Measuring Executive Ideology and its Influence

Seth Babson Warner
Penn State University

Executives are important elites, and ideology is important to elite behavior, but measurement challenges and a focus on the presidency have kept scholars from fully exploring executive ideology. This article proposes that students of executive politics may be well-served by turning to the gubernatorial context, where campaign finance-based ideology scores (Bonica 2014) provide a common-scale measure of executive preferences. These scores, or CFscores for short, are drawn from 103 million campaign contributions based on patterns of "who gives to whom?" A series of validation exercises show that CFscores provide valuable information about executive preferences, even among governors of the same party. I then use these scores to assess the centrality of executives in the policymaking process. Four models explain state policy liberalism as a function of executive, legislative, and citizen ideology. Gubernatorial preferences emerge as the most predictive of the three. Executive ideology is more explanatory than legislative preferences when Democrats are in office and dwarfs the predictive power of public opinion in all cases. A one standard deviation shift in executive ideology corresponds with 2.3 to 8.3 times more policy change than a similar shift in public opinion.

Ideology has long been understood to play an important role in executive behavior. Spatial models of legislative policymaking hold a large role for executive preferences (Krehbiel 1998), and at both the state and federal levels, there is reason to believe executives wield more influence still over the bureaucracy (Wood and Waterman 1991, Woods and Baranowski 2007, Clinton et al. 2014).

However, two realities have kept scholars from exploring the effect of executive ideology fully. Measurement challenges are the primary obstacle, rooted in the fact that executives make distinct decisions from other actors. They do not cast votes, for example, on the same bills as legislators do, decide on the same cases as judges do, or answer the same survey questions as constituents do. This makes placing them on a common ideological scale difficult. Recognizing this problem, a handful of scholars have proposed workarounds (Zupan 1992, Poole and Rosenthal 1995, Bailey 2007, Treier 2010). Nevertheless, the literature on executive ideology remains thin compared to parallel literatures on judicial and legislative preferences, and most studies omit ideology as a predictor of executive behavior, instead focusing on differences by party (see Barrilleaux and Berkman 2003, Leigh 2008, Kelly and Witko 2014). A governor's party may be a reasonable proxy, but it assumes away within-party ideological differences, which in turn may leave scholars with an incomplete picture of executive behavior.

The secondary reason for our lack of knowledge stems from the discipline's greater historical emphasis on the presidency as opposed to subnational executives (Squire and Fastnow 1994, Klarner and Karch 2008). While the president's power and profile may justify this emphasis, the fact that only one president serves at a time makes it difficult to counterfactually estimate the effect of ideology on executive behavior. As Kousser and Phillips (2012, 6) have previously argued, studying governors is an effective way of learning about executives in

general. Their offices are similar enough in nature to facilitate cross-sectional analysis, and high enough in profile that data on their behaviors and governing contexts are accessible to researchers. These features allow us to examine the correlates of executive behavior, and the conditions and issue domains under which executive power is greatest.

Bonica's (2014) campaign finance-based method of measuring ideology is a promising advance that could allow studies of gubernatorial behavior to account for their ideology. Yet it is an open question as to whether these ideology scores, or CFscores, are valid when applied to US governors. Derived from patterns of "who donates to whom," CFscores are unique in their ability to place governors on the same scale as one another. In the same way, they allow direct comparisons between governors and other pertinent actors, such as state legislators, elected judges, and local officials. To preview the results of my validation study, the scores prove to be quite useful as indicators of executive preferences. Although there is some noise in the measures, CFscores capture meaningful differences among governors of the same party, which is a high bar for measures of ideology in a polarized era (Tausonovitch and Warshaw 2018). I then use these scores to gauge the relative impact of executive ideology on policy, relative to the preferences of legislators and citizens. A preliminary exploration shows that, even when holding partisanship constant, a one standard deviation shift in gubernatorial ideology corresponds with 2.3 to 8.3 times more policy change than a similar shift in public opinion.

The article proceeds as follows. I begin by providing a description of CFscores, how they are calculated, and why they may or may not be useful indicators of ideology for state-level officials. I then conduct eight validation exercises, showing that CFscores converge with other measures of ideology, and predict relevant gubernatorial actions. Finally, I present models that explain state policy outcomes as a function of gubernatorial, legislative, and citizen ideology,

and find that even within parties, the preferences of the governor are highly determinative. These results suggest a large, independent role in policymaking for governors, and should encourage further research into their role in the policy process.

Measuring Ideology with CFscores

Ideology plays a central role in formal models of political behavior and policymaking (e.g., Arrow 1951, Downs 1957, Krehbiel 1998). Many of these models, however, require measurements of ideology for groups or actors in different parts of the political system. Whereas measures have been created for actors as disparate as legislators (McCarty, Poole, and Rosenthal 2001), judges (Bailey 2013), and the public at large (Berry et al. 1998, Enns and Koch 2013), scholars have had difficulty creating measures that provide information about many different sets of actors on the same scale.

Bonica's (2014) campaign finance-based ideology scores, or CFscores, provide a partial yet promising solution to this problem. CFscores draw on 103 million contributions to state and federal candidates as expressions of ideological preference, and use them to estimate ideology scores for politicians and donors alike based on patterns of "who donates to whom?" Because contributors often donate to candidates in different districts or at different levels of government, these scores theoretically place politicians across a variety of offices on the same ideological scale. Already, scholars have used CFscores in the gubernatorial context to explain how governors advertise themselves during elections (Fowler et al. 2019) and fill legislative vacancies (Cooper et al. 2018).

The construction of CFscores involved multiple rounds of scaling based on matrices that indicated how much donor i contributed to candidate j. The first scaling, which anchored subsequent steps, focused only on federal candidates, which raises the question of whether CFscores are as informative for governors as they are for members of Congress. For his part, Bonica (2014) demonstrated that CFscores reliably predict members of Congress' roll call votes across a variety of contexts, such as district competitiveness, member tenure and leadership status, and legislative chamber. In addition, Bonica and Woodruff (2015) also show that donors' CFscores do not change significantly when their judicial contributions are excluded, suggesting that donor behavior is consistent across offices. On the other hand, very few donors contribute to state-level races in multiple states, which led Bonica to scale ideal points for state-level politicians in relation to those of federal candidates from the same state.

In this light, the validity of CFscores applied to governors is an open question. The success of CFscores for other offices provides grounds for optimism. But the unique features of the governor office in the US and the relatively small number of donors to multiple governors raises questions. To address this uncertainty, the first part of this paper puts CFscores to the test, scrutinizing their validity from multiple perspectives.

A topline analysis of gubernatorial CFscores

I begin my validation study by creating a dataset that takes as its units the 223 individuals who served as governor of a U.S. state in the period between 1991 and 2013. This is the timeframe that overlaps with Bonica's (2014) data on contributions to state-level candidates. Then, I set to work pairing these governors with their CFscores.

Most governors in this timeframe have CFscores, but not all do. There are two reasons for this imperfect coverage. First, it was not until 2001 that all states required that candidates for state and local offices disclose their contributions. Without these disclosures, no CFscores are possible. Second, in some cases, governors are not elected to the office, but rather ascend to it after their predecessor dies or resigns. If those individuals do not then seek reelection, they are not assigned a CFscore. In addition, it is feasible that a fully self-funded candidate would not have a CFscore, but none exist among the governors I researched.

Table 1 presents the number of governors who fall into each of these groups, and the number that remain in my dataset after the former are omitted. To check that there are no systemic differences between governors with and without CFscores, I conducted a series of t-tests that can be found in the appendix. I find that the groups are the same in terms of partisanship, gender, and race. The only minor difference that exists is by region: governors from the American West are somewhat more likely to have CFscores than those from elsewhere.

Table 1. Availability of CFscores for Governors (1991 to 2013)

|	
Governors with CFscores	152
Disclosure not required	59
Ascended to office [1]	12
Fully self-funded	0
All individuals who served	**223**

[1] And did not run for reelection

CFscores for elites are scaled from -2 (most liberal) to +2 (most conservative), but the range observed among governors is narrower, from -1.45 (Howard Dean of Vermont) to +1.35

(Jack Dalrymple of North Dakota). The distribution of governors between these extremes is presented in the histogram in Figure 1. Blue bars indicate the number of Democratic governors whose CFscores fall in each interval of 0.25. Red bars indicate the number of Republican governors whose CFscores fall in each interval, and the yellow bars do the same for the four independent governors in this dataset.[1]

Looking at Figure 1, one notices that although ideology and partisanship correlate, they are not one and the same. The parties' modal intervals are equidistant from the centerpoint of 0. The largest number of Democrats have scores between -0.75 and -1.0, and the largest number of Republicans have scores between 0.75 and 1.0. However, many governors have CFscores well to the left or right of their party's modal interval. Finally, we observe some ideological overlap between the parties. Eight Republicans have more liberal scores than the most conservative Democrat, and 13 Democrats are more conservative than the most liberal Republican. These include GOP governors of traditionally Democratic states, such as George Pataki of New York and M. Jodi Rell of Connecticut, and Democratic governors of more conservative states, such as Kathleen Blanco of Louisiana and Phil Bresden of Tennessee.

---

[1] Governors are coded based on their party when they first ran for governor. As such, these four include Wally Hickel (Alaska) and Lincoln Chafee (Rhode Island), who ran as independents but registered with a major party while in office. The other independent governors are Angus King (Maine) and Jesse Ventura (Minnesota).

Figure 1. CFscores of US Governors by Party, 1991 to 2013
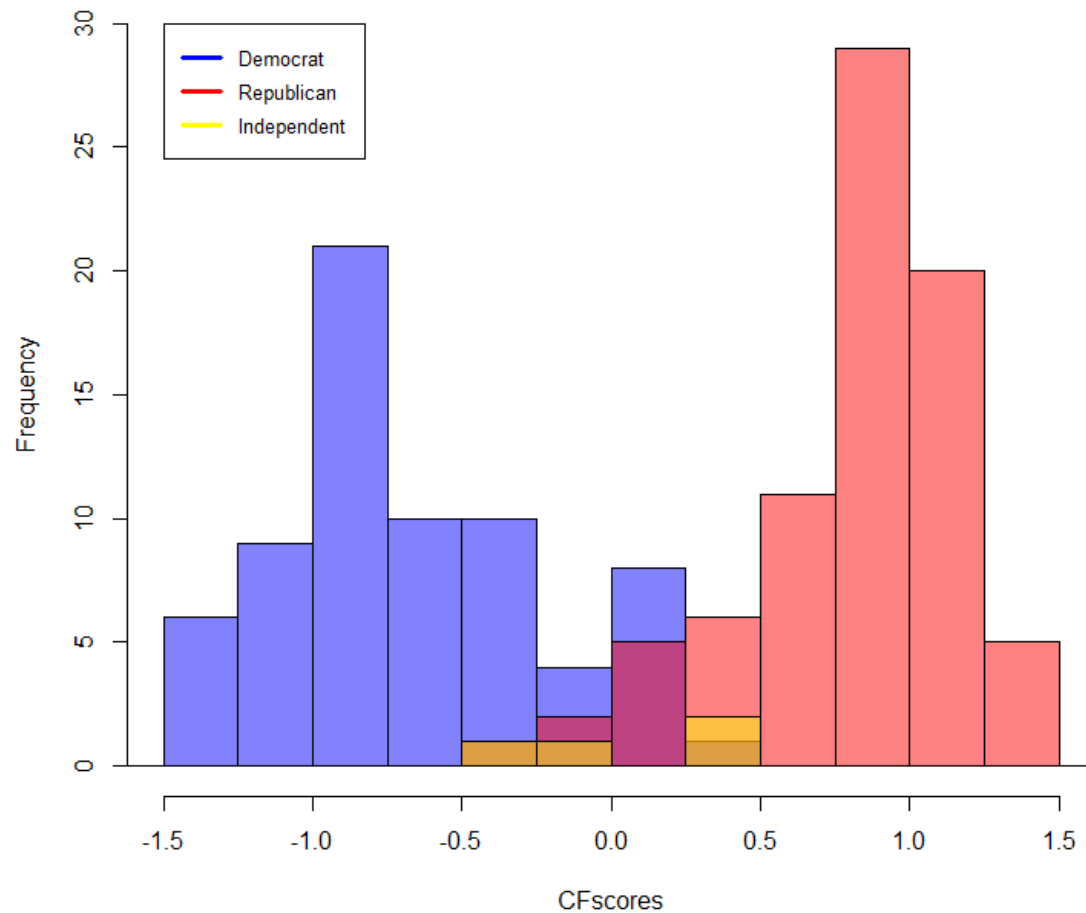
The question now is whether this observed variation is a valid reflection of differences in gubernatorial ideology. I respond to this question assessing the convergent and predictive validity of CFscores.

Do CFscores correlate with other measures of ideology?

One method of validating a measurement is to compare it with others that purport to measure the same concept. If the measures correlate, they provide evidence to one another that

the same concept is indeed being quantified. This is known as "convergent validation" (Campbell and Fiske 1959). Here, I assess convergent validity by calculating the extent to which CFscores correlate with other measures of ideology. These other measures are voter placements and legislative roll-call ideology.

A governor's voter placement is a measure derived from survey responses, in which citizens are asked to place a politician on an ideological scale. Scholars have found that these placements, when averaged together, come very close to measurements based on observed behavior (Erikson and Romero 1990). In 2006, 2010, 2012, and 2016, respondents to the Cooperative Congressional Election Study survey were asked to rate their governor on an ideological scale. When combined, their responses provide ideological estimates for 102 governors in my dataset.

The scale used by the CCES varies by year. The 2006 CCES asked respondents to place their governor and gubernatorial candidates on a scale of 0 (most liberal) to 100 (most conservative).[2] Later surveys asked respondents to place their governor on a seven-point ideology scale.[3] For my analysis, I therefore separate placements from the 2006 CCES from those taken in 2010, 2012, and 2016. To ensure that state-level samples are large enough to meaningfully estimate ideology, I calculated the standard error of each estimate. In the 2006 survey, the largest standard error was 3.5 and the mean estimate was 61.7. In later surveys, the largest standard error was 0.15 and the mean estimate was 4.30.

---

[2] For governors first elected in 2006, I use their candidate ratings in my validation exercise.
[3] The options are: (1) very liberal, (2) liberal, (3) somewhat liberal, (4) moderate, (5) somewhat conservative, (6) conservative, and (7) very conservative.

In addition to voter placements, I also draw on two measures of legislative ideology. Fifty-nine governors in my dataset served in either Congress, a state legislature, or both before entering the governor's mansion and were assigned either a NOMINATE score (Poole and Rosenthal 2011) based on their Congressional roll-call votes, or an NPAT score (Shor and McCarty 2011) based on their state legislative votes. Even though these estimates are based on politicians' behavior *before* they became governor, there is reason to believe their ideology should remain consistent over time. First, the ambition theory holds that politicians "act today in terms of the electorate [they] hope to win tomorrow" (Schlesinger 1966, Berkman 1994). Second, scholars have observed consistency in elite ideology; politicians are said to "die with their ideological boots on" (Poole 2007).

To compare these measures to CFscores, I calculate the Pearson's correlation coefficient between them, and the corresponding 95-percent intervals. The results of these exercises are presented in Table 2. The headings indicate the source and data used in each of the four exercises. The first row indicates the Pearson's *r* when calculated among all governors for whom data were available. We can see that the average correlation appears to be around $r = 0.9$, and even the smallest lower-bound estimate is 0.77. This indicates that CFscores distinguish between liberal and conservative governors very well.

The second row indicate the Pearson's *r* coefficients among only Democratic governors. The average correlation is $r = 0.58$, and despite the small sample sizes, three of the four measures significantly associate with CFscores. These results indicate that CFscores do a good job of distinguishing between moderates and extremists among Democratic governors. The third row focuses on Republican governors. This set of results is the weakest, with an average correlation

of $r = 0.43$, but still strong enough to say that CFscores capture some of the heterogeneity among these executives.[4]

Table 2. Pearson's correlations between governors' CFscores and alternative ideological measurements

	DWNOMINATE (Poole and Rosenthal 2007)	N	NPAT State Legislative (McCarty and Shor 2011)	N	Voter placements (CCES 2006)	N	Voter placements (CCES 2010-16)	N
All	0.92* [0.84, 0.96]	32	0.88* [0.77, 0.94]	31	0.93* [0.89, 0.96]	59	0.91* [0.87, 0.94]	75
Democrats	0.42 [-0.17, 0.79]	13	0.70* [0.22, 0.91]	13	0.60* [0.29, 0.80]	29	0.59* [0.30, 0.78]	32
Republicans	0.33 [-0.16, 0.69]	19	0.44^ [-0.03, 0.75]	18	0.73* [0.51, 0.87]	30	0.22 [-0.09, 0.49]	42

Notes: ^ $p < 0.10$, * $p < 0.05$. Brackets contain 95 percent confidence intervals.

Do CFscores predict gubernatorial behavior?

Another method of validating a measure is to determine whether it predicts future behavior in the way we would expect (Cronbach and Meehl 1955). CFscores are based on donation patterns that occur before a politician has assumed office.[5] Do they predict the politician's behavior once they become governor?

---

[4] It is worth noting also that lower correlations can be due to a low signal-to-noise ratio in *either* X *or* Y. A part of these weak results appears to be related to the greater uniformity in citizens' ideological placement of Republicans during the 2010s. The 90% most conservative governors received a mean score of between 5.2 and 6.1, while the similar range for Democrats was more than 50% wider (2.4 to 3.8).

[5] Bonica (2014) does not calculate new CFscores based on donations to gubernatorial reelection campaigns.

To answer this question, I draw upon two types of gubernatorial behavior. The first come from content analyses of gubernatorial State of the State (SOTS) speeches. Given at the beginning of each year to the state legislature, governors use these addresses to lay out their policy and fiscal agenda. Research has shown that governors' SOTS proposals tend to be pursued in earnest—and often successfully (Kousser and Phillips 2012).

Two scholarly treatments have analyzed SOTS speeches for their proportions of liberal to conservative proposals. Coffey (2005) analyzed 93 speeches given in 2000 and 2001, coding each proposal as a -1 if it was liberal or a 1 if it was conservative. The mean proposal for each speech represented its policy score, and I take the average score for each governor as a measure of their behavior. Kousser and Phillips (2012) coded speeches given by 48 governors in the years 2001 and 2006. I use their dataset of SOTS proposals to create a policy score similar to Coffey's (2005).[6]

Table 3 presents the Pearson's correlations between CFscores and SOTS policy scores for data from each source. The second row shows that, based on Coffey's (2005) scores, a governor's CFscore and the average ideological direction of their policy proposals correlates at $r = 0.50$. If we compare CFscores to proposals collected by Kousser and Phillips (2010), the correlation is similar, at $r = 0.42$. These results indicate that CFscores are reasonably predictive of the types of policies a governor will propose.

---

[6] Kousser and Phillips (2010) code each proposal as liberal, neutral, or conservative, and provide a score of 1 to 5 to indicate the proposal's significance. I subset the data to only include proposals of significance 3 or greater. I then assign values of -1, 0, and 1 to liberal, neutral, and conservative proposals respectively. Finally, I take the mean value for each governor as his or her policy score.

**Table 3. Pearson's correlations between CFscores
and State of the State policy scores**

	As calculated by Coffey (2005)	As calculated by Kousser and Phillips (2010)
	0.50*	0.42*
	[0.24, 0.69]	[0.15, 0.62]
N	45	49
Years	2000, 2001	2001, 2006

Notes: * $p < 0.05$. Brackets contain 95 percent confidence intervals. Policy score derived from mean of speech proposals, coded as -1 (liberal), 1 (conservative), and 0 (neutral).

My second source of data examines CFscores in comparison to a governor's actions on LGBT rights. Sellers' (2017) examined the conditions under which governors signed laws or orders to protect the LGBT community from employment discrimination.[7] His dataset provides an indication of gubernatorial action or inaction on the issue for each state-year from 1971 to 2013.

I compare CFscores to a dummy variable for whether (1) or not (0) a governor signed LGBT protections into law, given that none were in place when they took office. Sellers (2017) breaks these protections into two types, both of which protect citizens from employment discrimination. The first protect individuals on the basis of sexual orientation (LGB citizens), and the second protect them on the basis of gender identity (trans citizens). These issues provide useful examples for validating ideology because they confound partisan divides to a degree. A

---

[7] Governors generally have the power to enact these protections via executive order. Many of the first such protections came from gubernatorial decree. However, all else equal, executives prefer to make these changes via legislation (Mayer 2009).

large minority of Democrats declined to protect LGBT rights, while a small handful of Republicans and independents did so against partisan expectation.

Table 4 presents correlations between governors' CFscores and their actions to protect these groups. Because higher CFscores indicate greater conservatism, it is as expected that the correlations are negative. On the left side, we see that the correlation between CFscores and gubernatorial action to protect LGB citizens is very strong, at $r = -0.68$. On the right side, the correlation between CFscores and protections on the basis of gender identity is weaker, but still significant at $r = -0.44$.

Table 4. Pearson's correlations between CFscores and gubernatorial action against LGBT discrimination (1991 – 2013)

	Sexual Orientation	Gender identity
	-0.68*	-0.44*
	[-0.55, 0.79]	[-0.29, -0.57]
N	79	132

Notes: * $p < 0.05$. Brackets contain 95 percent confidence intervals. Higher CFscores indicate greater conservatism. Action coded as 1 if governor signed law or order banning employment discrimination, 0 if not.

Is executive ideology important for understanding policy?

In the previous section, I concluded my validation of CFscores by showing that they predict gubernatorial behavior in the form of policy proposals and (in)action on social issues under the governor's purview. However, does ideology have any effect when the role of partisanship is controlled for?

Spatial models of interbranch bargaining (e.g., Krehbiel 1998) understand policy as a function of the distance between actors' ideological preferences. But in Congress and state legislatures, the ideological polarization of elites and the decline of bipartisan lawmaking mean that partisanship can account for much of the relationship between ideology, legislative behavior, and policy outcomes they produce (Warshaw and Tausonovitch 2017). If the same is true among executives, scholarly attention to their ideological preferences may be more peripheral than essential to our understanding of modern policymaking. By contrast, if executive ideology has a strong, independent effect on policy apart from partisanship, then it should receive more attention from scholars than it has.

To test these possibilities, I turn to Caughey and Warshaw's (2016) measures of state policy liberalism from 1936 to 2014. Their scores are based on a scaling of 148 policies, separated into economic and social policy domains. As the authors note, year-over-year changes in policy are generally very small—and in all but two states, governors serve four-year terms—so I limit my analysis to even-numbered years between 1991 and 2013 for which data on gubernatorial, state legislative, and citizen ideology are available.[8] To assist with interpretability, I then standardize all policy and ideology scores to have a mean of 0 and standard deviation of 1, and reverse measures of gubernatorial and legislative ideology so that positive values indicate liberalism across all variables.

---

[8] Data on state legislative ideology are again taken from McCarty and Shor (2011), and I use the mean of chamber medians to represent lawmaker preferences. Measures of public ideology, broken into social and economic issue domains, are taken from Caughey and Warshaw's (2018) analysis of public opinion at the state level.

The result is a panelized dataset reflecting the policy and political contexts in 50 states. Because of missing data, the panel is unbalanced. Each state is represented by between three and nine timepoints, and the median has seven.

I estimate four models using random effects.[9] Each model represents a different pairing of policy type (economic or social) and gubernatorial partisanship (Democrat or Republican). Separating the models out by partisanship in this way allows me to assess the independent effect of executive ideology relative to that of the legislative branch and the mass public.

The results of my analyses are presented in Table 5. In all four models, executive ideology explains significant variation in public policy across the 50 states in this analysis. More importantly, executive preferences seem to dominate—at least in some contexts—those of the legislature and mass public. Across all four models, the governor's ideology is estimated to have between 2.3 and 8.3 times greater effect on policy than that of the mass public. Executive dominance over the legislature is more conditional. When Republicans are in office, executive and legislative ideology appear to wield roughly equal influence. But when Democrats control the governor's mansion, a one-standard deviation shift in executive ideology associates with twice as much policy change as a similar shift in the legislature.

---

[9] Analysts have a few options when specifying panel data. Here, a pooled OLS may also be appropriate. When the data are specified as such, the Breusch-Pagan test finds no heteroskedasticity in two models, and only marginal heteroskedasticity in the others. The substantive results of these models are the same. I use a random effects model to account for some but not all of the variation across units. According to Clark and Linzer (2015), this is more appropriate than a fixed-effects model, because for all X, the within-unit variation is much greater than for Y. In these cases, the bias incurred by not specifying separate intervals for all units is expected to be minimal.

**Table 5. The effects of executive, legislative and citizen ideology
on social and economic policy**

	Democratic Governors		Republican Governors	
	(1) Social Policy	(2) Economic Policy	(3) Social Policy	(4) Economic Policy
Executive Ideology	0.28* [0.08]	0.43* [0.09]	0.17* [0.10]	0.25* [0.11]
Legislative Ideology	0.13* [0.04]	0.16* [0.04]	0.28* [0.05]	0.22* [0.05]
Citizen Ideology	0.12* [0.04]	0.09* [0.04]	0.06 [0.05]	0.03 [0.04]
Constant	-0.25* [0.11]	-0.33* [0.12]	0.08 [0.13]	0.10 [0.13]
N (n)	143 (40)		174 (41)	
Adj. R^2	0.31	0.24	0.19	0.11

Notes: * $p < 0.05$, one-tailed. Brackets contain standard errors. Policy and ideology scores standardized to mean = 0 and standard deviation = 1, with positive scores indicating greater liberalism. Model estimated with random effects. Data represent policy, politics, and public opinion in 50 states, measured biannually from 1991 to 2013.

What do these models tell us, and what don't they? To the former, for the scores of

studies that have been published on measuring the effects of legislative and citizen ideology on

policy, there is at least topline evidence that the executive's preferences carry greater weight. But

to the latter, it is important to clearly state these models' limitations. Namely, the models are not

intended to represent a unified theory of state policy liberalism. They do not include other

forces—such as interest groups, institutional features, or citizen demographics—that are known

to a role in policy outcomes. Future studies may wish to incorporate some of these factors to see

how they affect the significance of executive ideology. If the governor's influence wanes when

other factors are considered, then it may provide evidence for interesting narratives about how

executives are constrained by constituent groups or coalition partners. If executive ideology

remains a strong force, or even zaps the significance of other factors, then the literature should devote greater attention yet to the role of executives over the policy process.

Conclusion

This study advocates for greater attention to the role of executives generally, and to their ideological preferences in particular, as a determinant of key policy and political outcomes. Previous efforts to study executive ideology have been limited by the challenge of measuring executive preferences, and by the fact that much of the literature focuses on presidents, who do not serve concurrently and of whom there are relatively few, making counterfactual and quantitative analyses difficult. To address these obstacles, I explore the potential for using CFscores in the gubernatorial context to learn more about how executive's preferences shape policy. A series of validations confirm that, while imperfect, CFscores are useful indicators of executive ideology. They meet *prima facie* assumptions (e.g., that Democrats are more liberal than independents, who are more liberal than Republicans), correlate with other items, and predict behavior. Then, I provide a preliminary comparison of how executive preferences compare to those of the legislature and the mass public in terms of predicting policy. Across four models, executives appeared to wield greatest influence relative to the other groups.

Future work on this subject may take two tracks. The first, proposed in the previous section, is to incorporate additional factors into models of policymaking to examine the conditions under which executive preferences wax or wane. Scholars may also wish to use CFscores to create more sophisticated models of interbranch relations, seeing as the scores provide measurements of legislative and judicial preferences as well. The second, which relate to

the limitations of CFscores, may seek to explore why they meet expectations or predict behavior better for some governors than others. Given the scores' limitations, there is also value in the development of additional measures. An expansion of State of the State-based policy scores (Coffey 2006, Kousser and Phillips 2010) can provide an alternative method of measuring executive preferences that is available for all governors over a long stretch of time.

Taken as a whole, this piece intends to facilitate further research into the executive preferences and power. Given their apparent impact and the growing availability of data regarding executives, it seems likely that political science will experience a windfall of insights about these politicians in the near future.

Works Cited

Arrow, K. J. (1951). *Social choice and individual values*. Yale university press.

Bailey, M. A. (2007). Comparable preference estimates across time and institutions for the court, congress, and presidency. *American Journal of Political Science*, *51*(3), 433-448.

Bailey, M. A. (2013). Is today's court the most conservative in sixty years? Challenges and opportunities in measuring judicial preferences. *The Journal of Politics*, *75*(3), 821-834.

Barrilleaux, C., & Berkman, M. (2003). Do governors matter? Budgeting rules and the politics of state policymaking. *Political Research Quarterly*, *56*(4), 409-417.

Bell, A., & Jones, K. (2015). Explaining fixed effects: Random effects modeling of time-series cross-sectional and panel data. *Political Science Research and Methods*, *3*(1), 133-153.

Berkman, M. B. (1994). *The state roots of national politics: Congress and the tax agenda, 1978–1986*. University of Pittsburgh Press.

Berry, W. D., Ringquist, E. J., Fording, R. C., & Hanson, R. L. (1998). Measuring citizen and government ideology in the American states, 1960-93. *American Journal of Political Science*, 327-348.

Bonica, A. (2014). Mapping the ideological marketplace. *American Journal of Political Science*, *58*(2), 367-386.

Bonica, A., & Woodruff, M. J. (2015). A common-space measure of state supreme court ideology. *The Journal of Law, Economics, and Organization*, *31*(3), 472-498.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological bulletin*, *56*(2), 81.

Caughey, D., & Warshaw, C. (2016). The dynamics of state policy liberalism, 1936–2014. *American Journal of Political Science*, *60*(4), 899-913.

Caughey, D., & Warshaw, C. (2018). Policy preferences and policy change: Dynamic responsiveness in the American states, 1936–2014.

Clark, T. S., & Linzer, D. A. (2015). Should I use fixed or random effects. *Political science research and methods*, *3*(2), 399-408.

Clinton, J. D., Lewis, D. E., & Selin, J. L. (2014). Influencing the bureaucracy: The irony of congressional oversight. *American Journal of Political Science*, *58*(2), 387-401.

Coffey, D. (2005). Measuring gubernatorial ideology: A content analysis of state of the state speeches. *State Politics & Policy Quarterly*, *5*(1), 88-103.

Cooper, C. A., Knotts, H. G., & Ragusa, J. (2016). The Constrained Governor: Exploring Gubernatorial Decision Making on Senate Appointments. *Political Research Quarterly*, *69*(3), 482-494.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological bulletin*, *52*(4), 281.

Dieleman, J. L., & Templin, T. (2014). Random-effects, fixed-effects and the within-between specification for clustered data in observational health studies: a simulation study. *PloS one*, *9*(10), e110257.

Downs, Anthony. "An economic theory of political action in a democracy." *Journal of political economy* 65.2 (1957): 135-150.

Enns, P. K., & Koch, J. (2013). Public Opinion in the US States: 1956 to 2010. *State Politics & Policy Quarterly*, *13*(3), 349-372.

Erikson, R. S., & Romero, D. W. (1990). Candidate equilibrium and the behavioral model of the vote. *The American Political Science Review*, 1103-1126.

Esarey, J., & Jaffe, J. (2017). A Direct Test for Consistency of Random Effects Models that Outperforms the Hausman Test. Working Paper.

Fowler, E. F., Franz, M. M., Martin, G. J., Peskowitz, Z., & Ridout, T. N. (2019). Political Advertising Online and Offline. In *Annual Meeting of the American Political Science Association Conference, August* (Vol. 29).

Grammich, C., Hadaway, K., Houseal, R., Jones, D. E., Krindatch, A., Stanley, R., & Taylor, R. H. (2019, February 10). Longitudinal Religious Congregations and Membership File, 1980-2010 (State Level).

Kelly, N. J., & Witko, C. (2014). Government ideology and unemployment in the US states. *State Politics & Policy Quarterly*, *14*(4), 389-413.

Klarner, C. E., & Karch, A. (2008). Why do governors issue vetoes? The impact of individual and institutional influences. *Political Research Quarterly*, *61*(4), 574-584.

Kousser, T., & Phillips, J. H. (2012). *The power of American governors: Winning on budgets and losing on policy*. Cambridge University Press.

Krehbiel, K. (1998). *Pivotal politics: A theory of US lawmaking*. University of Chicago Press.

Leigh, A. (2008). Estimating the impact of gubernatorial partisanship on policy settings and economic outcomes: A regression discontinuity approach. *European Journal of Political Economy*, *24*(1), 256-268.

McCarty, N. M., & Poole, K. T. (1995). Veto power and legislation: An empirical analysis of executive and legislative bargaining from 1961 to 1986. *JL Econ. & Org.*, *11*, 282.

McCarty, N., Poole, K. T., & Rosenthal, H. (2001). The hunt for party discipline in congress. *American Political Science Review*, 673-687.

Mayer K. R. (2009). Going alone: the presidential power of unilateral action . In Edwards G. C. III and Howell W. G. (eds.), *The Oxford Handbook of the American Presidency*. Oxford, NY : Oxford University Press , 427 -454 .

Mummolo, J., & Peterson, E. (2018). Improving the interpretation of fixed effects regression results. *Political Science Research and Methods*, *6*(4), 829-835.

Poole, K. T. (2007). Changing Minds? Not in Congress!. *Public Choice*, *131*(3-4), 435-451.

Poole, K. T., & Rosenthal, H. L. (2011). *Ideology and congress* (Vol. 1). Transaction Publishers.

Schlesinger, J. A. (1966). *Ambition and politics: Political careers in the United States*. Rand MacNally.

Sellers, M. D. (2017). Gubernatorial use of executive orders: unilateral action and policy adoption. *Journal of Public Policy*, *37*(3), 315.

Shor, B., & McCarty, N. (2011). The ideological mapping of American legislatures. *American Political Science Review*, 530-551.

Squire, P., & Fastnow, C. (1994). Comparing gubernatorial and senatorial elections. *Political Research Quarterly*, *47*(3), 705-720.

Tausanovitch, C., & Warshaw, C. (2017). Estimating candidates' political orientation in a polarized congress. *Political Analysis*, *25*(2), 167-187.

Wood, B. D., & Waterman, R. W. (1991). The dynamics of political control of the bureaucracy. *The American Political Science Review*, 801-828.

Woods, N. D., & Baranowski, M. (2007). Governors and the bureaucracy: Executive resources as sources of administrative influence. *International Journal of Public Administration*, *30*(11), 1219-1230.

Zupan, M. A. (1992). Measuring the ideological preferences of US presidents: A proposed (extremely simple) method. *Public Choice*, *73*(3), 351-361.

Appendix 1

Examining sample bias in CFscores

Two hundred thirty-six individuals served as governor of a U.S. state in the years between 1991 and 2015. (This count excludes governors who retired after the 1990 election cycle but served for a few days in 1991 before their successor was inaugurated.)

Only 150 of these governors are included in Bonica's (2014) Database on Ideology, Money, and Elections. To create his database, Bonica relied on the availability of state-level campaign finance data from the National Institute for Money in State Politics (NIMSP). States only began requiring disclosures in 1990, and the rollout of these requirements was uneven through that decade. Only by 2001 were state-level candidates in every state required to record and report their campaign contributions.

As a result, CFscores cover more of the total sample of U.S. governors in later decades. Only 53 of the 112 (47%) governors who served at any point from 1991 to 1999 had CFscores. However, this number improves to 88% (110 out of 125) in the period from 2000 to 2009, and 98% (81 out of 83) in the period from 2010 to 2013.

Aside from this limitation of the data, they are otherwise quite representative. This is important because, were governors with and without CFscores different based on some characteristic, scholars that use CFscores to explain gubernatorial behavior may introduce bias along those lines.

Table A1 presents the results of a series of t-tests meant to identify any statistical differences between the populations of governors with and without CFscores. The leftmost column contains characteristics that may differentiate governors from one another. The second

(and third) columns present the percentage of governors with (or without) CFscores that claim each characteristic. The final two columns present the t-statistic and corresponding p-value of a Welch two-sample t-test comparing the two populations.

In the first three rows, we see that roughly equal percentages of governors with and without CFscores are Democrats, Republicans, and Independents respectively. The fourth row shows that slightly more governors with CFscores are women than exist in the full population, but this difference is not significant. The fifth row shows that the percentage of non-white governors is equal between populations. Finally, the rows under the "Region" subheading show that a proportionate number of governors from the Northeast and Midwest have CFscores, but that governors from the South are slightly underrepresented, and governors from the West are slightly overrepresented.

These results are important because, to the extent scholars wish to use CFscores in their models of gubernatorial behavior, they necessarily limit their sample size. If this occurs in an unrepresentative way, the scholar may introduce bias into their results. From these results, however, it appears that governors with CFscores are generally representative of the entire population of governors. Scholars need only exercise extra caution if they believe their data-generating process is different in the 1990s than in later decades, or if it varies by region of the country.

Table A1. Characteristics of governors with and without CFscores (1991-2013)

	Pct of governors with CFscores	Pct of governors w/o CFscores	t	p-value
Partisanship				
Democrat	45.4%	46.5%	-0.15	0.88
Republican	52.0%	52.2%	-0.02	0.98
Independent	2.6%	1.4%	0.64	0.52
Demographics				
Female	13.2%	7.0%	1.49	0.14
Non-white	5.3%	4.2%	0.34	0.73
Region				
Northeast	18.4%	21.1%	-0.47	0.64
Midwest	21.7%	23.9%	-0.37	0.72
South	29.6%	40.8%	-1.62	0.11
West	30.3%	14.1%	2.89	0.00
N	152	71		