Dynamic Representation in the American States, 1960–2012

Devin Caughey*  
Department of Political Science  
MIT

Christopher Warshaw†  
Department of Political Science  
MIT

First draft: July 28, 2014  
This draft: September 27, 2015

Prepared for the Washington University in St. Louis  
American Politics Workshop (September 2015)

Abstract
One of the most fundamental assumptions of democratic theory is that the views of citizens should influence government policy decisions. Previous studies have found a strong cross-sectional relationship between public opinion and state policy outputs. But the ultimate test of responsiveness is the extent to which changes in popular preferences cause changes in public policies. In this paper, we reassess the quality of representation in the American states over the past half century using a large battery of historical evidence and new statistical techniques. We show that changes in the mass public’s policy views are associated with changes in state policy outputs. In addition, we evaluate the influence of institutions, such as direct democracy, term limits, and legislative professionalism. We find some preliminary evidence that term limits may increase responsiveness, but legislative professionalism and direct democracy have no consistent impact on responsiveness. Our findings have large implications for both the study of representation and institutions in the American states.

We thank Jeffrey Lax, Seth Hill, participants in the Columbia University American Politics Workshop and panelists at the 2014 American Political Science Association Conference for feedback on a previous version of this manuscript. We appreciate the excellent research assistance of Melissa Meek, Meg Goldberg, Kelly Alexander, Aneesh Anand, Tiffany Chung, Emma Frank, Joseff Kolman, Mathew Peterson, Steve Powell, Charlotte Swasey, Lauren Ullmann, and Amy Wickett. We also appreciate the willingness of Frederick Boehmke and Carl Klarner to generously share their data. We are grateful for research support from the Dean of the School of Humanities, Arts, and Social Sciences at MIT. All mistakes, however, are our own.

*Assistant Professor, Department of Political Science, Massachusetts Institute of Technology, caughey@mit.edu
†Assistant Professor, Department of Political Science, Massachusetts Institute of Technology, cwarshaw@mit.edu
Perhaps the most important hallmark of democracy is “the continuing responsiveness of the government to the preferences of its citizens, considered as political equals” (Dahl 1971, 1). Popular control of government requires, at the very least, that citizens’ preferences influence the laws and policies of the polity. In modern democracies, such influence is primarily mediated through elected representatives, who formulate policy on behalf of the public (Pitkin 1967). Nevertheless, although the representativeness of elected officials is a crucial question, the ultimate metric of responsiveness is the extent to which changes in popular preferences cause changes in public policies (cf. Achen 1978, 479).¹

Studying responsiveness thus requires measures of citizens’ policy preferences as well as of government policies themselves. Moreover, these indicators should ideally be available at multiple points in time so as to measure changes in preferences and policy. Only with time-varying measures is it possible to trace out the dynamic process of responsiveness and identify the policy effects of preferences.²

Although the limitations of cross-sectional data for studying representation have long been recognized (Lowery, Gray, and Hager 1989; Ringquist and Garand 1999; McIver, Erikson, and Wright 2001), scholars have still relied overwhelmingly on such data due to the difficulty of constructing dynamic measures. Even in U.S. state politics, where high-quality data on public opinion and state policies can be obtained over many decades, scholars have hardly taken advantage of variation over time. As a consequence, though there is evidence of strong cross-sectional correlations between public opinion and state policy outputs (Erikson, Wright, and McIver 1993; Gray et al. 2004; Lax and Phillips 2011), the static nature of the data make it difficult to

---

¹ Of course, as Achen (1978) emphasized and Matsusaka (2001) has recently reminded us, responsiveness is by no means the only or even the best criterion for evaluating the quality of representation. Nor, for that matter, is policy the only component of responsiveness; see Eulau and Karps (1977) for a discussion.

² As Haber and Menaldo (2011, 2) note with reference to the effect of national resources on authoritarianism, “when a process is hypothesized to occur over time, it is best to employ evidence and methods designed to see whether that time series process actually occurred.”
determine the causal relationship between opinion and policy.\(^3\)

This paper addresses these challenges by bringing to bear a wealth of historical data and new statistical techniques to examine responsiveness in the American states over the past half century. An important reference point for our work is that of Erikson, MacKuen, and Stimson on “dynamic representation” in U.S. national politics. We too focus on the dynamic relationship between the liberalism of the public and the liberalism of government policies across over half a century. While they consider a single cross-sectional unit, however, we examine the time series of 50 U.S. states, both increasing the effective sample size and enabling us to examine cross-sectional heterogeneity in dynamic representation.

Our opinion estimates for each state and year are derived from the survey responses of over 650,000 individuals to nearly 47 domestic policy questions fielded between 1960 and 2012. To summarize this wealth of opinion data, we estimate a dynamic group-level item-response model, producing yearly estimates of the average policy liberalism of the citizens of each state (Caughey and Warshaw 2015). To generate an analogous summary measure of the liberalism of state policies, we estimate a dynamic latent-variable model using an original yearly dataset of more than 140 state policies (Caughey and Warshaw, Forthcoming).

Based on these two measures, we find that state governments are clearly responsive to shifts in public opinion: when the state public moves to the left, the state adopts more liberal policies. We therefore conclude that dynamic representation is alive and well in U.S. state politics. We also evaluate whether the opinion–policy relationship is moderated by three factors thought to influence representation: term limits, direct democracy, and legislative professionalism (e.g., Matsusaka 2010; Lax and Phillips 2011). We find some evidence that term limits strengthen responsive-

---

\(^3\) Strong correlations between popular preferences and government policy might still be substantively important even if the causal relationship is unclear. At the very least, they provide a normative metric by which to evaluate the quality of democracy.
ness, but no consistent support for interactions with direct democracy or legislative professionalism.

The paper proceeds as follows. We first discuss the background literature and our hypotheses. We then describe the conceptualization and measurement of our two main constructs of interest: Citizen Policy Liberalism and Government Policy Liberalism. The penultimate section evaluates the state-level relationship between these two constructs over the past half-century, with an emphasis on dynamic representation and how it has varied across different conditions. Finally, we briefly conclude.

1 Background

One of the most fundamental assumptions of democratic theory is that the views of citizens should influence government policy decisions. Achen (1978), following Dahl (1957)’s classic definition of power, argues that citizens’ influence over the government can be measured by the expected difference in government officials’ behavior following a change in the average preference in the public—that is, the regression slope, which he labels responsiveness. As Stimson, MacKuen, and Erikson (1995, 543) emphasize, responsiveness is a process that is “inherently structured in time,” and studying it properly requires data that captures change over time.

Electoral democracy can induce responsiveness via two basic mechanisms: selection and anticipation (Stimson, MacKuen, and Erikson 1995; Fearon 1999; Lee, 2011). More precisely, Achen (1978) defines responsiveness as both the intercept and slope of the regression, where the intercept indexes the “bias” of the electoral system (492). Since the intercept in our application has no natural meaning, we focus only on the regression slope, as do most studies of responsiveness. Achen also focuses on the opinions of elected representatives rather than on policy outputs, but there is no difficulty in extending his conception of responsiveness to the latter. As noted by Achen and more recently by Matsusaka (2001), greater responsiveness does not necessarily translate into government outputs that are more proximate to or congruent with public preferences. Notwithstanding the heroic assumptions required to measure spatial proximity between citizens and officials or polices (Lewis and Tausanovitch 2013), studies that have done so have found evidence that excessive responsiveness can result in policies that “overshoot” or “leapfrog” the median voter (Erikson, MacKuen, and Stimson 2002, 373; Bafumi and Herron 2010; see also Lax and Phillips 2011).
Moretti, and Butler 2004). Legislators and other elected officials are more likely to win office when their views are broadly in line with the views of the public. To the extent that this selection mechanism is dominant, responsiveness should occur mainly through the replacement of incumbents with new officials. But responsiveness can also occur without any turnover if reelection-minded elected officials modify their behavior in anticipation of electoral sanction.

Over the past twenty-five years, a large body of work has found public opinion to be highly correlated with global indicators of state policy (Erikson, Wright, and McIver 1993; Gray et al. 2004) as well as with specific policy outputs (Lax and Phillips 2011). However, this cross-sectional literature has obvious limitations for determining the causal relationship between public opinion and public policy outputs (Lowery, Gray, and Hager 1989; Ringquist and Garand 1999). A smaller literature has examined the relationship between changes in public opinion and outputs in specific policy areas. For instance, Pacheco (2013) finds a strong correlation between changes in public opinion and education and welfare spending. Based on all of this previous theoretical and empirical work, we expect to find a strong relationship between changes in public opinion and state policy liberalism.

Institutional features of state governments might also affect the link between public opinion and policy outputs in state government. We focus on three institutions that previous work has argued influence responsiveness. One institution that might improve responsiveness is the presence of direct democracy (Gerber 1996). There are three ways that the presence of direct democracy might increase states’ responsiveness to public opinion. First, it gives citizens the ability to circumvent elected officials and enact their preferred policy through the ballot box (Matsusaka 2008). Second, the threat of the initiative may lead elected officials to change their behavior. For instance, they may seek to pass laws to preempt future ballot measures (Gerber 1996). Finally, even if elected officials do not actively seek to preempt future initia-
tives, the results of initiatives may give them more accurate information about voter preferences (Matsusaka 2005).

Despite sound theoretical reasons to expect that direct democracy might improve responsiveness, empirical studies of its effects have been ambiguous. Some studies find that direct democracy enhances responsiveness, at least in some policy areas (Arceneaux 2002; Bowler and Donovan 2004; Gerber 1996; Matsusaka 2010), while others find that direct democracy has no effect on responsiveness in state governments (Monogan, Gray, and Lowery 2009; Lascher, Hagen, and Rochlin 1996; Lax and Phillips 2009a, 2011). In our analysis, we use data on when direct democracy was enacted from Matsusaka (2005).

Another institution that may improve responsiveness to public opinion is the presence of term limits. Some scholars argue that term limits should induce greater turnover among legislators. As a result, they might lead to the election of legislators who better reflect constituents’ preferences. Others argue that term limits lead to less experienced legislators, which reduces their capacity to assess and respond to public opinion. Term limits may also reduce incentives to respond to public opinion by limiting the value of a seat in the legislature (Kousser 2005). There have been few empirical studies of the effect of term limits on representation. But one recent study finds that term limits improve cross-sectional responsiveness to public opinion (Lax and Phillips 2011). In our empirical analysis, we measure the presence of term limits based on when they were enacted.

Finally, legislator professionalism may affect state governments’ responsiveness to public opinion. Legislative professionalism varies dramatically across states. Some states have very professional legislatures that resemble the U.S. Congress (e.g., California, New York, and Wisconsin), while others have part-time legislators that meet for only a few weeks a year (e.g., Delaware or Vermont) (Squire 2007). Professionalized legislatures meet in lengthy, often year-round, sessions, are well compensated,
and employ a large, professional staff (Squire 1992). In contrast, citizen legislatures often meet for only a few weeks a year (or even every other year), the legislators have few staff, and compensation is low, which means that most legislators hold jobs outside the legislature. Professional chambers can use their resources to assess changes in public opinion. They also can shape legislation to fit these shifts in the views of the public. In addition, seats in professional legislatures are more valuable than seats in part-time chambers, so there are greater incentives for lawmakers to be responsive to the public in order to retain office (Maestas 2000). As a result, we expect states with more professionalized legislatures to be more responsive to public opinion. Two recent studies find that states with higher levels of legislative professionalism are more responsive to public opinion (Pacheco 2013; Lax and Phillips 2011).

There is no time-varying measures of legislative professionalism that extends back to the 1960s. However, Lax and Phillips (2011) find that “the length of legislative sessions is the key component of professionalization.” Therefore, we use data from the Book of the States on the amount of time that legislatures are in session as a crude measure of professionalism. We define professionalism as the percentage of the year that the legislature is in session. This measure is correlated at about 0.81 with Squire 2007 more comprehensive measure of legislative professionalism, which is available for 1979, 1986, 1996, and 2003.5

2 Citizen Policy Liberalism

Evaluating responsiveness requires that citizen preferences and government outputs be measured with respect to the same object. Since we are interested in policy responsiveness, that object is policy, and so we must construct measures of citizens’ policy preferences and of the policy outputs of (state) governments. One option would

---

5. Note: We are working on cleaning and improving our measure of legislative professionalism. We expect to improve this measure in future drafts of this paper.
be to examine each policy or policy domain separately, as do, for example, Gilens (2005), Lax and Phillips (2009a), and Matsusaka (2010). Following Erikson, Wright, and McIver (1993), Stimson, MacKuen, and Erikson (1995), and many others, we take the alternative route of summarizing preferences and policies on a single dimension, which we label *policy liberalism* to distinguish it from ideological self-identification. This section outlines our measurement strategy for *citizen policy liberalism*, and the following section does the same for *government policy liberalism*.

The lack of a valid, time-varying measure of citizen policy liberalism has been one of the main barriers to the study of representation. As a result, scholars have been forced to rely on proxies, such as ideological self-identification, presidential vote, or the roll-call records of the state congressional delegation (e.g., Erikson, Wright, and McIver 1993; Levitt 1996; Berry et al. 1998). Only in recent years have the available survey data permitted the use of techniques like item-response (IRT) models to estimate citizens’ policy liberalism (e.g. Tausanovitch and Warshaw 2013). But such techniques cannot be extended back in time, when survey data were much sparser and included many fewer policy questions.

To overcome this challenge, we apply the dynamic, hierarchical group-level IRT model developed by Caughey and Warshaw (2015), which estimates the average policy

---

6. Lax and Phillips (2009a) claim that “using... policy-specific estimates” allows them to “avoid problems of inference that arise when policy and opinion lack a common metric.” On a policy-by-policy basis this may be true. But evaluating responsive on state policy in general, or even just in the domain of gay rights, requires that the policy-specific estimates of responsiveness be weighted or otherwise mapped onto a single dimension. Thus, dimension reduction must occur at some point, whether at the measurement stage or later in the analysis.

7. While an important construct in its own right, “symbolic” ideological identification is conceptually and empirically distinct from “operational” ideology expressed in the form of policy preferences (Free and Cantril 1967; Ellis and Stimson 2012). We conceptualize liberalism not as a logically coherent ideology, but as a set of ideas and issue positions that, in the context of American politics, “go together” (Converse 1964). Relative to conservatism, liberalism involves greater government regulation and welfare provision to promote equality and protect collective goods, and less government effort to uphold traditional morality and social order at the expense of personal autonomy. Conversely, conservatism places greater emphasis on economic freedom and cultural traditionalism (Ellis and Stimson 2012, 3-6).

8. For concision, we sometimes refer to these concepts as “citizen liberalism” and “government liberalism.”
liberalism of population subgroups such as states.\footnote{Our approach bears a close relation to the literature on “public policy mood” (Stimson 1991). Works in this tradition use Stimson’s Dyad Ratios algorithm to estimate changes in public preferences for government activity (i.e., left-liberalism). Mood is an inherently aggregate concept, and most studies of mood are concerned only with change over time, not cross-sectional differences (e.g., Bartle, Dellepiane-Avellaneda, and Stimson 2011, 267). Recently, however, Enns and Koch (2013) have combined the Dyad Ratios algorithm with MRP to generate state-level estimates of policy mood. As McGann (2014) observes, though, the Dyad Ratios algorithm has several unappealing features, most notably its ideological asymmetry and its lack of a grounding in a coherent individual-level model. As an alternative, he proposes a group-level IRT model for national mood that is similar to the approach we take. However, our dynamic group-level IRT model, accommodates cross-sectional and over-time variation within a common framework.} This approach builds upon three important approaches to modeling public opinion: item-response theory, multilevel regression and poststratification, and dynamic measurement models. Crucially, the model does not require multiple questions per respondent, allowing the use of the vast number of historical surveys that do not meet this standard.

**Measurement Model**

The conventional probit IRT model characterizes individual $i$’s dichotomous response $y_{ij} \in \{0, 1\}$ to item $j$ as

$$y_{ij} \sim \text{Bernoulli}(\Phi[\beta_j \theta_i - \alpha_j])$$

where $\Phi$ is the standard normal CDF, $\alpha_j$ is the item’s “difficulty,” $\beta_j$ is its “discrimination,” and $\theta_i$ is individual’s score on some latent trait (Jackman 2009, 455; Fox 2010, 10). In our application, $\theta_i$ represents $i$’s policy liberalism, which (together with $\alpha_j$ and $\beta_j$) determines $i$’s probability of favoring liberal policy alternatives.

Accurate estimation of $\theta_i$ requires data on many subjects, each of whom answers many items (Lewis 2001, 277). Unfortunately, only a small minority of public opinion surveys contain enough items to make estimation of $\theta_i$ remotely plausible. However, it is often possible to make inferences about the distribution of $\theta_i$, such as the latent liberalism of African-Americans in Texas, even when individual-level estimation is
impossible (see Lewis 2001; Mislevy 1983; McGann 2014). The idea is to model the \( \theta_i \) in group \( g \) as distributed normally around the group mean \( \bar{\theta}_g \) and then marginalize over the distribution of abilities.\(^{10}\) This model can be written as:

\[
s^*_gjt \sim \text{Binomial}(n^*_gjt, p_gjt),
\]

where

\[
p_gjt = \Phi\left( \frac{\bar{\theta}_{gt} - \kappa_j}{\sqrt{\sigma^2_{\theta,t} + \sigma^2_j}} \right).
\]

The time-indexed hierarchical model for the vector of group means is

\[
\bar{\theta}_t \sim N(\xi_t + X_t \gamma_t, \sigma^2_{\theta,t}).
\]

IRT models must be identified using restrictions on the parameter space (e.g., Clinton, Jackman, and Rivers 2004). To fix the direction of the metric, we code all question responses to have the same polarity (e.g., higher values as more liberal), and restrict the sign of the discrimination parameter \( \beta_j \) to be positive for all items. We identify the location and scale by rescaling the item parameters \( \alpha \) and \( \beta \) (Fox 2010, 88–9).\(^{11}\) We also assume that the item parameters \( \kappa_j \) and \( \sigma_j \), are constant across time, which bridges the model temporally, allowing latent opinion estimates in different periods to be compared on a common metric. For most parameters, we employ weakly informative priors that are proper but provide relatively little information.\(^{12}\) We estimated the model using the program Stan, as called from R.

---

\(^{10}\) For evidence that voter preferences are distributed normally (though not necessarily homoskedastically) within states and congressional districts, see Kernell (2009).

\(^{11}\) In each iteration \( m \), we set the location by transforming the \( J \) difficulties to have a mean of 0: \( \tilde{\alpha}_{jm} = \alpha_{jm} - J^{-1} \sum_{j=1}^{J} \alpha_{jm} \). Similarly, we set the scale by transforming the discriminations to have a product of 1: \( \tilde{\beta}_{jm} = \beta_{jm} (\prod_{j=1}^{J} \beta_{jm})^{-1/J} \). The transformed parameters \( \tilde{\alpha}_j \) and \( \tilde{\beta}_j \) are then re-parameterized as \( \kappa_j \) and \( \sigma_j \), which enter into the group-level response model (see Equation 4).

\(^{12}\) The first-period priors for all standard deviation parameters are half-Cauchy with a mean of 0 and a scale of 2.5 (Gelman 2007; Gelman, Pittau, and Su 2008). The difficulty and discrimination parameters are drawn respectively from \( \mathcal{N}(0, 1) \) and \( \ln\mathcal{N}(0, 1) \) prior distributions and then trans-
Finally, we combine our model estimates with poststratification weighting to estimate each state’s latent policy liberalism over the past fifty years. As Park, Gelman, and Bafumi (2004, 2006) demonstrated and others (Lax and Phillips 2009b; Warshaw and Rodden 2012) have confirmed, weighting model-based group opinion estimates to match population targets can substantially improve estimates of average opinion in states and other geographic units.

Public Opinion Data

Our public opinion data consists of survey responses to 47 domestic policy questions spread across 350 public-opinion surveys fielded between 1960 and 2012. The questions cover traditional economic issues such as taxes, social welfare, and labor regulation, as well as topics like gun control, immigration, and environmental protection. For conceptual clarity and comparability with policy mood, this application includes only questions for which the “liberal” answer involved greater government spending or activity.14 The responses of over 650,000 different Americans are represented in the data.

We model opinion in groups defined by states and a set of demographic categories (e.g., race and gender). In order to mitigate sampling error for small states, we

---

13. Stan is a C++ library that implements the No-U-Turn sampler (Hoffman and Gelman, Forthcoming), a variant of Hamiltonian Monte Carlo that estimates complicated hierarchical Bayesian models more efficiently than alternatives such as BUGS. In general, 4,000 iterations (the first 2,000 used for adaptation) in each of 10 parallel chains proved sufficient to obtain satisfactory samples from the posterior distribution. Computation time depends on the number of groups, items, and time periods; run times for the models reported in this paper ranged between a day and several weeks.

14. For example, questions about restricting access to abortion were not included. Stimson (1999, 89–91) notes that the temporal dynamics of abortion attitudes are distinct from other issues, at least before 1990.
model the state effects in the first time period as a function of state Proportion Evangelical/Mormon. The inclusion of state attributes in the model partially pools information across similar geographical units in the first time period, improving the efficiency of state estimates (e.g., Park, Gelman, and Bafumi 2004, 2006). We drop Proportion Evangelical/Mormon after the first period because we found that the state intercept in the previous period tends to be much more predictive than state attributes.

State Estimates of Citizen Policy Liberalism

To generate annual estimates of average opinion in each state, we weighted the group estimates to match the groups’ proportions in the state population, based on data from the U.S. Census (Ruggles et al. 2010). Figure 1 maps our estimates of state
policy liberalism in 1965, 1980, 1995, and 2010 (See Appendix A for validation of our estimates). The cross-sectional patterns are generally quite sensible—the most conservative states are in the Great Plains, while New York, California, and Massachusetts are always among the most liberal states. Moreover, Figure 1 confirms that the states have remained generally stable in their relative liberalism, consistent with Erikson, Wright, and McIver (2006)’s finding that state publics have been stable in terms of ideological identification. According to our estimates, only a few states’ policy liberalism has shifted substantially over time. Southern states such as Mississippi and Alabama have become somewhat more conservative over time, while states in New England have become somewhat more liberal.

3 Government Policy Liberalism

Next, we need a measure of the policy outputs in each state. There are a number of approaches that scholars have used to conceptualize and measure state policy outputs. One common approach has been to focus on particular policies or policy domains (e.g., Besley and Case 2003; Lax and Phillips 2009a, 2011; Leigh 2008). However, while such policy-specific measures are appropriate when the research question is limited to a particular policy area, they are suboptimal as summary measures of state policy in general. Moreover, it is difficult to find individual policies that extend more than a few decades backward in time (McIver, Erikson, and Wright 2001).

As a result, a number of scholars have used indices, factor analysis, or other dimension-reduction methods designed to summarize state policies in terms of one or more (latent) dimensions.\(^\text{15}\) While different works have identified different traits or

\(^{15}\) Dimension reduction has several advantages over policy-specific measures. First, using multiple indicators for a latent trait usually reduces measurement error on the construct of interest, sometimes substantially (Hofferbert 1966; Ansolabehere, Rodden, and Snyder 2008). Secondly, many concepts require multiple indicators to adequately represent the full content or empirical domain of the concept. For example, the concept of liberalism, in its contemporary American meaning, encompasses policy domains ranging from welfare to environmental protection to civil rights. A final
dimensions underlying state policies, the state politics literature has primarily focused on a single left–right policy dimension (e.g., Hofferbert 1966; Klingman and Lammers 1984; Wright, Erikson, and McIver 1987; Gray et al. 2004). We label this dimension *government policy liberalism*.

**Measurement Model**

Like most previous work on the subject, we treat policy liberalism as a latent variable whose values can be inferred from observed policy indicators. We parameterize policy liberalism as a latent trait $\theta_{st}$ that varies across states and years. For each state $s$ and year $t$, we observe a mix of $J$ continuous and ordinal indicators of policy liberalism, denoted $y_{st} = (y_{1st}, \ldots, y_{jst}, \ldots, y_{Jst})$, whose distribution is governed by a corresponding vector of latent variables $y^*_{st}$. We model $y^*_{st}$ as a function of policy liberalism ($\theta_{st}$) and item-specific parameters $\alpha_t = (\alpha_{1t}, \ldots, \alpha_{jt}, \ldots, \alpha_{Jt})$ and $\beta = (\beta_1, \ldots, \beta_j, \ldots, \beta_J)$,

$$y^*_{st} \sim N_J(\beta\theta_{st}\alpha_t, \Psi),$$

where $N_J$ indicates a $J$-dimensional multivariate normal distribution and $\Psi$ is a $J \times J$ covariance matrix. Note that $\alpha_{jt}$, which is analogous to the “difficulty” parameter in the language of item-response theory, varies by year $t$, whereas the “discrimination” $\beta_j$ is assumed to be constant across time.

---

16. Our latent-variable model (LVM), however, offers several improvements over previous measurement strategies, most of which have relied on factor analysis applied to cross-sectional data. First, we use a Bayesian model, which unlike classical factor analysis provides straightforward means of characterizing the uncertainty of the latent scores and also easily handles missing data by imputing estimates on the fly (Treier and Jackman 2008; Jackman 2009, 237–8). Second, most of our policy indicators are dichotomous variables, a poor fit for a factor-analytic model, which assumes that the observed indicators are continuous. Third, our measurement model is dynamic, both in that it allows policy liberalism to vary by year and in that it specifies a dynamic linear model that links the measurement model between periods.

17. In this application, we assume $\Psi$ to be diagonal, but this assumption could be relaxed to allow for correlated measurement error across variables.
If policy indicator $j$ is continuous, $y_{jst} = y_{jst}^{*}$. That is, we assume that we observe $y_{jst}^{*}$ directly, just as in conventional factor analysis. If policy indicator $j$ is ordinal, however, we treat the observed $y_{jst}$ as a coarsened realization of $y_{jst}^{*}$ whose distribution across $K_j > 1$ ordered categories is determined by a set of $K_j + 1$ thresholds $\tau_j = (\tau_{j0}, \ldots, \tau_{jk}, \ldots, \tau_{jK_j})$. The conditional probability that $y_{jst}^{*} \sim N(\beta_j \theta_{st} \alpha_{jt}, 1)$ is observed as $y_{jst} = k$ is

\[
\Pr(\tau_{j,k-1} < y_{jst}^{*} \leq \tau_{jk} | \beta_j \theta_{st} \alpha_{jt}) = \Phi(\tau_{jk} | \beta_j \theta_{st} \alpha_{jt}) - \Phi(\tau_{j,k-1} | \beta_j \theta_{st} \alpha_{jt}),
\]

(6)

where $\Phi$ is the standard normal CDF (Fahrmeir and Raach 2007, 329). In the dichotomous case, where there are $K_j = 2$ categories ("0" and "1"), the conditional probability that $y_{jst}$ falls in the second category (i.e., "1") is

\[
\Pr(\tau_{j1} < y_{jst}^{*} \leq \tau_{j2} | \beta_j \theta_{st} \alpha_{jt}) = \Phi(\beta_j \theta_{st} \alpha_{jt}),
\]

(7)

which is identical to the conventional probit item-response model (Quinn 2004, 341).

We allow the $\alpha_{jt}$ to vary by year to account for the fact that many policies (e.g., segregation laws) trend over time towards universal adoption or non-adoptions. We use a local-level DLM, which models $\alpha_{jt}$ using a "random walk" prior centered on $\alpha_{j,t-1}$:

\[
\alpha_{jt} \sim N(\alpha_{j,t-1}, \sigma_{\alpha}^2).
\]

(8)

If there is no new data for an item in period $t$, then the transition model in Equation 8 acts as a predictive model, imputing a value for $\alpha_{jt}$ (Jackman 2009, 474).\footnote{Following convention, we define $\tau_{j0} \equiv -\infty$, $\tau_{j1} \equiv 0$, and $\tau_{jK_j} \equiv \infty$, and we set the diagonal elements of $\Psi$ that correspond to ordinal variables equal to 1 (Quinn 2004, 340). As in the conventional ordered probit model, $y_{jst}$ falls into category $k$ if and only if $\tau_{j,k-1} < y_{jst}^{*} \leq \tau_{jk}$.}

\footnote{The transition variance $\sigma_{\alpha}^2$ controls the degree of smoothing over time. Setting $\sigma_{\alpha}^2 = \infty$ is equivalent to estimating $\alpha_{jt}$ separately each year, and $\sigma_{\alpha}^2 = 0$ is the same as assuming no change over time. We take the more agnostic approach of estimating $\sigma_{\alpha}^2$ from the data, while also allowing it to differ between continuous and ordinal variables.}
We identify the location and scale of the model by post-processing the latent measure of state policy liberalism to be standard normal. For the prior on the innovation parameter $\sigma$, we use a half-Cauchy distribution with a mean of 0 and a scale of 2.5 (Gelman 2007; Gelman, Pittau, and Su 2008). The difficulty and discrimination parameters are drawn from normal distributions with a mean of 0 and a standard deviation of 10. We fix the direction of the model by constraining the sign of several item parameters (Bafumi et al. 2005). We further constrain the polarity by assigning an informed prior to the policy measure for four states in year $t = 0$ (Martin and Quinn 2002). We estimated the model using the program Stan, as called from R (Stan Development Team 2013; R Core Team 2013).

**Policy Data**

In order to measure state policy liberalism, we sought to include in our model all left–right policies between 1936 and 2012 on which we could obtain data over at least a five-year period. The policy data come from a variety of sources, including government documents, the Book of the States, interest groups, and various secondary sources. About 75% of the policies are dichotomous, while the remainder are ordinal.

---

20. Specifically, we constrain continuous measures of state spending to have a positive discrimination parameter, which implies that more liberal states spend more money. We also constrain the polarity of four dichotomous items. The discrimination of ERA ratification and prevailing wage laws are constrained to be positive, while the discrimination of right to work laws and bans on interracial marriage are constrained to be negative.

21. Note that we started the model in 1935 ($t = 0$) and discarded the first year of estimates. As a result, the informed priors on $\theta$ for four states in year $t = 0$ have little effect on the estimates of state policy liberalism that we report in our analysis. We assign a $N(1, 0.2^2)$ prior on $\theta_{st}$ to New York and Massachusetts, and a $N(-1, 0.2^2)$ prior for Georgia and South Carolina. Other states are given diffuse priors for $\theta_{st}$.

22. Stan is a C++ library that implements the No-U-Turn sampler (Hoffman and Gelman, Forthcoming), a variant of Hamiltonian Monte Carlo that estimates complicated hierarchical Bayesian models more efficiently than alternatives such as BUGS. We ran the model for 1,000 iterations (the first 500 used for adaptation) in each of 4 parallel chains, which proved sufficient to obtain satisfactory samples from the posterior distribution.

23. In general, we tried to obtain primary sources for each policy indicator. When this proved impossible, we obtained multiple secondary sources to corroborate the information about each policy in our database.
or continuous. Some of the policies move policy in a conservative direction, but in the majority of cases, adopting the policy is a liberal act. We adjusted all monetary expenditure or welfare benefit policies into 2012 dollars. We also adjusted for cost of living differences between states (Berry, Foding, and Hanson 2000).

In keeping with our broad definition of policy liberalism, our measure is based on a wide variety of policy domains. These domains include abortion (e.g., parental notification requirements for minors) criminal justice (e.g., the death penalty), drugs and alcohol (e.g., marijuana decriminalization), education (e.g., per-pupil education spending; ban on corporal punishment), the environment (e.g., protections for endangered species), gambling (e.g., casino bans), civil rights (e.g., fair employment laws; gay marriage), gun control (e.g., handgun registration), immigration (e.g., English-only laws), labor (e.g., right-to-work laws), social welfare (e.g., AFDC/TANF benefits), taxation (e.g., income tax rates), and miscellaneous other regulations (e.g., fireworks bans; no-fault divorce).

State Estimates of Government Policy Liberalism

Estimating our measurement model using the policy data just described produces estimates of the policy liberalism of each state between 1936 and 2012. Here, we focus on state policy liberalism between 1960 and 2012 since that is the period where we currently have public opinion data available (See Appendix B for validation of our estimates). The maps in Figure 2 demonstrate that in relative terms, the policy liberalism of most states stayed fairly consistent over the last fifty years. Across the entire

24. We standardized each continuous policy to ensure that the scales were comparable across policy areas.

25. It makes sense that not having a policy is usually the conservative position, given that at the most basic level, liberalism generally (though not always) involves government “doing more” (e.g., Stimson 1999).

26. When interpreting these estimates, one should bear in mind that the model allows the difficulty parameters $\alpha_i$ to evolve over time. As a result, aggregate ideological shifts common to all states will be partially assigned to the policy difficulties. We use a time-varying model because it helps avoid the interpretational difficulties of assuming that policies have the same substantive meaning across long stretches of time.
Figure 2: The evolution of state policy liberalism. Bluer colors indicate more liberal states, and redder colors indicate more conservative states. The estimates have been re-centered and standardized in each year to accentuate the color contrasts.

During this time period, the most conservative states are in the South, whereas New York, Massachusetts, and New Jersey are always among the most liberal states. Not all states, however, have been ideologically stable. For example, Idaho, Utah, and most other states in the Mountain West became much more conservative in this period. Other states exhibited shifts in the opposite direction. For instance, Maryland, Delaware, Maine, Vermont, and New Hampshire started the period with relatively conservative policies but had become quite liberal by the turn of the 21st century.

4  Dynamic Representation

Before we proceed to the dynamic relationship between citizen and government policy liberalism, we first examine their cross-sectional relationship across states over the
past half-century. Consistent with previous cross-sectional analyses (e.g., Erikson, Wright, and McIver 1993; Gray et al. 2004; Lax and Phillips 2011), Figure 3 shows that the policy liberalism of state publics is a strong predictor of the policy liberalism of state governments. The upper-left corner of each plot indicates the year-specific correlation coefficient, which is very close to the standardized regression coefficient because Government Liberalism has approximately unit-variance within years. The average standardized coefficient across years is 0.73, implying that a cross-state difference in Citizen Liberalism of one standard deviation is associated with a difference in Government Liberalism of nearly three-quarters of a standard deviation.

Unlike previous studies, our data also enable us to see how the cross-sectional relationship has evolved over time. Since the 1960s, states have polarized in terms of their Citizen Liberalism, greatly increasing the correlation between opinion and policy even as the unstandardized regression slope of their relationship has remained largely stable. In other words, the proportion of the variation in Government Liberalism explained by Citizen Liberalism has grown substantially over time. At this point, we are not certain to what degree this is an artifact of decreasing measurement error (imprecise measurement of citizen liberalism would attenuate the relationship) or an indication that policy responsiveness has actually increased over time. We suspect both factors are at play, and we hope to disentangle them in future analyses.

A major problem with cross-sectional analyses of representation is that it is very difficult to rule out the possibility that some third, unmeasured characteristic of states—the legacy of plantation slavery, for example—confounds the relationship between citizen liberalism and policy liberalism, or even that policy liberalism causes citizen liberalism. The advantage of our time-series cross-sectional (TSCS) dataset is that it enables us to exploit within-state variation in opinion and policy and thus control for persistent characteristics of states. The cross-sectional character of the data is useful too, because in contrast to time series with one cross-sectional unit
Figure 3: Relationship between Citizen Policy Liberalism and Government Policy Liberalism, 1964–2008. Correlation coefficients are indicated in the upper-left corner. The scale of the x-axis has been normalized within years. For space reasons, the plot excludes 2012, which looks very similar to 2008.

(e.g. Stimson, MacKuen, and Erikson 1995), it allows us to control for time-specific factors that affect all states, such as a stock-market crash.

We use two basic modeling frameworks to identify the effect of citizen liberalism on government liberalism. The first is a model with fixed effects (FEs) for state and year, which control for time-invariant state characteristics and for time-specific effects common to all states (Angrist and Pischke 2009). The primary threats to inference in a two-way FE model are interactions between time and state-specific omitted variables (Bai 2009, 1230). To address this threat, we augment the basic FE setup by controlling for time-varying state characteristics, such as Percent Urban,
and by interacting state intercepts with quadratic time trends, which accounts for
state-specific unobserved variables that vary smoothly through time.

Our second modeling approach dispenses with state FEs and instead uses lagged
dependent variables (LDVs) to capture unobserved omitted variables in each unit
(beck2011modeling: De Boef and Keele 2008). Because an LDV model estimates
the effect of change in $X$ on change in $Y$, this specification is the closest in spirit
to Stimson, MacKuen, and Erikson 1995 model of dynamic representation. Angrist
and Pischke (2009) argue that the FE and LDV approaches to modeling TSCS data
roughly bracket the true causal effect. We therefore report the results of a set of
alternative model specifications using both frameworks. Finally, we report results
using both standard OLS models, as well as models that take into account the mea-
surement error in our estimates of citizen and government policy liberalism (Gelman
and Hill 2006, 542).27

The results are summarized in Table 1. Across specifications, there is a robust
positive relationship between Citizen Liberalism and Government Liberalism. The
magnitude of the estimate is largest under the basic FE model (columns 1 and 4).
The fact that the coefficient is substantially reduced by the addition of quadratic time
trends (columns 2 and 5) suggests that the FE estimates are biased upward due to
common state-specific time trends in Citizen Liberalism and Government Liberalism.
The estimates for the LDV model, which controls for Government Liberalism $t-1$
and Citizen Liberalism $t-1$, are roughly the same magnitude as the model with
quadratic trends and still highly significant (columns 3 and 6).

The estimated effects are also substantively important, if not overwhelmingly so.
The within-state standard deviation of Government Liberalism is 0.28 across years
and that of (average) Citizen Liberalism is 0.21, so the coefficient estimates from
models 3 and 6 imply that a typical change in Citizen Liberalism causes a change in

---

27. We estimate the models that take into account measurement error using Zelig (Imai, King, and
Lau 2009).
<table>
<thead>
<tr>
<th></th>
<th>OLS</th>
<th>Errors in Variables</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
<td>(6)</td>
</tr>
<tr>
<td>Constant</td>
<td>0.59**</td>
<td>0.25</td>
<td>−0.08**</td>
<td>0.89**</td>
<td>0.30</td>
<td>−0.19**</td>
</tr>
<tr>
<td></td>
<td>(0.16)</td>
<td>(0.22)</td>
<td>(0.02)</td>
<td>(0.17)</td>
<td>(0.22)</td>
<td>(0.09)</td>
</tr>
<tr>
<td>Citizen Policy Liberalism</td>
<td>1.92**</td>
<td>0.48**</td>
<td>0.25**</td>
<td>1.03**</td>
<td>0.16</td>
<td>0.73**</td>
</tr>
<tr>
<td></td>
<td>(0.17)</td>
<td>(0.15)</td>
<td>(0.05)</td>
<td>(0.19)</td>
<td>(0.13)</td>
<td>(0.16)</td>
</tr>
<tr>
<td>Income Per Capita</td>
<td>0.03**</td>
<td>0.00</td>
<td>0.00</td>
<td>0.03**</td>
<td>0.00</td>
<td>0.01**</td>
</tr>
<tr>
<td></td>
<td>(0.00)</td>
<td>(0.01)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.01)</td>
<td>(0.00)</td>
</tr>
<tr>
<td>% Urban</td>
<td>−0.87**</td>
<td>−0.09</td>
<td>0.00</td>
<td>−1.02**</td>
<td>−0.10</td>
<td>0.16**</td>
</tr>
<tr>
<td></td>
<td>(0.23)</td>
<td>(0.49)</td>
<td>(0.02)</td>
<td>(0.23)</td>
<td>(0.50)</td>
<td>(0.06)</td>
</tr>
<tr>
<td>% Evang./Mormon</td>
<td>−1.11**</td>
<td>0.64</td>
<td>−0.07**</td>
<td>−1.35**</td>
<td>0.90</td>
<td>−0.48**</td>
</tr>
<tr>
<td></td>
<td>(0.38)</td>
<td>(0.62)</td>
<td>(0.03)</td>
<td>(0.39)</td>
<td>(0.62)</td>
<td>(0.08)</td>
</tr>
<tr>
<td>Lagged Cit. Lib.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.98**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.00)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FE for Year</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>FE for State</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quad. Time Trend by St.</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lagged Policy</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Account for Meas. Error</td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>2.650</td>
<td>2.650</td>
<td>2.650</td>
<td>2.650</td>
<td>2.650</td>
<td></td>
</tr>
<tr>
<td>R²</td>
<td>0.92</td>
<td>0.98</td>
<td>0.99</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adjusted R²</td>
<td>0.92</td>
<td>0.97</td>
<td>0.99</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: *p<0.1; **p<0.05

Zelig results take into account measurement error in our estimates of citizen and government policy liberalism.
Government Liberalism of somewhere between 0.19 and .55 standard deviations. In short, while our measure of Citizen Liberalism explains only a portion of the time-series variation in Government Liberalism, dynamic representation does appear to be alive and well in U.S. state politics.28

Next, in Table 2 we consider whether the strength of responsiveness varies depending on three factors: term limits for state legislators, the professionalism of the state legislature, and mechanisms for direct democracy. Of these three factors, the evidence is strongest for an interaction with Term Limits, though even this is not consistently significant across all specifications. Consistent with the cross-sectional results of Lax and Phillips (2011) but contrary to Matsusaka (2010), we find no evidence that direct democracy strengthens responsiveness. All of these results should be taken with a grain of salt, however, since we have little reason to believe that variation in these factors is uncorrelated with other determinants of responsiveness.

5 Conclusion

One of the most important assumptions of democratic theory is that the views of citizens should influence government policy decisions. A number of previous studies have found a strong cross-sectional relationship between public opinion and state policy outputs (Erikson, Wright, and McIver 1993; Gray et al. 2004; Lax and Phillips 2011). But the ultimate metric of responsiveness is the extent to which changes in popular preferences cause changes in public policies (Stimson, MacKuen, and Erikson 1995).

In this paper, we develop new time-varying measures of both citizen and government policy liberalism in the American states over the past half century. This enables

28. The substantial variation unexplained by Government Liberalism suggests that there is a lot of room for other factors in state policymaking. These include changes in the objective “need” for certain policies—spending cuts during an economic downturn, for example—as well as political factors other than the public’s policy preferences, such as scandal- or recession-induced shocks to the vote share of the president’s party.
<table>
<thead>
<tr>
<th></th>
<th>OLS</th>
<th>Errors in Variables</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>Constant</td>
<td>0.71**</td>
<td>0.35</td>
</tr>
<tr>
<td></td>
<td>(0.17)</td>
<td>(0.24)</td>
</tr>
<tr>
<td>Citizen Liberalism</td>
<td>1.66**</td>
<td>0.51**</td>
</tr>
<tr>
<td></td>
<td>(0.17)</td>
<td>(0.15)</td>
</tr>
<tr>
<td>Direct Democracy</td>
<td>−0.08</td>
<td>−0.07</td>
</tr>
<tr>
<td></td>
<td>(0.05)</td>
<td>(0.06)</td>
</tr>
<tr>
<td>Term Limits</td>
<td>0.13**</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>(0.04)</td>
<td>(0.03)</td>
</tr>
<tr>
<td>Leg. Professionalism</td>
<td>−0.09</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>(0.08)</td>
<td>(0.06)</td>
</tr>
<tr>
<td>Income Per Capita</td>
<td>0.03**</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>(0.00)</td>
<td>(0.01)</td>
</tr>
<tr>
<td>% Urban</td>
<td>−0.90**</td>
<td>−0.21</td>
</tr>
<tr>
<td></td>
<td>(0.23)</td>
<td>(0.50)</td>
</tr>
<tr>
<td>% Evang./Mormon</td>
<td>−0.73*</td>
<td>0.66</td>
</tr>
<tr>
<td></td>
<td>(0.38)</td>
<td>(0.62)</td>
</tr>
<tr>
<td>Lagged Policy</td>
<td>0.98**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.00)</td>
<td></td>
</tr>
<tr>
<td>Citizen Lib. × Dir. Dem.</td>
<td>−0.02</td>
<td>−0.04</td>
</tr>
<tr>
<td></td>
<td>(0.05)</td>
<td>(0.04)</td>
</tr>
<tr>
<td>Citizen Lib. × Term Lim.</td>
<td>1.53**</td>
<td>0.19</td>
</tr>
<tr>
<td></td>
<td>(0.18)</td>
<td>(0.13)</td>
</tr>
<tr>
<td>Citizen Lib. × Leg. Prof.</td>
<td>0.71**</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>(0.24)</td>
<td>(0.16)</td>
</tr>
<tr>
<td>FE for Year</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>FE for State</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Quad. Time Trend by St.</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Lagged Policy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Account for Meas. Error</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>2,649</td>
<td>2,649</td>
</tr>
<tr>
<td>R²</td>
<td>0.93</td>
<td>0.98</td>
</tr>
<tr>
<td>Adjusted R²</td>
<td>0.92</td>
<td>0.97</td>
</tr>
</tbody>
</table>

*Note:* *p<0.1; **p<0.05

Zelig results take into account measurement error in our estimates of citizen and government policy liberalism.
us to examine the dynamic relationship between public opinion and state policy outputs (Lowery, Gray, and Hager 1989; Ringquist and Garand 1999; McIver, Erikson, and Wright 2001). We find that state governments are clearly responsive to shifts in public opinion: when the state public moves to the left, the state adopts more liberal policies. We also evaluate whether the opinion–policy relationship is moderated by three factors thought to influence representation: term limits, direct democracy, and legislative professionalism. We find some evidence that term limits strengthen responsiveness, but no consistent support for the hypotheses that legislative professionalism and direct democracy strengthen responsiveness.

In future versions of this paper, we hope to extend our data on citizen and government policy liberalism backward in time, perhaps even to the dawn of the era of public opinion polling in 1936 (Berinsky et al. 2011). This will enable us to examine the quality of representation in the American states from the New Deal to the present. We also plan to evaluate the impact of additional institutions and other factors that might influence responsiveness.
References


——. 2008. For the many or the few: The initiative, public policy, and American democracy. University of Chicago Press.


Appendix A - Validation: Citizen Policy Liberalism

In the main paper, we use estimates U.S. citizen policy liberalism from 1960-2012. Here, we focus on validating the estimates from 1972-2012. This section is based on the discussion in Caughey and Warshaw (2015). Eventually, we will conduct a more extensive validation of the entire time series.

Cross-Validation

The use of multilevel modeling to smooth subnational opinion estimates across cross-sectional units has been well validated (Lax and Phillips 2009b; Warshaw and Rodden 2012; Tausanovitch and Warshaw 2013). A more innovative aspect of our model is the DLM for the parameters of the hierarchical model, which pools information across time in addition to cross-sectionally. Although a number of political science works have employed similar temporal smoothing methods (e.g., Martin and Quinn 2002; Jackman 2005; Park 2012; Linzer 2013; Wawro and Katznelson 2013), their application to dynamic public opinion has not been validated as extensively as multilevel modeling has. One noteworthy potential concern about our approach to dynamics is that even though the $\tilde{\theta}_{gt}$ are re-estimated in each period, smoothing the hierarchical coefficients across periods dampens the estimates’ sensitivity to rapid opinion changes (e.g., a sharp conservative turn in a specific state), especially in years when the data are thin.

To investigate this possibility, we designed a cross-validation study that compared the performance of our approach (the pooled model) to one in which the intercept and coefficients of the hierarchical model are estimated separately in each period (the separated model).\footnote{To keep the comparison transparent and minimize computation time (which was still very lengthy), we defined groups by state only, with no demographic covariates. We also restricted the time period covered to 1976-2010.} Specifically, we took a validation set approach (James et al. 2013, 176-8) in which 25% of respondents in each group-year were sampled to create a
training dataset.\textsuperscript{30} We used the training data to estimate both the pooled model and the separated model. Based on the parameter estimates from each model, we calculated the predicted proportion of liberal responses to each item in each group-year:

\[
\hat{p}_{gjt} = \Phi[(\hat{\theta}_{gt} - \hat{\kappa}_j)/\sqrt{\hat{\sigma}^2_{\theta,t} + \hat{\sigma}^2_j}].
\] (9)

To evaluate the out-of-sample performance of each model, we compared each predicted proportion with the proportion of liberal responses in the other 75\% of the data, generating the prediction error for each of the \(N\) item-group-year triads:

\[
\hat{e}_{gjt} = \frac{s^*_g}{n^*_g} - \hat{p}_{gjt}.
\] (10)

We contrasted the two models in terms of three metrics: bias \((N^{-1} \sum \hat{e}_{gjt})\), mean absolute error \((N^{-1} \sum |\hat{e}_{gjt}|)\), and root-mean-square error \((\sqrt{N^{-1} \sum \hat{e}^2_{gjt}})\). We replicated the whole process 10 times, thus producing 10 out-of-sample estimates of bias, MAE, and RMSE for each model.

As Table 2 indicates, the pooled model is clearly superior to the separated model in terms of bias, MAE, and RMSE. Though the differences (expressed in percentage points) are not large, the pooled model strictly dominates the separated in every replication but one. The improvement in efficiency is to be expected given that the pooled model borrows strength from adjacent periods. That the pooled model exhibits less bias—in fact, is nearly unbiased when averaged across replications, in contrast to the liberal bias of the separated model—is perhaps more surprising, given that Bayesian smoothing shrinks estimates away from the (unbiased) maximum likelihood estimate. The explanation is that the coefficient estimates in the separated model are shrunk as well, but towards the cross-sectional mean rather than towards their value.

\ \textsuperscript{30} We sampled 25\% rather than splitting the sample equally because we wanted to compare the models’ performance when data are relatively sparse, and secondly to leave enough out-of-sample data to generate precise estimates of bias, MAE, and RMSE.
<table>
<thead>
<tr>
<th>Rep.</th>
<th>Separated Model</th>
<th>Pooled Model</th>
<th>Diff. in Magnitude</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Bias MAE RMSE</td>
<td>Bias MAE RMSE</td>
<td>Bias MAE RMSE</td>
</tr>
<tr>
<td>1</td>
<td>0.50 13.37 18.71</td>
<td>0.21 13.08 18.36</td>
<td>0.29 0.29 0.35</td>
</tr>
<tr>
<td>2</td>
<td>0.43 13.31 18.58</td>
<td>0.11 13.02 18.23</td>
<td>0.31 0.29 0.35</td>
</tr>
<tr>
<td>3</td>
<td>0.26 13.46 18.80</td>
<td>0.00 13.17 18.44</td>
<td>0.26 0.29 0.36</td>
</tr>
<tr>
<td>4</td>
<td>0.13 13.44 18.76</td>
<td>−0.26 13.13 18.40</td>
<td>−0.13 0.31 0.36</td>
</tr>
<tr>
<td>5</td>
<td>0.53 13.35 18.64</td>
<td>0.24 13.07 18.32</td>
<td>0.28 0.29 0.32</td>
</tr>
<tr>
<td>6</td>
<td>0.36 13.37 18.79</td>
<td>0.01 13.11 18.46</td>
<td>0.35 0.26 0.33</td>
</tr>
<tr>
<td>7</td>
<td>0.22 13.50 18.86</td>
<td>−0.10 13.20 18.49</td>
<td>0.12 0.30 0.37</td>
</tr>
<tr>
<td>8</td>
<td>0.50 13.43 18.76</td>
<td>0.28 13.17 18.44</td>
<td>0.22 0.26 0.32</td>
</tr>
<tr>
<td>9</td>
<td>0.15 13.40 18.78</td>
<td>−0.14 13.11 18.42</td>
<td>0.01 0.30 0.35</td>
</tr>
<tr>
<td>10</td>
<td>0.42 13.36 18.72</td>
<td>0.21 13.06 18.37</td>
<td>0.21 0.30 0.35</td>
</tr>
<tr>
<td>Mean</td>
<td>0.35 13.40 18.74</td>
<td>0.06 13.11 18.39</td>
<td>0.19 0.29 0.35</td>
</tr>
</tbody>
</table>

Table 3: Out-of-sample bias, MAE, and RMSE of the separated and pooled models across 10 cross-validation replications. The rightmost panel reports the difference in magnitude between the models (e.g., \(|\text{Bias}_{\text{separated}}| - |\text{Bias}_{\text{pooled}}|\)). All values are expressed in terms of percentage points.

In summary, the cross-validation results corroborate the value of pooling the hierarchical coefficients over time via a dynamic linear model. Temporal smoothing results not only in greater efficiency but also in less bias than estimating the hierarchical model separately by period, at least in this application. Thus for the general purpose of measuring opinion over time, pooling appears to be the better choice. Nevertheless, the separated model may be preferable in certain circumstances, such as when one wishes to estimate abrupt opinion changes within a demographic group or geographic unit.

**Construct Validation**

The split-sample validation approach shows that our pooled model dominates a separated model where the intercept and coefficients of the hierarchical model are estimated separately in each period. However, it only partially speaks to the ability of our model to accurately estimate state and national-level policy liberalism. To
further assess our estimates’ validity as a measure of policy liberalism, we examine their correlation with measures of several theoretically related constructs (a procedure Adcock and Collier 2001 refer to as “construct validation”).

First, we examine the cross-sectional correlation between our measure of policy liberalism and Democrats’ presidential vote share. While presidential election results are not a perfect measure of citizens’ policy preferences (Levendusky, Pope, and Jackman 2008; Kernell 2009), a variety of previous scholars have used presidential election returns to estimate state and district preferences (Ansolabehere, Snyder, and Stewart 2001; Canes-Wrone, Brady, and Cogan 2002). Thus, to the extent that policy attitudes predict presidential partisanship, a high correlation with Democratic presidential vote share would suggest that our estimates are accurate measures of states’ policy preferences. Figure 4 shows that there is indeed a strong cross-sectional relationship between our estimates of state policy liberalism and presidential vote share between 1972 and 2012.\footnote{We find a similarly strong relationship between our estimates of state policy liberalism and estimates of state ideology from exit polls.} Moreover, the relationship increases in strength over time, mirroring the growing alignment of policy preferences with partisanship and presidential voting at the individual level (Fiorina and Abrams 2008, 577–82).

While the strong relationship with presidential vote share demonstrates the cross-sectional validity of our measure, it does not provide information about the ability of our model to detect changes in the mass public’s preferences over time. Presidential votes are ill-suited for this task since partisan vote shares could ebb and flow for reasons unrelated to changes in the policy liberalism of the American public. For instance, parties could nominate a low-valence candidate, or there could be an incumbency advantage for presidents running for a second term. To validate the over time validity of our estimates, we turn to Stimson’s “public policy mood”, which is explicitly designed to measure changes in the mass public’s policy preferences over
time (Stimson 1991). Of course, we should not expect a perfect correlation between policy liberalism and mood since they are measuring different concepts. Mood is focused on whether the government should be doing “more” or “less” than it currently is (e.g., Stimson 2012, 31). In contrast, our measure of policy liberalism is an absolute measure of the public’s preferences for government spending or activity that is not explicitly tied to the status quo.

Despite these theoretical differences between policy liberalism and Stimson’s mood, the national trends in both measures look very similar. The most liberal period for both mood and policy liberalism was around 1990, while the most conservative period was around 1980. Moreover, both mood and our measure of policy liberalism show a marked shift to the ideological right after 2008. The only major divergence between the two scales is in the early 2000s. However, note that Stimson’s mood estimates

---

32. Enns and Koch (2013) use a similar validation strategy for their measures of state-level mood.
Figure 5: Relationship between national policy liberalism and policy mood, 1972–2012.

are quite inefficiently estimated during this period. Overall, the correlation between policy liberalism and Stimson’s mood is 0.67, which further validates the ability of our model to detect over time changes in latent public opinion.
Appendix B – Validation: Govt. Policy Liberalism

In this appendix, we provide more systematic evidence for the validity of our measure of state government policy liberalism. We do so by documenting our estimates’ empirical relationship with alternative measures of policy liberalism, what Adcock and Collier (2001) refer to as “convergent” validation. Then we examine their association with other, theoretically related concepts (“construct” validation, in their terminology). Finally, we provide evidence that a one-dimensional model adequately captures the systematic variation in states’ policies. Overall, we find strong evidence that our estimates are valid measures of state policy liberalism.

Convergent Validation

In order to evaluate the convergent validity of our measure of state policy liberalism, we examine its correlation with other measures of policy liberalism (Adcock and Collier 2001). If our estimates accurately capture the evolution of state policy liberalism, they should be highly correlated with cross-sectional measures of state policy liberalism developed by other scholars. Figure 6 compares our dynamic measure with six cross-sectional measures of state policy liberalism from previous studies.

First, we compare our measure of state policy liberalism to several measures developed by scholars in the 1960s. In the upper-left panel of Figure 6, we compare our estimates of state policy liberalism to Hofferbert 1966 index of welfare orientation, which he interprets as an indicator of state liberalism.\(^{33}\) The correlation between our measure of policy liberalism and Hofferbert 1966 measure is 0.77. On both indices, Mississippi is the most conservative state and New York is the most liberal state. Next, we compare our estimates of state policy liberalism to Sharkansky and Hoffer-
Figure 6: Validation of our Policy Measure: Correlation with Previous Policy Indices

Sharkansky and Hofferbert 1969 index of state liberalism on welfare and education policies in 1962 (upper-right panel of Figure 6). Our measure and Sharkansky and Hofferbert (1969)’s are correlated at 0.82. Once again, Mississippi is the most conservative state on both indices, while California, Massachusetts, New Jersey, and New York are among the most liberal states on both indices.

Next, we compare our measure of state policy liberalism to several measures developed by scholars in the 1970s and 1980s. The middle-left and upper-right panels of Figure 6 plot our estimates’ relationship with Klingman and Lammers 1984 and Erikson, Wright, and McIver 1993 factor-analytic measures of policy liberalism. The

34. This index is based on about twenty education and welfare policies. Note, however, that this index also includes several social outcomes, such as school graduation rates.

35. Klingman and Lammers 1984 policy index is based on data measured at a variety of points.
correlations between their measures and ours are 0.91 and 0.92, respectively.

Lastly, we compare our estimates of state policy liberalism to several measures developed by scholars in the 2000s. The lower-left panel of Figure 6 compares our policy liberalism estimates for 2000 with those of Gray et al. (2004), which are correlated with our measure at 0.87. Finally, the lower-right panel of Figure 6 compares our measure of state policy liberalism in 2006 to the index from Sorens, Muedini, and Ruger (2008). Their model of state policy liberalism is based on over 100 distinct policy indicators and is the most analogous to our index. Our measure of policy liberalism is correlated with Sorens, Muedini, and Ruger (2008)’s measure at 0.85. On both indices, Mississippi, South Dakota, Wyoming, and Alabama are among the most conservative states and California and New York are the most liberal states.

Overall, our measure of state policy liberalism is highly correlated with a variety of previous cross-sectional measures. Each of these previous measures of states policy liberalism uses slightly different data. Moreover, they are based on data measured at different points in time, sometimes covering a decade or more. The strong empirical relationships between our estimates of state policy liberalism and previous measures of the same concept provide reassuring evidence for the validity of our measure.

between 1961 and 1980 on state innovativeness, anti-discrimination policies, monthly payments for Aid to Families with Dependent Children (AFDC), the number of years since ratification of the Equal Rights Amendment for Women, the number of consumer-oriented provisions, and the percentage of federal allotment to the state for Title XX social services programs actually spent by the state. We compare Klingman and Lammers 1984 scale with our measure of state policy liberalism in 1973 since this is the midpoint of the years they include in their index. Erikson, Wright, and McIver 1993 measure is based on state education spending, the scope of state Medicaid programs, consumer protection laws, criminal justice provisions, whether states allowed legalized gambling, the number of years since ratification of the Equal Rights Amendment for Women, and the progressivity of state tax systems. We compare Erikson, Wright, and McIver 1993 scale with our measure of state policy liberalism in 1980 since this is roughly the midpoint of the years they include in their index.

36. This index is based on state firearms laws, state abortion laws, welfare stringency, state right-to-work laws, and the progressivity of state tax systems.
Construct Validation

We provide further evidence for the validity of our measure by demonstrating its association with measures of concepts theoretically related to policy liberalism, a procedure Adcock and Collier (2001) refer to as “construct validation.” First, we examine the relationship between mass political attitudes and state policy liberalism. Previous work shows that the liberalism of state publics have a strong cross-sectional association with state policy liberalism (Wright, Erikson, and McIver 1987; Erikson, Wright, and McIver 1993; Lax and Phillips 2011). Unfortunately, there is no extant survey-based measure of state ideology that extends back to 1936, so we instead use Democratic presidential vote share to proxy for mass liberalism (see, e.g., Ansolabehere, Snyder, and Stewart 2001; Canes-Wrone, Brady, and Cogan 2002). Consistent with past work, we focus on non-southern states.

Figure 7 shows the correlation of our dynamic measure of policy liberalism with the Democratic candidate’s state-level vote share in every presidential election year from 1936 to 2012. As expected, the two measures are highly correlated across the entire time period. Moreover, the relationship between public opinion and policy liberalism increases in strength over time, mirroring the growing alignment of policy preferences with partisanship and presidential voting at the individual level (Fiorina and Abrams 2008, 577–82).

Next, we examine the relationship between state policy liberalism and the liberalism of state legislatures. Previous work suggests that policy outcomes should be tightly linked to the ideological location of the median in the legislature, at least over the long term (e.g., Krehbiel 1998). To measure the liberalism of legislative medians, we use Shor and McCarty 2011 estimates of state legislators’ ideal points, which map ideal points from different chambers onto a common scale. As Figure 8 demonstrates for state houses of representatives, more conservative state legislatures have more conservative state policies. The figure also shows that the link between the liberalism
Figure 7: Relationship between State Policy Liberalism and Democratic Presidential Vote Share in the Non-South.
Figure 8: Correlation between Median Legislator in State House and Policy Liberalism of state legislatures and state policy has strengthened considerably over the period for which data are available (1996–2012). In sum, as one would expect theoretically, state policy liberalism is strongly—but not perfectly—correlated with mass liberalism and the liberalism of state legislatures.

**Dimensionality**

Our one-dimensional model of state policies implies that a single latent trait captures systematic policy variation across states. This is not to say that it captures *all* policy differences, but it does imply that once states’ policy liberalism and policies’ difficulty and discrimination are accounted for, any additional variation in state policies is essentially random. This assumption would be violated if there were instead multiple dimensions of state policy, as some scholars have claimed. Given that for much of the 20th century, roll-call alignments in the U.S. Congress were substantially two-dimensional (Poole and Rosenthal 2007), it is not unreasonable to suspect that state
policies might be as well. As we demonstrate, however, a one-dimensional model captures state policy variation surprisingly well, and there is little value to increasing the complexity of the model by adding further dimensions.

In order to perform an initial evaluation of the dimensionality of state policies, we first compare our one-dimensional measure of state policy liberalism with four single-issue scales from various points in time. The upper left panel of Figure 9 shows the correlation between our scale and NARAL’s grades for each state on their pro-choice abortion policies in 2011 (NARAL 2012). The upper right panel shows the correlation between our scale and the Green Index of Environmental Innovation in 1991–92 (Hall and Kerr 1991; Ringquist and Garand 1999). The lower left panel shows the correlation between our scale and the number of liberal gay rights policies in 2009 from Lax and Phillips (2009b). Finally, the lower right panel shows the correlation between our scale and average AFDC benefits per family in each state (Moffitt 2002). Overall, there is a very high correlation between our index and each of these single-issue policy measures. This provides some evidence that one-dimension adequately characterizes the policies in our data.

For further evidence regarding dimensionality, we also compare the percent of policies correctly predicted by models with different numbers of latent dimensions. Overall, we find little evidence of higher-dimensional structure in this data. On average, a one-dimensional model correctly classifies 82.1% of all responses. A two-dimensional model increases the percent correctly predicted (PCP) to 83.5%, an increase of less than 1.5 percentage points (adding dimensions cannot decrease PCP). Moreover, the improvement in model fit from adding a second dimension varies little across time. This improvement in model fit is less than the increase in fit that is used.

---

37. We evaluated the dimensionality using the ideal function from the R package plsc (Jackman 2012). We estimated separate cross-sectional IRT models in each year. Since ideal only handles binary data, we dichotomized all our continuous data using random cut points. Finally, we ran 10 separate models in each year and report the mean percent correctly predicted to evaluate the improvement in model fit from including additional dimensions.
in the congressional literature as a barometer of whether roll-call voting in Congress has a one-dimensional structure (Poole and Rosenthal 2007, 33–4). Finally, the highest discriminating items on our scale span a wide variety of substantive policy

38. Some other papers have found modest evidence of higher dimensionality. However, they all use factor analytic methods with different sets of data. For instance, Sorens, Muedini, and Ruger (2008) find two dimensions of state policy liberalism in their factor analytic model using data from the 2000s. They find that the second dimension consists of “drug, gambling, and weapons offense arrest rates (higher in urbanist states), fiscal decentralization (urbanist states are more decentralized), death penalty (more likely in urbanist states), strictness of seat belt laws, laxity of campaign finance laws (urbanist states generally have higher contribution limits), per capita police (higher in urbanist states), lower tobacco taxes, less state role in alcohol distribution, lack of medical marijuana laws, lower wages for public employees, lower corporate income taxes, and a laissez-faire environment for private schools (fewer registration and approval requirements).” We do not include drug, gambling, and weapons offense arrest rates in our model since they are policy outcomes rather than policy decisions made by state governments. We also do not include per capita police spending since police spending is much higher in urban than rural areas. Most of the other policies that load on the second dimension in their model appear to load on the first dimension in our model.
Figure 10: Dimensionality of State Policies. Percent of policies correctly predicted by one-dimensional (blue dots) and two-dimensional (red triangles) models.

areas. The fact that these highly discriminating policies come from a variety of different policy areas further supports our finding that one-dimension is adequate to describe most of the variation in state policy liberalism.

39. For instance, the most highly discriminating items include policies about racial discrimination, women’s rights (jury service for women), gun control, labor law, energy policy, criminal rights, and welfare policy.