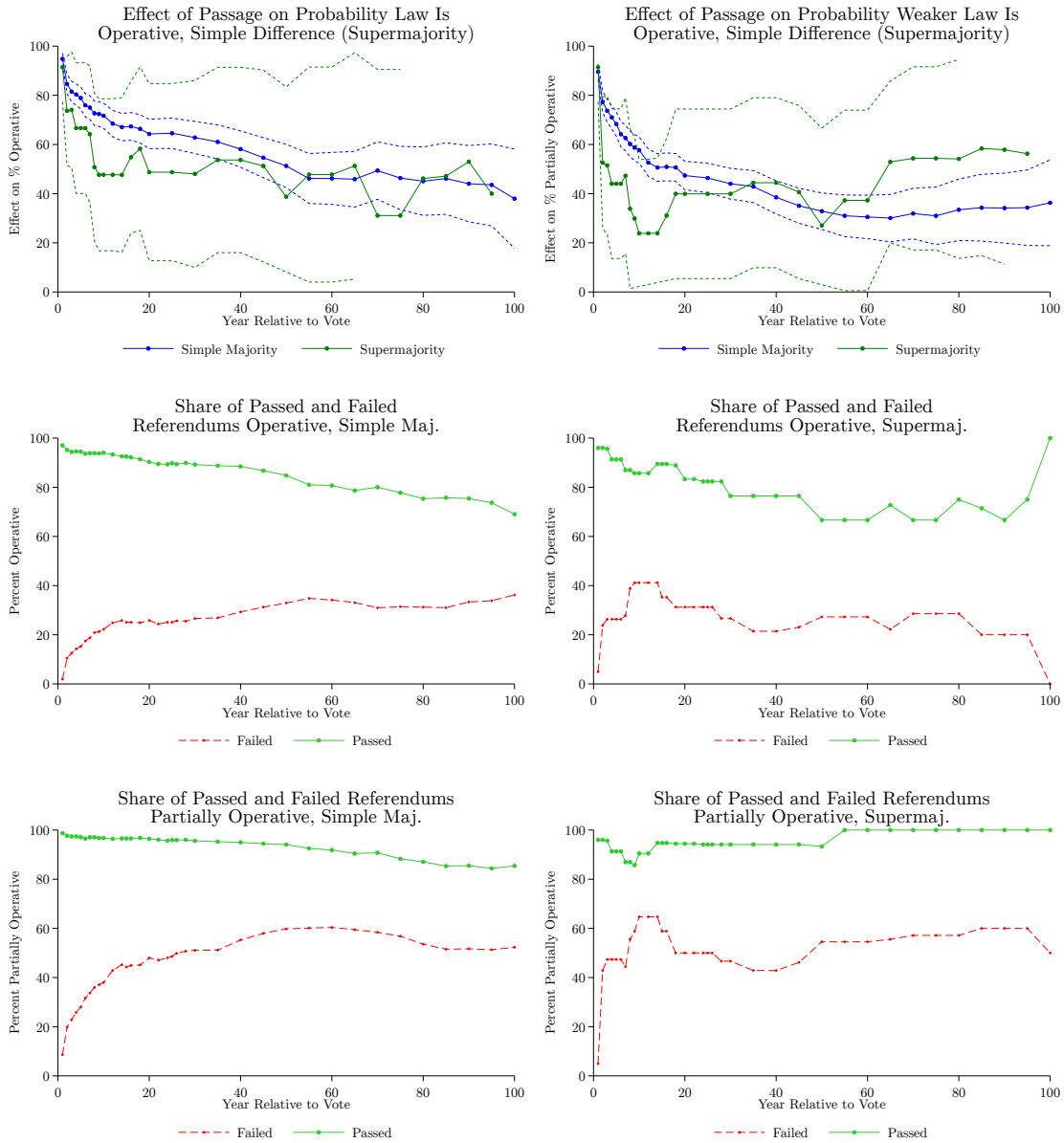
*Notes:* The left panel in each row shows the coefficient on passage from a simple difference regression, or equation (1) without the vote share terms. The outcomes are indicators, from top to bottom, for whether any, possibly weaker, version of a law is in place, whether a stronger version is in place, and whether an opposing law is in place as described in Section 2.2. These correspond to the RD estimates in Figure 5

Appendix Figure E14: Campaign Spending by Referendum Proponents and Opponents



*Notes:* The histogram displays the distribution of spending for and against all referendums in the sample since 2002 using data from Open Secrets, a nonprofit that documents U.S. campaign contributions.

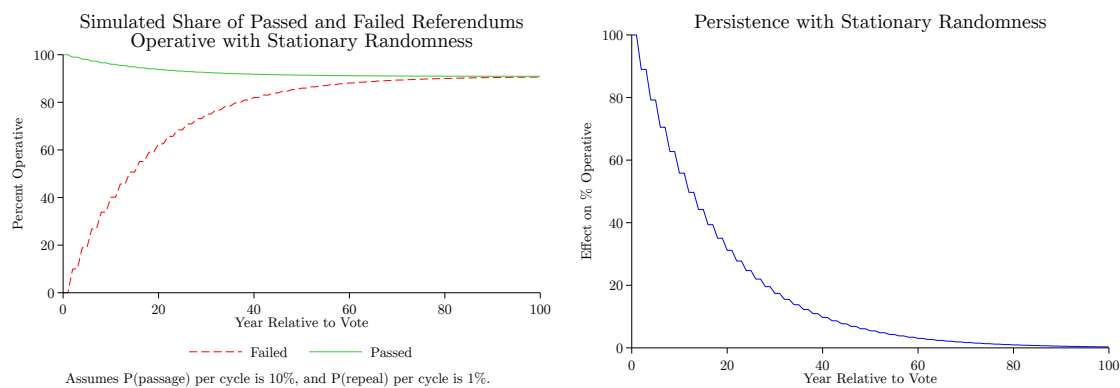
Appendix Figure E15: Similar Persistence with Supermajority Thresholds



*Notes:* The top two panels show persistence, for full and at-least weaker versions of a policy, by whether a referendum was subject to a supermajority requirement, estimated using equation (3). The middle row shows the share of passed and failed referendums operative over time with and without having initially been subject to a supermajority requirement, and the bottom row shows the same for the effect on whether an at-least weaker version of a policy is operative.

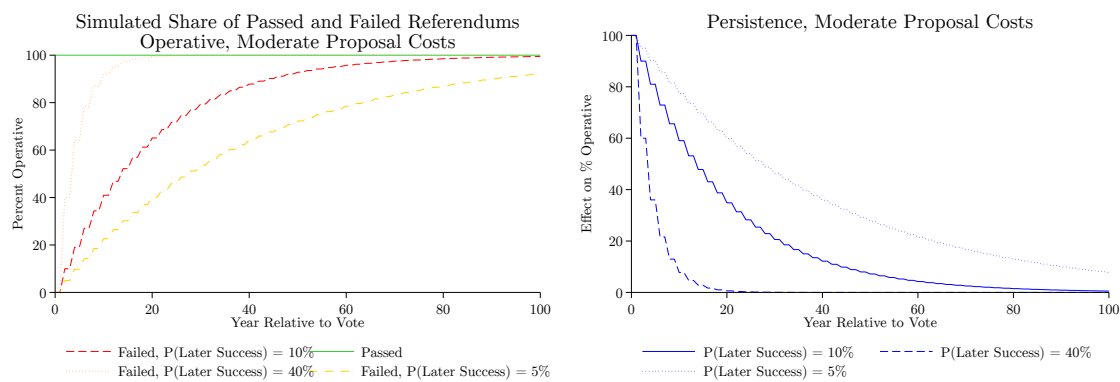


Appendix Figure E16: Simulated Policy Histories—Constant Hazard Rate with Stationary Stochastic Pay-offs



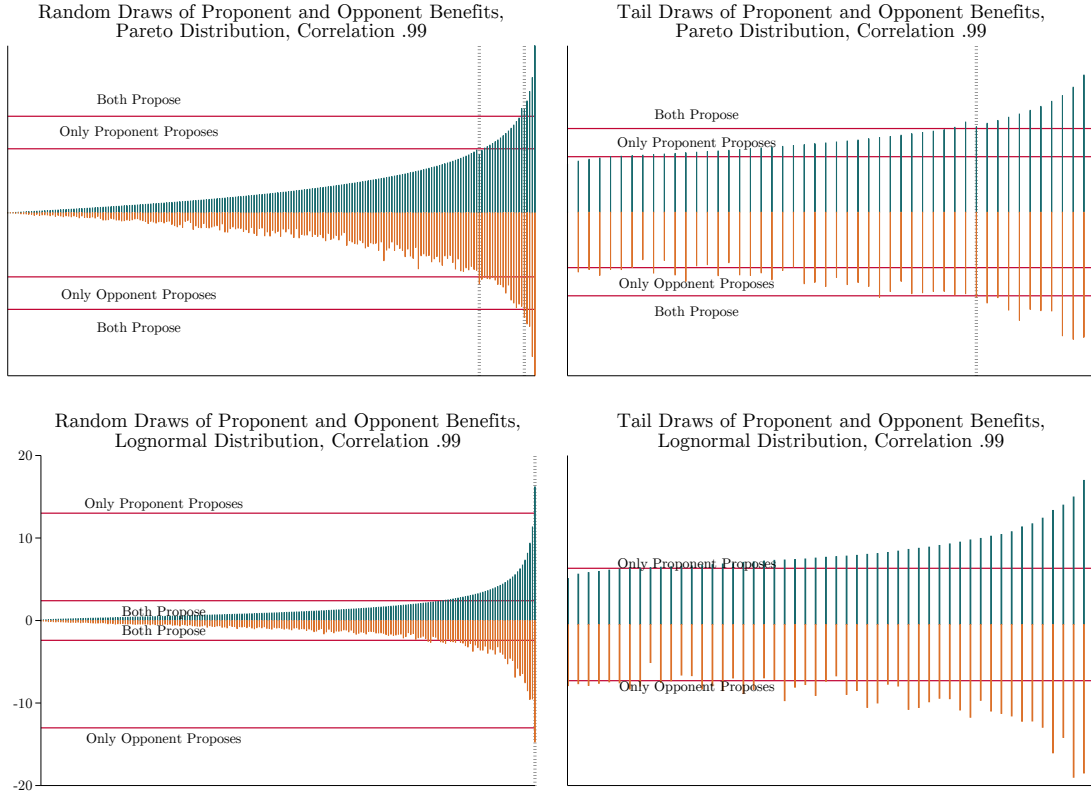
Notes: Left panel shows the likelihood of a policy being in place over time based on whether it passes or fails in an initial period if proponents' and opponents' willingness to propose varies over time according to a stationary distribution. Right panel shows persistence as the difference between the two lines on the left.

Appendix Figure E17: Simulated Policy Histories—Constant Hazard Rate with Stationary Stochastic Pay-offs



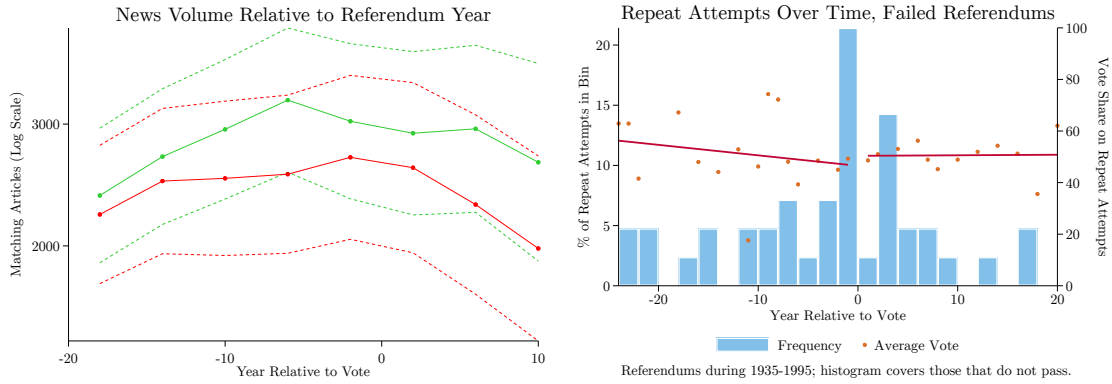
Notes: Each panel depicts the same quantities as in Figure 7 but for three different successful repeat attempt frequencies.

Appendix Figure E18: Simulating Repeals—Selection Implies Few Repeals Under Most Plausible Assumptions



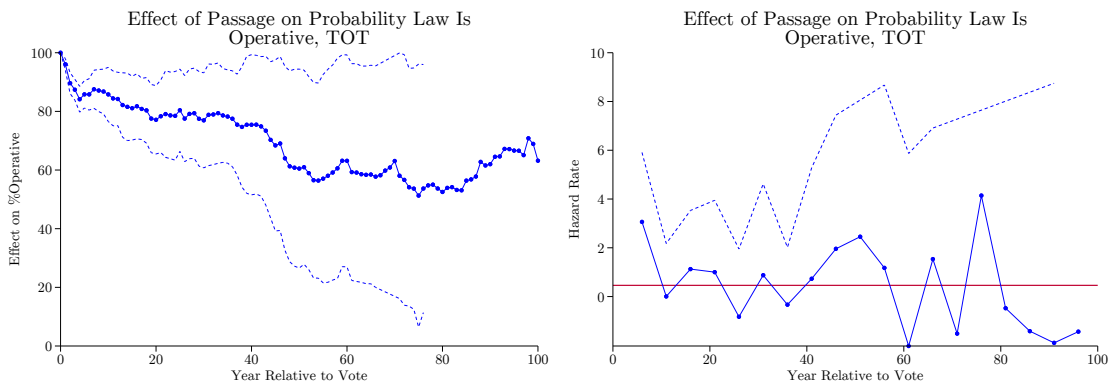
*Notes:* The top two panels show random draws of proponent and opponent willingness to pay for a policy change ( $b_{00}$  and  $b_{10}$  in the framework of Section 4) under alternative assumptions. The draws are ordered by whether there is an equilibrium where both propose, where only the opponent proposes, where only the proponent proposes, and finally by the proponent’s willingness to pay. The dotted horizontal lines indicate the levels above and below which there is an equilibrium where one side proposes but not the other and the levels above and below which there is an equilibrium where both sides propose. The latter does not appear for the lognormal plots because it is out of the range of the plots. The vertical line in each Pareto plot demarcates the draws where there is an equilibrium where both propose. I assume in each that the discount rate is 0.95 and that there is a 50% chance of a referendum going either direction.

Appendix Figure E19: Issue Salience Rises and Falls Over Time



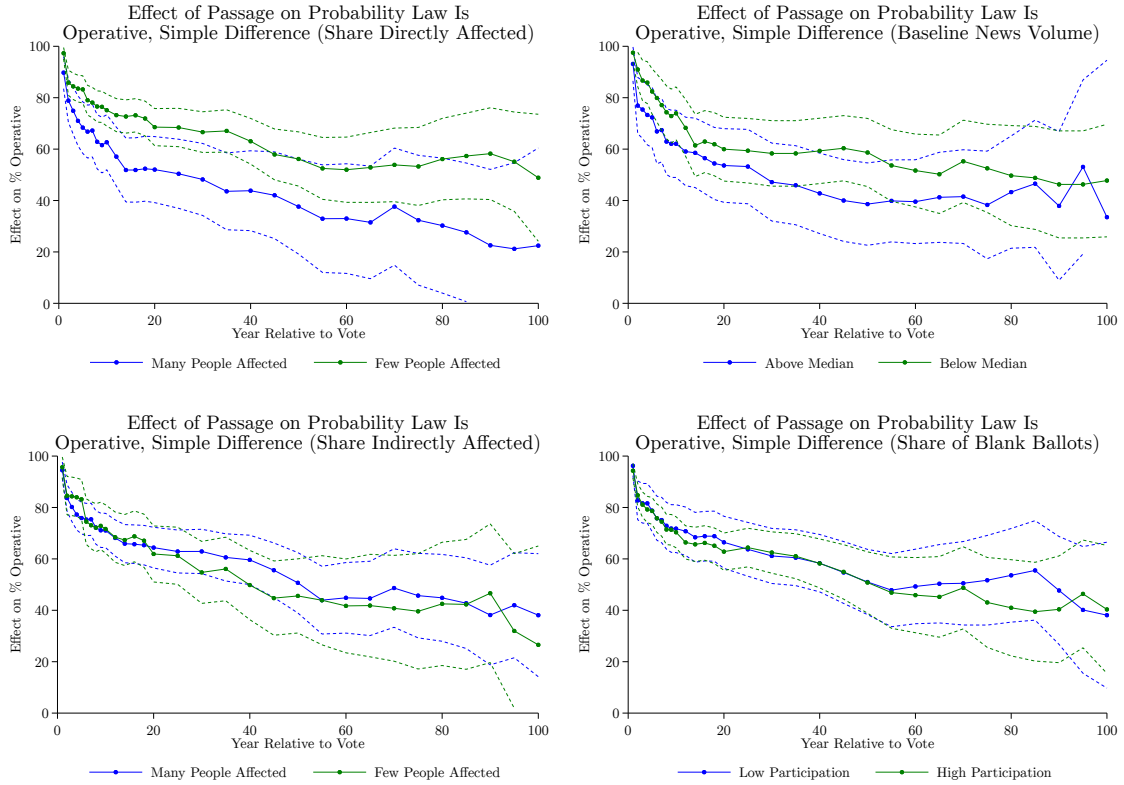
Notes: Left panel displays the average log number of news articles in the Library of Congress’s “Chronicling America” historical news database that match selected keywords for four-year bins before and after the referendum year, as described in Section 2.2.4 and shown across all measures in Figure 9. Right panel plots the distribution of attempts to pass a referendum over time. The horizontal axis is the year relative to the year of the referendum in our sample, and the height of the bars indicates the share of all repeat attempts for referendums in our sample that happen in each two-year bin against the left-vertical axis. Red dots indicate the average vote share in each relative year according to the right-vertical axis. The histogram only includes referendums that do not pass within 25 years.

Appendix Figure E20: Persistence of Successful Campaigns—Successfully Proposing a Policy Has a Large Effect Decades Later



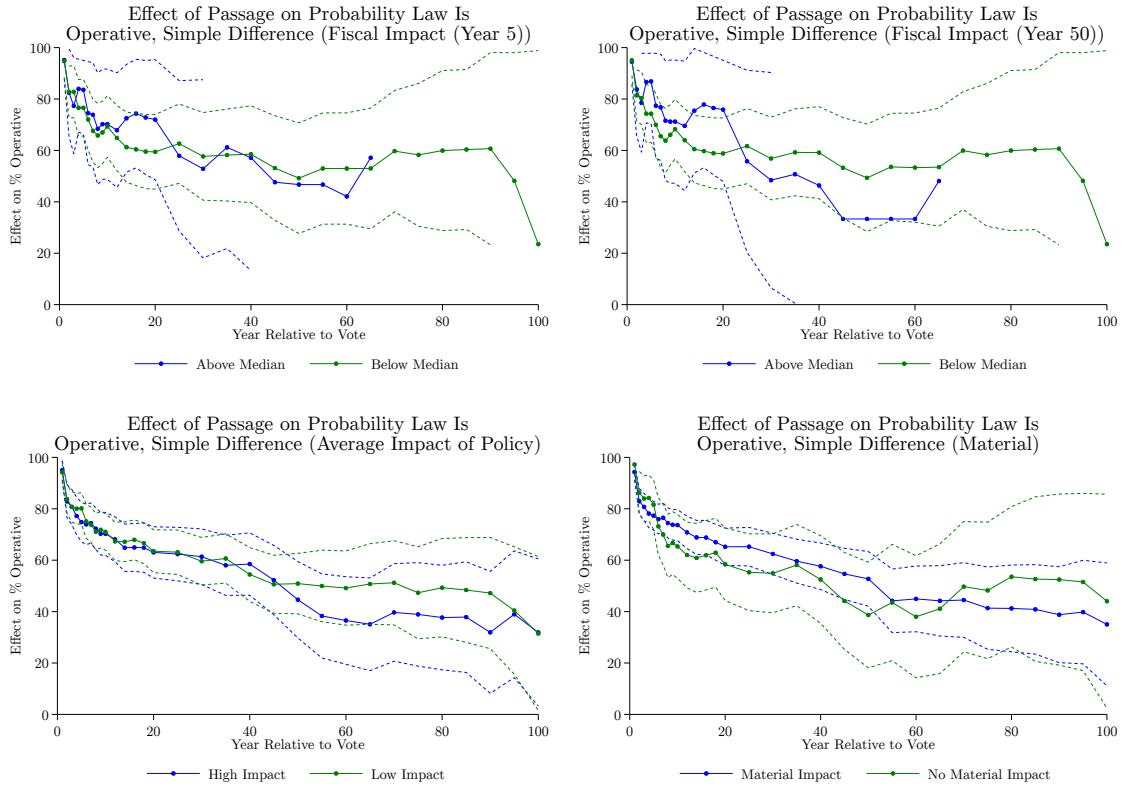
Notes: Left panel shows the estimated persistence of successful policy campaigns, using the recursive “treatment on the treated” estimate from equation 7 as described in Appendix Section A.3.1. Right panel shows the estimated hazard rate, computed by estimated equation 2 but with  $\beta_{1t} = TOT_t$ .

Appendix Figure E21: Heterogeneity by Measures of Salience



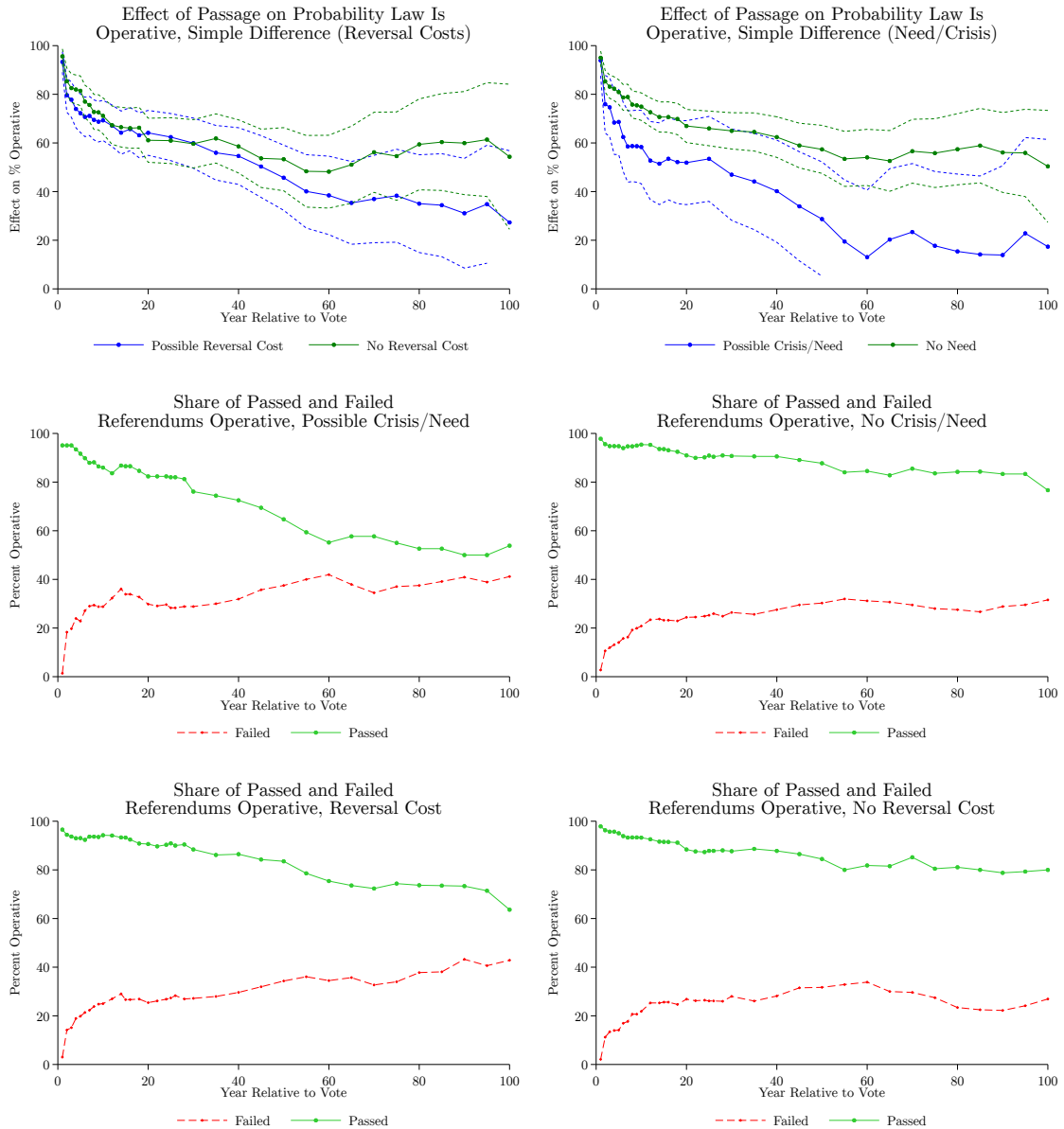
*Notes:* The top-left panel shows Figure 9 with 95% confidence intervals. The top-right panel shows persistence for referendums based on prior news interest, measured as the difference between the log of the total number of related articles in the twenty years before a referendum minus the log of the total number of related articles in the four years starting in the year of a referendum. The bottom-left panel shows the persistence of referendums based on the total number of people indirectly affected as described in Section 2.2.4, and the bottom-right panel shows persistence based on the share of voters in that election who cast votes on the referendum.

Appendix Figure E22: Heterogeneity by Measures of Importance



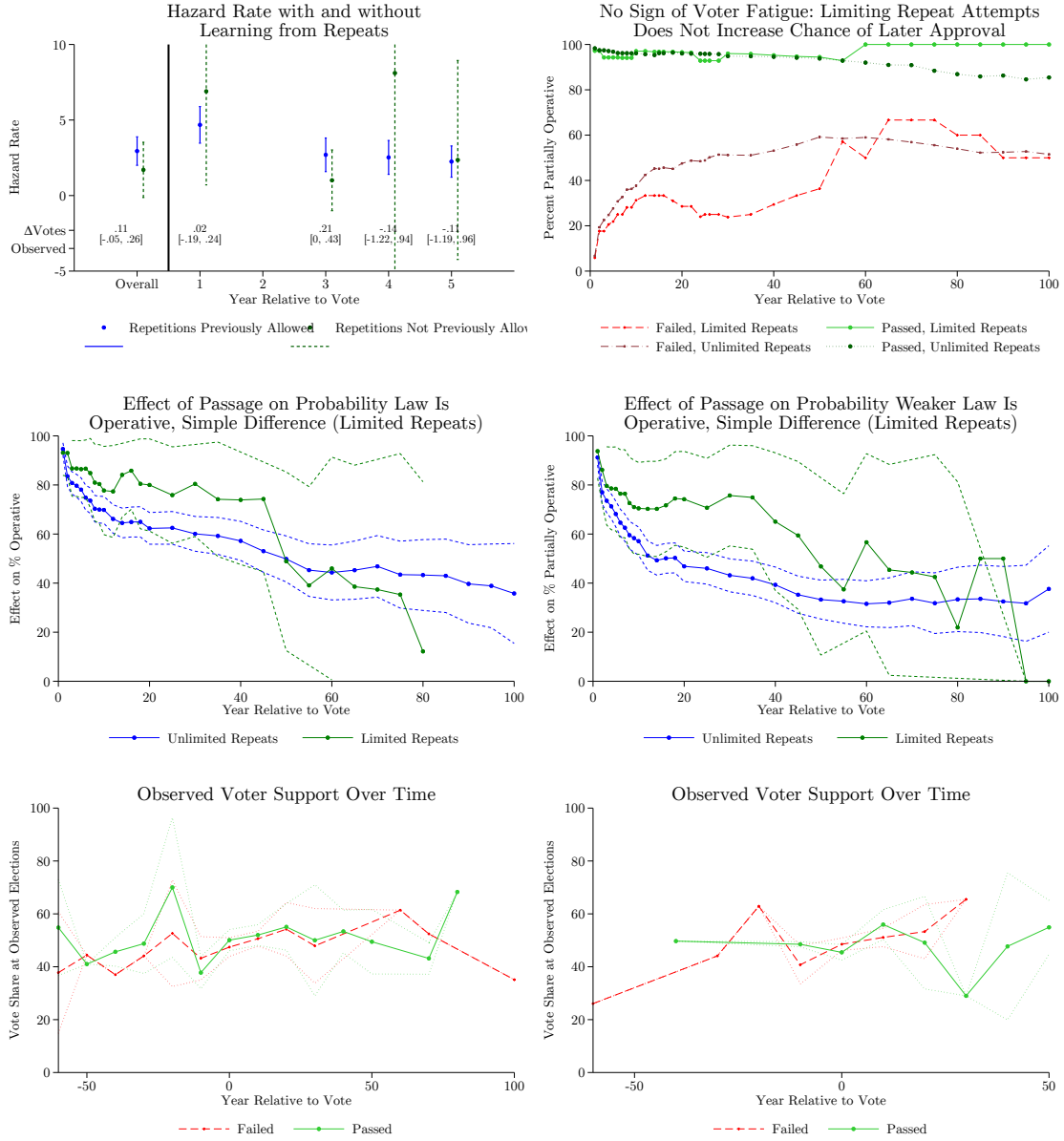
*Notes:* The top row shows heterogeneity in persistence by states' estimates of policies' fiscal impacts in five and fifty years, respectively. The bottom-left panel shows heterogeneity based on a subjective indicator for the scale of a policy's impact as described in Section 2.2.4. The bottom-right panel shows heterogeneity with regard to an indicator for whether a policy has a material impact (as opposed to a procedural or symbolic impact).

Appendix Figure E23: Persistence by Perceived Crises and Reversal Costs



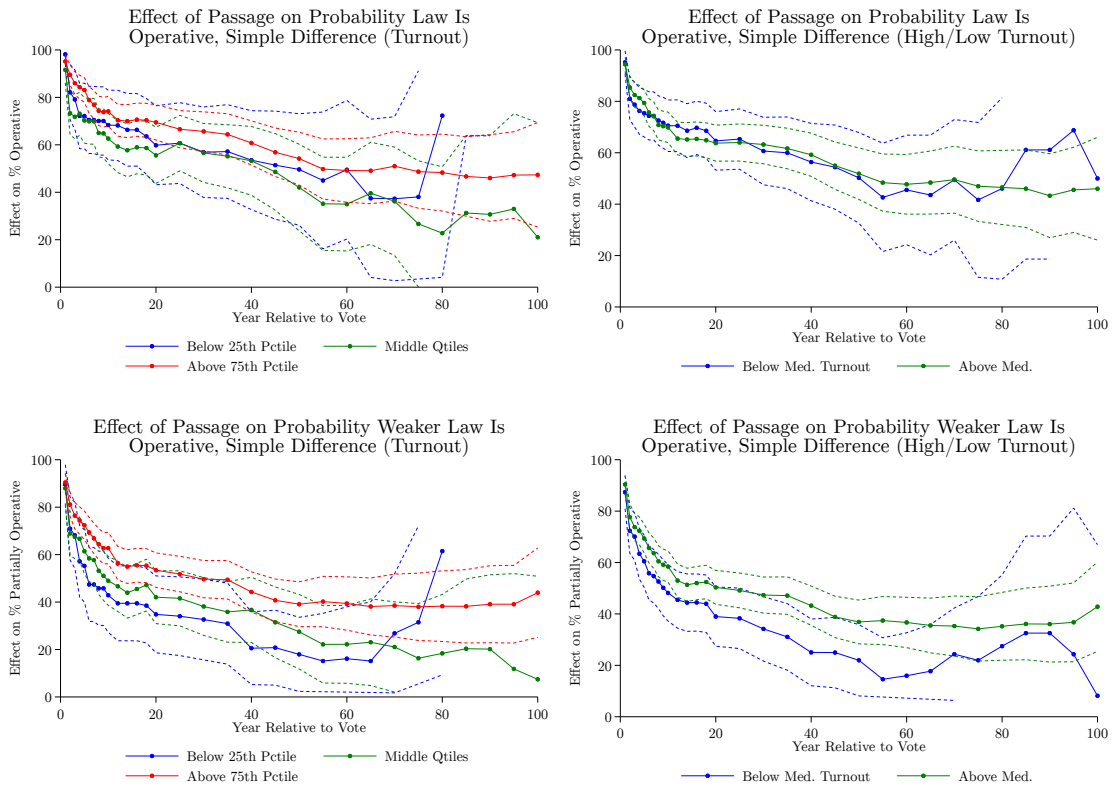
*Notes:* The top row shows heterogeneity in persistence by whether a policy plausibly would be costly to reverse and whether it is in response to a perceived need or crisis. The next two rows show the share of passed and failed policies that are in place for each type, where the left is the relevant subgroup (responses to crises or costly costly to reverse), and the right is everything else.

Appendix Figure E24: Reactions to Prior Votes



Notes: The top-left panel shows the five-year average hazard rate of persistence beginning in the year labeled on the  $x$ -axis for referendums in states that allowed or did not allow prior repeat attempts; the left of the black line indicates the variance-weighted average of the hazard rates across all years. The numbers below each confidence interval indicate the difference between the number of repeat attempts in the states without a limit compared to the states with a limit. The limit is imperfect, and sometimes a nonidentical referendum is essentially a repeat attempt. The top-right panel shows an analogous figure to the left panel of Figure 8 but for whether a partial version of the proposed policy is operative. The middle row shows estimates of persistence for states with and without limits on prior repeat attempts. The bottom panels show observed vote share over time as in Figure 8 with confidence intervals. The bottom-right panel restricts the vote shares to only consider attempts at repeal or revisiting a referendum without any modification.

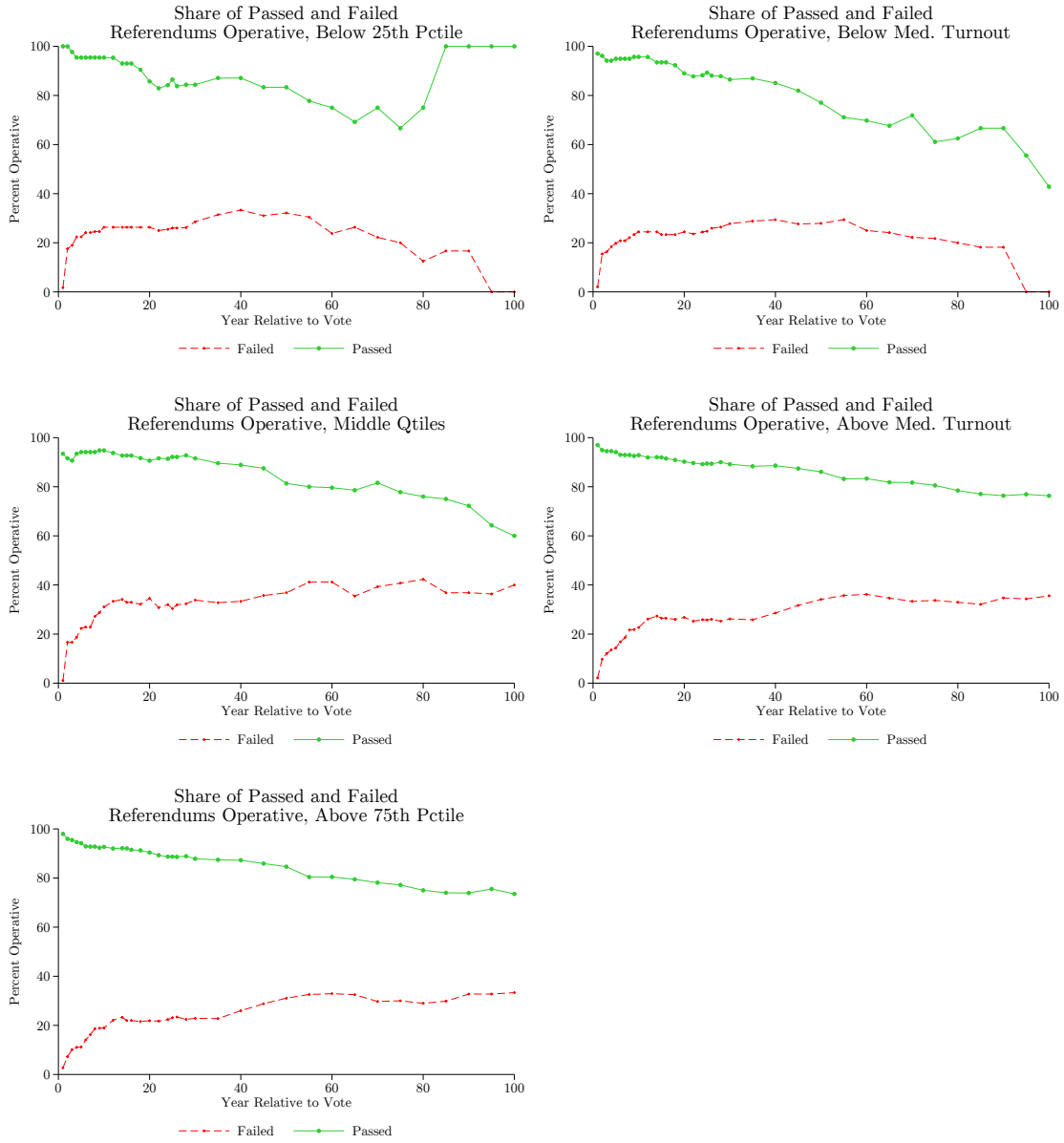
Appendix Figure E25: Heterogeneity of Persistence by Turnout



*Notes:* The panels display heterogeneity by turnout, estimated using equation (3). The top panels display heterogeneity using my main measure of persistence, i.e., the effect of passage on whether the proposed policy is operative. The bottom panels display heterogeneity of the effect on whether a weaker version of the policy is in place.

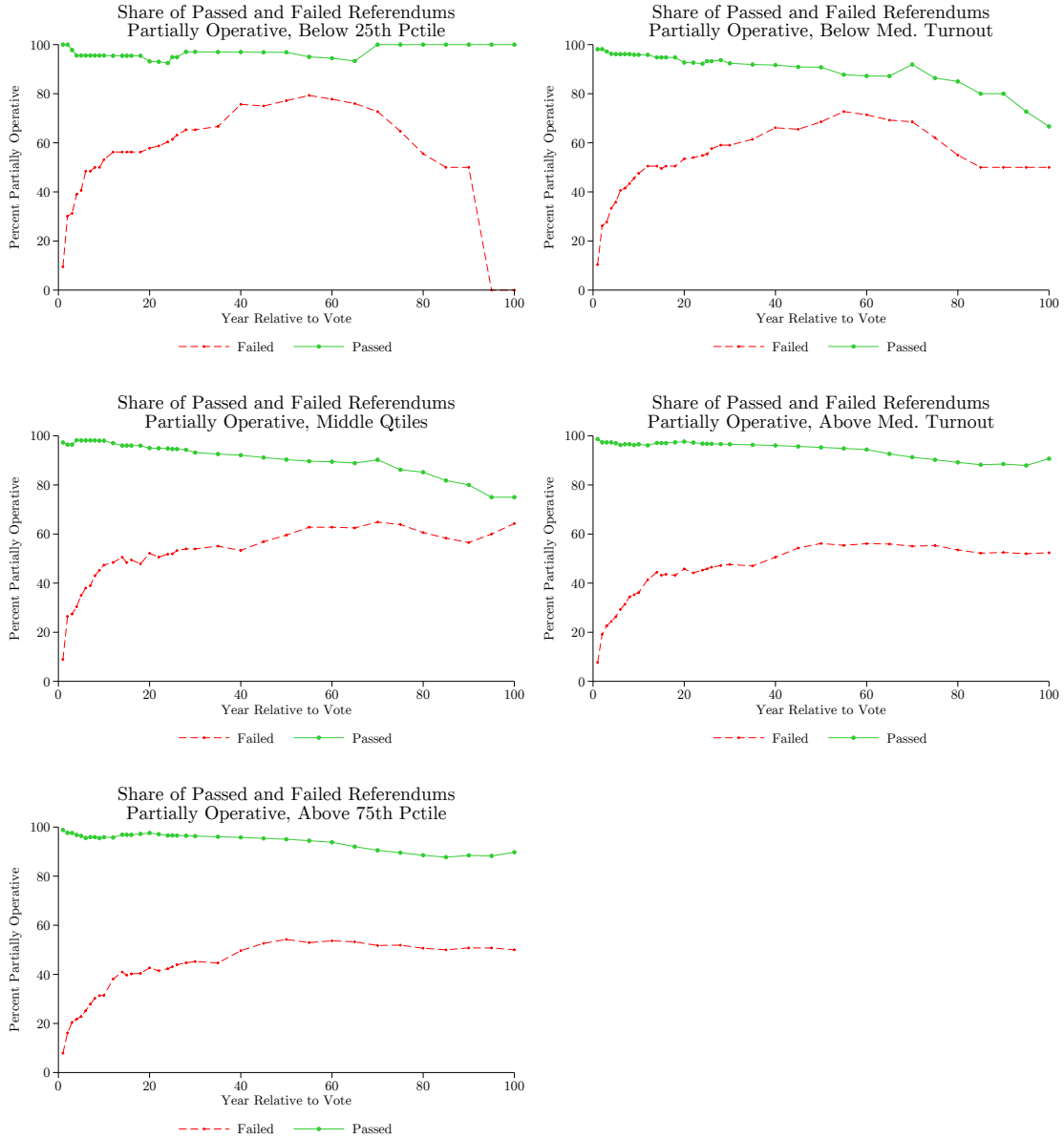


Appendix Figure E26: Statutory Histories of Close Referendums by Turnout



*Notes:* The panels display the share of passed and failed referendums whose policies are in place over time by turnout at the initial election. These are the same outcomes used to estimate E25

Appendix Figure E27: Statutory Histories (Weaker Policies) of Close Referendums by Turnout



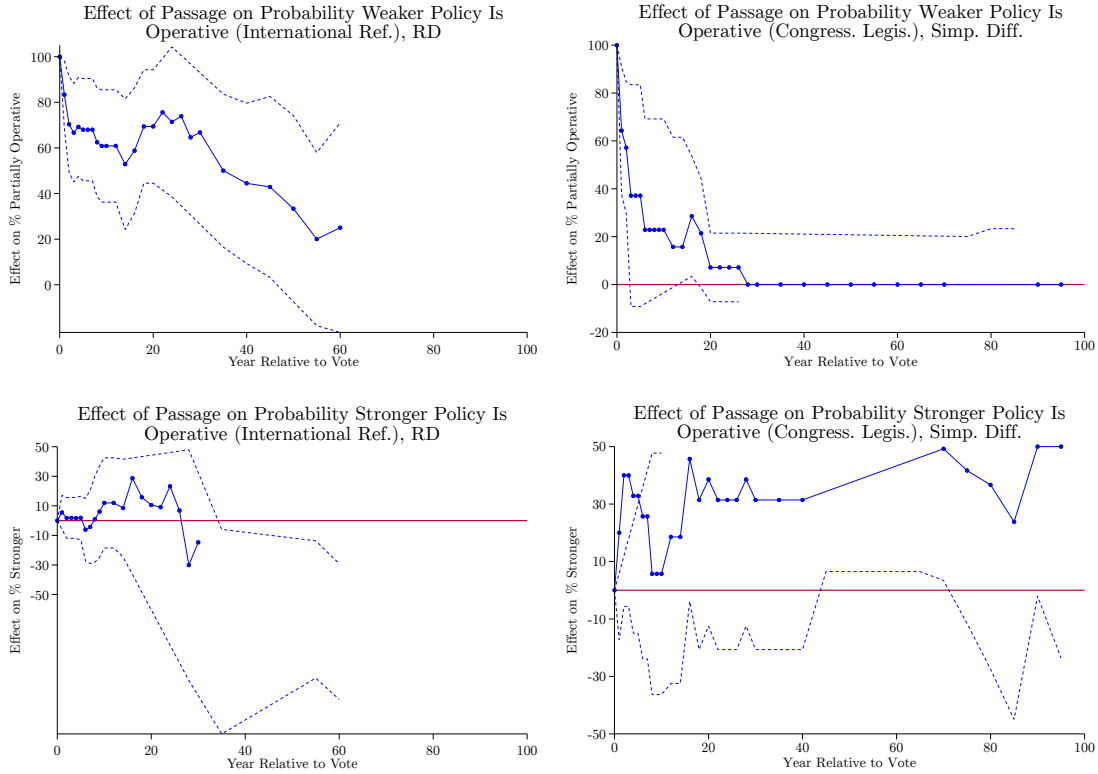
*Notes:* The panels display the share of passed and failed referendums whose policies are in place over time by turnout at the initial election, with the  $y$  variable the share of proposed policies that are in place, at least in weaker form. These are the same outcomes used to estimate E25

Appendix Figure E28: Persistence and the Costs of Policy Proposals



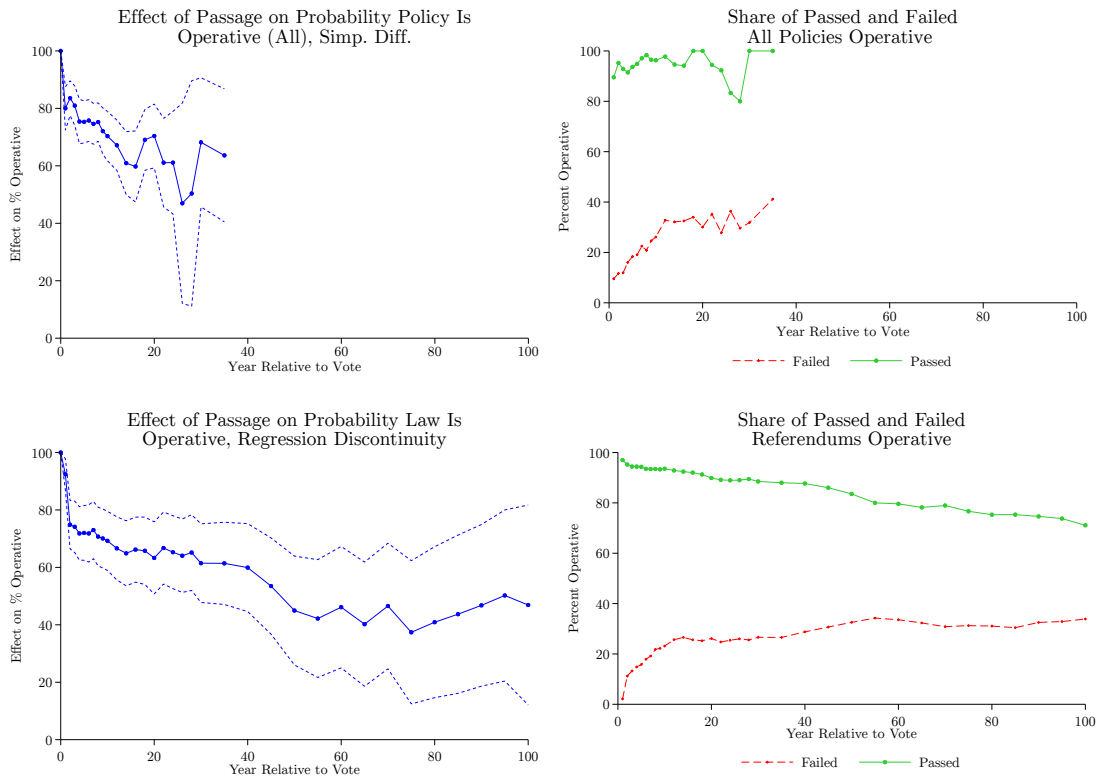
*Notes:* The top-left panel shows persistence of initiatives with high and low signature requirements, estimated using equation (3). The top-right panel shows the number of initiatives in a given state over that state's required number of signatures. The bottom row shows the share of passed and failed referendums that are operative over time by the baseline signature requirement.

Appendix Figure E29: Intensive Margin Effects for Legislation and International Referendums



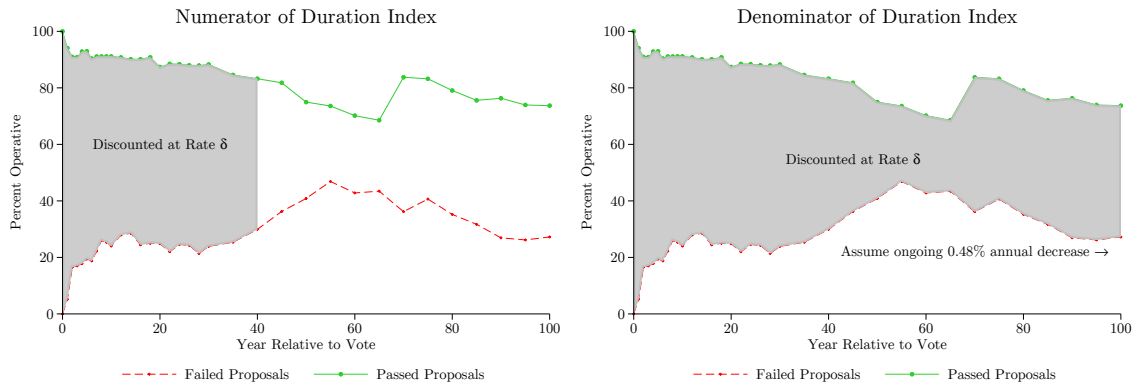
*Notes:* The left panels show the coefficient on passage on whether an at-least-partial version of a policy is in place (top) and whether a stronger version is in place (bottom) for non-U.S. national-level referendums, estimated as a simple difference. The right panels show the coefficient on passage for the same outcomes compared to comparison states matched according to their Mahalanobis distance, estimated using equation (3) with data from Caughey and Warshaw (2016) and Grossmann et al. (2021). The third panel shows the coefficient on passage on whether proposed congressional legislation that narrowly passed or failed is in place, selected using a narrative approach similar to Romer and Romer (1989, 1994) and estimated as a simple difference.

Appendix Figure E30: Mahalanobis and Regression Discontinuity Methods Yield Similar Results



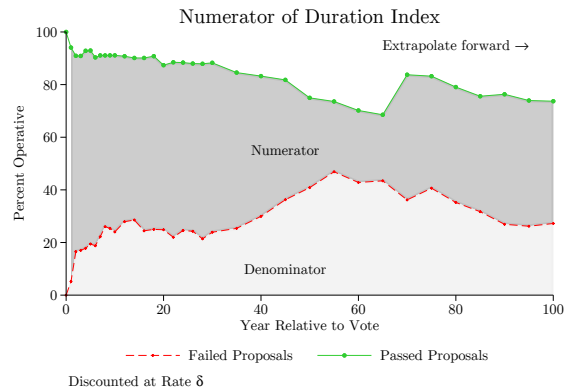
Notes: The graph estimates the persistence of state referendums using Mahalanobis matching and outcomes from Caughey and Warshaw (2016), replicating the result from the regression discontinuity design presented in Figure 4, repeated in the lower panels.

Appendix Figure E31: Construction of Duration Index



Notes: The policy duration index divides the area in the left panel by the area in the right panel, extended based on the assumed ongoing persistence hazard rate.

Appendix Figure E32: Calculating the Threshold for Marginally Imperfect Reforms



Notes: The value of supporting an imperfect policy when a voter expects the later version to be marginally better is the ratio of the area between the two curves to the area under the lower curve.