


ORIGINAL ARTICLE

A Validation and Extension of State-Level Public Policy Mood: 1956–2020

Julius Lagodny¹, Rebekah Jones², Julianna Koch³ and Peter K. Enns⁴ 

¹Data Science Lab, Hertie School, Berlin, Germany

²Department of Political Science, University of California, Berkeley, CA, USA

³buzzback Market Research, New York, NY, USA

⁴Department of Government and Brooks School of Public Policy, Cornell University, Ithaca, NY, USA

Corresponding author: Peter K. Enns, email: peterenns@cornell.edu

(Received 13 January 2021; revised 28 June 2021; accepted 15 December 2021)

Abstract

To fully understand state policy outcomes or elections in the US, we need valid over-time measures of state-level public opinion. We contribute to the research on measuring state public opinion in two ways. First, we respond to Berry, Fording, Hanson, and Crofoot's (BFHC) critique of Enns and Koch's measure of state policy mood. We show that when BFHC's analysis is performed using the same states and examining annual change, it validates the Enns and Koch measure and raises questions about the Berry, Ringquist, Fording, and Hanson measure. Second, we generate a new measure of state policy mood building on Enns and Koch's approach. The new measure has even better properties than the previous measure and relates to state presidential vote and state policy liberalism in similar ways to Caughey and Warshaw's measure of state economic liberalism. We conclude with recommendations for using the various direct measures of state public opinion.

Keywords: public opinion; issue preferences; voting behavior; political behavior; survey research; methodology

Introduction

Whether we consider policies or elections in the US, the states are influential. To understand these political outcomes, we need valid over-time measures of state-level public opinion. Fortunately, many researchers have developed such measures (e.g., Caughey and Warshaw 2018a; Enns and Koch 2013; Pacheco 2011; Shirley and Gelman 2015). In this article, we contribute to the state public opinion literature in two ways. First, we respond to Berry, Fording, Hanson, and Crofoot's (BFHC) critique of Enns and Koch's (2013; 2015) measure of state policy mood. We show

© The Author(s), 2022. Published by Cambridge University Press and State Politics & Policy Quarterly. This is an Open Access article, distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike licence (<http://creativecommons.org/licenses/by-nc-sa/4.0>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the same Creative Commons licence is used to distribute the re-used or adapted article and the original article is properly cited. The written permission of Cambridge University Press must be obtained prior to any commercial use.

that when BFHC's analysis is performed using the same states and examining annual change, it actually validates the Enns and Koch measure and raises questions about the Berry, Ringquist, Fording, and Hanson (BRFH) measure. Second, we build on Enns and Koch's methods to develop a new and extended measure of state policy mood. The new measure ranges from 1956 to 2020, includes data from more than 140 additional surveys that were not previously available, and modifies their multi-level regression and poststratification (MRP) model to improve the cross-sectional properties of the estimates. We show the new measure has even better properties than the Enns and Koch measure and relates to relevant state-level outcomes in similar ways to Caughey and Warshaw's (2018a, 2018b) measure of state economic liberalism. We conclude by highlighting the benefits of both the Caughey and Warshaw measure and our new measure and by offering recommendations for when these two direct measures of state public opinion are most applicable and when researchers might also rely on the original Enns and Koch measure.

BFHC's Analysis Validates Enns and Koch's Measure

BFHC aim to assess Berry, Ringquist, Fording, and Hanson's (BRFH) *indirect* state-level measure of citizens' preference for more or less government on a range of policy issues (often referred to as policy mood [Stimson 1999]) and Enns and Koch's (2013; 2015) measure, which *directly* measures the same concept. BRFH's measure is indirect because it relies on the ideological ratings of members of Congress (not the general public). The Enns and Koch measure is direct because it is based on attitudes expressed in public opinion surveys. BFHC's evaluation of the two measures seems straightforward. They use the General Social Survey, a high-quality national probability-based survey that began in 1972, to estimate policy mood in southern states. They then compare the over-time relationship of southern policy mood based on the GSS with southern policy mood based on the BRFH and Enns and Koch measures.

They find the BRFH measure corresponds more closely with the GSS measure. This is a surprising result for several reasons. First, although no measure is perfect, a variety of analyses have validated the Enns and Koch (2013; 2015) measure. Second, the general approach of MRP—which Enns and Koch use—has been widely used to generate valid over-time state-level estimates (e.g., Caughey and Warshaw 2015, Pacheco 2011, Shirley and Gelman 2015) and Enns and Koch's specific approach has been validated in other contexts, such as generating over-time state-level measures of the public's punitiveness (Enns 2016) and correctly forecasting the 2020 presidential winner in 49 states plus Washington, DC (Enns and Lagodny 2021a, 2021b). Third, and perhaps most important, the Enns and Koch measure includes the same GSS data that BFHC use in their validation. Why would the Enns and Koch measure correlate so weakly with the data used to generate the measure?

There are two reasons for this weak correlation. First, although BFHC focus on southern states, the states compared across measures are not the same (see their footnote 3). Because they relied on the GSS coding for southern states, their GSS measure includes five states (Delaware, Kentucky, Maryland, Oklahoma, and West Virginia) and Washington DC that are *not* included in the BRFH or Enns and Koch measures of southern public opinion. Different research questions will undoubtedly suggest different definitions of the south, but when comparing measures it is crucial to keep the states defined as southern consistent. Otherwise, it is impossible to know if

differences result from different measurement strategies or from the different states analyzed. To solve this problem, we obtained a version of the GSS from NORC at the University of Chicago that includes state identifiers. Thus, we were able to generate a new measure of Southern policy mood based on the GSS data that exactly matches BRFH and Enns and Koch's coding.

Second, despite their interest in over-time relationships, BFHC limit their analysis to comparing linear trends instead of evaluating year-to-year variation. Not only does year-to-year variation offer a more rigorous test of relationships, but this is the concept of interest in time series analysis. In fact, to avoid the risk of identifying spurious relationships, scholars must *remove* or control for linear trends in their data.¹ To solve these concerns, we analyze the exact same states across measures and we analyze the relationship in the year-to-year *change* for each series. If the measures are related, we expect an increase (or decrease) in one measure to correspond with an increase (or decrease) in the other measure, yielding a positive correlation.

Although BFHC generate 22 different series from the GSS to compare with their measure and the Enns and Koch measure, we focus our analysis on the GSS series that BFHC refer to as "Stimson Items Index (11 Items)." Enns and Koch explicitly aim to create a state-level measure of Stimson's Policy Mood. Thus "Stimson Items Index" is the appropriate comparison.² Table 1 presents the bivariate correlation between changes in the GSS Southern policy mood measure and the Enns and Koch and BRFH measures for the exact same southern states. When we compare the correct states and analyze change, the relationship between the BRFH and GSS measure is not statistically significant and the correlation is *negative* (−0.11). By contrast, the correlation between the Enns and Koch measure and the GSS measure is 0.40 and significant. When comparing the same states and examining annual change, BFHC's analysis validates the Enns and Koch measure and raises concerns about the validity of the BRFH measure.

While the results for the BRFH measure are concerning, we do not want to place too much emphasis on this analysis. Using the GSS to generate an estimate of Southern policy mood seems intuitive, but there are several reasons BFHC's approach

Table 1. The bivariate correlation between *change* in the GSS Southern policy mood and the Enns and Koch and BRFH measures, 1973–2010

	Δ GSS: Stimson items index (11 items)
Δ BRFH citizen ideology	−0.11
Δ Enns and Koch policy mood	0.40*

Note. Southern states include AL, AR, FL, GA, MS, LA, NC, SC, TN, TX, and VA. $N = 26$.

* $p < 0.05$.

¹More specifically, the series (or a linear combination of the series) must be stationary for time series analysis (e.g., Banerjee et al. 1993; Enders 2004; Enns and Wlezien 2017).

²Since Stimson's policy mood is the concept of interest, it is not clear why BFHC present 22 different comparisons, including abortion rights and support for gay rights, which do *not* relate to Stimson's policy mood theoretically or empirically (Stimson 1999, Ch. 4). To generate the "Stimson Items Index (11 Items)" we follow BFHC and standardize each of the 11 items to have a mean of 0 and a variance of 1. We then calculate the annual average of the 11 items, which include support for education spending, environmental spending, welfare spending, healthcare spending, drug addiction spending, spending on big cities, paying higher taxes, government doing more, redistribution, government helping the poor, and government aid for healthcare. We weight the data using the *wtsall* variable.

is not an ideal validation test. First, although the questions analyzed above were asked in the GSS from 1973 to 2010, the GSS was not conducted in 11 of those years. Second, the GSS was designed to be nationally representative. While the GSS represents a relatively large national sample, the analysis of southern states includes an annual average of about 560 respondents (median equals 446) and BFHC's approach does not reweight these to be representative of the South. For these reasons, we want to be cautious about making too much of BFHC's proposed validation. We are also mindful, however, that others have also raised questions about the BFRH measure, particularly the reliance on legislator behavior to measure public preferences (e.g., Brace *et al.* 2006; Carsey and Harden 2010; Erikson, Wright, and McIver 2007).³ Further, BRFH have expressed caution about their measure in the past, advising against using it prior to the mid-1970s (Berry *et al.* 2015, 10). Given these concerns and with further evidence of the validity of the Enns and Koch measure in Table 1, in the next section, we build on the Enns and Koch approach to generate a new and extended measure of state policy mood.

A New Measure Using the Enns and Koch Approach

Enns and Koch use a two-step approach to generate estimates of the public's policy mood in each state. First, they use MRP to estimate state-level public opinion for each of the questions for which individual-level data are available that Stimson uses to estimate policy mood at the national level (Enns and Koch 2013; 2015).⁴ Then, for each state, they use Stimson's Dyad Ratios Algorithm, which uses a factor analytic approach to combine the questions into a measure of policy mood (Stimson 1999; 2004).⁵

We make three changes to the Enns and Koch measures. First, in recent years, Stimson has limited which types of questions he includes in his measure of policy mood. He now only includes survey questions that directly measure, "scope of government issues, items on business, labor, economic policy, healthcare, welfare, income equality, cities, education, taxes and the like" (Stimson 2018b). Survey questions about race, guns, and social and cultural issues, such as abortion, immigration, and gay marriage, are no longer included. Historically, the inclusion of these other policy questions did not have a substantial influence on Stimson's measure of

³Because the BRFH measure relies on interest group ratings of legislative votes in the U.S. Congress, anything beyond public opinion that influences federal and state legislators from the *same state* in the same direction would introduce bias, inflating the estimated relationship between public opinion and state policy. Unfortunately, there are many potential sources of this bias. For example, any organization that lobbies or donates at the federal and state level is trying to push legislative behavior in the same direction across both levels of government. Events in a state, such as an economic shock, natural disaster, or a political scandal could also push state and federal legislators from that state in the same direction. Because BRFH's measure is based on the behavior of members of Congress, any of these scenarios would bias the estimated relationship between public opinion and state policy upwards.

⁴MRP estimates a multilevel model of the probability of a particular survey response. Predicted responses are then estimated for each chosen demographic type (e.g., African American females, age 30–44, with some college education, in South Carolina) and these estimates are weighted (poststratified) to match population benchmarks. This approach allows the estimation of state-level public opinion from national surveys (Gelman and Little 1997; Lax and Phillips 2009; Pacheco 2014). To estimate the MRP models and the uncertainty around the resulting estimates, we use the R-package "autoMrp" (Broniecki, Leemann, and Wüest 2021).

⁵Stimson's Dyad Ratios Algorithm can be downloaded here: <http://stimson.web.unc.edu/software/>.

policy mood, and in fact, loaded on the second dimension. However, the increase in polling on cultural issues and the unique behavior of issues like racial attitudes (which have been trending more liberal) have led Stimson to focus on survey questions that are most theoretically related to the concept (Stimson 2018b). Since our goal is to generate measures of policy mood, we follow Stimson's lead and only retain the scope of government questions he uses in his most recent estimates.⁶

Second, we add data from 2011 to 2020. These 10 additional years of data include 169 questions from 75 different surveys for an additional 86,307 respondents. We were also able to add almost 70 additional surveys to Enns and Koch's previous estimates that have become available through the Roper Center for Public Opinion Research and through Gallup Analytics. Thus, the new estimates include a substantial increase in data and a narrower focus on the questions most directly related to policy mood.

Third, we have added state-level variables for the percent of African Americans and the percent of Democrats in each state to the MRP model. The estimates of state partisanship are based on more than 1.3 million individual responses.⁷ State-level variables can improve MRP estimates (Buttice and Highton 2017; Leemann and Wasserfallen 2020; Warshaw and Rodden 2012) and these two state-level variables are likely to influence policy preferences. Given the relationship between state racial/ethnic composition and state political environment (Hero and Tolbert 1996), an individual in a state with very few African Americans, like Idaho or Montana, is likely to hold different views than someone who shares the same demographics in a state with a high percentage of African Americans, like Mississippi or Georgia. Similarly, we might expect an individual in a majority Republican state to hold different policy preferences than an individual with the same demographic characteristics in a swing state or majority Democratic state.

Evaluating the New Measure of Policy Mood

To get an initial sense of our new measure (Lagodny et al. 2022), we generate a national-level estimate of policy mood based on the population-weighted average of the state estimates. This national estimate (based on our state-level estimates) allows us to compare our measure with Stimson's national policy mood. The two measures should be closely related, but they should not be identical, because our measure is limited to questions for which individual-level data are available, while Stimson's measure only requires survey marginals (i.e., percentages from topline reports). Thus, his measure includes more questions than we are able to include. Our comparison also includes a national measure based on the population-weighted average of Enns

⁶The one exception is questions that ask about spending on race, because these relate to the size of government.

⁷To estimate the percent of Democrats in each state, we estimated another MRP model using Democratic and Republican partisanship as the dependent variable. Because partisanship is relatively slow to change and to include as much data as possible in our state-level estimates, each year's estimate includes data from that year and the previous two years. We only rely on prior years so we do not introduce partisan information from future years into our final estimates. Because nearly all surveys ask about partisanship, our estimates of state partisanship include more data than our policy mood estimates—more than 1.3 million individual responses in total. The percent of African Americans in each state comes from the Census and the American Community Survey accessed through IPUMS (Ruggles et al. 2021).

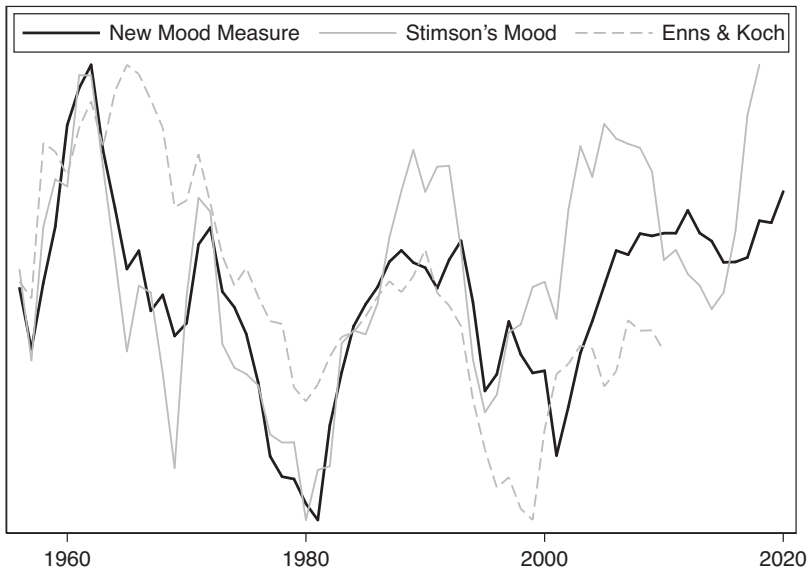


Figure 1. Comparison of Stimson's policy mood with our new measure and the Enns and Koch state-level measure. *Note.* Each state was weighted by its population size to generate a national-level estimate.

and Koch's measure. [Figure 1](#) presents these three series.⁸ As expected, they share substantial over-time variation, though our new measure appears to track Stimson's measure more closely.

The correlation between Stimson's policy mood and our updated measure (based on the state measures) is $r = 0.74$. The correlation of the first differences is $r = 0.59$.⁹ The relationship between our measure and Stimson's mood exceeds comparable correlations between the BRFH measure and Stimson's mood, offering further validation to our approach.¹⁰

To gain a further sense of the new estimates, we return to the analysis of southern states and examine the relationship between annual changes in the new measure and

⁸Of course, uncertainty exists around these estimates (As noted above, we use the R-package "autoMrP" [Broniecki, Leemann, and Wüest 2021] to estimate this uncertainty). Our measures are based on the percent offering a liberal response, while Stimson's estimates are based on the percent offering a liberal response out of those who offered a liberal or conservative response (i.e., middle categories are omitted from the denominator). The result is that Stimson's series is approximately 15% points more liberal on average, so the series in [Figure 1](#) have been rescaled to a common minimum and maximum value to allow over-time comparison.

⁹The Enns and Koch measure correlates with Stimson's series at $r = 0.29$, which is much lower than the correlation Enns and Koch (2013, 356) report. The reason for the difference is that Enns and Koch used a previous (2012) version of Stimson's mood for their comparison. As Section 5 of the Supplementary Material shows, while estimates of Stimson's policy mood from different years show remarkably similar over-time trajectories, the addition of new data does produce some year-to-year fluctuations across different releases of Stimson's policy mood.

¹⁰Berry *et al.* (2010) report a correlation of $r = 0.45$ between 1960 and 1999. Our series correlates with Stimson's at $r = 0.87$ during these years. Between 1976 and 1998, Berry *et al.* (2007) report a correlation of $r = 0.85$. Our measure correlates with Stimson's at $r = 0.90$ during these years.

changes in the GSS measure. The correlation is $r = 0.29$, slightly lower than that reported for the Enns and Koch measure in Table 1. This result shows that if studying the south or southern states, our new measure and the Enns and Koch measure are both preferable to BRFH's measure—at least if annual change is of interest. The bivariate correlation (in levels) among southern states between our new measure and the GSS measure is $r = 0.56$.

As a final assessment of our new measure, we examine the relationship with state presidential vote measured as the percent Democrat out of the two-party vote.¹¹ State presidential vote provides an advantageous validation check because we expect policy mood to correlate with presidential vote (Erikson, MacKuen, and Stimson 2002) and we observe presidential vote at the state level every four years. We analyze the estimated relationship between state public opinion and presidential vote share using three measures of state opinion—our new measure, the Enns and Koch measure, and Caughey and Warshaw's (2018a) measure of state economic liberalism.

The Caughey and Warshaw measure has excellent properties (Caughey and Warshaw 2018b), but differ from the Enns and Koch approach in several ways. First, while policy mood reflects preferences for more or less government activity, Caughey and Warshaw focus on generating an absolute measure of policy liberalism that explicitly avoids survey questions that include a relative frame, such as "Should federal spending on Child Care be increased, decreased or kept about the same?" These relative questions are a major contributor to Enns and Koch's policy mood. Caughey and Warshaw's measure also extends back to 1936 and includes more data. The estimation strategy is also different, as it relies on a dynamic group-level IRT model to estimate annual average liberalism in groups defined by state, race, and urban residence and then poststratifies the group estimates to match the groups' proportions in the state population (Caughey and Warshaw 2018a, 255). Due to endogeneity concerns, we do not include the BRFH measure in this analysis.¹²

Figure 2 reports the estimated relationship between the three opinion measures and state presidential vote.¹³ To facilitate comparison, all measures of state opinion were rescaled to a mean of zero and a standard deviation of one and the analyses all include the same years (full regression results appear in Section 3 of the Supplementary Material). We see that all three opinion measures correlate with presidential vote

¹¹The presidential vote data come from Dave Leip's Atlas of U.S. Presidential Elections.

¹²The BRFH measure weights the incumbent candidate ideology by the proportion of votes that candidate received and the challenger ideology by the proportion of votes the challenger received (since the ideology of the challenger is not directly observable, BRFH use the average ideology score of all incumbents from the state from the challenger's party as a proxy [Berry et al. 1998, 331]). Thus, the BRFH measure is a direct function of the proportion of votes received by the Democratic and Republican congressional candidates, meaning that an analysis of the relationship between presidential vote share and the BRFH measure would put vote share on both sides of the equation. The use of vote share in the BRFH measure would not be a problem if vote share at the congressional level was *not* correlated with presidential vote share. But this is not the case. A strong relationship exists between congressional vote and presidential vote, and this relationship has been getting stronger. In 2020, just 4% of congressional districts (16 out of 435) supported a House and presidential candidate from different parties (Nir 2020; Skelley 2021). This straight-ticket voting means it would be impossible to determine how much any relationship between the BRFH measure and state presidential vote results because BRFH use congressional vote share to calculate their measure.

¹³The models include the lagged dependent variable and state-fixed effects. A Levin, Lin, and Chu (2002) test rejects the null of a unit root.

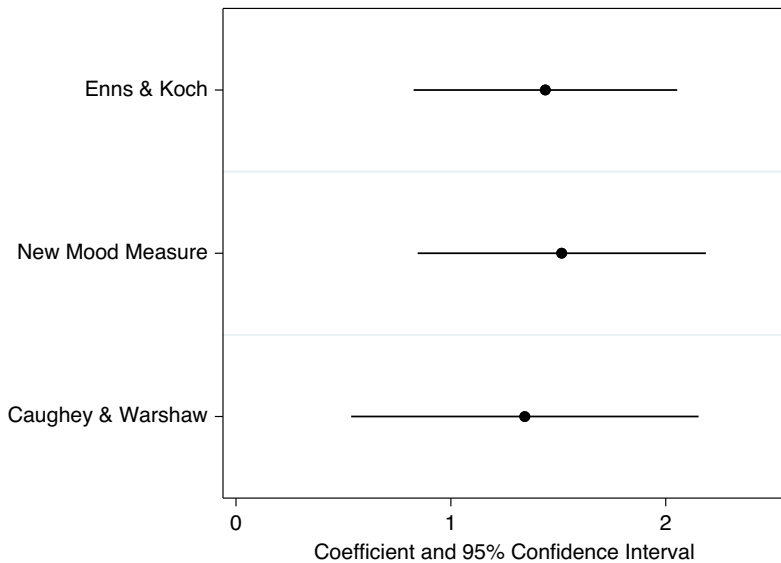


Figure 2. Comparison of the estimated over-time relationship between three measures of state public opinion and state presidential vote, 1956–2008.

to roughly the same extent. Focusing on our new measure, we would expect a standard deviation shift in state opinion to correspond with a about a 1.5% shift in presidential vote share in that state. (The Supplementary Material presents a similar analysis with similar conclusions using Caughey and Warshaw’s (2016) measure of state policy liberalism.)

A second question is whether our new measure of state policy mood provides additional information *beyond* including Stimson’s national measure of mood. To evaluate this question, we replicate the analysis of our new measure reported in Figure 2, adding Stimson’s national measure to the model.¹⁴ Column 1 of Table 2 shows that even when Stimson’s national measure is included in the model, the estimated relationship between our new state-level measures and state presidential vote share is substantial and statistically significant. In this model, a one standard deviation shift in state policy mood corresponds with more than a 1.5% shift in presidential vote share. The estimated relationship for Stimson’s national measure is close to zero (0.20) and not statistically significant. Column 2 repeats the analysis, this time only including Stimson’s national measure. The estimated coefficient is significant and close to one. Consistent with past research (Erikson, MacKuen, and Stimson 2002), Stimson’s policy mood is indeed related to presidential vote outcomes, but our state-level measure corresponds more closely with state presidential vote. While it is not surprising that a state-level opinion measure corresponds more closely with state votes than a national measure, this exercise further validates the utility of our measure.

¹⁴We use Stimson’s 2018 measure. Both variables are standardized to a mean of zero and standard deviation of one, and we analyze a presidential election years for which both measures are available.

Table 2. The estimated relationship between our new measure and state presidential vote, controlling for stimson's measure of national policy mood, 1956–2016

	Presidential vote	
	(1)	(2)
Presidential vote _{t-1}	0.19* (0.04)	0.22* (0.04)
New mood measure	1.66* (0.39)	
Stimson's national measure	0.20 (0.34)	0.92* (0.30)
Constant	38.43* (1.72)	37.42* (1.73)
N	748	748

Note. The state opinion measures were standardized to a mean of 0 and standard deviation of 1. Models include state-fixed effects.

* $p < 0.05$.

So far, we have focused on over-time validity. BFHC critique the cross-sectional patterns of the Enns and Koch measure, and argue that including state-fixed effects, which Enns and Koch recommend, is not sufficient. It is common practice in time series analysis to include such fixed effects—after all, the focus is often on isolating and understanding over-time relationships—so we find Enns and Koch's recommendation to be quite sensible. But, as noted above, we have updated the model to improve the cross-sectional behavior in our new estimates. Figure 3a evaluates these patterns by plotting the cross-sectional relationship between the three state opinion measures analyzed above and state presidential vote from 1956 to 2010 (the years we have data for all series). For most of the values of policy mood, the Caughey and Warshaw measure (top right panel) shows a slightly stronger relationship with presidential vote share than our new measure (top middle panel), though both appear to have superior cross-sectional properties than the Enns and Koch measure (top left panel).

The marker labels in the top middle panel of Figure 3 show that in 1964, support for Johnson, the Democratic presidential candidate, was much less in Alabama and Mississippi than we would expect given the relatively liberal policy mood in these states. This pattern reflects Republican Barry Goldwater's popularity in these states, which resulted, in part, because of his vote against the Civil Rights Act of 1964.¹⁵ We would not expect our measure of policy mood to capture racial attitudes. To further evaluate the possibility that support for Goldwater is influencing the observed relationship, the bottom panel of Figure 3 plots the same data from 1968 to 2010. As expected, the cross-sectional relationship between our new measure of policy mood and Democratic vote share is even stronger after 1964.¹⁶ Interestingly, during these

¹⁵Goldwater received 69.5% of the vote in Alabama and 87.1% of the vote in Mississippi.

¹⁶Although George Wallace proclaimed, "segregation now...segregation tomorrow...segregation forever" in his 1963 inaugural address as Governor of Alabama (<https://digital.archives.alabama.gov/digital/collection/voices/id/2952>), his popularity in Southern states in 1968 does not affect the analysis the way Goldwater's candidacy did. The difference results because Goldwater ran as a Republican in 1964 and Wallace ran as a third-party candidate in 1968. We analyze the percent of the Democratic vote share out of Democrat and Republican votes, which minimizes the influence of his third-party candidacy on our analysis.

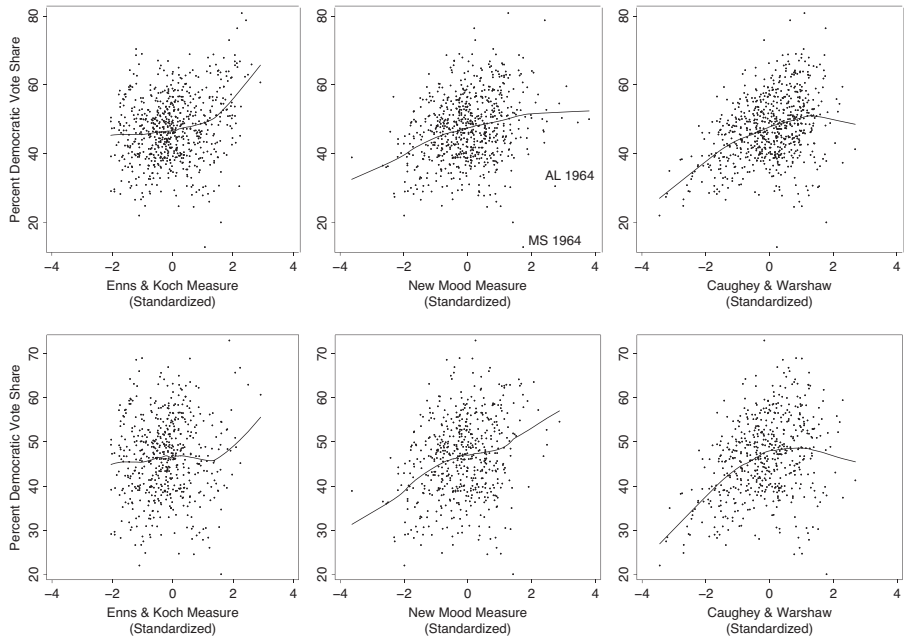


Figure 3. Comparison of the cross-sectional relationship between three measures of state public opinion and state presidential vote, 1956–2008 (top) and 1968–2008 (bottom). *Note.* The line represents the locally weighted regression with a bandwidth of 0.8. Democratic vote share is based on the two-party vote.

years the Enns and Koch measure is only positively related to vote share for the most liberal states and the Caughey and Warshaw measure is negatively related to Democratic vote share among states with the highest values of liberalism.

Conclusions and Recommendations

We began by reconsidering BFHC’s critique of the Enns and Koch (2013; 2015) measure of policy mood. When comparing the same states and examining annual change, BFHC’s analysis validates the Enns and Koch measure and calls the BRFH measure into question. Changes in the BRFH measure are *negatively* correlated with changes in the GSS measure they use. BRFH’s original article (Berry et al. 1998) has been cited nearly 2,000 times according to Google Scholar and researchers continue to utilize their measure of citizen preferences. Yet, even they have cast doubt on their measure. In 2015 (footnote 12, emphasis added) they wrote, “There is a weak correlation between the nationally aggregated BRFH measure and Stimson’s measure over periods including observations in years prior to 1974. This leads us to advise *against* using the BRFH measure in studies of periods extending before the mid-1970s” (Berry et al. 2015). They have also said, “If, at some point, a better direct annual measure of policy mood (based on survey responses reflecting attitudes about public policy issues or some other methodology not yet envisioned) is developed, we would favor using this measure over our less direct proxy” (Berry et al. 2007, 127).

Several direct measures of state public opinion now exist. While much of the previous analysis validates the Enns and Koch measure, we have found that our new measure, which builds on their methods, performs even better. We have also seen that the Caughey and Warshaw measure performs very well. In light of these results, scholars may wonder which measure they should utilize. Our new measure and the Caughey and Warshaw measure are moderately correlated ($r = 0.58$), but theoretically they are distinct. Our measure follows Stimson's approach as closely as possible in an attempt to directly measure the public's policy mood, while Caughey and Warshaw's measure is designed to measure absolute liberalism. For these reasons, the data and statistical models differ (Stimson 2018a). Ideally, theory will guide researchers' decisions. However, it may not always be clear whether a relative or absolute measure is more appropriate.

Of course, pragmatic constraints also matter. If the analysis focuses on years prior to 1956, the Caughey and Warshaw measures should be used (our new measures start in 1956). It may also be that in practice, the distinction does not matter (Stimson 2018a). In the analyses above (also see Supplementary Table A-3) the two measures yield very similar insights about the relationship between state public opinion, presidential vote, and policy. Instead of choosing between the two, the two measures may help scholars evaluate the robustness of their findings, allowing them to assess whether two alternate conceptions of state opinion lead to the same conclusions. In other words, the different measurement strategies and data may actually offer advantages to researchers.

There may also be analyses for which the Enns and Koch measure is a better fit. If the goal is to proxy Stimson's policy mood, our new measure is better, both in terms of empirical properties and length of time series. However, for some applications, the broader inclusion of questions by Enns and Koch could be advantageous. For example, analyses of judicial behavior (e.g., Owens and Wohlfarth 2017) might theoretically prefer a broader measure of state public opinion given the breadth of cases that are considered by the courts. By contrast, analyses of state vote or economic policy, which should theoretically align more closely with Stimson's policy mood, should utilize our updated measure.

An implicit assumption of this exchange is that the concept of national-level policy mood should have similar implications when analyzed at the state level. Several considerations support this conclusion. First, as our analyses highlight, presidential election outcomes reflect national policy mood (Erikson, MacKuen, and Stimson 2002), but due to the Electoral College, state opinion is what drives these outcomes. Thus state-level measures of policy mood can aid our understanding of national politics. For similar reasons, state policy mood can enhance understanding of Senate elections. Further, state policy mood, even if the measure includes questions that focus on the federal government, likely taps attitudes toward government at the state level. It would be surprising (though not impossible) for individuals to prefer more government spending at the state level and less at the federal level. As long as preferences for more or less government do not consistently differ at the state and national levels, our measure of national policy mood at the *state* level should offer a valid proxy for state policy mood. Consistent with this expectation, Section 4 of the Supplementary Material finds evidence of a relationship between our new measure of policy mood and state policy outputs.

Researchers should keep in mind, however, that if a particular analysis does *not* find a relationship between our new measure of state policy mood and policy

outcomes, it is impossible to know if the lack of relationship reflects states ignoring the public's will in that context or whether state policy preferences diverge from national policy preferences in that domain. For this reason, it would be preferable if sufficient data existed to generate state-level measures based on survey questions about *state* policy. The Roper Center for Public Opinion Research now includes nearly 11,000 survey questions from state-level surveys,¹⁷ so at least for some states such a measure may be possible. However, when the research requires analyzing all states over extended periods of time, we recommend our measure or the Caughey and Warshaw measure. If both measures support the same conclusion, researchers will have even more evidence of the robustness of their results.

Supplementary Materials. To view supplementary material for this article, please visit <http://doi.org/10.1017/spq.2021.26>.

Data Availability Statement. Replication materials (Lagodny *et al.* 2022) are available from the SPPQ Dataverse at <https://doi.org/10.15139/S3/QW8QGA>.

Funding Statement. The authors received no financial support for the research, authorship, and/or publication of this article.

Conflict of Interest. The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

References

- Banerjee, Anindya, Juan Dolado, John W. Galbraith, and David F. Hendry. 1993. *Co-Integration, Error Correction, and the Econometric Analysis of Non-Stationary Data*. Oxford: Oxford University Press.
- Berry, William D., Richard C. Fording, Evan J. Ringquist, Russell L. Hanson, and Carl E. Klarner. 2010. "Measuring Citizen and Government Ideology in the U.S. States: A Re-appraisal." *State Politics and Policy Quarterly* 10 (2): 117–35.
- Berry, William D., Evan J. Ringquist, Richard C. Fording, and Russell L. Hanson. 1998. "Measuring Citizen Ideology and Government Ideology in the American States, 1960–1993." *American Journal of Political Science* 42 (1): 327–48.
- Berry, William D., Evan J. Ringquist, Richard C. Fording, and Russell L. Hanson. 2007. "The Measurement and Stability of State Citizen Ideology." *State Politics and Policy Quarterly* 7 (2): 111–32.
- Berry, William D., Evan J. Ringquist, Richard C. Fording, and Russell L. Hanson. 2015. "Assessing the Validity of Enns and Koch's Measure of State Policy Mood." *State Politics and Policy Quarterly* 15 (4): 425–35.
- Brace, Paul, Kevin Arceneaux, Martin Johnson, and Stacy Ulbig. 2006. "Rejoinder to Berry, Ringquist, Fording, and Hanson 'Comment'." *Political Research Quarterly* 59 (4): 655.
- Broniecki, Philipp, Lucas Leemann, and Reto Wüest. 2021. "Improved Multilevel Regression with Post-Stratification Through Machine Learning (autoMrP)." *Journal of Politics* 84 (1): 597–601.
- Buttice, Matthew K., and Benjamin Highton. 2017. "How Does Multilevel Regression and Poststratification Perform with Conventional National Surveys?" *Political Analysis* 21 (4): 449–67.
- Carsey, Thomas M., and Jeffrey J. Harden. 2010. "New Measures of Partisanship, Ideology, and Policy Mood in the American States." *State Politics and Policy Quarterly* 10 (2): 136–56.
- Caughey, Devin, and Christopher Warshaw. 2015. "Dynamic Estimation of Latent Opinion Using a Hierarchical Group-Level IRT Model." *Political Analysis* 23 (2): 197–211.
- Caughey, Devin, and Christopher Warshaw. 2016. "The Dynamics of State Policy Liberalism, 1936–2014." *American Journal of Political Science* 60 (4): 899–913.

¹⁷<https://ropercenter.cornell.edu/ipoll/>.

- Caughey, Devin, and Christopher Warshaw. 2018a. "Policy Preferences and Policy Change: Dynamic Responsiveness in the American States, 1936-2014." *American Political Science Review* 112 (2): 249–66.
- Caughey, Devin, and Christopher Warshaw. 2018b. "Supplementary Information for "Policy Preferences and Policy Change: Dynamic Responsiveness in the American States, 1936–2014"." *American Political Science Review* 112 (2): A1–66.
- Enders, Walter. 2004. *Applied Econometric Time Series*, 2nd ed. Hoboken: John Wiley & Sons.
- Enns, Peter K. 2016. *Incarceration Nation: How the United States Became the Most Punitive Democracy in the World*. New York: Cambridge University Press.
- Enns, Peter K., and Julianna Koch. 2013. "Public Opinion in the U.S. States: 1956 to 2010." *State Politics and Policy Quarterly* 13 (3): 349–72.
- Enns, Peter K., and Julianna Koch. 2015. "State Policy Mood: The Importance of Over-Time Dynamics." *State Politics and Policy Quarterly* 15 (3): 436–46.
- Enns, Peter K., and Julius Lagodny. 2021a. "Forecasting the 2020 Electoral College Winner: The State Presidential Approval/State Economy Model." *PS: Political Science & Politics* 54 (1): 81–5.
- Enns, Peter K., and Julius Lagodny. 2021b. "Using Election Forecasts to Understand the Potential Influence of Campaigns, Media, and the Law in U.S. Presidential Elections." *University of Miami Law Review* 75 (2): 509–46.
- Enns, Peter K., and Christopher Wlezien. 2017. "Understanding Equation Balance in Time Series Regression." *The Political Methodologist* 24 (2): 2–12.
- Erikson, Robert S., Michael B. MacKuen, and James A. Stimson. 2002. *The Macro Polity*. New York: Cambridge University Press.
- Erikson, Robert S., Gerald C. Wright, and John P. McIver. 2007. "Measuring the Public's Ideological Preferences in the 50 States: Survey Responses versus Roll Call Data." *State Politics and Policy Quarterly* 7 (2): 141–51.
- Gelman, Andrew, and Thomas C. Little. 1997. "Poststratification into Many Categories Using Hierarchical Logistic Regression." *Survey Methodology* 23 (2): 127–35.
- Hero, Rodney E., and Caroline J. Tolbert. 1996. "A Racial/Ethnic Diversity Interpretation of Politics and Policy in the States of the U.S." *American Journal of Political Science* 40 (3): 851–71.
- Lagodny, Julius, Rebekah Jones, Julianna Koch, and Peter K. Enns. 2022. "Replication Data for: A Validation and Extension of State-Level Public Policy Mood: 1956 to 2020." UNC Dataverse. V2. <https://doi.org/10.15139/S3/QW8QGA>.
- Lax, Jeffrey R., and Justin H. Phillips. 2009. "How Should We Estimate Public Opinion in The States?" *American Journal of Political Science* 53 (1): 107–21.
- Leemann, Lucas, and Fabio Wasserfallen. 2020. "Measuring Attitudes - Multilevel Modeling with Post-Stratification (MrP)." In *The SAGE Handbook of Research Methods in Political Science and International Relations*, eds. L. Curini and R. Franzese, 371–84. Thousand Oaks: Sage.
- Levin, Andrew, Chien-Fu Lin, and Chia-Shang James Chu. 2002. "Unit Root Tests in Panel Data: Asymptotic and Finite-Sample Properties." *Journal of Econometrics* 108 (31): 1–24.
- Nir, David. 2020. "Daily Kos Elections' Presidential Results by Congressional District for 2020, 2016, and 2012." *Daily Kos*.
- Owens, Ryan J., and Patrick C. Wohlfarth. 2017. "Public Mood, Previous Electoral Experience, and Responsiveness Among Federal Circuit Court Judges." *American Politics Research* 45 (6): 1003–31.
- Pacheco, Julianna. 2011. "Using National Surveys to Measure Dynamic U.S. State Public Opinion: A Guideline for Scholars and an Application." *State Politics and Policy Quarterly* 11 (4):415–539.
- Pacheco, Julianna. 2014. "Measuring and Evaluating Changes in State Opinion Across Eight Issues." *American Politics Research* 42 (6): 986–1009.
- Ruggles, Steven, Sarah Flood, Sophia Foster, Ronald Goeken, Jose Pacas, Megan Schouweiler, and Matthew Mobek. 2021. "IPUMS USA: Version 11.0 [dataset]." Minneapolis, MN: IPUMS.
- Shirley, Kenneth E., and Andrew Gelman. 2015. "Hierarchical Models for Estimating State and Demographic Trends in US Death Penalty Public Opinion." *Journal of the Royal Statistical Society* 178 (1):1–28.
- Skelly, Geoffrey. 2021. "Why Only 16 Districts Voted for a Republican and a Democrat in 2020." *FiveThirtyEight*.
- Stimson, James A. 1999. *Public Opinion in America: Moods, Cycles, and Swings*, 2nd ed. Boulder, CO: Westview Press.

- Stimson, James A. 2004. *Tides of Consent: How Public Opinion Shapes American Politics*. New York: Cambridge University Press.
- Stimson, James A. 2018a. "The Dyad Ratios Algorithm for Estimating Latent Public Opinion: Estimation, Testing, and Comparison to Other Approaches." *Bulletin de Méthodologie Sociologique* 137–138: 201–18.
- Stimson, James A. 2018b. "Estimation Notes for Mood (1952–2018)." *Unpublished Manuscript*.
- Warshaw, Christopher, and Jonathan Rodden. 2012. "How Should We Measure District-Level Public Opinion on Individual Issues?" *Journal of Politics* 74 (1): 203–19.

Author Biography. Julius Lagodny is a Research Fellow at the Hertie School Data Science Lab. His research focuses on public opinion, election forecasting, and political behavior in Europe and the US.

Rebekah Jones is a PhD student in Political Science at the University of California, Berkeley.

Julianna Koch is a Senior Research Director at buzzback. She specializes in primary market research in the healthcare industry.

Peter K. Enns is a Professor in the Department of Government and the Brooks School of Public Policy at Cornell University and the Robert S. Harrison Director of the Cornell Center for Social Sciences. He is also Co-founder and Chief Data Scientist at Verasight.

Cite this article: Lagodny, Julius, Rebekah Jones, Julianna Koch, and Peter K. Enns. 2022. A Validation and Extension of State-Level Public Policy Mood: 1956–2020. *State Politics & Policy Quarterly* 1–14, doi:10.1017/spq.2021.26