The Concentric Circles of Constituency: Geographic and partisan representation in the U.S. Senate, 1989-2006

John H. Aldrich*
Michael Brady
Scott de Marchi
Ian McDonald
Brendan Nyhan
David Rohde
Duke University

Michael W. Tofias
University of Wisconsin - Milwaukee

Abstract

Research on representation in Congress has typically focused on the relationship between constituent preferences and legislative behavior. But as Fenno points out (1977), legislators may represent other groups than their geographic constituencies. In particular, constituents from the legislator’s party may be particularly influential (Clinton 2006, Bafumi and Herron 2007). Using a broader dataset than previous studies, we measure the linkages between Senate roll-call voting behavior and state geographic and primary constituencies for the 101st, 103rd, 107th, and 109th Congresses. Our findings indicate that the characteristics and preferences of same-party constituents are a much better predictor of Senate voting than those of state residents as a whole and that this relationship appears to have strengthened over time.

* Correspondence can be sent to John Aldrich (aldrich@duke.edu), Department of Political Science, Duke University, 326 Perkins Library, Box 90204, Durham, NC 27708-0204.
Introduction

In his seminal research on the “home style” of members of Congress, Richard Fenno (1977:883) famously asked: “What does an elected representative see when he or she sees a constituency?” His answer was rich and complex. Fenno defined four types of constituencies: the geographic constituency, i.e., all residents in the representative’s district or state; the re-election constituency, comprising all the representative’s election supporters; the primary constituency\(^1\), consisting of the representative’s most reliable and active supporters; and a personal constituency of the representative’s acquaintances and confidants (1977, 1978). The choices representatives make in attending to the desires of these different constituencies are a key aspect of the process of electoral representation in a democracy.

To date, research on representation in Congress has typically focused on the relationship between the geographic constituency and legislative behavior. However, constituents from the legislator’s party may also be particularly influential, especially in the contemporary period (Clinton 2006, Bafumi and Herron 2007). This paper takes a quantitative approach to Fenno’s question for the case of U.S. Senators by measuring the linkages between their roll-call voting behavior and the characteristics of their geographic (i.e. state) and primary constituencies.

Specifically, we estimate statistical models predicting roll call voting behavior in the U.S. Senate during the 101\(^{st}\), 103\(^{rd}\), 107\(^{th}\), and 109\(^{th}\) Congresses using aggregated survey data from the 1988-1992 Senate Election Study (SES) and the National Annenberg Election Study of 2000 and 2004 (NAES). Using a broader dataset than previous studies, we find that the characteristics and preferences of those constituents who identify with the winning candidate’s party (the contemporary equivalent of the primary constituency) are an increasingly better predictor of ideological Senate voting than the characteristics and preferences of state residents as a whole.

\(^1\) Fenno credits the term “primary constituency” to Leo Snowiss (1966).
Theoretical approach

To date, research on constituency and representation has typically focused on the relationship between what Fenno called a legislator’s geographic constituency and their behavior in Congress (e.g. Turner 1951, Miller and Stokes 1963, Fenno 1973, Fiorina 1974, Bailey and Brady 1998, Uslaner 1999). In particular, several studies have attempted to measure legislator “shirking” by examining the errors (i.e. the residuals) from regressions of interest group scores of legislators’ liberalism on the characteristics of state and districts (e.g. Kalt and Zupan 1984). However, we cannot infer from such data that individual legislators are shirking (Goff and Grier 1993). Social choice theory has demonstrated that no unique policy platform generally exists that can beat all others when there are multiple policy issues in question. Without such a uniquely defined winning platform, we cannot measure “shirking,” which suggests deviation from a preferred platform (and which Kalt and Zupan interpret as deviating from the wishes of the constituency to vote for the legislator’s personal policy preferences).

In previous research (Aldrich et al. 2008), we took a different approach, testing the relationship between state demographics (i.e. a senator’s geographic constituency) and voting in the Senate. We found that demographics were more closely related to voting on cross-cutting issues (i.e. the second dimension of DW-NOMINATE) than to party-line voting or voting on ideological issues (i.e. the first dimension of DW-NOMINATE). This tendency was particularly strong during the twenty years or so after World War II when race and civil rights issues split the Democratic Party, but has declined considerably in the period since.

As Fenno clarifies, however, legislators may represent other groups than their geographic constituency. In particular, members of the legislator’s party in the district may be especially
influential in shaping legislative behavior.\textsuperscript{2} There is good reason to believe that representatives attend more closely to this group (their partisan constituency) than to their geographic constituency as a whole. In Fenno’s terms, these supportive partisans can be thought of as the “primary constituency,” which he describes as “[s]trong supporters [who] … tender their support ‘through thick and thin,’ regardless of who the challenger may be” (1978:18).

Two recent studies have shown that this group has a significant influence on legislative behavior in contemporary politics. Using 1999-2000 data from Knowledge Networks and the National Annenberg Election Survey, Clinton (2006) finds that the ideology of same-party constituents has a disproportionate influence on the estimated ideal points of and key votes by members of the 106\textsuperscript{th} House of Representatives (1999-2000). Similarly, Bafumi and Herron (2007) use a novel survey design in the Cooperative Congressional Election Study to show that “members of Congress, both Senators and Representatives, are more representative of state partisans than they are of state medians” in the 109\textsuperscript{th} and 110\textsuperscript{th} Congresses (2005-2008).\textsuperscript{3}

However, these studies – which rely on new datasets that did not exist before 1999 are snapshots that do not allow us to consider changes in the association between constituency characteristics and legislative behavior. This study therefore takes a broader perspective to measure how those associations have changed during the contemporary period. To test this proposition, we compare the performance of models of legislative voting using state-level constituency characteristics with those estimated using only the characteristics of a senator’s fellow partisans. We also employ a broader array of constituency characteristics than previous

\textsuperscript{2} The shift in attention toward the partisan constituency at the district level corresponds to the increasing number of studies focusing on the role of ideological and partisan activists in national politics (Aldrich 1995, Fiorina 1999, Jacobson 2000, Layman and Carsey 2000, Crespin, Gold, and Rohde 2006).

\textsuperscript{3} One mechanism for this level of partisan influence is suggested by Brady, Han, and Pope (2007), who find that members who are ideologically extreme relative to their districts receive a higher proportion of the primary vote and are less likely to be defeated in a primary.
studies, which rely on self-reported ideology (Clinton 2006) or a scaling of voters’ issue preferences (Bafumi and Herron 2007).

**Data**

*Dependent variables*

Our primary dependent variables are first and second dimension DW-NOMINATE scores (Poole and Rosenthal 1991, 1997, 2007), which summarize roll-call voting behavior by members of Congress. Scores on the first dimension typically capture the primary axis of ideological conflict between the parties, which centers on economic policy. These scores account for most of the variance in Congressional voting (typically ninety percent or more in recent years). The second dimension captures a variety of issues that do not line up with the economic policy divide. Most prominently, the issue of race cut across party lines in the mid-20th century (Poole and Rosenthal define its peak as the period 1940-1966) before being integrated into the first dimension during the civil rights era (Carmines and Stimson 1989). Since then, the second dimension has captured other cross-cutting issues such as abortion in the 1970s and trade and immigration more recently.

*Independent variables*

The independent variables required for this analysis need to measure the demographics and policy preferences of residents of each state as well as for identifiers from each of the two major parties. Such data have not been readily available to political scientists until recently. To create a dataset for a longer time period than has previously been available, we aggregate data from several different surveys. For this approach to be valid, our data must meet two key standards.
First, our surveys must ask similar questions in order to make it possible to perform comparisons over time. In addition, there must be a sufficient numbers of respondents in each state to allow us to draw valid inferences about its characteristics.

We collected individual-level constituency data from two sources. First, we used data from the National Election Study’s Pooled Senate Election Study, which ran in 1988, 1990, and 1992. This study consisted of three waves with identical questions that each included approximately 2500-3000 respondents. Moreover, unlike other NES surveys, the SES used the state as a unit of analysis, permitting inferences to that level. Because of our particular interest in studying issues that do not align with the primary dimension of partisan conflict, we dropped the 1990 cross-section, which omitted a question about international trade (one of the few questions tapping the second dimension of DW-Nominate) that was present in the 1988 and 1992 versions. Second, we used data from the 2000 and 2004 National Annenberg Election Study (NAES), which surveyed 75,000 and 85,000 respondents, respectively. The Annenberg studies comprise a large set of individual surveys conducted over the duration of the two presidential election campaigns, including surveys of specific states with primary elections, small panel studies, national cross-sections, and national panels. It, too, permits valid inferences based on state samples.

We identified a set of similar questions across all four studies that covered respondent demographics, party identification, and issue preferences. (See Appendix A for a list of all variables.) We then computed mean values of individual-level data by state and Congress, giving us state-level data for 1988, 1992, 2000, and 2004, which we linked to the corresponding

---

4 For panel respondents, we used their survey response in the first wave.

It is important to note that the sample size of the responses used to construct our independent variables varied by state, survey, and variable. Smaller states tended to have fewer respondents; the NES surveys had fewer respondents than the NAES surveys; and the number of respondents asked a particular NAES question often varied dramatically. For the NES 1988-1992 data, almost every question in the 1988 and 1992 waves had at least 40 respondents per state (with an overall state mean of approximately 65 respondents per question). The NAES data for 2000 and 2004 produced at least 100 respondents per demographic question for each state. The minimum number of respondents for the issue questions was 41; other questions had a minimum of between 62 and 146 respondents per state. Finally, because the NAES surveys omitted Alaska and Hawaii, we dropped respondents from those states in the SES survey. Respondents from the District of Columbia, who are ineligible to vote for senators, were also dropped. (See Appendix A for further details.)

**Empirical approach**

To measure the strength of the electoral connection, we compare the performance of simple OLS models predicting senators’ scores on the two dimensions of DW-NOMINATE using state-level

---

5 We do not estimate any pooled statistical models to avoid any problems of direct comparison between measures on our individual surveys.

6 The worst question was the one on trade policy, which had as few as 17 respondents in one case, but this question had substantially fewer respondents than any of the others.

7 Clinton (2006) uses analytical weights in his OLS models to account for potential heteroskedasticity due to differing sample sizes by district. Supplementary analyses (available upon request) show virtually identical results using analytical weights in our survey data models to those reported below – in almost every case, the resulting model fits correlate at .9 or greater with those generated without weights. As such, we do not employ them. Similarly, Clinton also uses errors-in-variables regression to account for possible measurement error using reliabilities calculated according to the methodology described in Wright, Erikson, and McIver (1985). However, we cannot calculate an overall reliability measure for each variable without neglecting the panel structure of our data.
measures of demographics and issue preferences calculated from the NES and NAES surveys.

Our goal is to evaluate the explanatory power of demographics and issue preferences in explaining voting in Congress over time. To do so, we estimate the following models for each of the four sessions of Congress in our data:

“Demographics-only” model:

\[
\text{DWNOM} = f(\% \text{ seniors}, \% \text{ black}, \text{ median income}, \% \text{ union}, \% \text{ Latino})
\]  \hspace{1cm} (1)

“Issues-only” model:

\[
\text{DWNOM} = f(\text{ abortion, defense, education, health care, trade})
\]  \hspace{1cm} (2)

We first compare the explanatory power of our survey-based demographics-only model for senators’ geographic constituencies with an equivalent demographics-only model estimated using Census data (Aldrich et al. 2008). This step is intended to serve as a validation check of both analyses. If results are consistent across datasets, our confidence in them will increase. We then compare the relative power of geographic and partisan constituency characteristics to explain roll-call voting behavior. To do so, we compute separate values of each state-level demographic and issue variable for party identifiers from our survey data. We then contrast the

---

8 Given survey respondents' well-known difficulties understanding ideological labels, we prefer to enter issue variables directly into the statistical model rather than relying on self-reported ideology (as in Clinton 2006).

9 Additional analyses showed little difference in model fit between combined demographics/issues models and the demographics-only and issues-only models reported here. They are thus omitted but available upon request.

10 Appendix B tabulates the independent variables used in the previous paper and those used in this chapter. In this paper, we used presidential favorability/approval as a survey-based proxy for the Democratic presidential vote variable used in our previous analysis.

11 In particular, it is important to validate our results using Census data due to the possibility that differing sample sizes or question wording in the NAES data will lead to spurious differences in model fit.
predictive power of these variables for senator(s) from the same party with measures calculated for the state population as a whole.

To evaluate model performance, we rely on adjusted $R^2$ – a well-known model fit metric that includes a penalty term for the use of additional explanatory variables. Because of our interest in model fit and the large number of models estimated for this chapter, we summarize our results in a series of figures below. (Results for all models are available upon request.)

**Statistical results**

Figure 1 presents a plot of adjusted $R^2$ values by Congress of our “demographics-only” model for each dimension of DW-NOMINATE using our survey and Census data. To put the results from our survey data in perspective, we plot the Census results since 1973, the first Congress after the 1970 Census was completed. (We chose this period because it roughly corresponds to the post-civil rights era.)

[Figure 1]

We see that the model fit for the first dimension using the aggregated survey data improved dramatically between 1989-1990 (the 101\textsuperscript{st} Congress) and 2005-2006 (the 109\textsuperscript{th}), with an adjusted $R^2$ over .4 for the 109\textsuperscript{th} Congress. Reassuringly, these results are highly consistent with the Census data, which also show increased predictive power for the first dimension over the 1989-2004 period. By contrast, both datasets show a dramatic decline in model fit for the second dimension (a trend that the Census data suggests started after the 103\textsuperscript{rd} Congress of 1993-1994).
To compare the relative influence of the party and geographic constituencies, we next calculate the independent variables above separately for identifiers with the senator’s party in our survey data.\textsuperscript{12} We then estimate models of DW-NOMINATE scores using a “demographics-only” model for both all respondents and party identifiers.\textsuperscript{13} Figure 2 plots the results for both dimensions of DW-NOMINATE.

![Figure 2]

In this figure, we see that the party demographics model clearly outperforms the geographic constituency demographics, suggesting that the first dimension of legislative voting in the Senate has become increasingly well explained by the demographics of identifiers with the senator’s party in their state. By contrast, senators’ second dimension voting is less well explained. However, we do observe better model fits for the partisan constituency model than the geographic constituency model in the 107\textsuperscript{th} and 109\textsuperscript{th} Congresses. This finding may indicate that partisan responsiveness on the second dimension has increased in recent years.

Next, Figure 3 compares model fits using the issue preferences of party identifiers with the issue preferences of constituents as a whole for both dimensions of DW-NOMINATE.

![Figure 3]

\textsuperscript{12} To maximize the sample size of our partisan constituencies, we include leaners from the party ID questions.  
\textsuperscript{13} We also ran models of DW-NOMINATE using independent variables calculated only for those respondents who reported voting or an intention to vote; the results were largely indistinguishable from those for all respondents and are therefore omitted (but are available upon request).
We find that partisan identifiers’ issue preferences are a much stronger predictor of the primary dimension of Congressional roll-call voting than are the issue preferences of geographic constituencies as a whole for all four Congresses in the sample. As in the previous analysis, second dimension model fits are again lower in general than first dimension scores. However, we again observe a somewhat improved fit in recent sessions of Congress for the model using the issue preferences of party identifiers relative to state constituencies as a whole.\footnote{When we estimated a combined model of demographics and issues for both all respondents and party identifiers, we also found that model fit on the first dimension of DW-NOMINATE was much higher for party identifiers than all respondents (results available upon request).}

To put the results from Figures 2 and 3 in perspective, we compare model fit between the geographic and partisan constituency models of DW-NOMINATE and a combined model in which both sets of measures are included as predictors. Our goal in doing so is to assess whether the geographic measures provide additional explanatory power compared to the demographic measures model. Figure 4 presents the results for the first dimension of DW-NOMINATE.

As in Figures 2 and 3, the partisan constituency model has substantially higher levels of model fit for all four Congresses in our sample. More importantly, however, we find that the fit of the combined model is very similar to that of the partisan constituency model for both the demographic (left panel) and issue variables (right panel), indicating that the geographic variables provide little additional explanatory power when added to the partisan model.\footnote{It is true that we can reject the null hypothesis that the geographic variables are jointly zero at the $p < .05$ level of greater in each Congress for the combined models. However, this is a much weaker test.}

Figure 5 presents comparable results for the second dimension of DW-NOMINATE.
Echoing Figures 2 and 3, we observe that the partisan constituency model provides greater levels of model fit than the geographic model in the 107th and 109th Congresses. But in contrast to Figure 4, we also observe that the combined model fits substantially better than either the partisan or geographic constituency model. Second dimension voting appears to reflect the characteristics and preferences of both geographic and partisan constituencies.

Overall, we find that the ideological roll call voting records of members of the Senate (as measured by the first dimension of DW-NOMINATE) are more closely associated with the demographic characteristics and issue preferences of party identifiers in their states than with the corresponding aspects of their state constituencies as a whole. These results provide support for the Clinton (2006) and Bafumi and Herron (2007) findings over a much wider time period. In particular, the association between the demographics of partisan identifiers and DW-NOMINATE first-dimension scores appears to have increased dramatically in recent years. By contrast, the association between constituency characteristics and roll-call voting on cross-cutting issues (as measured by the second dimension of DW-NOMINATE) is relatively weaker. The differences in model fit between partisan identifiers and the state constituency are also less clear for the second dimension. However, the partisan identifier model appears to fit the data better for both the 107th and 109th Congresses, which may indicate that responsiveness to the primary constituency has increased in recent years.
Conclusion

This paper uses survey and Census data to distinguish between the views of the general public and party identifiers, an empirical application of Fenno’s distinction between the geographic and primary constituencies. We find that the issue preferences and demographics of a senator’s fellow partisans are a significantly better predictor of first-dimension DW-NOMINATE scores than the state constituency as a whole and that this gap has widened over time. The differences are much less pronounced on the second dimension, though there is some indication of increased responsiveness to partisan identifiers in recent years.

These findings are consistent with the idea that partisan polarization may extend to the level of demographics (Abramowitz and Saunders 2005; Aldrich et al. 2006a, 2006b; Jacobson 2006). As the explanatory power of the second dimension of roll-call voting in Congress declines (Poole and Rosenthal 2007), the demographics of party identifiers are becoming more closely aligned with the main dimension of partisan conflict. The challenge for political scientists is to determine how this process works and to generate improved measures of the characteristics and preferences of supporting partisans. As Fenno reminds us, elected representatives do not see only the geographic constituency – a statement that is even more true today than when he first wrote it more than thirty years ago.
Bibliography


Figure 1: Performance of demographics-only models of Senate DW-NOMINATE scores 1973-2006
Figure 2: Performance of partisan and overall demographics models of Senate DW-NOMINATE scores 1989-2006
Figure 3: Performance of partisan and overall issue preference models of Senate DW-NOMINATE scores 1989-2006
Figure 4: Performance of geographic and partisan constituency models of first dimension
Senate DW-NOMINATE scores 1989-2006

[Diagram showing performance of geographic and partisan constituency models over time, with plots for both Demographics and Issues, and lines for All constituents, Partisan identifiers, and Combined model.]
Figure 5: Comparison of geographic and partisan constituency models of second dimension Senate DW-NOMINATE scores 1989-2006

Demographics

Issues

Adjusted R²

Congress start year

Adjusted R²

Congress start year

All constituents
Partisan identifiers
Combined model
### Appendix A

<table>
<thead>
<tr>
<th>Type</th>
<th>NES: Responses by state</th>
<th>NAES: Responses by state</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Type</td>
<td>Mean</td>
</tr>
<tr>
<td>Age 65+</td>
<td>Binary</td>
<td>60.0</td>
</tr>
<tr>
<td>Black</td>
<td>Binary</td>
<td>59.8</td>
</tr>
<tr>
<td>Hispanic</td>
<td>Binary</td>
<td>60.3</td>
</tr>
<tr>
<td>Income</td>
<td>Continuous</td>
<td>60.3</td>
</tr>
<tr>
<td>Unemployed</td>
<td>Binary</td>
<td>60.1</td>
</tr>
<tr>
<td>Union member</td>
<td>Binary</td>
<td>59.8</td>
</tr>
<tr>
<td>Anti-abortion</td>
<td>Binary/categorical (3,5)</td>
<td>57.4</td>
</tr>
<tr>
<td>Pro-defense spending</td>
<td>Categorical (3,4)</td>
<td>58.5</td>
</tr>
<tr>
<td>Pro-school spending</td>
<td>Categorical (3,4)</td>
<td>58.8</td>
</tr>
<tr>
<td>Pro-health spending</td>
<td>Categorical (3,4)</td>
<td>58.1</td>
</tr>
<tr>
<td>Opposition to trade</td>
<td>Binary/categorical (4,5)</td>
<td>38.6</td>
</tr>
</tbody>
</table>

---

16 As noted in the text, the widely varying number of responses for NAES variables is not an indication of high levels of non-response. Most issue questions were simply asked much less frequently than the standard demographic questions.

17 Income measures were constructed from the median income level by state and transformed to the midpoint of the range of that category. For instance, a median response of $30,000-$40,000 was coded as $35,000. Income values were not adjusted for inflation because we did not estimate a pooled model comparing states across time.

18 101st: 3 point; 103rd: 3 point; 107th: binary; 109th: 5 point.
19 101st: 3 point; 103rd: 3 point; 107th: 4 point; 109th: 4 point.
20 101st: 3 point; 103rd: 3 point; 107th: 4 point; 109th: 4 point.
21 101st: 3 point; 103rd: 3 point; 107th: 4 point; 109th: 4 point.
22 101st: Binary; 103rd: Binary; 107th: 4 point; 109th: 5 point.
Appendix B: Variables in NES/NAES and Census models

Figure 1 presents a comparison with previously estimated models predicting senators’ scores on the two dimensions of DW-NOMINATE (Aldrich et al. 2008). The following table describes the variables used in this analysis.23

<table>
<thead>
<tr>
<th>Variable</th>
<th>Both models</th>
<th>NES/NAES</th>
<th>Census</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age 65 and over</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black residents</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Farmers and farm workers</td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Finance workers</td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Foreign born</td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Government workers</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Hispanic</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Income24</td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Manufacturing workers</td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Population density</td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Total population (log)</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Unemployed</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Union members</td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Urban population</td>
<td></td>
<td></td>
<td>X</td>
</tr>
</tbody>
</table>

23 All demographic variables representing subpopulations (number of African Americans, farmers, etc.) were transformed to proportions of total state population.

24 In the Census data, income is recorded as the log of per capita income. The NES/NAES data used the raw value of median income by state.