

“Congruence” and “Responsiveness” in the Study of Representation

*John G. Matsusaka**

University of Southern California

Research on representation employs two measures: “congruence”, whether a roll-call vote reflects constituent opinion, and “responsiveness”, the correlation between roll-call votes and opinion. While sometimes viewed as two ways to capture the same phenomenon, theoretically the two measures are not necessarily connected, and responsiveness can produce counterintuitive conclusions. This paper assesses to what extent responsiveness and congruence capture the same thing empirically. Using 3,983 roll-call votes on state laws that were subsequently challenged in a referendum, I calculate both congruence and responsiveness for the same set of votes. The main finding is that responsiveness is an unreliable predictor of congruence. Our empirical approaches to measuring representation may be on less secure footing than believed.

January 16, 2022

* Comments, corrections, and suggestions are welcome: matsusak@usc.edu. I thank USC for support and participants at the Midwest Political Science Association meetings for helpful comments.

“Congruence” and “Responsiveness” in the Study of Representation

1. Introduction

The study of representation has taken great strides in recent years with development of new empirical methods to measure preferences and rich new data sets, yet there is an unresolved tension at the heart of the literature about how to measure representation.¹

Two approaches are commonly used, typically based on availability of data.

“Congruence”, which dates back at least to Achen (1977, 1978), measures representation as the “distance” between what legislators do and what their constituents would like them to do. A general definition is

$$(1) \quad CONG(Y_n - Y_n^*),$$

where Y is a legislator n 's action (e.g., a roll-call vote), Y_n^* is the action preferred by his or her constituents, and $CONG(x)$ is a function with a maximum at $Y = Y^*$, that is decreasing in distance from $x = 0$. Examples are $CONG(x) = -|x|$ and $CONG(x) = -x^2$. Recent studies estimating congruence include Matsusaka (2010), Lax and Phillips (2012), and Simonovitz et al. (2019).

The other approach employs a regression of the form

$$(2) \quad Y_n = \alpha + \beta \cdot Y_n^* + e_n,$$

where n indexes a legislator, α and β are coefficients to be estimated, and e_n is an error term. The coefficient β is referred to as “responsiveness”. If legislators ignored citizen preferences, then $\beta = 0$, so we can use $\beta > 0$ as a diagnostic for the existence of representation, as in Erikson et al. (1993), Tausanovitch and Warshaw (2014), and Caughey and Warshaw (2018).

The tension comes from the fact that although congruence and responsiveness are viewed as different ways to measure essentially the same object – the degree of representation – theory suggests they do not capture the same thing.² Theory shows that high responsiveness

¹ More broadly, representation is a multidimensional concept that might also involve process and descriptive issues (Pitkin, 1967).

² See Romer and Rosenthal (1979), Matsusaka (2001, 2010), and Golder and Stramski (2010).

does not predict congruence, and moreover that responsiveness does not necessarily order situations the same way that our intuitions do. There is no guarantee that responsiveness correctly detects high versus low congruence situations (Romer and Rosenthal, 1979; Erikson et al., 1993, Ch. 4; Matsusaka, 2001). This matters because a common research strategy is to estimate β for subgroups, and use it to draw conclusions about their relative representation.³ For example, we could estimate β separately in states with and without direct democracy, and use the two coefficients to infer which states are more representative; we could estimate β separately for Democratic and Republican legislators to determine which represent their constituents better; or we could test if the rich or the poor are better represented by comparing the responsiveness of policy to each group's opinion.

Of course, saying responsiveness *might* be an inaccurate proxy for representation does not mean that it actually is. Both congruence and responsiveness have an intuitive appeal, and feel like they should be related to the object we want to capture. It would give us much more confidence in our metrics if turned out empirically that they were highly correlated in practice. The purpose of this study is to investigate whether such an empirical link exists. The research strategy, which I believe is novel, is to compare measures of responsiveness and congruence for votes on exactly the same policy proposals.

The main empirical obstacle is measuring congruence, which requires data on legislator roll-call votes and constituent preferences on the same issues. We have an abundance of polling data about issue preferences, but polling data at the level of an individual legislative district on a specific law for which we have roll call votes is rare. My solution is to use roll-call votes on laws that were subsequently challenged in a referendum: in these cases, we observe the votes of legislators and constituents on the same law. I construct a data set that includes 3,983 roll-call votes, covering 31 laws in nine states. For each law, a legislator's roll-call vote is counted as congruent if it corresponds with the district's majority position in the referendum. I then calculate responsiveness by regressing roll call votes on constituent opinion, and examine how well the responsiveness coefficients predict congruence.

³ The responsiveness approach has been employed for decades. Recently, for example, Gilens (2005) uses it to assess if the rich are better represented than the poor; Mian et al. (2010) use it to test if politicians represent their constituents better in competitive than noncompetitive districts, and Monogan et al. (2009) use it to determine if policies represent public opinion better in initiative than noninitiative states.

The main finding is that the responsiveness coefficient is a poor predictor of congruence. I conduct several comparisons between different groups (male versus female legislators, votes in upper versus lower chambers, etc.) and find that the responsiveness coefficients sometimes correctly indicate which group is most congruent, but sometimes bears no relation to congruence, or points in the wrong direction. The weak connection between our two primary metrics – congruence and responsiveness – suggests that our empirical methods may be in a less settled state than we believe, and raise questions about what β is capturing.

2. Theoretical Connection between Responsiveness and Congruence

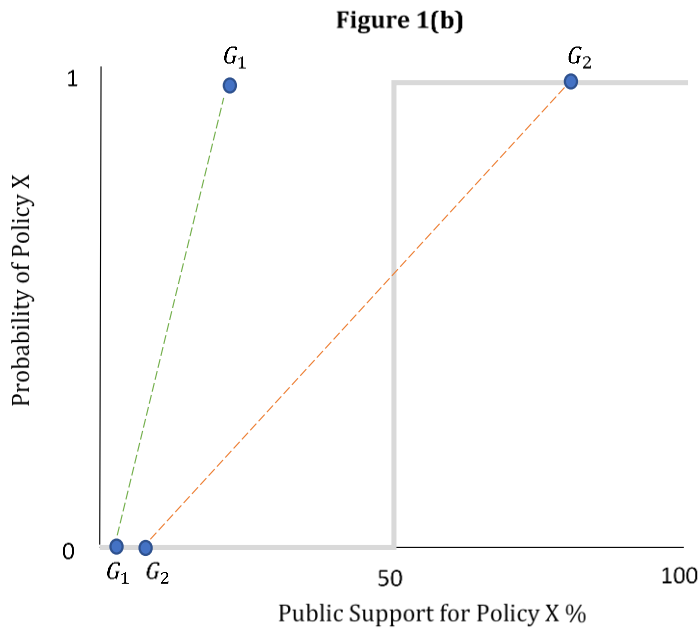
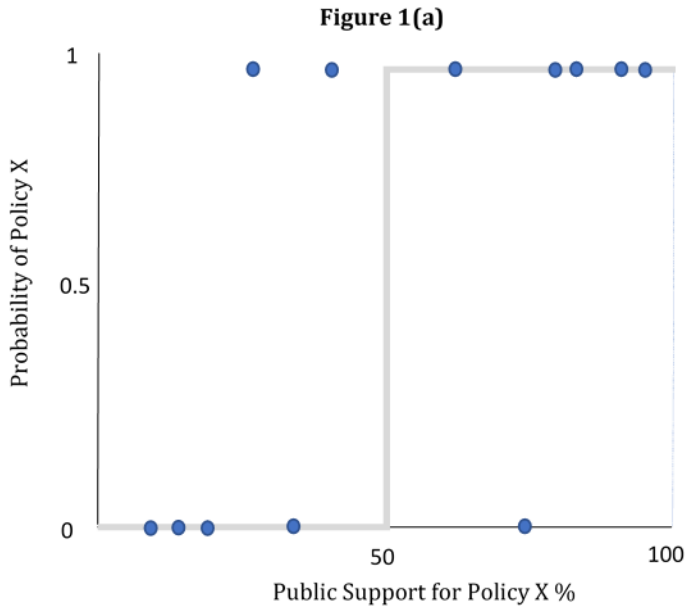
The theoretical distinction between congruence and responsiveness is well established so a brief review will suffice. Conceptually, the starting point for representation is (1) above, the extent to which a legislator votes (Y) in accordance with the wishes of his or her constituents (Y^*). A common operationalization is to treat Y^* as the preferred policy of the majority, but in principle it could be measured as the preference of any constituent, such as the median, the educated, the poor, etc. Specific to the application at hand, if legislator n 's vote on a law is $Y_n \in \{0,1\}$ and the law preferred by the district majority is $Y_n^* \in \{0,1\}$, then congruence takes the simple form:⁴

$$(3) \quad CONG_n = \begin{cases} 1 & \text{if } Y_n = Y_n^*; \\ 0 & \text{if } Y_n \neq Y_n^*. \end{cases}$$

To see the difference between congruence and responsiveness, consider Figure 1a. The figure shows a distribution of data by legislative districts (dots). The x -axis is the percent of the public in a district that supports a policy and the y -axis is the probability that their legislator votes for the policy. Instead of districts, one could study states, cities, counties, etc.; instead of looking across jurisdictions one could study the same jurisdiction over time, which each observation being a different year; and instead of votes one could study policy outcomes. If congruence is defined in terms of the majority's preference, then the gray line shows observations that are congruent. In this figure, 9 observations are congruent and 3 are noncongruent, for an overall congruence of 75 percent.

Responsiveness is calculated by running a regression line through the data, with the slope of the line denoted "responsiveness." (With dichotomous data, one might prefer a logistic regression; the argument that follows goes through either way.) Note the conceptual difference

⁴ That is, "distance" takes on one of two values, 0 ("far") and 1 ("near").



here: *congruence* is the propensity of roll-call votes to match the preferences of citizens; *responsiveness* is the **slope** of the relation between roll-call votes and preferences. While congruence can be estimated for individual observations, responsiveness is a parametric characterization that can only be estimated for a set of observations.

If policies are congruent more often than not, then the responsiveness line will slope up; if legislators completely ignore citizen preferences, then the line will be flat. Thus, a responsiveness coefficient $\beta > 0$ is evidence of some degree of representation – policy choices are not completely divorced from public opinion. The more demanding use of responsiveness

coefficient is to compare the quality of representation across different groups of observations. Such studies, roughly speaking, estimate two regression lines, one for each group (say female vs. male legislators), and compare the slope coefficients. The problem is that the relative slopes do not necessarily capture our intuition about whether a situation is more or less representative.

Figure 1b gives a simple example. There are two groups, G_1 and G_2 , each with two observations. Intuitively, it seems plain that the G_2 observations capture what we think of as “representation” better than the G_1 observations. Policies in G_1 are congruent for only one of two observations, while those in G_2 are congruent for both observations. However, consider the estimated regressions indicated by dashed lines between the two points in each group. The slope is larger for G_1 than G_2 , so G_1 is more “responsive” than G_2 . This seems to be misleading us about which situation is more representative. One can easily concoct examples in which the responsiveness coefficient points in the correct direction, as well as examples in which the slopes are the same but congruence is different, and conversely. The general point is that there is no logical connection between β and the amount of representation, as we intuitively envision it. Therefore, we lack a theoretical basis for using β 's to compare representation across groups.

This theoretical point is broader than it might seem at first glance. First, there is nothing special about using linear estimates; the point would stand for more general functional forms. Second, there is nothing special about the policy being dichotomous; essentially the same arguments hold if policy is a continuous variable (see Matsusaka (2001) for an example). Third, as mentioned above, nothing hinges on defining congruence in terms of majority rule – the same argument applies to any definition of representation that hinges on citizen preferences.

At first glance, we might wonder if congruence and responsiveness simply capture different aspects of representation. But that does not seem an acceptable interpretation: to maintain it we would have to say that G_1 is in this other aspect more representative than G_2 – but it is hard to see what that other aspect that might be – G_1 seems less representative than G_2 according to any plausible definition of representation.

It nevertheless remains possible that in practice congruence and representation are highly correlated, even though we can construct theoretical arguments where they are not correlated or negatively correlated. Indeed, given the appeal of both metrics, we might suspect this to be case. The rest of this paper examines whether such an empirical connection exists.

3. Data and Methods

My approach is to calculate congruence directly using (3), calculate responsiveness using (2), and then assess how well responsiveness proxies for congruence. In order to calculate both congruence and responsiveness, I examine a set of state laws for which we can observe roll call votes as well as citizen preferences in each district on exactly the same policy.

At present, 23 American states allow citizens to use the referendum process to challenge state laws approved by elected officials (passed by the legislature and signed by the governor). Implementation details differ, but in these states, if citizens collect a predetermined number of signatures from fellow citizens, an election is held involving the electorate at large in which voters can approve or repeal the law.⁵ I use district level referendum election returns to measure the preference of the majority of voters (which is also the preference of the median) in each legislator's district.⁶

To construct the sample, I began by identifying all state-level referendums during the period 2000-2020 using the Initiative and Referendum Institute database. From this list of ballot measures, I searched for district-level returns from each state's election division; the necessary data were available for 31 referendums.⁷ The bill associated with each referendum was then identified, and the roll call votes on that bill were drawn from legislative records. The final sample after dropping abstentions contains 3,983 roll-call votes associated with 31 laws in nine states.⁸

⁵ For example, in California, petitioners have 90 days after approval of a law to collect signatures from eligible voters equal to 5 percent of the number of votes cast in the previous gubernatorial election (as of 2015, roughly 505,000 signatures). For institutional details, see Gerber (1999).

⁶ By construction, my measures of congruence and responsiveness are based on preferences expressed by those that turned out to vote. Alternatively, representation could be based on the opinion of the entire population. Both approaches have precedents in the literature.

⁷ Referendums were excluded if a state did not report sufficiently disaggregated data, if a state changed its district lines between the time of the roll call vote and the time of the referendum election, or if a referendum was abandoned by its sponsors after qualifying for the ballot.

⁸ Referendums are held if some group is unhappy enough about a law to launch a petition campaign against it. As discussed below, in some cases these laws are unpopular laws; in other cases they are popular and the petitioner is an unpopular fringe group. Although the level of congruence overall may

The laws are listed in Table 1, with the year of legislative approval indicated. The referendums took place in Alaska, California, Maine, Maryland, Michigan, North Dakota, Ohio, South Dakota, and Washington. These states represent a mix of urban and rural, and include both “blue” and “red” states. The laws covered fiscal, political, and social issues, and included hot-button topics of national interest such as same-sex marriage, as well as issues of local interest such as Alaska’s law allowing aerial hunting of wolves and North Dakota’s law allowing the state university to discontinue use of the “Fighting Sioux” nickname for its mascot. The ideological orientation of the laws was also mixed, some proposing to move policy in a liberal direction (e.g., allowing same-sex marriage or allowing the state to use affirmative action/racial preferences) and others proposing to move policy in a conservative direction (e.g., allowing charter schools or limiting collective bargaining by public employees). Voters repealed 45 percent of the laws.

4. Findings

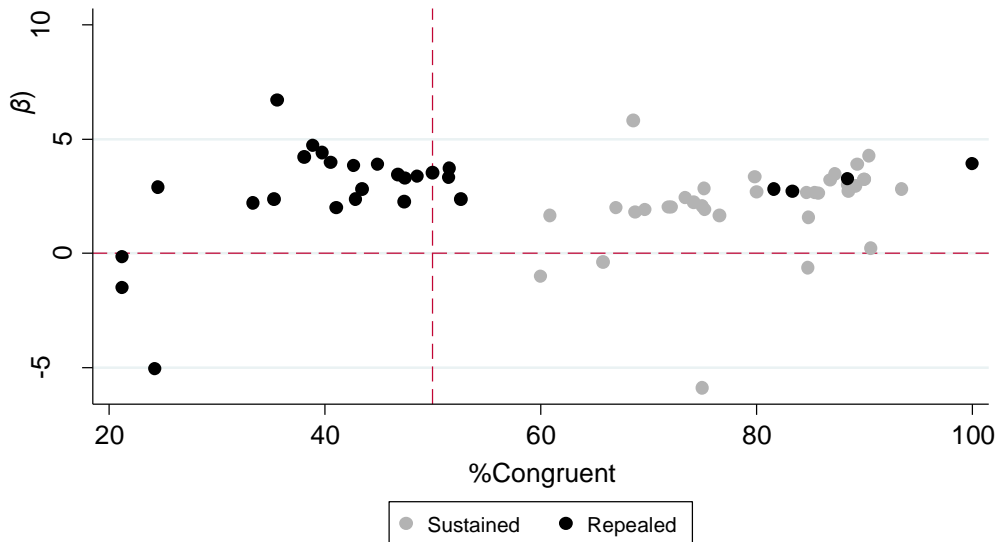
Congruence was 67 percent for the entire sample of 3,983 observations, meaning that two-thirds of roll-call votes were cast in accordance with majority opinion in the district. Since congruence would be 50 percent with random voting, anything over this threshold indicates some degree of representation. Responsiveness for the full sample was 1.17, meaning that a 1 percent increase in public support for a law was associated with a 1.17 percent chance that a legislator votes in support. Again, this indicates that some representation was taking place.

I next turn to the question of how well responsiveness tracks congruence by estimating congruence and responsiveness separately for each law and each chamber of each state, producing a total of 62 points of comparison. Figure 2 is a scatterplot with congruence on the horizontal axis and responsiveness on the vertical axis. If responsiveness predicts congruence, we would expect to see a positive relation between the two measures, and the responsiveness coefficient would be negative when congruence is less than 50 percent and positive when congruence exceeds 50 percent. No connection between congruence and responsiveness is visible. A regression that uses β to predict congruence for these data has an R^2 of 0.006.⁹ If a

be different on these laws than others, there is no reason to expect that responsiveness regressions should be less able to detect the level of congruence on them.

⁹ The coefficient on responsiveness is a statistically insignificant -0.85 (standard error 1.38). The story is the same with a logistic instead of linear probability regression.

Figure 2. Congruence and Responsiveness



Note. Each a point is a chamber-law. The horizontal axis is the percent of roll-call votes congruent with majority opinion in a district. The vertical axis is the responsiveness coefficient β from the regression: $Y = \alpha + \beta \%Yes$, where $Y = 1$ if a legislator voted yes and $\%Yes$ is the percent of yes votes in the district.

researcher was interested in which legislatures were more likely to vote in accordance with majority opinion in their districts, knowing the responsiveness coefficient would provide no insight.

A few examples may help illustrate the point: The chamber with the highest congruence (100 percent), California’s senate in 2003, had responsiveness $\beta = 3.93$, while the chamber with one of the lowest congruences (38.9 percent), California’s senate in the second insurance law of 1999, had a *higher* responsiveness coefficient $\beta = 4.71$. The chamber with the highest responsiveness coefficient of $\beta = 5.82$, Maryland’s lower chamber in 2006, had a middling congruence of 68.6 percent, while the chamber with the lowest responsiveness coefficient of $b = -5.88$, California’s senate in 1998, had a higher congruence of 75.0 percent.

We might expect the laws that voters approved to have been more representative than those they repealed.¹⁰ Figure 2 shows the two groups in gray and black, respectively. Consistent with this intuition, all of the laws that voters sustained were more than 50 percent congruent, and most of the laws they repealed were less than 50 percent congruent. In this sense,

¹⁰ This is plausible but not necessary. If districts are gerrymandered, even a fully congruent legislature (that is, a legislature in which every member votes according to majority opinion in his or her district) can choose a policy that is not supported by a majority of voters statewide (Gilligan and Matsusaka, 2006).

congruence “works” as a measure of representation. Responsiveness, on the other hand, does not. There is no reliable relation between a law’s responsiveness coefficient and whether or not voters approved it.

I next provide comparisons between several natural groups. For each comparison, I calculate congruence and responsiveness for each group and compare them. The question is: to what extent do differences in responsiveness coefficients predict congruence differences between the groups?

Table 2 reports comparisons using groups that are state- or chamber-specific. Panel A formalizes the comparison between laws that were approved and repealed, which was just discussed. Congruence was 79.2 percent for approved laws and 48.0 percent for repealed laws, a difference of 31 percent that is statistically different from zero.¹¹ Responsiveness was positive for both groups; the latter pattern is somewhat misleading, as noted, because congruence is less than 50 percent. Responsiveness was higher for approved than repealed votes, but at the edge of statistical significance.

Panel B of Table 2 distinguishes between the lower and upper chambers of the legislature. This is an interesting comparison because lower chambers are designed to be more closely connected to popular opinion than upper chambers by requiring members to stand for election more frequently and by having fewer constituents per district.¹² The congruence between lower upper chambers was essentially identical; the 0.5 percent difference is not statistically different from zero. Similarly, the responsiveness coefficients were essentially identical and not statistically different. In this case, the responsiveness coefficients tell the same story as congruence – that representation was similar in lower and upper chambers.

In Panel C of Table 2 observations are distinguished based on whether the state had term limits on legislators.¹³ There are offsetting theoretical predictions about how term limits affect representation. Proponents argue that term limits bring about the election of “citizen legislators”, whose views are more consonant with their constituents, instead of career

¹¹ Here and below, when comparing congruence or responsiveness, I estimate a regression that allows the measurement to vary by group. Standard errors are calculated after clustering by state-chamber.

¹² All states in the sample have more members in the lower than upper chamber. Alaska, California, Michigan, Ohio, and Washington have longer terms for members of the upper than lower chamber.

¹³ California, Maine, Michigan, Ohio, and South Dakota had legislative term limits during the sample period.

politicians, leading to superior representation. Skeptics argue that term-limited legislators have less incentive to follow constituent preferences when they no longer face re-election incentives.¹⁴ For the votes studied here, congruence was 11.4 percent lower in term-limit states than other states, although the difference is not statistically distinguishable from zero. The responsiveness coefficients point in the same direction, with a lower β coefficient for term-limit than non-term-limit states, although they are not statistically distinguishable. The magnitude of the differences for both measures seems sizeable but are noisy.

Table 3 reports comparisons based on legislator-specific or district-specific variation. Panel A compares congruence of female and male legislators. A growing literature investigates whether female or male legislators behave differently. Anzia and Berry (2011) find that female members of Congress deliver more federal spending to their districts than male members; Stadelmann et al. (2014) report that female legislators vote more congruently than male legislators in Switzerland; while Ferreira and Gyourko (2014) find no difference in the policy choices of female and male mayors. For the votes in my sample, female legislators were 2.2 percent more likely to cast a congruent vote than male legislators, a small and statistically insignificant difference. The responsiveness coefficient is also slightly larger for female than male legislators, although the difference is not different from zero statistically.

Panel B of Table 3 compares Democratic (plus Green) and Republican legislators. While there is no theoretical reason to expect one party to be more representative than the other, some evidence suggests there may be partisan differences in how accurately legislators perceive constituent preferences (Broockman and Skovron, 2018). Democratic legislators voted congruently with their constituents 12.9 percent of the time more than Republican legislators, although the difference is not statistically significant. Democratic legislators were more than three times as responsive as Republican legislators, although that difference too is not statistically significant. In this case, responsiveness tells the same story as congruence, albeit with substantial noise in the estimates.

Panel C of Table 3 distinguishes legislators according to whether they belonged to the majority or minority party. A body of theory argues that the majority party forms a coalition or cartel to advance the party's interest, and the coalition functions by ensuring that its members vote in accordance with the party agenda, not necessarily in accordance with the interests of the individual members' constituents (e.g., Cox and McCubbins (1993, 2005)). If members of the

¹⁴ See Besley and Case (2003), Kousser (2008), and Mooney (2009).

majority are under more pressure to conform to the party's goals, they would be less likely to cast congruent votes. The congruence estimates provide some support for this idea by showing that members of the majority party were 22.4 percent less likely to cast a congruent vote, a quantitatively large and statistically significant number. The responsiveness regressions produce the same pattern: the responsiveness coefficient is quite small and statistically insignificant for the majority party and positive and statistically significant for the minority party, with the difference statistically different from zero at the 1 percent level. Responsiveness predicts the congruence differences.

Panel D of Table 3 distinguishes legislators according to whether they were in their last term under the state's term limits law. The sample is restricted to the five states that had term limits during the sample period. The congruence numbers do not show a meaningful difference between last-term and other legislators, nor do the responsiveness coefficients.

As a final exercise, Panel E of Table 3 compares legislators that represented competitive versus noncompetitive districts. Competition is a holy grail for some activists, who believe it will bring about better representation. I classify a district as competitive if the margin of victory between the top two candidates in the previous election was less than 10 percent (the patterns are similar with other cutoffs such as 5 percent or 33 percent). These should not be taken as causal estimates. Legislators representing noncompetitive districts were actually 7.1 more congruent than those representing competitive districts, a difference near the border of statistical significance. The responsiveness coefficients point in the same direction, with a higher β for legislators in noncompetitive than competitive districts, again near the border of statistical significance.

5. Discussion

This study seeks to shed light on the thorny question of how we measure representation. There is a tension in the literature concerning the two leading measures, "congruence" and "responsiveness." Although they are often viewed as interchangeable, or two different ways to capture aspects of the same phenomenon, in theory responsiveness – the correlation between roll-call votes and constituent opinion – is not necessarily correlated with congruence or to our basic intuitions about representation. As Achen (1978, p. 475) put it (speaking theoretically): "correlations are virtually uninterpretable within the framework of democratic theory."

Despite their shaky theoretical foundation, responsiveness coefficients might work in practice. Certainly, they appeal to our sense that representation should consist of a correlation

between roll-call votes and constituent preferences. This paper offers what I believe is the first attempt to empirically test whether responsiveness is a good predictor of congruence in practice. Such an assessment requires measuring both responsiveness and congruence for the same sample of votes. This paper calculates congruence using roll-call votes on laws that later came to a vote of the general public in a referendum.

I find that responsiveness is an unreliable predictor of congruence. Responsiveness sometimes correctly identifies roll-call votes that are more congruent than others, but sometimes fails. The most dramatic failure in my sample is laws that voters approved compared to laws they repealed: congruence was much higher on approved than repealed laws, as one would expect, but the responsiveness coefficients were essentially the same.

These findings suggest that the empirical foundations of our research on representation may be less secure than we think. Contrary to what seems to be a popular view, congruence and responsiveness may not be two different ways to capture the same object, the quality of representation. Congruence stands on fairly solid ground theoretically, and with the development of new methods to measure public opinion at the micro level (including multilevel regression and poststratification and referendum votes), a growing number of studies calculate congruence.¹⁵ The use of responsiveness would seem to merit a closer, more critical look. It would be useful to have a stronger theoretical foundation for the concept, to know under what conditions the correlations accurately predict situations that we consider to be representative. In the meantime, in light of the potential for responsiveness coefficients to mislead, researchers may wish to cross-check their estimates with other metrics.

¹⁵ Examples: Matsusaka (2010), Lax and Phillips (2012), Portmann et al., (2012), Brunner et al. (2013), Stadelmann et al. (2013, 2014), Krimmel et al., (2016), Kousser et al. (2018), Becher and Menendez Gonzalez (2019), Lax et al. (2019), Simonovitz et al. (2019), Matsusaka (2020). There is also a literature that uses ideal-point estimates to measure representation. See Tausanovitch and Warshaw (2013) and Broockman (2016) for reviews and a critique.

References

Achen, Christopher H., "Measuring Representation: Perils of the Correlation Coefficient," *American Journal of Political Science*, November 1977, Vol. 21(4), 805-815.

Achen, Christopher H., "Measuring Representation," *American Journal of Political Science*, August 1978, Vol. 22(3), 475-510.

Anzia, Sarah F. and Christopher R. Berry, "The Jackie (and Jill) Robinson Effect: Why Do Congresswomen Outperform Congressmen?," *American Journal of Political Science*, July 2011, Vol. 55(3), 478-493.

Becher, Michael and Irene Mendendez Gonzalez, "Electoral Reform and Trade-offs in Representation," *American Political Science Review*, 2019, Vol. 113(3), 694-709.

Broockman, David E., "Approaches to Studying Policy Representation," *Legislative Studies Quarterly*, February 2016, Vol. 41(1), 181-215.

Broockman, David E. and Christopher Skovron, "Bias in Perception of Public Opinion among Political Elites," *American Political Science Review*, August 2018, Vol. 112(3), 542-563.

Brunner, Eric, Stephen L. Ross, and Ebonya Washington, "Does Less Income Mean Less Representation?," *American Economic Journal: Economic Policy*, May 2013, Vol. 5(2), 53-76.

Caughey, Devin and Christopher Warshaw, "Policy Preferences and Policy Change: Dynamic Responsiveness in the American States," *American Political Science Review*, 2018, 112(2), 249-266

Cox, Gary W. and Mathew D. McCubbins, *Legislative Leviathan: Party Government in the House*, Berkeley, CA: University of California Press, 1993.

Cox, Gary W. and Mathew D. McCubbins, *Setting the Agenda: Responsible Party Government in the U.S. House of Representatives*, Cambridge: Cambridge University Press, 2005.

Erikson, Robert S., Gerald C. Wright, and J. P. McIver, *Statehouse Democracy: Public Opinion and Policy in the American States*, Cambridge, UK: Cambridge University Press, 1993.

Ferreira, Fernando and Joseph Gyourko, "Does Gender Matter for Political Leadership? The Case of U.S. Mayors," *Journal of Public Economics*, 2014, Vol. 112, 24-39.

Gerber, Elisabeth R., *The Populist Paradox: Interest Group Influence and the Promise of Direct Legislation*, Princeton, NJ: Princeton University Press, 1999.

Gilens, Martin, "Inequality and Democratic Responsiveness," *Public Opinion Quarterly*, 2005, Vol. 69(5), 778-796.

Gilligan, Thomas W. and John G. Matsusaka, "Public Choice Principles of Redistricting," *Public Choice*, December 2006, Vol. 129(3-4), 381-398.

Golder, Matt and Jacek Stramski, "Ideological Congruence and Electoral Institutions," *American Journal of Political Science*, January 2010, Vol. 54(1), 90-106.

Kousser, Thad, "Term Limits and State Legislatures," in *Democracy in the States: Experiments in Election Reform*, edited by Caroline Tolbert, Todd Donovan, and Bruce Cain, Washington D.C.: Brookings Institution Press, 2008.

Kousser, Thad, Justin Phillips, and Boris Shor, "Reform and Representation: A New Method Applied to Recent Electoral Changes," *Political Science Research and Methods*, October 2018, Vol. 6(4), 1-19.

Krimmel, Katherine, Jeffrey R. Lax, and Justin H. Phillips, "Gay Rights in Congress: Public Opinion and (Mis)representation," *Public Opinion Quarterly*, December 2016, Vol. 80(1-4), 888-913.

Lax, Jeffrey R. and Justin H. Phillips, "The Democratic Deficit in the States," *American Journal of Political Science*, January 2012, Vol. 56(1), 148-166.

Lax, Jeffrey R, Justin H. Phillips, and Adam Zelizer, "The Party or the Purse? Unequal Representation in the US Senate," *American Political Science Review*, 2019, Vol. 113(4), 917-940.

Matsusaka, John G., "Problems with a Methodology Used to Test the Responsiveness of Policy to Public Opinion in Initiative States," *Journal of Politics*, November 2001, Vol. 63, 1250-1256.

- Matsusaka, John G., "Popular Control of Public Policy: A Quantitative Approach," *Quarterly Journal of Political Science*, 2010, Vol. 5, 133-167.
- Matsusaka, John G., *Let the People Rule: How Direct Democracy Can Meet the Populist Challenge*, Princeton, NJ: Princeton University Press, 2020.
- Mian, Atif, Amir Sufi, and Francesco Trebbi, "The Political Economy of the US Mortgage Default Crisis," *American Economic Review*, December 2010, Vol. 100(5), 1967-1998.
- Monogan, James, Virginia Gray, and David Lowery, "Public Opinion, Organized Interests, and Policy Congruence in Initiative and Noninitiative States," *State Politics and Policy Quarterly*, Fall 2009, Vol. 9(3), 304-324.
- Mooney, Christopher Z., "Term Limits as a Boon to Legislative Scholarship: A Review," *State Politics and Policy Quarterly*, Summer 2009, Vol. 9(2), 204-228.
- Pitkin, Hanna Fenichel, *The Concept of Representation*, Berkeley, CA: University of California Press, 1967.
- Romer, Thomas and Howard Rosenthal, "The Elusive Median Voter," *Journal of Public Economics*, October 1979, Vol. 12, 143-170.
- Simonovits, Gabor, Andrew M. Guess, and Jonathan Nagler, "Responsiveness without Representation: Evidence from Minimum Wage Laws in the U.S. States," *American Journal of Political Science*, April 2019, Vol. 63(2), 401-410.
- Stadelmann, David, Marco Portmann, and Reiner Eichenberger, "Quantifying Parliamentary Representation of Constituents' Preferences with Quasi-Experimental Data," *Journal of Comparative Economics*, 2013, Vol. 14, 170-180.
- Stadelmann, David, Marco Portmann, and Reiner Eichenberger, "Politicians and Preferences of the Voter Majority: Does Gender Matter?," *Economics and Politics*, November 2014, Vol. 26(3), 355-379.
- Tausanovitch, Chris and Christopher Warshaw, "Measuring Constituent Policy Preferences in Congress, State Legislatures, and Cities," *Journal of Politics*, April 2013, Vol. 75(2), 330-342.

Tausanovitch, Chris and Christopher Warshaw, "Representation in Municipal Government,"
American Political Science Review, August 2014, Vol. 108(3), 605-641.

Table 1. List of Laws

State	Description	Outcome	Year
Alaska	Permits hunters to use airplanes to hunt wolves	Repealed	2000
California	Permits Pala tribe to operate video lottery terminals	Approved	1998
California	Allows third parties to sue insurance companies	Repealed	1999
California	Allows third parties to sue insurance companies (alt)	Repealed	1999
California	Requires companies to provide health care coverage	Repealed	2003
California	Authorizes gambling compact with Pechanga tribe	Approved	2007
California	Authorizes gambling compact with Morongo tribe	Approved	2007
California	Authorizes gambling compact with Sycuan tribe	Approved	2007
California	Authorizes gambling compact with Agua Caliente tribe	Approved	2007
California	Allows North Folk tribe casino in Central Valley	Repealed	2013
California	Bans plastic bags in grocery stores	Approved	2014
Maine	Replaces health insurance claims tax with beverage tax	Repealed	2008
Maryland	Changes voting procedures	Approved	2006
Maryland	Allows illegal immigrants to pay in-state tuition rates	Approved	2011
Maryland	Congressional redistricting plan	Approved	2011
Maryland	Allows same-sex marriage	Approved	2012
Michigan	Allows hunting of mourning doves.	Repealed	2004
North Dakota	Discontinues "Fighting Sioux" college nickname	Approved	2011
North Dakota	Allows corporate dairy and swine farms	Repealed	2015
Ohio	Limits interest rate charged by payday lenders	Approved	2008
Ohio	Limits collective bargaining by public employees	Repealed	2011
South Dakota	Bans smoking in restaurants and bars	Approved	2009
South Dakota	Increases barriers for independent candidates for office	Repealed	2015
South Dakota	Reduces minimum wage for teenagers	Repealed	2015
Washington	Increases taxes for unemployment insurance	Repealed	2002
Washington	Allows charter schools	Repealed	2004
Washington	Prohibits insurers from denying certain claims	Approved	2007
Washington	Grants domestic partners same rights as married people	Approved	2009
Washington	Allows same-sex marriage	Approved	2012
Washington	Allows government to use racial preferences	Repealed	2019
Washington	Requires schools to provide sex education	Approved	2020

Table 2. Congruence and Responsiveness by Outcome, Chamber, and Term Limits

	(1)	(2)	Difference	
Panel A	Approved by Voters	Repealed	(1) – (2)	
Congruence (%)	79.2	48.0	31.1	[.001]
Responsiveness (β)	2.09 (0.23)	1.40 (0.28)	0.69	[.062]
N	2,284	1,699		
Panel B	Lower House	Upper House	Difference	
Congruence (%)	65.7	66.2	-0.5	[.757]
Responsiveness (β)	1.04 (0.18)	0.99 (0.19)	0.06	[.556]
N	2,786	1,197		
Panel C	Term Limits	No Term Limits	Difference	
Congruence (%)	59.9	71.3	11.4	[.113]
Responsiveness (β)	1.03 (0.24)	1.26 (0.36)	0.23	[.598]
N	1,910	2,073		

Note. Standard errors are in round parentheses. Differences are calculated from regressions, with standard errors clustered at the law \times chamber level, and p -values reported in square brackets.

Table 3. Congruence and Responsiveness by Legislator Gender, Party, Last Term

	(1)	(2)	Difference (1) – (2)	
Panel A	Female	Male		
Congruence (%)	67.4	65.2	2.2	[.438]
Responsiveness (β)	1.07 (0.18)	1.00 (0.20)	0.07	[.709]
N	1,139	2,844		
Panel B	Democratic	Republican	Difference	
Congruence (%)	71.8	58.9	12.9	[.197]
Responsiveness (β)	1.12 (0.27)	0.36 (0.47)	0.76	[.250]
N	2,158	1,824		
Panel C	Majority party	Minority party	Difference	
Congruence (%)	57.8	80.1	-22.4	[.015]
Responsiveness (β)	0.20 (0.13)	1.80 (0.37)	-1.60	[.001]
N	2,535	1,448		
Panel D	Last term	Not last term	Difference	
Congruence (%)	61.9	59.3	2.6	[.434]
Responsiveness (β)	1.12 (0.22)	1.01 (0.28)	0.11	[.670]
N	491	1,419		
Panel E	Competitive district	Noncompetitive	Difference	
Congruence (%)	60.1	67.2	-7.1	[.056]
Responsiveness (β)	0.68 (0.22)	1.07 (0.18)	-0.39	[.052]
N	726	3,257		

Note. Standard errors are in round parentheses. Differences are calculated from regressions, with standard errors clustered at the law \times chamber level, and p -values reported in square brackets. A district is defined to be “competitive” if the margin between the top two candidates in the previous election was less than 10 percent.