Measuring Constituency Ideology in U.S. House Districts: A Top-Down Simulation Approach

Phillip J. Ardoin
Appalachian State University
James C. Garand
Louisiana State University

One of the most intractable problems associated with studying representation in the U.S. House of Representatives involves the measurement of district-level constituency opinion. In measuring constituency opinion in House districts, scholars have relied on a number of alternative approaches, including the use of demographic variables, small-sample estimates of public opinion, presidential election results, referenda data, and “bottom-up” simulated opinion. In this article we develop an innovative “top-down” simulation of House district opinion that provides more reliable and valid measures of House district ideology. We model state-level ideology (as measured by Erikson, Wright, and McIver 1993) as a function of various demographic and political variables found at both the state and House district levels, and then use the estimates from the state-level model to generate predicted ideology scores for each House district during the 1980s and 1990s. Our findings suggest that the top-down simulated measure is a valid indicator of House district ideology that can be used in a number of research venues.

Representation is particularly important to democratic theory. According to Eulau et al. (1959), representation is what makes democratic legislatures both legitimate and authoritative decision-making institutions. Likewise, Key (1961) notes, the representation of public opinion is widely regarded as one of the distinguishing characteristics of democracy. Dahl (1989) further contends that the primary justification for democratic governments is found in their ability to provide an orderly and peaceful process for representation to occur.

Considering the important role that representation plays in democracies, the study of representation has occupied a central place in the field of political science and has long captivated the interest of both theorists and empirical analysts. Although scholars have explored a range of representational behaviors, the focus of much of this research has been on what Eulau and Karps (1978) have referred to as policy responsiveness, which we define as the degree to which legislators reflect in their roll-call behavior the policy views of their constituents.

The concept of policy representation requires at least some level of congruence between the policy views of constituents and the policy behavior of repre-
sentatives. Using a variety of data sources and analytical approaches, scholars have mapped out some of the basic contours of the linkage between constituency policy preferences and legislative behavior (Erikson 1978; Glazer and Robbins 1985; McCrone and Kuklinski 1979; Miller and Stokes 1963). Moreover, other scholars have gone a bit further by exploring variation in representational behavior across legislators and constituency contexts (Ahuja 1994; Glazer and Robbins 1985; Herrick, Moore, and Hibbing 1994; Hibbing 1984; Jackson and King 1989; McCrone and Kuklinski 1979; Thomas 1985).

Clearly, studying the relationship between constituency policy preferences and legislative roll-call behavior requires one to be able to measure these two concepts. Legislative roll-call data are readily available over time and across a range of legislative bodies, so the dependent variable in studies of policy congruence can usually be obtained without difficulty. Unfortunately, the same cannot be said about the key independent variable, constituency ideology. For the most part, data on constituency preferences are either unavailable or very difficult to collect. While reliable data on constituency policy preferences at the state level are available (Berry et al. 1998; Erikson and Palfrey 1993), data on constituency policy preferences in legislative districts below the state level are almost impossible to obtain. This has meant that legislative scholars have had a difficult time in studying policy responsiveness in the U.S. House of Representatives, state legislatures, and local legislatures such as city councils and school boards.

Of course, the fact that policy responsiveness is difficult to study does not mean that legislative scholars have not tried, particularly with regard to the U.S. House of Representatives. Over the years legislative scholars have struggled to measure the ideological and policy dispositions of House members’ constituencies. They have utilized a number of techniques and data sources to estimate constituency opinion in U.S. House districts, including small-sample estimates of public opinion (e.g., Miller and Stokes 1963), demographic variables (Erikson 1978; Jackson and Kingdon 1992; Page et al. 1984; Pool, Abelson, and Popkin 1965; Weber and Shaffer 1972; Wink, Livingston, and Garand 1996), presidential election results (Fleisher 1993; Glazer and Robbins 1985; Johannes 1984; LeoGrande and Jeydel 1997; Nice and Cohen 1983), congressional roll-call votes (Holbrook-Provow and Poe 1987), referenda voting (Erikson, Luttbeg, and Holloway 1975; Hedlund and Friesema 1972; Kuklinski 1977; McCrone and Kuklinski 1979) and “bottom-up” simulated opinion based on the extension of estimates from individual-level models to the aggregate district level (Erikson 1978; Sullivan and Minns 1976; Sullivan and Uslaner 1978; Uslaner and Weber 1979; Weber et al. 1972). Due to questions of validity, these measures of constituency opinion have proven to be considerably less satisfying than what one would obtain if large-sample estimates of opinion were available across all districts. Unfortunately, the great expense of conducting district-by-district surveys means that such large-sample estimates of opinion are not available for House districts, nor are they likely to become available in the foreseeable future.
In this article we suggest a potential solution to this problem. In particular, we modify the traditional simulation approach used by scholars such as Erikson (1978) and Uslaner and Weber (1979) in order to develop valid and more reliable estimates of constituency liberalism in House districts. However, rather than the traditional “bottom-up” simulation, in which data from a lower level of aggregation (i.e., from individual-level surveys) are used to simulate opinion at a higher level of aggregation (i.e., the district level), we develop a “top-down” simulation approach, which involves using data from a higher level of aggregation (i.e., the American states) to simulate opinion at a lower level of aggregation (i.e., the House district level). Based on data on mass state ideology compiled by Erikson, Wright, and McIver (1993), we estimate the relationship between state opinion ideology, on one hand, and various demographic variables, presidential election results, and other independent variables, on the other. Once we have estimated the parameters of the state-level model of constituent ideology, we substitute analogous data from House districts into the model to yield predicted levels of congressional district ideology. We suggest that the result is valid estimates of House district ideology that reflect the observed relationships between constituent ideology and various independent variables at the state level, as well as the values on these independent variables in U.S. House districts.

Measuring Constituent Ideology in U.S. House Districts

Given the substantial amount of attention paid to policy responsiveness in the U.S. House, it is clear that legislative scholars have had to develop surrogate measures of constituency preferences at the House district level. Most scholars agree that the prevailing measures in the literature are less than optimal. Ideally, one would want to have data drawn from large-scale samples in each of the 435 congressional districts and then use these data to derive point estimates of constituency preferences. Unfortunately, such an approach would strain the limited resources available for social-science research in the United States, and to date no such data collection effort has been made. Given this, scholars have had to rely on available data to provide rough estimates of House district ideology.

Small Sample Estimates of District Opinion

In lieu of large-sample estimates of district opinion, scholars have employed national samples with relatively small numbers of respondents from each of the sampled congressional districts. For instance, in their seminal work on legislative representation Miller and Stokes (1963) utilized data from the 1958 American National Election Study survey to estimate policy preferences at the district level and explore how those preferences were related to legislators’ roll-call behavior. However, the mean sample size for House districts was merely 11 usable respondents, which is considerably below the sample sizes used in conventional surveys to estimate public opinion. Consequently, the resulting estimates of
constituency opinion were extremely inefficient. While Erikson (1978) does note that point estimates derived from such small random samples are unlikely to be biased—i.e., such small samples are equally likely to overestimate opinion as underestimate opinion—the sampling error associated with measures based on such small samples is likely to be exceedingly high. Accordingly, the estimates of constituent opinion derived from small samples are likely to be relatively far removed from the actual (unobserved) opinion of the constituents. All in all, while the use of small-sample estimates helped to establish the literature on legislative responsiveness to constituency policy preferences, the significant limitations of this approach render it inadequate for studies of legislative policy responsiveness.

Demographic Variables

One of the most common approaches used in studies of policy responsiveness in the U.S. House is to measure constituency policy preferences using surrogate demographic variables. Usually this involves estimating a model in which legislative roll-call behavior is depicted as a function of a wide range of district demographic characteristics obtained from the U.S. Census. The demographic variables employed in such studies typically include indicators of racial composition, education, income, age, social class, occupational distribution, urbanization, home ownership, and family composition (Pool, Abelson, and Popkin 1965; Sinclair-Deckard 1976; Weber and Shaffer 1972).

Scholars adopting such an approach make some important assumptions about the political meaning of demographic characteristics. In particular, one assumes that (1) individuals’ demographic characteristics are related systematically to their policy preferences, (2) legislators are aware of the demographic composition of their districts and take those characteristics (or at least how they interpret those characteristics) into account when making roll-call decisions, and (3) such a relationship holds when one moves across levels of analysis (i.e., from the individual level to the aggregate level).

The first assumption is quite reasonable. Numerous studies have documented the demographic underpinnings of public opinion and political behavior; citizens’ general ideology and their views on public policy matters are often found to be related to their demographic characteristics. Such a relationship may be due to the degree to which self-interest is reflected in citizens’ demographic characteristics, or else demographic characteristics might represent how different groups in society attain different sets of symbolic attitudes through the socialization process.

Second, are legislators aware of the demographic characteristics of the constituents that they represent, and do they interpret these characteristics in such a

\footnote{Erikson (1978) also points out that the relationships between known district demographic characteristics and estimates derived from small samples are typically weak, which suggest that the small-sample estimates of constituency opinion are similarly inefficient.}
way as to permit the demographic flavor of a district to affect their roll-call decisions? Although there is only limited systematic evidence on this point (Erikson, Luttbeg, and Holloway 1975; Hedlund and Friesema 1972), Fenno (1978) suggests that members of the U.S. House think about their geographic constituencies in terms of a number of attributes, including demographic characteristics. It is reasonable to think that legislators are aware of the demographic characteristics of their districts and consider those attributes in understanding their districts. And, it should be noted, it is entirely plausible for House members and other legislators to consider politically relevant demographic characteristics as they cast roll-call votes.

The final assumption—i.e., that the relationship between aggregate demographic characteristics and aggregate policy preferences is a reflection of the same relationships at the individual level—is less certain, since making such an assumption has the potential of violating classic notions of the ecological fallacy. Simply, processes that operate at the individual level need not be in effect at the aggregate level. Although relationships found at the individual level often persist at the aggregate level, one must clearly take great care in making inferences about political processes across levels of analysis.

Ultimately, studies that rely on demographic variables to represent constituency influences are quite limited. Demographic characteristics are only imperfectly related to policy preferences among individual citizens. Although demographic variables might have significant impact on individuals’ policy preferences, they typically explain only a small amount of the variance in such preferences, and this means that roll-call models that simply rely on demographic variables are missing a substantial portion of the effect of constituency preferences. Moreover, the uncertainty surrounding the policy implications of demographic variables means that the policy signals directed at legislators by their constituents’ demographic characteristics are somewhat ambiguous. Knowing, for instance, that a district has a high proportion of its citizens with a college education does not necessarily give a legislator clear, unambiguous signals about the policy preferences of constituents, since this demographic characteristic, like others, is not perfectly related to policy preferences.

Presidential Election Results

Numerous scholars have also relied on presidential election results as a surrogate measure of district ideological orientation (Fleisher 1993; Glazer and Robbins 1985; Johannes 1984; LeoGrande and Jeydel 1997; Nice and Cohen 1983). The logic underlying this is grounded in standard spatial models of electoral choice. Arguably, many citizens cast their votes in presidential elections by comparing their own ideological positions with those of the competing candidates. Insofar as aggregate presidential election results reflect ideological voting in the electorate, scholars should be able to utilize presidential election results at the district level as a proxy measure of district ideology.
One of the problems with this approach is that not all presidential elections are equally ideological in nature. Most observers agree that certain presidential elections are highly ideological and that the presidential election results from those elections reflect the ideological characteristics of constituencies; the 1964, 1972, and 1988 elections come immediately to mind as elections in which support for the Democratic and Republican presidential candidates was differentiated by ideological considerations. On the other hand, some elections are known for their detachment from ideology; the 1968 and 1976 elections are often characterized as being somewhat less ideological than other elections. Clearly, not all presidential elections are equally ideological, and this affects the degree to which scholars can use district-level presidential elections results as a surrogate for district ideology.

Several scholars have investigated the ideological nature of aggregate presidential election results. Rabinowitz, Gurian, and MacDonald (1984) and Rabinowitz and MacDonald (1986) have utilized factor-analytic techniques to determine the dimensionality of presidential election results at the state level. They uncover two dimensions, one of which is an ideological dimension, the other a partisan dimension. Given these results, it would appear that presidential election results in general are not exclusively ideological in nature, though some elections are more ideological than others.

More recently, LeoGrande and Jeydel (1997) explore the possibility of utilizing presidential election results as a surrogate for district ideology. They find only moderate correlations for presidential election results between adjacent elections, suggesting that the reliability of the aggregate presidential vote is not extremely high. Ultimately, this leads them to conclude that the presidential vote is not a consistent proxy for the long-term effects implied by constituency ideological orientations. Moreover, LeoGrande and Jeydel (1997) confirm the findings of Rabinowitz, Gurian, and MacDonald (1984) and Rabinowitz and MacDonald (1986) that presidential election results can best be represented by two dimensions, one partisan and the other ideological. While the ideological dimension is the more important of the two, LeoGrande and Jeydel (1997) contend that the consistently significant partisan dimension suggests that ideology is not the only long-term factor that affects presidential voting.

Unfortunately, one cannot readily utilize factor scores derived from a factor analysis of district-level presidential election results to measure ideology at the district level. House districts change after every redistricting period, so it is impossible to utilize district-level presidential election data across decades in a factor analysis designed to create a general district ideology score. Even between census years and redistricting, it is difficult to utilize this approach in measuring district ideology. A factor analysis of presidential election results during, say, the 1970s would include one election (i.e., 1972) that is considered highly ideological, but would also include a nonideological election (i.e., 1976). Some presidential elections (i.e., 1964, 1972, 1988, and 1992) load more heavily on the ideological dimension, and LeoGrande and Jeydel (1997) suggest that scholars
should only utilize data from these elections as a proxy measure of district ideology.

**Referenda Results**

One of the most innovative approaches to measuring district ideology involves the use of referenda data. In referenda elections, voters are confronted with one or more policy positions on which they can express their preferences. A number of states hold referenda elections on a regular basis, and scholars have found it possible to utilize district-level data on referenda election results to estimate the policy preferences and/or ideological orientation of a given constituency.

The use of referenda data as a surrogate measure of constituency policy preferences is best represented by the work of Kuklinski (1977) and McCrone and Kuklinski (1979). In both studies the authors utilize data from California referenda to estimate the positions of district constituencies on three dimensions that emerge from a factor analysis of the referenda data. While these scholars find that referenda data can provide quite reliable measures of district ideology, unfortunately such data are available for only a limited number of states. While referenda data measures can be used successfully in the limited number of states that utilize referenda elections, scholars must rely on other sources of data to measure constituent ideology for the vast majority of districts.

**“Bottom-Up” Simulations**

Another innovation in the measurement of district opinion and constituency policy preferences is the use of simulated district opinion, a technique developed by Weber and Shaffer (1972) and subsequently utilized by several legislative scholars (Erikson 1978; Sullivan and Minns 1976; Sullivan and Uslaner 1978; Uslaner and Weber 1979). This approach is designed to take advantage of demographic data that are available at the district level, as well as knowledge concerning the relationship between individuals’ demographic characteristics and their policy positions. In traditional simulations of constituency opinion, scholars utilize what we refer to as a “bottom-up” simulation—i.e., data from a lower level of aggregation (i.e., from individual-level surveys) are used to simulate opinion at a higher level of aggregation (i.e., the district level). This involves the following steps. First, survey data are used to estimate the relationship at the microlevel between individuals’ various demographic characteristics (e.g., race, income, etc.) and their opinions or attitudes (e.g., liberalism). Care is taken to select survey items that are analogous to variables found at the aggregate (district) level. Second, once the individual-level regression estimates are obtained, the mean district characteristics on these independent variables are substituted into the regression model to yield predicted district-level opinion. The implication of this approach is that, if individual-level variables are related to individual-level opinion, then aggregations of those individual-level variables should be related to aggregations of (district-level) opinion.
On the face of it, this approach appears to be quite reasonable. The logic underlying the approach seems to be sensible, and simulated measures of opinion have been found to have a stronger association with roll-call behavior than measures based on small-sample estimates (Erikson 1978). Most importantly, the general availability of demographic and political variables with which to simulate public opinion means that this approach can be applied to estimating district opinion across a wide range of districts and across time.

Perhaps the most important concern that one might have with this approach is that the individual-level regressions upon which the simulations are based often exhibit exceedingly low levels of fit to the data. With adjusted R² levels that often fall below .20, measures of simulated district-level opinion have a significantly large amount of random error associated with them. This is not necessarily a surprise, since the level of measurement error in individual-level survey data is often much higher than that found in aggregate-level data. Ultimately, while bottom-up simulated measures may be an improvement over those obtained from other analytical approaches, they remain somewhat imprecise indicators of constituency opinion.

An Alternative Approach: “Top-Down” Simulation

Clearly, scholars have struggled to measure constituency opinion in U.S. House districts. Unfortunately, the various approaches found in the literature are less than satisfactory, either due to their availability for only a limited set of districts or years or because of their high levels of measurement or sampling error. Given that previous measures of district opinion have proven to be inadequate, we develop a more reliable and useful measure of district opinion by employing an alternative approach to simulating constituency ideology. Specifically, we use a “top-down” simulation approach, which involves using data from a higher level of aggregation (i.e., the American states) to simulate opinion at a lower level of aggregation (i.e., congressional districts).

First, we use state-level data to estimate the relationship between state ideology, as measured by Erikson, Wright, and McIver (1993), and various demographic and political variables. Once we estimate the parameters of the model of state ideology, we substitute analogous data from U.S. House districts into the model to yield predicted levels of opinion liberalism for House districts. The results provide improved estimates of House district ideology that reflect the observed relationships between constituent ideology and various independent variables at the state level, as well as the values on these independent variables in U.S. House districts.

It should be noted that we estimate separate models of state ideology for the 1980s and 1990s.² Individual models for each decade are necessary due to possible changes in the relationship between various independent variables and

²The districts demographics employed for the 1980s model were based on the districts recognized by the U.S. Congress for the 98th–102nd Congresses (1982–1990) and for the 1990s model the districts recognized by the U.S. Congress for the 104th–107th Congresses (1994–2000).
constituency ideology over time and to the unavailability of presidential election results across redistricting periods.

**Modeling State-Level Constituency Ideology**

The first step in developing a “top-down” simulated measure of U.S. House district ideology is to establish a model of constituency ideology at the state level that can be used to simulate U.S. House district constituency ideology. The model must include a range of independent variables measured at the state level that are analogous to variables available for congressional districts.

The dependent variable in our model is state ideology, as measured originally by Wright, Erikson, and McIver (1985) and Erikson, Wright, and McIver (1993). Utilizing data from the CBS News/New York Times surveys from 1976 to 1988, Erikson, Wright, and McIver (1993) created large-sample estimates of state political ideology. By aggregating respondents’ ideological self-reports across surveys, Erikson and his colleagues created samples for each state that were sufficiently large to provide quite reliable estimates of state political ideology. Utilizing questions measuring respondents’ ideological identity, they coded respondents as liberal, moderate, or conservative, and then measured the ideological orientation for each state as the mean ideology score for respondents residing in each state.3

We include in our model several independent variables that one might expect to be related to state ideology and for which data are available for U.S. House districts. First, following LeoGrande and Jeydel’s (1997) argument that presidential election results from highly ideological elections can be used as surrogates for district ideology, we include in our model a variable to represent the effects of state-level presidential election results. For the 1980s model of state ideology, Democratic presidential vote is measured as the average percentage of the vote for Democratic candidates in the 1984 and 1988 presidential elections. For the 1990s model of state ideology Democratic presidential vote is measured as the average percentage of the vote for Democratic candidates in the 1992 and 1996 presidential elections.4 We hypothesize that the coefficient for Democratic

---

1 While Erikson, Wright, and McIver’s published measures of state and partisan ideology only include data for the period 1976–1988, they have provided us with their updated database of CBS News/New York Times surveys through 1992. Utilizing this database and their methodology, we have updated the measures of state and partisan ideology through 1992. While the updated data do provide measures of state ideology that are more up-to-date and reliable, it is noteworthy that the updated measures of state ideology are highly correlated with the original measures of state ideology ($R^2 = .901$). The updated variable is coded so that high scores indicate more conservative states.

2 These four presidential elections are selected using a principal components factor analysis that included all presidential elections between 1976 and 1996 and Erikson, Wright, and McIver’s (1993) state ideology scores. The four selected elections loaded most heavily on an ideological dimension with the state ideology scores. A principal components factor analysis conducted with the four presidential elections included in our model and Erikson, Wright, and McIver state ideology measures confirm the ideological nature of each election. Each of the elections loads heavily on the single ideological dimension.
presidential vote, $b_1$, will be negative, indicating that states in which the Democratic presidential candidate makes a strong showing will be less conservative than those states in which the Democratic candidate runs poorly.

We also include a series of demographic variables found in previous research to be related to mass state ideology. These include the proportion of blue-collar workers in the state workforce, the proportion of the population living in urban areas, the proportion of homeowners among all households, and a dichotomous variable representing the Deep South region. We hypothesize that the size of the blue-collar workforce and the number of urban residents will be negatively related to state conservatism, while home ownership and the Deep South region will be positively related to state conservatism.\(^5\)

Given this, our first-stage model upon which our estimates of House district ideology are based is as follows:

\[
\text{State Ideology} = a + b_1 (\text{Average Democratic Presidential Vote}) + b_2 (\% \text{ Blue Collar Workers}) + b_3 (\% \text{ Urban Population}) + b_4 (\% \text{ Home Ownership}) + b_5 (\text{Deep South})
\]

\textit{Estimating Ideology in U.S. House Districts}

It is important to note that the major reason for estimating a model of state-level ideology is to permit us to estimate the ideological orientations of U.S. House district constituencies. Once we estimate our model of state ideology, we substitute U.S. House district-level values on each of the independent variables in the model and multiply each value by its associated unstandardized regression coefficient. The resulting predicted values represent the estimated ideological orientation of each U.S. House district, based on (1) the observed relationship between these variables and state ideology and (2) the values of these variables for each U.S. House district.

\(^5\) It should be noted that data were collected for a wide range of state and district-level characteristics that could conceivably serve as theoretically plausible independent variables in a model of state or district ideology. In addition to the variables listed in our model, data on over a dozen alternative state and district level characteristics were collected and tested in our model, including industry/occupation categories, education levels, individual/family income, home ownership, marital status, family size, and race and ethnicity. These variables were not included in our model for one of three reasons: (1) their coefficients were statistically nonsignificant or they did not otherwise improve the fit of the model; (2) the relationship between the variables and constituency ideology differed between states and House districts; or (3) reliable measures were not available at both the state and House district levels. Considering that the fundamental goal of our model is to develop reliable measures of simulated House district ideology, we utilize only demographic variables that improve the fit of the model and that exhibit a relationship with the Democratic presidential vote that can be shown to be similar at both the state and House district levels. Bivariate correlations are employed to confirm similar relationships between the demographic variables employed in our model and the proportion of constituents who voted for the Democratic presidential candidate, which is a relatively reliable measure of constituent ideology that is available at the state and House district level. (Interested readers can obtain a copy of these correlations by writing the lead author.)
A Caveat: The Race Problem

Any casual observer of American politics is aware of the relationship between race and ideology. Black and other minority citizens are, on average, more liberal than nonminority citizens (Tate 1993). In the aggregate, it is not unreasonable to expect that districts or states with large black or minority populations will be significantly more liberal than districts or states without large black or minority populations. Moreover, black and other minority legislators are generally more liberal than their nonminority colleagues, though they often depend on large black or minority populations in their districts to be elected to Congress (Lublin 1997). Given this, one could argue that any model of mass state ideology—or, for that matter, district ideology—should include as an independent variable a measure of the proportion of black or minority citizens in the population. It is reasonable to think that higher proportions of black and other minorities in state electorates would be related to higher levels of roll-call liberalism among U.S. Senators and that, knowing this, the same relationship would be observed in the U.S. House.

Unfortunately for our purposes, the American states are much more homogeneous in terms of racial composition than are U.S. House districts. In terms of percent black population, the states range from a low of .22% (Montana) to a high of 36% (Mississippi), and the standard deviation across states is 9.27 in the 1980s, 9.24 in the 1990s, and 9.21 for the two decades combined. This compares to a range of black population proportions in House districts of .07% to 92.07%, and a standard deviation of 14.87 in the 1980s, 16.12 in the 1990s, and 15.50 for the two decades combined. Most importantly, in no state does the minority and/or black proportion exceed (or even approach) 50%. This is a crucial issue, for it appears that the relationship between minority population and district ideology is nonlinear. For those districts or states with racial minority population proportions below 50%, increases in the minority population are associated with greater aggregate ideological conservatism. For those districts or states with racial minority population proportions above 50%, the opposite is true; increases in minority population proportions are associated with lower levels of ideological conservatism. Consequently, the direction of the relationship between racial composition and constituent ideology is different for states and congressional districts. These different relationships for house districts and states are displayed vividly in Figure 1, using the Democratic presidential vote as a surrogate for constituent ideology.

It should be noted that in a preliminary analysis we estimated the state ideology model to include the black and/or minority populations in each state as one of the independent variables. Surprisingly, the coefficients for black or minority population proportions are all positive, suggesting that increases in black or minority populations are associated with higher levels of ideological conservatism. One possible explanation for this pattern is the hostility, or threat, hypothesis (Lublin 1997). The central idea underlying this hypothesis is that white (and presumably more conservative) citizens respond to the racial composition of their
districts. When blacks or other ethnic groups are in a minority, white citizens begin to perceive an increased threat as the proportion of minority group members begins to approach 50% of the overall population. At relatively high black or minority population proportions, white citizens perceive the greatest threat, and they respond with increased levels of ideological conservatism to counterbalance the liberalism of black citizens. This creates a higher level of ideological polarization that benefits the white (conservative) majority population, and the result is that legislators representing these districts are often more conservative than those with small black populations. On the other hand, in House districts which have black and/or minority population proportions in excess of 50%, black and/or minority voters have the upper hand, and the ideological liberalism of the constituency increases as that percentage increases. Hence, the positive relationship between black or minority population and constituency conservatism observed below the 50% mark becomes a negative relationship above that mark. What this means is that, in the population of states (where the black and/or minority populations are always below 50%), a negative relationship between racial composition and ideological conservatism is observed. In the population of House districts, the opposite relationship is expected.

Finally, since we are using relationships observed in the states to estimate House district ideology, we assume that the relationships observed in the states are similar to those found for House districts. With regard to race, this is clearly not the case. To utilize the relationship between racial composition and con-
stituency ideology at the state level to estimate U.S. House district ideology would lead one to expect House districts with a significantly large (70–100%) black population to be the most conservative! Clearly, the relationship between racial composition and district ideology is nonlinear, and the truncated variance on the racial composition variable among the states prevents a precise enough estimation of the nonlinear relationship to apply the relationship to House districts. Ultimately, we are left with no other alternative than to discard the race variable from our models. Fortunately, at least some of the effects of race are likely to be captured in other variables such as the Democratic presidential vote, urban population, and home ownership.6

Empirical Results

State-Level Estimates

We begin with the OLS regression estimates for our model of state-level mass ideology, estimated separately for the 1980s and 1990s; these results are presented in Table 1. For the most part, the models fit the data quite well. The adjusted R² for our model is above .70 for both the 1980s and 1990s analyses, suggesting that nearly three-fourths of the variance in mass state ideology is explained by the independent variables included in our model.7 The strong fit of the model is

6 Evidence of this assertion is seen in the statistically significant (p < .001) correlation coefficients between proportion black population and Democratic presidential vote (.607), urban population (.234), home ownership (−.422), and ultimately our simulated measure of House district ideology (.833).

7 It is important to contrast the goodness-of-fit levels for our top-down simulation model with those from models used in bottom-up simulations of district opinion. For the purpose of comparison we develop several models of individual-level ideology using data the American National Election Studies from 1976 to 2000. The highest adjusted R² attained through these models is .20, which is in line with results of previous bottom-up simulations. (Model results can be obtained from the lead author.) The weak fit of the models is largely due to the substantial amount of measurement error found in survey data upon which bottom-up simulations are based. Consequently, the estimates of the impact of demographic and other variables on individual-level opinion are relatively inefficient, the fallout of which is that the resulting simulated opinion has a fair amount of random measurement error associated with it.

On the other hand, because of the strong goodness-of-fit for our aggregate-level model, the predicted values derived from that model closely track the observed value for state ideology. (For confirmation, see Figure 2A and 2B.) Assuming that similar relationships between the independent variables and the (unobserved) ideological orientation of House districts are in place, the strong goodness-of-fit in the state model suggests that the amount of random measurement error that often characterizes simulations based on a bottom-up approach will be much lower. In other words, the estimates of House constituency ideology based on the top-down simulation should be more efficient than those based on the bottom-up simulation approach. Because the R² values for models employing aggregate data will naturally be higher than those that employ individual-level data, direct comparisons between standard bottom-up models of state ideology (which typically report R² values of less than .20) and
graphically displayed in Figures 2A and 2B, in which we present for the 1980s and 1990s scatter plots of observed state ideology and the values of state ideology predicted by our models. Clearly, our model does a good job in explaining state ideology. Based on this, as well as our assumptions stated above, we would expect our model to do a reasonably good job at simulating constituency opinion in U.S. House districts.

The coefficients for the model are, with a few exceptions, consistent with expectations. As has been suggested in recent research (e.g., LeoGrande and Jeydel 1997), the Democratic presidential vote is negatively and significantly related to our measure of mass state ideology; simply, states that provided Michael Dukakis, Walter Mondale, and Bill Clinton with relatively high-vote proportions exhibit the lowest level of mass ideological conservatism. As expected, the proportion of residents living in urban areas is also significantly and inversely related to mass state conservatism. Conversely, the coefficients for the variables representing states of the Deep South and the proportion of state

the top-down models reported here (which have R^2 values in excess of .70) must be made only with caution.

As further evidence of the weakness of the bottom-up simulation methods, we explore the relationships among the bottom-up simulated state ideology measures, Wright, Erikson, and McIver’s state ideology measures, and the average Democratic presidential vote in ideological elections throughout the period. While the Pearson correlations between the presidential vote and the top-down simulated measures are .35 and significant at the .001 level, the correlations between the bottom-up simulated measures and Wright, Erikson, and McIver measures of state ideology are only .10 and not statistically significant.
FIGURE 2A
1980s State Ideology and Predicted State Ideology Scatterplot

FIGURE 2B
1990s State Ideology and Predicted State Ideology Scatterplot
residents who are homeowners are both significant and positive, indicating that states from the Deep South and states with a large proportion of homeowners have significantly higher levels of mass ideological conservatism. Finally, although the proportion of blue collar workers appears to be inversely associated with mass ideological conservatism, the coefficients for this variable are not statistically significant in either decade.

**Simulating House District Opinion**

Having developed and estimated our model of state mass ideology, we utilize the coefficients from the model to generate constituent ideology predictions in each House district for the 1980s and 1990s. This is done by substituting values from each House district and for each variable in the model into the prediction equation and solving the equation to generate a predicted value for each House district.8

In Table 2 we report the descriptive statistics for the simulated measure of House district ideology and include the descriptive statistics for the Erikson et al. measure of state ideology as a point of comparison. First, the mean score for the state mass ideology measure is 13.79, as compared to the mean House district ideology of 10.61 for the 1980s and 10.78 for the 1990s. This is not a large difference, but it does suggest that constituents from House districts are, on average, marginally less conservative than state constituencies. This may be due in part to the fact that more populous states are, on average, somewhat more liberal than smaller states; the larger number of House districts from these more populous (and more liberal) states results in an average House district ideology that is more liberal than the average (unweighted) state. Clearly, there are some very liberal House districts but no very liberal states.

Second, as can be seen in Table 2 the dispersion of the House district ideology measure is much greater than for the Erikson, Wright, and McIver state ideology measures. The range of the state ideology variable is from −1.42 (for Rhode Island) to 27.32 (for Idaho), and the standard deviation is 7.68. On the other hand, for the estimates of House district ideology, the range is much greater, from −46.71 (16th District in New York) to 32.44 (9th District in Georgia) for the 1980s and −42.50 (16th District in New York) to 35.47 (10th District in North Carolina) for the 1990s. Moreover, the standard deviations are 13.67 for the 1980s and 13.71 for the 1990s model.9 In one real sense, this is not surprising. While state

---

8For example, the ideology score for the 1st District of California (1990s) is based on the following political and demographic district characteristics: (1) 47.5% average Democratic presidential vote; (2) 63.2% home owners; (3) 26.4% blue collar workers; (4) 38.98% urban residents; and (5) 0 for a non-Deep South state. Given the regression coefficients presented in Table 3 and the values on the independent variables, the predicted value is 10.32: CA House District 1 = 25.570 + [−0.674*(47.5)] + [0.342*(63.2)] + [−0.057*(26.4)] + [−0.087*(38.98)] + [8.297*(0)].

9The liberal 16th District of New York is located in the South Bronx. It is described in Congressional Quarterly’s *Politics in America* as a solidly democratic district and as “one of the most
boundaries are fixed, and while the populations of states change relatively slowly, House districts are often drawn to represent common interests that are somewhat narrow in nature. Furthermore, while states are large enough to avoid relatively extreme ideo-logical orientations, House districts are small enough that some will be on the extreme ends of the ideological distribution. For instance, while no state has a black population that is in the majority, there are numerous House districts that have very large black populations, and these districts are likely to have populations that are more liberal than the most liberal state. Moreover, the enhanced variance in the mass ideological orientations of House districts represents not only between-state differences, but also differences that exist within states. For instance, even the most conservative (liberal) state will have House districts that are more conservative (liberal) than the state mean; this means that the range for constituency ideology will be greater than that for the states.

Validity Issues

How valid is our measure of House district ideology? One way of ascertaining the validity of our measure is to compare our measure of House district ideology with “known” ideological characteristics. If known liberal (conservative) districts exhibit consistently low (high) scores on our district ideology measure, then it would appear that our measure has some level of validity. Of course, the problem is that dependable measures of district ideology are not readily available.

REFERENDA DATA. Referenda data have proven to be an indirect but dependable measure of district ideology. While referenda data are available for only a very limited number of congressional districts, such data are available for the 52 U.S. House districts of California. In order to demonstrate the validity of our measure of simulated district ideology, we compare it to district voting returns on several economically devastated areas in the United States.” The conservative 10th District of North Carolina is described in Politics in America as “the most rock-ribbed Republican district in North Carolina.” In 1990 GOP Senator Jesse Helms posted his best showing in the 10th Congressional District.
referenda faced by California voters in the 1996 general election. These referenda include votes on four traditional ideological issues, which include Proposition 210 (increasing the state’s minimum wage), Proposition 214 (healthcare consumer protection), Proposition 217 (increasing the state income tax), and Proposition 218 (limits on property tax).10

As a starting point, we consider the bivariate correlations between our measure of simulated district ideology and the district voting returns on the four ideological propositions (results not shown). Overall, the correlations strongly support the validity of our simulated measure of district ideology. Our measure of district ideology is significantly correlated ($p < .001$) with each of the referenda votes, with the mean correlation equal to .841 and a minimum correlation of .704. Moreover, the simulated district ideology measure exhibits a correlation of at least .850 with the referenda votes on an increase in the minimum wage and limits on property tax, two classic ideological issues in American politics.

We also conduct a principal components factor analysis of the intercorrelations among the four propositions and our measure of district ideology (results not shown). As expected, voting support for the four propositions and our measure of district ideology all load heavily on a single factor, indicating that they represent a single ideological dimension. Clearly, if one accepts that these four referenda issues represent the ideological preferences of California voters, one can accept that our measure of district ideology provides a consistent gauge of the ideological preferences of California congressional districts.11

PREDICTIVE RELIABILITY. Another way of testing the validity of our measure of simulated House district ideology is to explore the degree to which the measure predicts outcomes that should be theoretically related to House district ideology. Specifically, to what extent is our measure related to congressional roll-call behavior? If our measure is valid, it should serve as a good predictor of and be strongly related to roll-call behavior. Moreover, if the relationship between the simulated measures of House district ideology and House members’ roll-call behavior is at least somewhat similar in magnitude and significance to the relationship between established measures of state ideology and Senate roll-call behavior, this would reinforce the view that our district ideology measure has high validity.

In order to consider the relationship between roll-call voting and constituency ideology, we estimate models that depict roll-call behavior as a function of our measure of district ideology, as well as the political party of each member, coded

10 A full description of the propositions employed in this analysis is available from the lead author or from the California Secretary of State Web Site http://www.ss.ca.gov. While a total of fifteen propositions were voted on during the 1996 general election, we employ content and factor analyses to select the most ideological issues.

11 Tables reporting the bivariate correlations between our measure of simulated district ideology and district voting results on the four ideological propositions, as well as results of the principal components analysis, are available and can be obtained by writing the authors.
1 for Republicans and 0 for Democrats.\textsuperscript{12} We utilize a measure of roll-call ideology based on Poole and Rosenthal’s (1997) spatial analysis of roll-call voting from 1789 to 1996. The Poole-Rosenthal scores range from $-1.00$ (strong liberal) to $+1.00$ (strong conservative).\textsuperscript{13} Moreover, we estimate our models in two ways: (1) with separate pooled analysis using data for all senators and House members, respectively, during the 98th to 104th Congresses; and (2) with separate cross-sectional analysis for each Congress individually. For the cross-sectional models, we report ordinary least-squares (OLS) coefficients and for the pooled models we report generalized estimation equation (GEE) coefficients.\textsuperscript{14}

In Table 3 we report the GEE estimates for the pooled analysis of House and Senate roll-call voting for the 98th to 104th congresses, with roll-call conservatism depicted as a function of party and constituency ideology. First, for U.S. senators the relationship between state ideology and roll-call voting is strong and significant ($b = .017$, $t = 9.025$), even controlling for the effects of party. This coefficient suggests that a six-point increase in state constituency conservatism will result in approximately a one-tenth-point shift in the Poole-Rosenthal scale. Since the scale ranges only from $-1.00$ to $+1.00$, this is not a trivial effect. Second, for House members the effect of constituency ideology is about one-half the magnitude of the Senate effect, but here again the effect is highly significant. ($b = .008$, $t = 19.428$). For House members, approximately a 12-point increase in constituency conservatism translates into a one-tenth-point shift in roll-call conservatism. Given this, over the almost 80-point range of House district ideology, one

\textsuperscript{12} This is a relatively simple specification of a model of roll-call behavior. One could include a wider range of constituency variables in our model, as well as other characteristics of House members. However, in a general model of roll-call behavior, other constituency variables will be expected to reflect the general ideological character of each district. Since we include a measure of ideology that reflects many of the demographic and political variables that one might use as surrogates of district ideology in the first place, we see no need to control for these other variables in this simple model of roll-call behavior.

\textsuperscript{13} Poole and Rosenthal utilize a dimensional scale technique that yields a dominant single dimension for roll-call votes that is applicable to roll-calls cast since 1789. Based on this analysis, each member of Congress from 1789 to the present is assigned a score on this dimension. According to Poole and Rosenthal, this single dimension can be viewed as a party loyalty dimension, insofar as it separates Democrats and Republicans; however, because partisan cleavages that separate Democratic and Republican members of Congress are based largely on ideological differences, these scores are widely interpreted as a liberal-conservative dimension (Poole and Rosenthal 1997; Lublin 1997).

\textsuperscript{14} OLS coefficients are not reported for the pooled models because they often involve violations of OLS assumptions of homoskedasticity and uncorrelated error terms (Gujarati 1995, Kementa 1986). While OLS estimates are unbiased in the presence of autocorrelation, these estimates are not efficient, and the variability of OLS coefficients contaminates tests of statistical significance. Given this, we estimate our pooled cross-sectional time-series models using the generalized estimation equation (GEE) extension of the generalized least-squares (GLS) random-effects estimator (Diggle, Liang, and Zeger 1994). This procedure is appropriate in cross-sectional dominant data sets, i.e., when the number of cross-sections (members) is larger than the number of time points, and it can be used with unbalanced data. Most importantly, this procedure yields parameter estimates that are uncontaminated by the effects of autocorrelated and heteroskedastic errors.
would predict that a House member representing the most liberal district would be almost two-thirds of a point more liberal in roll-call behavior than a House member representing the most conservative district, even controlling for the considerable effects of party. So, while the effect of district ideology is somewhat smaller in magnitude than the effect of state ideology, there remains strong evidence that our measure of district ideology behaves, as one would expect, as a strong predictor of roll-call ideology.

This conclusion is reinforced in Table 4, in which we present the constituency ideology coefficients for Senate and House members, calculated separately for each Congress. We also report separate sets of coefficients for Republican and Democratic Senators and House members. As one can readily see, the finding of strong constituency effects on roll-call behavior is quite robust. In each Congress the state ideology coefficients are positive and significant, indicating that in each year state ideology has a strong effect on the roll-call behavior of U.S. Senators. But more importantly, the coefficients for House district ideology are also positive and statistically significant in each Congress, suggesting that our measure of district ideology has substantial ability to predict House members’ roll-call ideology.

In addition, it appears that Republican and Democratic Senators and House members respond in predictable ways to constituency ideology. Among Republican Senators and House members, the coefficients for the state ideology measure are positive and highly significant, suggesting that Republicans are responsive to the broad policy preferences of their constituents. Much the same can be said for Democratic House member and Senators, with all of the coefficients in the expected positive direction and easily achieving conventional levels of statistical significance. Overall, our findings suggest that all members—

<table>
<thead>
<tr>
<th></th>
<th>U.S. Senate</th>
<th>U.S. House</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>−.701</td>
<td>−.407</td>
</tr>
<tr>
<td></td>
<td>−23.092***</td>
<td>−49.766***</td>
</tr>
<tr>
<td>Party (Republican = 1)</td>
<td>.820</td>
<td>.683</td>
</tr>
<tr>
<td></td>
<td>28.928***</td>
<td>57.653***</td>
</tr>
<tr>
<td>Constituency Conservatism</td>
<td>.017</td>
<td>.008</td>
</tr>
<tr>
<td></td>
<td>9.025***</td>
<td>19.428***</td>
</tr>
<tr>
<td>N</td>
<td>700</td>
<td>3036</td>
</tr>
<tr>
<td>R-Square</td>
<td>.779</td>
<td>.789</td>
</tr>
</tbody>
</table>

*** $p < .01$, one-tailed test.
** $p < .05$, one-tailed test.
* $p < .10$, one-tailed test.

Note: T-statistics reported in this table are based on heteroskedastic robust standard errors.
### TABLE 4

**OLS Estimates for House and Senate Models of Constituency Influence on Roll-Call Conservatism, by Congress, 98th Congress to 104th Congress**

<table>
<thead>
<tr>
<th>All Members</th>
<th>U.S. Senate</th>
<th>U.S. House</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>b</td>
<td>t</td>
</tr>
<tr>
<td>100th Congress (1987–1988)</td>
<td>.015</td>
<td>5.948***</td>
</tr>
<tr>
<td>104th Congress (1995–1996)</td>
<td>.014</td>
<td>5.245***</td>
</tr>
<tr>
<td>Overall</td>
<td>.017</td>
<td>9.025***</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Republicans</th>
<th>U.S. Senate</th>
<th>U.S. House</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>b</td>
<td>t</td>
</tr>
<tr>
<td>98th Congress (1983–1984)</td>
<td>.014</td>
<td>3.394***</td>
</tr>
<tr>
<td>100th Congress (1987–1988)</td>
<td>.016</td>
<td>3.435***</td>
</tr>
<tr>
<td>101st Congress (1989–1990)</td>
<td>.024</td>
<td>5.092***</td>
</tr>
<tr>
<td>103rd Congress (1993–1994)</td>
<td>.019</td>
<td>3.965***</td>
</tr>
<tr>
<td>Overall</td>
<td>.019</td>
<td>6.076***</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Democrats</th>
<th>U.S. Senate</th>
<th>U.S. House</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>b</td>
<td>t</td>
</tr>
<tr>
<td>99th Congress (1985–1986)</td>
<td>.017</td>
<td>8.556***</td>
</tr>
<tr>
<td>100th Congress (1987–1988)</td>
<td>.014</td>
<td>5.088***</td>
</tr>
<tr>
<td>101st Congress (1989–1990)</td>
<td>.016</td>
<td>4.886***</td>
</tr>
<tr>
<td>103rd Congress (1993–1994)</td>
<td>.008</td>
<td>4.405***</td>
</tr>
<tr>
<td>Overall</td>
<td>.015</td>
<td>6.785***</td>
</tr>
</tbody>
</table>

***p < .01 level, one tail test.
**p < .5 level, one tail test.
*p < .10 level, one tail test.

Republican and Democrat, House members or Senators—are responsive to constituency policy preferences.

We suggest that these similar results provide support for the validity of our simulated measure of House district ideology. Ultimately, given the weaknesses found in available measures of district constituent ideology, the top-down
simulated measure appears to provide a significantly improved measure of district ideology.

**Conclusion**

Studies of policy representation are often based on the idea that legislators respond to the policy preferences of their constituents. Estimating what those constituency policy preferences are has been a major problem for legislative scholars. While there are some adequate measures of the ideological preferences of Senate electorates (i.e., the states), adequate measures of constituency preferences of U.S. House members has thus far eluded legislative scholars.

Considering the significant weaknesses of available measures of district constituency ideology, we develop in this work an improved measure of constituency ideology in U.S. House districts. Utilizing a “top-down” simulation of House district opinion, we (1) model state-level ideology (as measured by Erikson, Wright, and McIver 1993) as a function of the Democratic presidential vote and various demographic variables found at both the state and congressional district levels, and then (2) use the coefficient estimates from the state-level model and characteristics of House districts to generate predicted ideology scores for each House district during the 1980s and 1990s.

Preliminary tests of our simulated district ideology scores suggest the validity of our measures. Specifically, the validity of our measure is supported by the results of the analysis of the relationship between our simulated measures of U.S. House district ideology and several ideological referenda faced by California voters in the 1996 general election. Moreover, we are encouraged by analyses of the relationship between our simulated measures and House roll-call voting that are roughly comparable in magnitude and statistical significance to the relationship found for roll-call voting among U.S. Senators and the established and reliable measures of mass state ideology developed by Erikson, Wright, and McIver (1993).

We believe that our approach provides an improved measure of U.S. House district ideology that may be used in a number of research areas. In particular, we suggest that many studies of representation in the U.S. House have been limited by the quality of measures of constituency policy preferences. With improved measures, the study of legislative policy responsiveness could reach new heights. What is the relationship between constituency preferences and roll-call behavior? More importantly, what explains variation across legislators and contexts in how legislators translate constituency preferences into roll-call behavior? Are some legislators (e.g., those elected from marginal districts, those with higher or lower levels of seniority) more responsive to their constituents than others? Do some kinds of districts facilitate policy responsiveness? For instance, are legislators who represent homogenous districts or districts with politically engaged populations more responsive to their constituents than those who represent diverse districts or districts with inactive, inattentive constituents?
Finally, it is worth noting that the approach that we have outlined here has broad applicability for the estimation of constituency policy preferences in other legislative settings. Is it possible, for instance, to estimate constituency ideology for state senate and house districts, which are even less likely to be the subject of systematic surveys? Could our approach be adjusted to provide estimates of constituency preferences for state legislative districts? We believe that the answer to this question is yes, and we suggest that the use of our approach has the potential for permitting scholars to extend studies of legislative policy representation to a wider range of legislative venues.

Acknowledgements

An earlier version of this article was presented at the 2000 Annual Meeting of the American Political Science Association. Constituent ideology scores and demographic data used to create scores of constituent ideology will be archived with the Inter-University Consortium for Political and Social Research (ICPSR) and are also available from the lead author. We are indebted to three reviewers and the editor of this journal for their helpful comments on this article. We are responsible for any remaining errors.

Manuscript submitted 14 February 2001
Final manuscript received 21 November 2002

References


Phillip J. Ardoin is assistant professor of political science, Appalachian State University, Boone, NC 28608-2107 (ardoinpj@appstate.edu). James C. Garand is Emogene Pliner Distinguished Professor of political science, Louisiana State University, Baton Rouge, LA 70803-5433 (pogara@lsu.edu).