

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

*John G. Matsusaka**
University of Southern California
matsusak@usc.edu

This paper compares two common methods used to measure representation in empirical research: “congruence,” whether a roll-call vote is aligned with constituent opinion; and “responsiveness,” the correlation between roll-call votes and constituent opinion. While sometimes viewed as different ways of capturing the same underlying concept of “representation,” theoretically the measures are not connected. This paper assesses to what extent responsiveness and congruence capture the same thing empirically. Using 4,192 roll-call votes on 33 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, and therefore cannot be taken as a proxy for congruence.

December 17, 2025

Keywords: representation; congruence; responsiveness; roll-call votes; referendums

JEL code: D7

* I received helpful comments from anonymous referees and participants at the Midwest Political Science Association meetings.

1. Introduction

Understanding representation is a core issue in public choice research. Democracies operate under the assumption that periodic elections give elected officials the incentive to represent their constituents, an idea formalized in the foundational models of Barro (1971) and Ferejohn (1986). While theory is clear, measuring the actual degree of representation in practice is challenging. In recent years, researchers have made significant progress on this issue, developing new empirical methods and creating rich new data sets. Yet an unresolved tension remains at the heart of the literature in the way that representation is measured.

One measure of representation is “congruence,” dating back at least to Achen (1977, 1978), which measures representation as the “distance” between what legislators do and what their constituents would like them to do.¹ A general definition is

$$(1) \quad \text{CONGRUENCE} = f(Y_n - Y_n^*),$$

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 $f(x)$ is a function with a maximum at $Y = Y^*$ that is decreasing in distance from $x = 0$. Examples are $f(x) = -|x|$ and $f(x) = -x^2$.

Another measure is “responsiveness,” essentially a correlation derived from 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 the “responsiveness.” If legislators ignored citizen preferences,

¹ Studies estimating congruence include Matsusaka (2010, 2020), Lax and Phillips (2012), Krimmel et al. (2016), Kousser et al. (2018), Becher and Menéndez-González (2019), Lax et al. (2019), Simonovitz et al. (2019), Giger et al. (2020), Caughey and Warshaw (2022). Caughey and Warshaw (2022) is the most comprehensive examination of congruence and responsiveness in recent years, and a good starting point for understanding the issues.

then $\beta = 0$, so $\beta > 0$ is a diagnostic for the existence of some amount of representation. Responsiveness can be estimated either in a cross-section or across time.²

The tension comes from the fact that although congruence and responsiveness are viewed as different ways of capturing the same underlying object – the amount of representation – theoretically they do not capture the same thing.³ Higher responsiveness is not logically associated with greater congruence, and conversely. This doesn't matter if a researcher is testing a theory that generates predictions about responsiveness per se, but it is important if the researcher is using responsiveness *as a proxy for congruence*. For example, a researcher may want to compare the cross-sectional β coefficients of two groups of legislators in order to draw conclusions about which group is better reflecting constituent preferences.

The fact that responsiveness is not a reliable proxy for congruence in theory does not mean that it doesn't work in practice. Both congruence and responsiveness have an intuitive appeal, and feel like they should be correlated. This paper investigates 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 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 4,192 roll-call votes, covering 33 laws in 10 states. For each law, a legislator's roll-call vote is counted as congruent if it corresponds to the district's majority position in the referendum. I calculate responsiveness by regressing roll call votes on constituent opinion, and examine how well the responsiveness coefficients predict congruence. My purpose, then, is not to use these data to explore the determinants of representation (something I

² Studies estimating responsiveness include Erikson et al. (1993), Tausanovitch and Warshaw (2014), and Caughey and Warshaw (2018, 2022).

³ See Romer and Rosenthal (1979), Matsusaka (2001, 2010), and Golder and Stramski (2010).

pursue elsewhere (Matsusaka, forthcoming)), but to investigate the validity of the methods researchers use to measure representation.

The main finding is that the responsiveness coefficient is not a reliable 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, sometimes bear no relation to congruence, and sometimes point in the wrong direction. The unreliable connection between responsiveness and congruence calls for a deeper theoretical justification for responsiveness – if it is not meant to capture the alignment between citizen preferences and legislator votes, then what is it trying to capture? – and implies that responsiveness should not be used as a proxy for congruence.

2. Theoretical Connection between Representation and Its Measures

The word “representation” has a multitude of meanings (Pitkin 1967), and even within the narrow category of *political* representation, there is a healthy debate about whether representatives should act as delegates/agents or as trustees. When it comes to empirical research, many researchers appear to have in mind a delegate model: representation “means that the representative must vote as a majority of his constituents would,” as Pitkin (1967, pp. 144-145) put it. This paper focuses on studies that seek to measure representation in this delegate sense, the extent to which representatives’ votes mirror those that would have been cast by their constituents. For expositional clarity, I work with the following definition:

Definition. *A group of legislators A are more “representative” than a group of legislators B if the roll-call votes cast by legislators A better mirror the preferences of their constituents than the roll-call votes cast by legislators B mirror the preferences of their constituents.*

The simplest way to implement this definition is to directly compare a legislator’s vote (Y) on a law with the vote preferred by a majority of the legislator’s constituents (Y^*). In the case of roll-call voting, where the choice is either yes or no, if legislator n ’s vote is

$Y_n \in \{0,1\}$ and the preferred vote by a majority of constituents is $Y_n^* \in \{0,1\}$, then “congruence” is:⁴

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

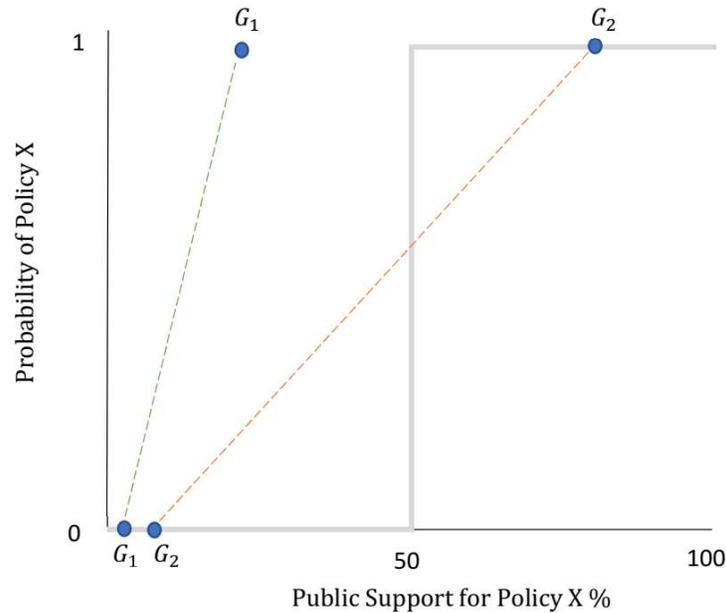
The congruence of legislators A is then $CONG(A) = \left(\frac{1}{N_A}\right) \sum_{n \in A} CONGRUENCE_n$, where N_A is the number of legislators in group A . We can compare congruence between two groups A and B , such as Democrats vs. Republicans voting on the same bill, or between the same group at different points in time voting on different bills. Note that congruence directly maps into the stated definition of representation.

Observation 1. *If $CONG(A) > CONG(B)$ then A is more representative than B according to the definition given above.*

To illustrate the theoretical difference between congruence and responsiveness, consider Figure 1, which shows the roll-call votes and constituent preferences of four legislators, divided into two groups G_1 and G_2 . The x -axis is the percent of constituents that support the policy and the y -axis is the probability that the legislator votes for the policy

⁴ With dichotomous outcomes, the preferred vote of the majority is also the preferred vote of the median. Alternatively, one could measure congruence as $|Y_n - Y_n^*|$ where Y_n^* is the percentage of constituents in favor of the policy, in which case congruence is a continuous variable between 0 and 1. The analysis in this section applies to that case as well.

Figure 1



(the probability is 1 if yes, and 0 if no). With congruence defined in terms of the majority's preference, the gray step function shows the points that would be congruent. For example, the topmost legislator in group G_1 votes yes on the bill, which less than 50 percent of the legislator's constituents support, a noncongruent vote. Both legislators in G_2 are congruent, $CONG(G_2) = 1$, but only one legislator in G_1 is congruent, $CONG(G_1) = 0.5$.

Now consider responsiveness, 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.) The slope is larger for G_1 than G_2 , so G_1 is more "responsive" than G_2 . However, G_1 is *less* congruent and *less* representative than G_2 according to the working definition of representation. (And it is not obvious what alternative definition of representation would justify describing G_1 as more representative than G_2 .) One can construct 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.

Observation 2. *If $\beta_A > \beta_B$ then A could be more, less, or equally representative than B.*

This theoretical point is fairly general. 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.

The limitations of responsiveness can also be illustrated with a verbal example motivated by Caughey and Warshaw (2022, p. 97): “In a perfectly responsive democracy, a given change in citizens’ preferences (e.g., from a \$10 minimum wage to an \$11 one) would be followed by a policy change of the same magnitude (a \$1 increase in the minimum wage).” This is true in a perfectly representative democracy, but the intuition does not carry through in a partially representative democracy. Suppose that the minimum wage is \$11; initially citizens prefer a minimum wage $W^* = \$10$ and their preferred policy changes to $W^* = \$11$. In fact, it would be more representative in terms of the definition above to keep the minimum wage at \$11 than to increase it to \$12. Responsiveness in this case *reduces* the representativeness of the policy.

The same logic extends to comparisons between groups. Suppose the initial minimum wage is \$11; there are two group of citizens *A* and *B* that initially prefer $W_A^* = \$10$ and $W_B^* = \$8$; and each group experiences a +1 change in its preferences, that is, to $W_A^* = \$11$ and $W_B^* = \$9$. Now suppose the government of *A* does not respond and holds the minimum wage to \$11, while the government of *B* increases it to \$12. Here the government of *A* has been entirely “unresponsive” ($\beta_A = 0$) while the government of *B* has been perfectly “responsive” ($\beta_B = 1$) but the government of *A* is better representing voter preferences than the government of *B*.

Given that “responsiveness” lacks a consistent theoretical link to congruence and to the underlying concept of representation, the natural question is how to justify its use as a measure of representation? One possible justification is empirical: even though there is no necessary theoretical link, there may be an empirical link: responsive legislators may be

congruent in practice. The rest of this paper examines whether such an empirical connection exists in a large sample of roll-call votes. Relying on the observation the congruence reliably captures representation (as defined above), I examine if responsiveness is correlated with congruence.

Observation 3. *If, in practice, responsiveness captures representation as defined above, then we would observe empirically that groups with a high responsiveness would have a high congruence, and conversely.*

3. Data and Methods

My approach is to calculate responsiveness using (2), congruence using (3), and then assess how well responsiveness proxies for congruence. The challenge in executing such a comparison is the need for data on both legislator roll-call votes and the preferences of their constituents on the same laws. My research strategy is to focus on laws that were challenged by a referendum, for which we can observe roll call votes and 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 (i.e., passed by the legislature and signed by the governor). Implementation details differ, but in all cases if sponsors collect a predetermined number of signatures from fellow citizens, an election is held in which voters approve or repeal the law.⁵ I use district-level referendum election returns to measure the preference of the majority (= median) of voters in each legislator's district.

To construct the sample, I began by identifying all state-level referendums during the period 2000-2024 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.⁶

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 (546,651 signatures as of 2024). For institutional details, see Gerber (1999).

⁶ 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, if individual roll-call votes were unavailable, or if a referendum was abandoned by its sponsors after qualifying for the ballot.

The bill associated with each referendum was then identified, and roll call votes on that bill were drawn from legislative records. The final sample after dropping abstentions contains 4,192 roll-call votes on 33 laws in 10 states.⁷

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 a conservative change (e.g., allowing charter schools or limiting collective bargaining by public employees). Voters approved 55 percent of the laws.

An alternative approach would be to measure district preferences by opinion surveys instead of referendum election returns.⁸ A significant limitation of referendum data is that the number of such elections is limited; opinion data are much more abundant, and using modern MRP (multilevel regression and poststratification) methods can be extended to jurisdictions that were not surveyed in the first place. A relative advantage of referendum data is that they reveal voter preferences on exactly the law approved by the legislature, while opinion surveys usually ask about a law in general terms – and poll results can vary dramatically with minor changes in question wording (Caughey and Warshaw, 2022, p. 17). Another advantage of referendum data is that referendum elections make law – citizen opinions may be more informed when they have real consequences and have been exposed to campaigns than when they respond to a pollster. Both referendum

⁷ Matsusaka (forthcoming) provides more descriptive information about the data.

⁸ Studies using opinion data include Krimmel et al. (2016), Lax et al. (2019), and Caughey and Warshaw (2022). Portmann et al. (2012), Stadelmann et al. (2013, 2014), and Giger et al. (2020) use ballot measure returns to calculate congruence.

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
California	Ends use of cash bail	Repealed	2018
Colorado	Joins state in national popular vote compact	Approved	2019
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

data and opinion surveys suffer from having a nonrepresentative sample of issues – the issues that come to a referendum or that pollsters ask about tend to be more controversial than others. Having said this, the two approaches are complementary, and having evidence from both sources helps form a full picture.

Congruence is calculated for each legislator, 1 if the legislator's vote matched the preference of a majority of his or her constituents, and 0 otherwise. Group congruence is

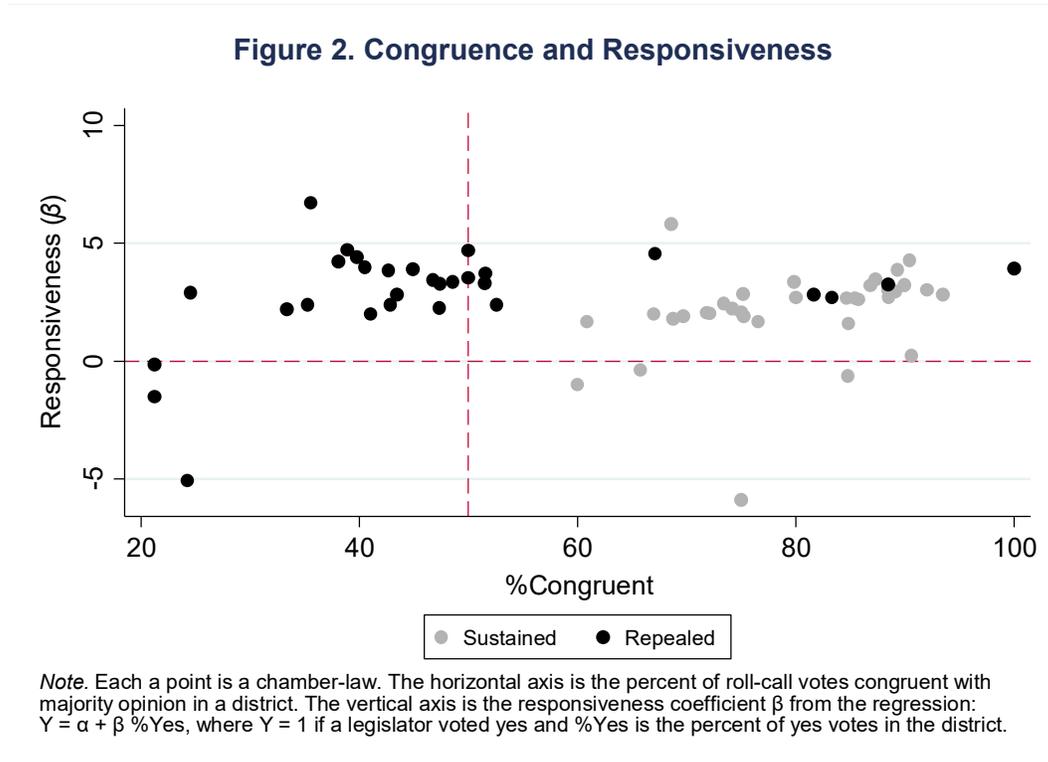
calculated as the average congruence of its members. Responsiveness is calculated from a linear regression on a group of votes, with the dependent variable equal to 1 if a legislator voted yes, and 0 if the legislator voted no, and the explanatory variable is the percentage of constituents in favor of the bill. Note that responsiveness here is entirely a cross-sectional estimate.

Note that representation is measured by comparing a legislator's roll-call vote and the expressed preferences of his or her constituents on the referendum that came some months after the roll-call vote. It is not unusual to use public opinion data, collected both before and after a vote or policy decision, to calculate congruence, but in my context, it raises the question whether the legislator's vote might influence public opinion. This possibility cannot be ruled out, but it's unlikely that citizens know their representative's votes on most of these bills, many of which received scant media coverage (Matsusaka, forthcoming). It's also possible that public opinion changed as a result of the referendum campaign; representation should then be understood in terms of the alignment between a legislator's vote and the informed preferences of his or her constituents after hearing campaign arguments.

4. Findings

Congruence was 66 percent for the entire sample of 4,192 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 1.08, meaning that a 1 percent increase in public support for a law was associated with a 1.08 percentage point greater chance that a legislator voted in support.

To investigate how well responsiveness tracks congruence, I begin by estimating responsiveness and congruence separately for each law and each chamber of each state, producing a total of 66 points of comparison. Figure 2 is a scatterplot of congruence against responsiveness. If responsiveness predicts congruence, we should see a positive relation between the two measures, and the responsiveness coefficient should be negative when congruence is less than 50 percent and positive when congruence exceeds 50 percent. Neither of these patterns is visible in the figure. A regression that uses β to predict



congruence for these data has an R^2 of 0.006.⁹ If a researcher was interested in which legislatures were more likely to vote in accordance with majority opinion in their districts, knowing the responsiveness coefficient would not help.

A few examples to illustrate: The chamber with the highest congruence (100 percent), California’s senate in 2003, had responsiveness of $\beta = 3.93$, while the chamber with one of the lowest congruences (38.9 percent), California’s senate in the second insurance law of 1999, had *higher* responsiveness of $\beta = 4.71$. The chamber with the highest responsiveness $\beta = 6.71$, North Dakota’s Senate in 2015, had a low congruence of 35.6 percent, while the chamber with the lowest responsiveness of $\beta = -5.88$, 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 This provides an opportunity for a basic “proof-of-concept” test.

⁹ The coefficient on responsiveness is a statistically insignificant $\beta = 0.83$ (standard error 1.34).

¹⁰ This is not hard-wired into the data. It is possible to have 100 percent congruence on a law that voters repealed, or zero percent congruence on a law they upheld if districts are gerrymandered (Gilligan and Matsusaka, 2006).

Figure 2 shows the two groups in gray and black, respectively. As one would expect, all of the laws that voters sustained were more than 50 percent congruent, and most of the laws that they repealed were less than 50 percent congruent. In this sense, congruence passes the basic test. 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 report comparisons of congruence between several natural groups that have been studied in the representation literature. The question is: to what extent do differences in responsiveness coefficients predict differences in congruence between the groups? Congruence numbers are calculated by summing over all legislators in a given group. Responsiveness coefficients are calculated by estimating model (2) with chamber-law fixed effects using all legislators in the group. Fixed effects allow the propensity to support a law to vary by chamber and law.

In Table 2, Panel A compares approved and repealed laws, to quantify the pattern seen in Figure 2. Congruence was 79.6 percent for approved laws and 48.8 percent for repealed laws, a difference of 30.8 percent that is statistically different from zero in a regression.¹¹ Responsiveness was positive for both groups, but higher for votes on the repealed than the approved laws. The two responsiveness coefficients are statistically different from each other at the 5 percent level. Responsiveness incorrectly implies that voting on the repealed laws was more representative in the sense used here.

Panel B 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 Despite the intention, congruence was essentially identical between lower and upper chambers; the 0.1 percent difference is not statistically different from zero. Similarly, the responsiveness

¹¹ Here and below, comparisons of congruence or responsiveness between groups is done by estimating a single a regression pooling all observations with different coefficients for each group. Standard errors are clustered by chamber-law.

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.

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

Panel A	Approved by Voters	Repealed	Difference	
Congruence (%)	79.6	48.8	30.8	[.001]
Responsiveness (β)	2.69 (0.12)	3.20 (0.18)	-0.51	[.019]
N	2,382	1,810		
Panel B	Lower House	Upper House	Difference	
Congruence (%)	66.3	66.4	0.1	[.992]
Responsiveness (β)	2.93 (0.13)	2.66 (0.15)	0.27	[.191]
N	2,922	1,270		
Panel C	Term Limits	No Term Limits	Difference	
Congruence (%)	61.4	71.3	9.9	[.057]
Responsiveness (β)	3.17 (0.14)	2.67 (0.13)	0.50	[.010]
N	2,119	2,073		
Panel D	Years 1998-2010	Years 2011-2020	Difference	
Congruence (%)	63.9	69.0	-5.2	[.330]
Responsiveness (β)	2.68 (0.21)	2.95 (0.09)	-0.28	[.236]
N	2,186	2,006		

Note. Congruence is calculated using all legislators in a group. Responsiveness is calculated from regression (2) with the addition of chamber-law fixed effects. Standard errors, clustered at the chamber-law level, are in round parentheses. Differences are calculated from regressions that pool all observations, with standard errors clustered at the chamber-law level; *p*-values are shown in square brackets.

coefficients are similar and not statistically distinguishable. In this case, the responsiveness coefficients tell the same story as congruence – that representation was similar in the two chambers.

Panel C compares states that did and did not have term limits on legislators.¹³ There are offsetting theoretical predictions about how term limits affect representation.¹⁴ Term limit advocates argue that term limits bring about the election of “citizen legislators” whose views are more consonant with their constituents, replacing career politicians, leading to superior representation. Skeptics argue that term-limited legislators have less incentive to

¹³ The term limit states were California, Colorado, Maine, Michigan, Ohio, and South Dakota.

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

follow constituent preferences when they no longer face re-election incentives. For the votes studied here, congruence was 9.9 percent lower in term-limit states than other states, with a p -value of 0.057. The responsiveness coefficients point in the opposite direction, with a lower β coefficient for term-limit than non-term-limit states, although they are not statistically distinguishable. Again, responsiveness does not proxy for congruence.

Panel D compares representation across time. With the growth of polarization, one might wonder if representation has fallen over time. I calculate congruence and responsiveness for two periods, 1998-2010 and 2011-2020, that approximately divide the observations in half. Congruence increased 5.2 percentage points between the periods, although the difference is not statistically significant. The responsiveness coefficients point in the same direction: larger in the second period but statistically insignificant. In this case, responsiveness tracks congruence.

Table 3 reports additional comparisons, including some in which the groups are defined based on legislator-specific characteristics. Panel A compares Democratic (plus Green) vs. Republican legislators. While there is no theoretical reason to expect one party to be more representative than the other, recent evidence suggests there may be partisan differences in how accurately legislators perceive constituent preferences (Broockman and Skovron 2018). Democratic legislators voted congruent with their constituents 11.1 percent more often than Republican legislators, a sizeable but not statistically significant difference. The responsiveness coefficients are similar for the two parties. In a regression pooling all observations, with a different coefficient on congruence in the two groups, the responsiveness difference is tiny and statistically insignificant. Responsiveness and congruence are consistent in this case.

Panel B compares congruence of female and male legislators. A growing literature investigates whether female or male legislators behave differently. Anzia and Berry (2011) found that female members of Congress delivered more federal spending to their districts than male members; Stadelmann et al. (2014) reported that female legislators voted more congruently than male legislators in Switzerland; while Ferreira and Gyourko (2014) found 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

Table 3. Congruence and Responsiveness by Legislator Party, Gender, Vote Margins

Panel A	Democrat	Republican	Difference	
Congruence (%)	71.4	60.3	11.1	[.257]
Responsiveness (β)	1.06 (0.24)	1.09 (0.25)	-0.55	[.263]
N	2,291	1,900		
Panel B	Female	Male	Difference	
Congruence (%)	67.9	65.7	2.2	[.234]
Responsiveness (β)	2.49 (0.14)	2.93 (0.11)	-0.10	[.457]
N	1,213	2,979		
Panel C	Roll-call margin < 10%	Roll-call margin > 10%	Difference	
Congruence (%)	68.8	65.7	3.1	[.662]
Responsiveness (β)	3.20 (0.11)	2.72 (0.13)	0.48	[.006]
N	872	3,320		
Panel D	District vote margin < 10%	District vote margin > 10%	Difference	
Congruence (%)	60.4	67.7	-7.3	[.016]
Responsiveness (β)	2.80 (0.31)	2.82 (0.11)	0.18	[.251]
N	762	3,430		

Note. Congruence is calculated using all legislators in a group. Responsiveness is calculated from regression (2) with the addition of chamber-law fixed effects. Standard errors, clustered at the chamber-law level, are in round parentheses. Differences are calculated from regressions that pool all observations, with standard errors clustered at the chamber-law level; *p*-values are shown in square brackets.

legislators, a small and statistically insignificant difference. The responsiveness coefficient is larger for female than male legislators, but not statistically different from zero.

Panel C distinguishes legislators according to whether the roll-call vote was close or lopsided, defined by a vote margin less or greater than 10 percent (the patterns are similar with other cutoffs). Snyder and Groseclose (2000) argue that legislative leaders are more likely to pressure members on close than lopsided roll-call votes because it's costly to persuade a legislator to vote against constituent interests, and there is no need to buy votes when the outcome is a foregone conclusion. Contrary to this argument, congruence was

essentially the same for close and lopsided roll-call voters.¹⁵ The responsiveness coefficient is larger for close than lopsided votes, and statistically different from zero in the pooled regression, contrary to the congruence evidence and the theoretical prediction.

Finally, Panel D compares legislators that represented competitive versus noncompetitive districts. Some activists believe that competition brings about better representation; on the other hand, it may be easier to determine the majority opinion in a one-sided district and thus easier to cast a congruent vote in a noncompetitive district. 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). Legislators representing noncompetitive districts were actually 7.4 percent more congruent than those representing competitive districts, a difference significant at the 5 percent level, consistent with the idea that one-sided districts are easier to represent. The responsiveness coefficients reveal no meaningful differences, again failing to detect a difference that is picked up by congruence.

5. Discussion

This study aims to advance the study of representation by probing empirically two leading methods used to measure representation, “congruence” and “responsiveness.” Although sometimes viewed as interchangeable ways to capture the same underlying concept, previous research has shown that this is not true theoretically. This paper’s core contribution is to evaluate empirically if responsiveness can be used as a proxy for congruence. In a large sample of roll-call votes, I find that responsiveness is not a reliable predictor of congruence; sometimes it works but sometimes it does not. Perhaps the most dramatic failure in my sample comes from comparing laws that voters approved with laws they repealed: congruence was much higher on approved than repealed laws, as one would expect, but the responsiveness coefficients were lower for approved than repealed laws.

¹⁵ Matsusaka (forthcoming), in what is arguably a more parsimonious test, finds that congruence was higher on close votes if the sample is restricted to legislators who were ideologically opposed to their constituents on an issue.

A virtue of congruence is its clear conceptual grounding as an indicator of whether a legislator voted according to the wishes of his or her constituents. Responsiveness, on the other hand, lacks a conceptual grounding as a measure of representation; Achen (1978, p. 475) argued that “correlations are virtually uninterpretable within the framework of democratic theory.” The finding here that responsiveness is not a reliable empirical proxy for congruence underscores the need for research that provides a theoretical grounding for responsiveness as a measure of representation.

One implication of the findings is that empirical researchers should be clear how they are defining representation, and choose the empirical construct that corresponds to their definition. If the definition is similar to the one stated in this paper, then congruence is an appropriate measure and responsiveness is not. There may be cases where responsiveness is the appropriate object of interest, such as in studies of dynamic adjustment, but that should follow from an explicit definition of representation. In cases where it’s not clear what definition to apply, researchers are advised to calculate both congruence and responsiveness for robustness, as in Caughey and Warshaw (2022).

As for previous research that used responsiveness as a proxy for congruence, those findings ought to be interpreted cautiously. It would be informative for researchers to revisit those findings using congruence as the measure of representation. A good example: Bashir (2015), Enns (2015), and Branham et al. (2017) revisited the well-known finding of Gilens and Page (2014) that policy was more responsive to rich than poor Americans, and found that congruence was essentially the same if not reversed for the two groups.

References

- Achen, Christopher H. 1977. Measuring representation: Perils of the correlation coefficient. *American Journal of Political Science* 21: 805-815.
- Achen, Christopher H. 1978. Measuring representation. *American Journal of Political Science* 22: 475-510.
- Anzia, Sarah F. and Christopher R. Berry. 2011. The Jackie (and Jill) Robinson Effect: Why do Congresswomen outperform Congressmen? *American Journal of Political Science* 55: 478-493.
- Barro, Robert J. 1973. The control of politicians: An economic model. *Public Choice* 14: 19-42.
- Bashir, Omar S. 2015. Testing inferences about American politics: A review of the 'oligarchy' result. *Research and Politics* 2: 1-7.
- Becher, Michael and Irene Mendéndez González. 2019. "Electoral reform and trade-offs in representation. *American Political Science Review* 113: 694-709.
- Branham, J. Alexander, Stuart N. Soroka, and Christopher Wlezien. 2017. "When do the rich win? *Political Science Quarterly* 132: 43-62.
- Broockman, David E. and Christopher Skovron. 2018. "Bias in perception of public opinion among political elites. *American Political Science Review* 112: 542-563.
- Caughey, Devin and Christopher Warshaw. 2018. "Policy preferences and policy change: Dynamic responsiveness in the American states. *American Political Science Review* 112: 249-266

- Caughey, Devin and Christopher Warshaw. 2022. *Dynamic democracy: Public opinion, elections, and policymaking in the American states*. Chicago: University of Chicago Press.
- Enns, Peter K. 2015. Relative policy support and coincidental representation. *Perspectives on Politics* 13: 1053-1064.
- Erikson, Robert S., Gerald C. Wright, and J. P. McIver. 1993. *Statehouse democracy: Public opinion and policy in the American states*. Cambridge, UK: Cambridge University Press.
- Ferejohn, John. 1986. Incumbent performance and electoral control. *Public Choice* 50: 5-25.
- Ferreira, Fernando and Joseph Gyourko. 2014. Does gender matter for political leadership? The case of U. S. mayors. *Journal of Public Economics* 112: 24-39.
- Gerber, Elisabeth R. 1999. *The populist paradox: Interest group influence and the promise of direct legislation*. Princeton: Princeton University Press.
- Giger, Nathalie, Heike Klüver, and Christopher Witko. 2020. Electoral vulnerability, party affiliation, and dyadic constituency responsiveness in U. S. legislatures. *American Politics Research* 48: 484-491.
- Gilens, Martin and Benjamin I. Page. 2014. Testing theories of American politics: Elites, interest groups, and average citizens. *Perspectives on Politics* 12: 564-581.
- Gilligan, Thomas W. and John G. Matsusaka. 2006. Public choice principles of redistricting. *Public Choice* 129: 381-398.
- Golder, Matt and Jacek Stramski. 2010. Ideological congruence and electoral institutions. *American Journal of Political Science* 54: 90-106.

- Kousser, Thad. 2008. Term limits and state legislatures. In *Democracy in the states: Experiments in election reform*, ed. Caroline Tolbert, Todd Donovan, and Bruce Cain. Washington D.C.: Brookings Institution Press.
- Kousser, Thad, Justin Phillips, and Boris Shor. 2018. Reform and representation: A new method applied to recent electoral changes. *Political Science Research and Methods* 6: 1-19.
- Krimmel, Katherine, Jeffrey R. Lax, and Justin H. Phillips. 2016. Gay rights in Congress: Public opinion and (mis)representation. *Public Opinion Quarterly* 80: 888-913.
- Lax, Jeffrey R. and Justin H. Phillips. 2012. The democratic deficit in the states. *American Journal of Political Science* 56: 148-166.
- Lax, Jeffrey R, Justin H. Phillips, and Adam Zelizer. 2019. The party or the purse? Unequal representation in the US Senate. *American Political Science Review* 113: 917-940.
- Matsusaka, John G. 2001. Problems with a methodology used to test the responsiveness of policy to public opinion in initiative states. *Journal of Politics* 63: 1250-1256.
- Matsusaka, John G. 2010. Popular control of public policy: A quantitative approach. *Quarterly Journal of Political Science* 5: 133-167.
- Matsusaka, John G. 2020. *Let the people rule: How direct democracy can meet the populist challenge*. Princeton: Princeton University Press.
- Matsusaka, John G. Forthcoming. When do legislators represent their constituents? Evidence from roll-call and referendum votes. *Journal of Politics*.

- Mooney, Christopher Z. 2009. Term limits as a boon to legislative scholarship: A review. *State Politics and Policy Quarterly* 9: 204-228.
- Pitkin, Hanna Fenichel. 1967. *The concept of representation*. Berkeley: University of California Press.
- Romer, Thomas and Howard Rosenthal. 1979. "The elusive median voter. *Journal of Public Economics* 12: 143-170.
- Simonovits, Gabor, Andrew M. Guess, and Jonathan Nagler. 2019. Responsiveness without representation: Evidence from minimum wage laws in the U. S. states. *American Journal of Political Science* 63: 401-410.
- Stadelmann, David, Marco Portmann, and Reiner Eichenberger. 2013. Quantifying parliamentary representation of constituents' preferences with quasi-experimental data. *Journal of Comparative Economics* 14: 170-180.
- Stadelmann, David, Marco Portmann, and Reiner Eichenberger. 2014. Politicians and preferences of the voter majority: Does gender matter? *Economics and Politics* 26: 355-379.
- Tausanovitch, Chris and Christopher Warshaw. 2014. Representation in municipal government. *American Political Science Review* 108: 605-641.