State-Level Opinions from National Surveys:
Poststratification using Multilevel Logistic Regression

David K. Park[*], Andrew Gelman[+], Joseph Bafumi[±]

Formatted: Bullets and Numbering

---

[*] Department of Political Science, Washington University, St. Louis, dpark@artsci.wustl.edu
[+] Department of Statistics and Department of Political Science, Columbia University, New York
[±] Department of Political Science, Columbia University, New York

1.Introduction

One of the first projects to simulate state-level opinions using national data was undertaken by Pool, Abelson, and Popkin (1965). When the three MIT scholars began their work, computer simulation was only about twenty years old. It was primarily used by engineers who tackled military problems, bridge designs, flight characteristics of new aircraft, and work assignment rules in factories.[1] Pool et al. aggregated 64 survey datasets of national respondents from 1952 to 1960. They then used poll, voting and census data to design 480 voter types based on such factors as income, religion, party, population density, region, race and sex.[2] Differences across states were not attributed to state-level factors but to differences in the proportion of voter types. With this assumption and the data in hand, they determined the percent of each voter type who held an opinion of interest, weighted it according to the number of each type in each state and contrived state-level results. Their 'best-fit' simulation did quite well with respect to vote choice in 1960. Their data produced a 2.5% median percentage point difference in the Kennedy simulation vote versus the election vote across 48 states.

Weber, Hopkins, Mezey, and Munger (1972-1973) undertook a similar project. Their paper took issue with research that emphasized socio-economic variables as determinates of policy output in the American states. They attribute those results to invalid measures of state opinion and instead proposed to estimate state level opinion in much the same way as Pool et al. But considering a wider range of models and settling on a method including 960 categories of voter types, they were able to estimate and report state-level opinions on favorability toward the

---

[1] Pool et al. received the support of the Democratic Party and the Kennedy campaign to employ computer simulation for the purposes of bolstering their electoral odds. Highlighting the controversy surrounding the project, Theodore Sorenson would insist that no information from the simulations was used to craft Kennedy's campaign message (Sorenson 1965).
[2] The project was also the basis for a satirical novel by Eugene Burdick titled *The 480*.

death penalty and toward teacher unionization. They note, "The voter-type approach to creating synthetic electorates could also be employed to calculate such electorates for congressional districts, metropolitan areas, counties or cities" (Weber et al., 565). In a follow-up article, Weber and Shaffer (1972) tested for congruence between state-level opinions and state-policy outputs. They found opinion to be a more important determinant of policy than other factors.

Erikson (1976) provided further evidence in support of these conclusions. He cleverly used forgotten large sample opinion data from the 1930's and showed that state policy reflected the opinions of their populace for capital punishment, child labor and female juror laws. Wright, Erikson, and McIver (1987) and Erikson, Wright, and McIver (1993) extended this analysis to generate state-level measures of partisanship and ideology and tested for representation across policy outcomes that fall on the left/right continuum. Their study was in part prompted by the deficiencies they found in Weber et al. (1972-1973) and Weber and Shaffer (1972). Erikson et al. found their method to inadequately tap state opinion apart from socioeconomic factors (despite Weber et al.'s intention) since voter types were largely based on demographic variables of the socioeconomic sort. They showed high congruence between state policy and state opinion even in the face of socioeconomic variables that had been deemed such powerful predictors. Their partisanship and ideology scores have become standard predictors in many state-level regression models (i.e. Hill and Leighley 1992, Hill, Leighley, and Hinton-Anderson 1995).

Brace, Sims-Butler, Arceneaux, and Johnson (2002) sought to use the same methodology as Erikson et al. (1993) to estimate state-level attitudes for tolerance, racial integration, abortion, religiosity, homosexuality, feminism, environmentalism, welfare and capital punishment. They collapsed General Social Survey (GSS) data from 1974 to 1998 for these items (some of which

2

are composite), obtained adequate sample sizes for most states, applied appropriate weights and calculated frequencies.[3] As with Erikson et al., the method requires strong stability over time and reliability in the scores. Otherwise, it cannot be said that welfare opinions in New York, for example, are the same from 1974 to 1998.

However, as Table 1 indicates, Erikson et al. and Brace et al. aggregated over 13 and 22 years, respectively, to produce state-level public opinion for less populous states, such as Montana, Vermont, Wyoming, etc.[4] This simple aggregation creates a significant tradeoff when producing state-level public opinion, namely, the researcher must focus on state-level opinions that do not vary over time, or else, accept that one is implicitly estimating time-averaged opinions.[5]

[Insert Table 1 Here]

In order to overcome the limitations of Erkison et al. and Brace et al., we construct a multilevel logistic regression model for a binary response variable conditional on poststratification cells to estimate state-level opinions from national surveys. This approach combines the modeling often used in small-area estimation with the population information used in poststratification (Gelman and Little 1997), which is the standard method for adjusting for nonresponse in political polls (Voss, Gelman, and King 1995).

Section 2 of this paper presents the model – both poststratification and multilevel logistic regression. Section 3 presents the data used in the analysis – the national surveys and US Census

---

[3] State codes for respondents can be purchased from the GSS

[4] Erikson, et al. do not include Alaska and Hawaii and Brace et al. do not include Hawaii, Idaho, Maine, Nebraska, Nevada, and New Mexico in their analysis.

[5] See Brace et al.'s discussion of stability of the state-level measures.

data. Section 4 evaluates the model by comparing state-level estimates of candidate choice with actual election outcomes for 1988 and 1992. Section 5 evaluates the model by comparing state-level estimates of partisanship and ideology with Erikson, Wright, and McIver. Section 6 considers further research to improve and test the model. Section 7 discusses the implications of the results.

2. Model[6]

2.1 Postratification

The standard practice for weighting in pubic opinion polls is based entirely (or primarily) on poststratification, which is generally referred to any estimation scheme that adjusts to population totals. There is a fundamental difficulty in setting up poststratification categories. It is desirable to divide up the population into many smaller categories, but if the number of respondents is small, it is difficult to accurately estimate the average response within each category. One can improve efficiency of estimation by fitting a multilevel model.

Consider subsets of the population defined by R categorical variables, where the r-th variable has J_r levels, for a total of $J = \prod_{r=1}^{R} J_r$ categories (cells), which are labeled $j = 1, \ldots, J$. Assume that N_j, the number of individuals in the population in category j, is known for all j. Let y be a binary response of interest. For each j, let n_j and N_j be the number of individuals in category j in the sample and the population, respectively. The R variables should include all information used to construct survey weights, as well as other variables that might be informative about y. For example, the population of adults in the 50 U.S. states plus the District

---

[6] We follow Gelman and Little's (1997) notation for postratification and the multilevel logistic regression.

4

of Columbia are categorized by R = 5 variables: state of residence, sex, ethnicity, age, and education, with $(J_1, \ldots, J_5)$ = (51 x 2 x 2 x 4 x 4).

The J = 3,264 categories range from "Alabama, Male, Not Black, 18-29, Not High School Graduate" to "Wyoming, Female, Black, 65 and over, College Graduate," and the U.S. Census, provides good estimates of N_j in each of these categories.[7] For example, in 1990 there were 66,177 adults who lived in Alabama, were male, not Black, between the ages 18-29, and did not have a high school diploma. ~~I~~We shall consider adult population estimates (summing over all 3,264 categories) and also estimates within individual states (separately summing over the 64 categories for each state).

2.2 Regression Modeling in the Context of Poststratification

One can set up a logistic regression model for the probability π_j of a "yes" for respondent in category *j*:

$$\text{logit}(\pi_j) = X_j\beta \qquad (1)$$

where X is a matrix of indicator variables (such as age, education, gender and sex), and X_j is the *j*-th row of X. If a uniform prior distribution on β is assumed, then Bayesian inference (for different choices of X) under this model corresponds closely to various classical weighting schemes (Gelman and Little 1997).

---

[7] Census of Population and Housing, 1990: Subject Summary Tape File (SSTF) 6, Education in the United States [machine-readable data files] / prepared by the Bureau of the Census. --Washington: The Bureau [producer and distributor], 1993.

2.3 Multilevel Logistic Regression Model[8]

The multilevel model allows for partial pooling across the cells by modeling exchangeable batches of coefficients. The model then can be written in the standard form of a multilevel logistic regression as:

$$y_i \sim \text{Binomial}(p_i) \qquad (2)$$

$$\text{logit}(p_i) = (X\beta)_i \qquad (3)$$

$$\beta \sim N(0, \textstyle\sum_\beta) \qquad (4)$$

$$\gamma_{km} \overset{ind}{\sim} N(0, \sigma_m^2), k = 1, \ldots, K_m. \qquad (5)$$

We write the vector β as (α, $\gamma_1,\ldots,\gamma_M$), where α is a subvector of unpooled coefficients and $\sum_\beta^{-1}$ is a diagonal matrix with 0 for each element of α, followed by σ_m^{-2} for each element of γ_m, for each m. Each γ_m, for m = 1,…,M, is a subvector of coefficients (γ_{km}) to which we fit a multilevel model. We use the notation p_i, for the probability corresponding to the unit *i*, as distinguished from π_j, the aggregate probability corresponding to the category *j*. A constant term has been included as part of the unmodeled coefficients α, and so we can give each γ_{km} a prior mean of 0 with no loss of generality. The group-level standard deviations σ_m are given independent noninformative prior distributions:

---

[8] For introduction to multilevel models see chapters 5 and 13 of Gelman, Carlin, Stern, and Rubin (1995). Also see the following: Gill (2002), Goldstein (1987), Heck et al. (2000), Kreft and de Leeuw (1998), Leyland and Goldstein (2001), Raudenbush and Bryk (1992), and Reise and Duan (2001).

$$\sigma_m \sim \text{Uniform}(0, 100), \qquad m = 1, \ldots, M \qquad (6)$$

This essentially noninformative prior distribution allows each σ_m to be estimated from the data. This can be contrasted to two extremes that correspond to classical analyses. Setting σ_m to 0 corresponds to excluding a set of variables, i.e., complete pooling; setting σ_m to ∞ corresponds to a noninformative prior distribution on the γ_{km} parameters, i.e., no pooling.

2.4 Estimates Under the Model

To obtain the quantities of interest, the following strategy was used:

1. Perform Bayesian inference for the regression coefficients, β, and the hyperparameters, σ_m, given the data y.

2. For each of the J categories of person in the population, compute $p_j = \text{logit}^{-1}(X\beta)_j$. This is done for all categories j including those, such as Black male, college graduate, 18-29, Wyoming, that are not represented in the sample.

3. Compute inferences for the population quantities by summing $N_j p_j$'s.

The model is estimated via Bayesian MCMC (Markov chain Monte Carlo) methods simulation, using WinBUGS (Spielgelhalter et al. 1997) as called from R. If fact, these simulations are used to compute uncertainties and standard errors.

3. Data

3.1 National Survey Data

3.1.1   1988 Survey Data

All respondents from seven pre-election national tracking polls conducted by CBS News/New York Times during the nine-days preceding the 1988 U.S. Presidential election were used (N = 13,544).[9] The outcome variable is support for the Republican presidential candidate where $y_i = 1$ is assigned to supporters of Bush, $y_i = 0$ to supporters of Dukakis, and NA for respondents who expressed "other" or were missing (we follow the standard practice and count respondents who "lean" toward one of the candidates as full supporters).[10] Predictor variables include sex, ethnicity, age and education (respondents were excluded if sex, ethnicity, age, or education were missing).

Even though no data were included from Hawaii and Alaska, they are included in the model.[11] Also, it is common practice to exclude Washington, D.C. because (i) its voting preferences are so different from the other states that a model that fit the 50 states would not fit D.C. and (ii) it would unduly influence the results for the other states. However, D.C. is included, and in order to mitigate against such problems, D.C. is given its own region code.

3.1.2   1992 Survey Data

---

[9] ICPSR numbers 9152, 9153, 9154, 9155, 9156, 9157, 9158.

[10] There are 1,978 respondents who expressed "other" or were missing. WinBUGS, when encountering missing data on a outcome variable in a regression-type setting, will make multiple imputations. See Jackman (2000) for a more detailed explanation of how WinBUGS handles missing data on the outcome variable.

[11] WinBUGS will estimate 'phantom parameter' for these two states. It will treat them as nuisance parameters or missing data in the Bayesian sense.

All respondents from 18 national surveys, conducted by CBS/NY Times and ABC/Washington Post from March to November 1992, were used (N=24,072). However, due to the strong showing of a third party candidate, Ross Perot, three subsets of the overall survey were taken to estimate support for Bush among the two major party candidates: first subset included all respondents who were registered or expected to vote (N=24,072); second subset included all respondents who expected to vote for either the Republican or Democratic presidential candidate (N=18,460); and the third subset included all respondents who expected to vote for Bush or Clinton (N=15,497).

For the third subset of the data, and similar to 1988, $y_i = 1$ is assigned to supporters of Bush and $y_i = 0$ to supporters of Clinton, and NA for respondents who were missing.[12] Again, respondents were excluded if sex, ethnicity, age, or education were missing. Even though no data were included from Hawaii and Alaska, they are again included in the model, and D.C. is again given its own region code.

3.2 National Census Data

Survey organizations create weights on the following variables (Voss, Gelman King 1995):

Census Region:	Northwest, South, North Central, West
Sex:	Female, Male
Ethnicity:	Black, Non-Black
Age:	18-29, 30-44, 45-64, 65+
Education:	Not High School Grad, High School Grad, Some College, College Grad

---

[12] There are 2,717 respondents who expressed "other" or were missing.

This includes all main effects plus the interactions of 'sex x ethnicity' and 'age x education.' Sex, ethnicity, and their interaction are included as fixed effects in the multilevel logistic regression model, and as mentioned previously, respondents with nonresponse in any of the predictor variables are excluded.

The model goes beyond the traditional analysis by survey organizations by including indicators for the 50 states (plus D.C.) clustered into five batches corresponding to the four census regions plus a separate region for D.C. In order to improve the estimates, the model is run with the average vote for the Republican presidential candidate for the past three presidential elections for the 50 states (plus D.C.) as a group-level predictor for the state coefficients.[13] In addition, 'sex x ethnicity,' education, age, and 'age x education' are treated as separate batches of coefficients in the multilevel model. The performance of the model is checked by comparing estimates to the actual Presidential returns for each state (plus D.C.).

In order to poststratify on all the variables listed above, along with the state, we need the joint population distribution of the demographic variables within each state: that is, population totals N_j for each of the 2 x 2 x 4 x 4 x 51 cells of 'sex x ethnicity x age x education x state.' As an approximation to that distribution the Census of Population and Housing, 1990 is used: Subject Tape File (SSTF) 6, Education in the United States. SSTF 6 contains the joint population distribution of the demographic variables within each state weighted to represent the total population. SSTF 6 contains sample data weighted to represent the total population. In addition, the file contains 100-percent counts and unweighted sample counts for total persons and total housing units.

---

[13] It would be slightly preferable to adjust these for home-state and home-region effects..

4. Results

The method can be applied for any yes/no survey response. We use presidential choice because it can be compared with the actual election outcomes.[14] For 1988, we compare the multilevel model with three other models: CBS/NY Times, no pooling, and complete pooling. For 1992, we compare the multilevel model with pooled surveys from CBS/NY Times and ABC/Washington Post. We expect the multilevel model to perform the best because of the flexibility of the multilevel logistic regression model and because poststratification uses the population numbers N_j (Gelman and Little 1997).[15]

Surveys organizations assign weights to each respondent as the inverse of the probability of selection, modified by a series of ratio estimates. The first stage estimate is essentially a non-interview adjustment within geographic region. For each region, an adjustment is made to approximate the number of adults in that region. In the next stage, ethnicity by sex is the ratio estimate characteristics, and the final stage is age by education. On occasion, because of a small number of sample cases in some cells at the final stage, some educational groups are collapsed within age categories. To obtain state-level estimates we perform weighted averages within each state, using the weights provided by the survey organization.

The "no pooling model" estimates use demographic variables and also indicators for the states, regions, with no multilevel framework. The "complete pooling model" estimates use only age, education, ethnicity, and sex, with state and region effects set to zero. This model allows

---

[14] Jackman and Rivers (2001) use a dynamic Bayesian multilevel model to produce forecasts of the Electoral College outcome by aggregating state-level forecast.
[15] In order to improve our election-forecasting model, we can adjust the N_j's to represent registered and/or likely voters instead of the voting age population as a whole.

the average responses within states and regions to differ only because of demographic variables.[16]

For the no pooling, complete pooling, and multilevel logistic models, we fit the regression models to the survey data, obtain a posterior simulation draws for each coefficient, and reweight based on the 1990 Census of Population and Housing data to obtain the poststratified estimates for the proportion of voters in each state (including D.C., Alaska, and Hawaii) who supported Bush for President for 1988 and 1992.

4.1 **1988 Election Results[17]**

Table 2 gives the mean national popular vote and mean absolute error of the states. At the national level, the five methods offer very similar results. The actual result was 53.9% support for Bush, the four models produced results that ranged from 54.8% to 58.2%. The real efficiency gained from the model-based estimates occurs in estimating individual states. The reduction in the mean absolute error of the states from 5.4% (CBS/NYT) to 3.7% (Multilevel I) can be attributable to the poststratification and multilevel modeling. Furthermore, Table 3 shows that the uncertainty (average width of the 50% interval) from Multilevel I is relatively small and with slightly less than 24 of the 51 states estimates falling inside the 50% intervals.

---

[16] The "multilevel logistic model" (Multilevel I) estimates use demographic variables, with the 50 state (plus D.C.) effects, and five region effects within a multilevel framework. Multilevel II includes all the variables used in Multilevel I, but adds the average vote share for the Republican presidential candidate for the previous three presidential elections for the 50 states, plus D.C.

[17] Results are based on two chains each running for 3000 iterations, with a 2000 iteration burn-in. Gelman-Rubin convergence diagnostic showed convergence for all estimated coefficients.

12

[Insert Table 2 Here]

Table 3 presents the same estimates as Table 2, but includes states with sample sizes less than 100. Table 3 clearly illustrates the gain from the multilevel models for small population estimation. Again, at the national level, the five models offer similar results (for states with sample size less than 100, the mean national popular vote was 53.9% and the five models produced estimates ranging from 50.7% to 58.7%). However, the reduction in the mean absolute error of the states falls from 10.3% (CBS/NYT) to 4.9% (Multilevel I) to 2.1% (Multilevel II) and uncertainty from 0.227 (No Pooling) to 0.085 (Multilevel I) to 0.075 (Multilevel II) and with 8 out of 15 state estimates falling inside the 50% intervals for Multilevel I and 9 out of 15 states for Multilevel II.

[Insert Table 3 Here]

Figure 1 plots, by state, the actual election returns (vertical axis) versus the CBS/NY Times estimates and the posterior medians (horizontal axis) for the three models (+ indicates states with sample size larger than 100 and ■ indicates states with sample size smaller than 100. As the results from Table 4 indicate, the multilevel model reduces variance, and thus estimation error, especially for states with sample sizes less than 100.

[Insert Figure 1 Here]

13

For example, in Nevada (N=32), 59% supported Bush in the 1988 election, CBS/NY Times estimated 51% while the multilevel model estimated 59%; Vermont (N=12), 51% supported Bush, CBS/NY Times estimated 80% while the multilevel model estimated 55%.

4.2 **1992 Election Results**[18]

Table 4 gives the mean national popular support vote and the mean absolute error of the states. At the national level, the actual result was 46.5% support for Bush, while the weighted survey estimated 43.3% and the multilevel model 45.5%. As mentioned previously, the efficiency gained from the model occurs in estimating individual states. The reduction in the mean absolute error of the states falls from 7.9% for the weighted survey to 4.9% for the multilevel model. The average width of the 50% interval from the multilevel is 0.036 with 51 states (including D.C.) falling inside the interval.

[Insert Table 4 Here]

Figure 2 plots, by state, the actual election returns (vertical axis) versus the CBS/NY Times and ABC/Washington Post estimates and the posterior medians (horizontal axis). Again, (+ indicates states with sample size larger than 100 and ■ indicates states with sample size smaller than 100).

[Insert Figure 2 Here]

---

[18] Results are based on two chains each running for 5000 iterations, with a 4000 iteration burn-in. Gelman-Rubin convergence diagnostic showed convergence for all estimated coefficients.

For example, in Delaware (N=49), 45% supported Bush in 1992 while the national surveys estimated 35% and the multilevel model estimated 48%; Idaho (N=66) Bush received 66%, national survey 36% and multilevel model 52%; Nevada (N=66) Bush received 48%, national survey 37% and multilevel model 48%; and Rhode Island (N=56) Bush received 38%, national survey 26% and multilevel model 36%.

4.3 Partisanship[19]

Table 5 gives the mean ideology estimates of the states. The multilevel model estimates the mean Republican identification to be approximately 4.0% higher than Erikson et al.'s estimates. Recall that Erikson et al.'s estimates come from pooled surveys that span from 1976 to 1988; and the multilevel estimates are from 1988. Therefore, as Erikson et al. state, "[i]n terms of party identification, the electorate did indeed change. As recently as the Carter presidency, the Democratic Party dominated the Republicans in terms of identification, with an average edge of 14.2 percentage points in the states. By Reagan's second term, this lead had shrunk to an average edge of only 2.8 points. Clearly the Republican party had made major gains (Erikson et al. 1993, 30)."

[Insert Table 5 Here]

Figure 2 shows the plots of Erikson et al. (horizontal axis) versus the multilevel model (vertical axis) for the Republican, Independent, and Democrat and the difference between Republican-

---

[19] Results are based on two chains each running for 3000 iterations, with a 2000 iteration burn-in. Gelman-Rubin convergence diagnostic showed convergence for all estimated coefficients.

Democrat self-identifiers (+ indicates non-Southern states and ■ indicates Southern). The Republican graph clearly shows the expected Republican bias for the multilevel results. For example, Mississippi increased their Republican identification from 27% to 40%, South Carolina from 28% to 37%, Tennessee from 27% to 40% and Virginia from 29% to 42%. Therefore, relying on Erikson et al.'s scores of partisanship for 1988 would not capture the dramatic shift toward the Republican Party for many of the states.

[Insert Figure 3 Here]

4.4 Ideology[20]

Unlike the results for Republican identification, the mean conservative scores for the multilevel model are relatively close to Erikson et al. As table 6 shows, the multilevel model estimates 36% of the population identify as conservative, while Erikson et al. estimate 35%. Again, at the national level, this fits with our expectations regarding the change (or lack thereof) in conservatism during 1976 to 1988. As Erkison et al. state, "[m]ean ideology was quite stable over the 13 year period. In other words, the popular notion that the electorate became more conservative during the Reagan presidency is not supported by CBS/NYT survey responses on ideological self-identification. While the electorate certainly preferred the conservative label under Reagan, it was no more conservative than before. For example, the lead of conservative identifiers over liberals was virtually the same during Reagan's second term as during the Carter presidency!" (Erikson et al., 1993, 30).

---

[20] Results are based on two chains each running for 3000 iterations, with a 2000 iteration burn-in. Gelman-Rubin convergence diagnostic showed convergence for all estimated coefficients.

[Insert Table 6 Here]

Figure 3 shows the plots of Erikson et al. (horizontal axis) versus the multilevel model (vertical axis) for conservative, moderate, liberal and difference between conservative-liberal self-identification (+ indicates non-Southern states and ■ indicates Southern). The conservative results show the Erikson et al. and multilevel model estimates to be relatively close (clustering around the 45-degree line). However, not all states remained the same. According to the multilevel estimates, several states were estimated as more conservative, such as Alabama from 41% to 50%, Arkansas from 37% to 45%, and Connecticut from 29% to 41%. Other states were estimated as less conservative, such as Arizona from 37% to 29% and Idaho from 43% to 34%. Furthermore, as the liberal graph shows, the estimate of liberal self-identification decreased. For example, Nevada estimated liberalism went from 29% to 20%, West Virginia from 23% to 14%, and Oregon from 25% to 17%. Again, relying on Erikson et al.'s pooled estimates would produce a different ideological picture of the American states.

5. Conclusion

In the past twenty years, both historical and institutional scholars have pushed the discipline to reconnect the study of institutions and policymaking (and policy outcomes) to address the broader questions raised by democratic theory. Public opinion scholars have tended to narrowly focus on individual opinion formation and change. As Erikson et al. state, "[a]n unfortunate by-product of this emphasis on the individual over the aggregate has been the disconnection of the study of mass behavior from the larger body of democratic theory" (Erikson et al. 1993).

17

When public opinion scholars have emphasized the aggregate over the individual, they have tended to focus on the "national" electorate and the presidency. Even though the House and Senate represents a third of the federal system (and some would argue the most important institution in regards to understanding issues of representation and democratic theory), there has been a dearth of research examining the congruence, or lack thereof, between the "state" or "congressional" electorate and their representatives.

The few statistical studies of political representation, beginning with the classic analysis of Miller and Stokes (1963), have found high degrees of representation between constituents' policy preferences and their representatives' policy choices in Congress (for example, Achen 1978; Page, et al. 1984; Bartels 1991, 2002; Erikson et al. 1993; Stimson, MacKuen and Erikson 1995). However, these previous works were limited by the small sample size, either at the state or congressional district level. The method presented here does not completely solve the problem, but it does allow the researcher to potentially produce more reliable estimates of constituency preferences than previous methods.

Tables

Table 1. Years, Surveys and Sample Size to Produce State-Level Opinion

Study	Years	Surveys	N
Erikson, Wright & McIver	13	122	167,460
Brace, Sims-Butler, Arceneaux, Johnson	25	22	35,000[*]
Multilevel Model (1988)	9 days	7	13,544
Multilevel Model (1992)	9 months	18	24,072

*approximate

Table 2. Summary statistics for CBS/NY Times, No Pooling, Complete Pooling and Multilevel. Summaries given are the estimated mean of the 50 state vote (plus the D.C.) for Bush; the average width of the 50% intervals for the states (plus D.C.); and the number of the 50 states (plus D.C.) whose true values fall within the 50% intervals.

Summary	Actual Result	CBS/NYT	No Pooling	Complete Pooling	Multilevel Model I	Multilevel Model II*
Mean of national popular vote	0.539	0.568	0.548	0.582	0.539	0.549
Mean absolute error of states compared to actual election estimate	--	0.054	0.068	0.059	0.037	0.031
Average width of 50% interval	--	--	(0.112)	(0.012)	(0.053)	(0.048)
Number of states contained in 50% interval	--	--	38/51	4/51	24/51	24/51

*Includes the average vote share for the Republican presidential candidate for the previous three presidential elections for the 50 states (plus D.C).

Table 3. Summary statistics for CBS/NY Times, No Pooling, Complete Pooling and Multilevel for states with sample sizes less than 100. Summaries given are the estimated mean of the 50 state vote (plus the D.C.) for Bush; the average width of the 50% intervals for the states (plus D.C.); and the number of the 50 states (plus D.C.) whose true values fall within the 50% intervals.

Summary	Actual Result	CBS/NYT	No Pooling	Complete Pooling	Multilevel Model I	Multilevel Model II*
Sample Size[21] ≤ 100						
Mean of national popular vote	0.531	0.520	0.509	0.587	0.507	0.529
Mean absolute error of states compared to actual election estimate	--	0.103	0.139	0.075	0.049	0.021
Average width of 50% interval	--	--	(0.227)	(0.014)	(0.085)	(0.075)
Number of states contained in 50% interval	--	--	8/15	2/15	8/15	9/15

* Includes the average vote share for the Republican presidential candidate for the previous three presidential elections for the 50 states (plus D.C).

---

[21]States with sample size ≤ 100 are: AK, DE, DC, HI, ID, ME, MT, NV, NH, ND, RI, SD, UT, VT, WY.

Table 4. Summary statistics for weighted survey estimate and multilevel model. Summaries given are the estimated mean of the 50 state vote (plus the D.C.) for Bush; the average width of the 50% intervals for the states (plus D.C.); and the number of the 50 states (plus D.C.) whose true values fall within the 50% intervals.

Summary	Actual Result	National Surveys	Multilevel *
Mean of national popular vote	0.465	0.433	0.455
Mean absolute error of states compared to actual election estimate	--	0.073	0.049
Average width of 50% interval	--	--	(0.036)
Number of states contained in 50% interval	--	--	51/51

* Includes the average vote share for the Republican presidential candidate for the previous three presidential elections for the 50 states (plus D.C).

Table 5. Mean partisanship estimates for Erikson, Wright and McIver and Multilevel Model and their Correlations

Mean Partisanship	Erikson, Wright & McIver	Multilevel Model[22]	Correlations
Republican	0.29	0.33 (0.04)	0.72
Democrat	0.37	0.39 (0.03)	0.66
Independent	0.34	0.29 (0.04)	0.69

---

[22] Does not add up to 1.0 due to rounding.

Table 6. Mean ideology estimates for Erikson, Wright and McIver and Multilevel Model and their Correlations

Mean Ideology	Erikson, Wright & McIver	Multilevel Model	Correlations
Conservative	0.35	0.36 (0.03)	0.54
Liberal	0.21	0.19 (0.03)	0.71
Moderate	0.44	0.45 (0.03)	0.21

Figures

Figure 1. 1988 Election results by state vs. (a) CBS/NY Times weighted means and posterior median estimate for (b) no pooling (c) complete pooling and (d) multilevel model for state, region, and previous presidential vote effects.
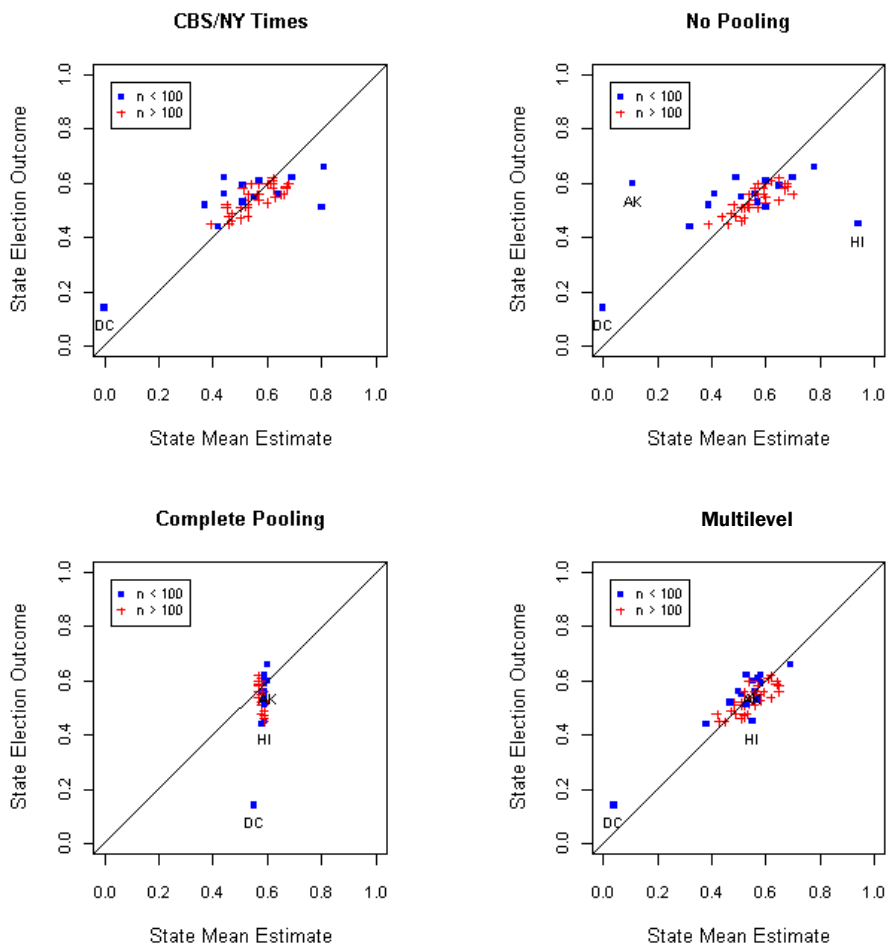
Figure 2. 1992 Election results by state vs. (a) survey weighted means and (b) posterior median estimate for multilevel model for state, region, and previous presidential vote effects.
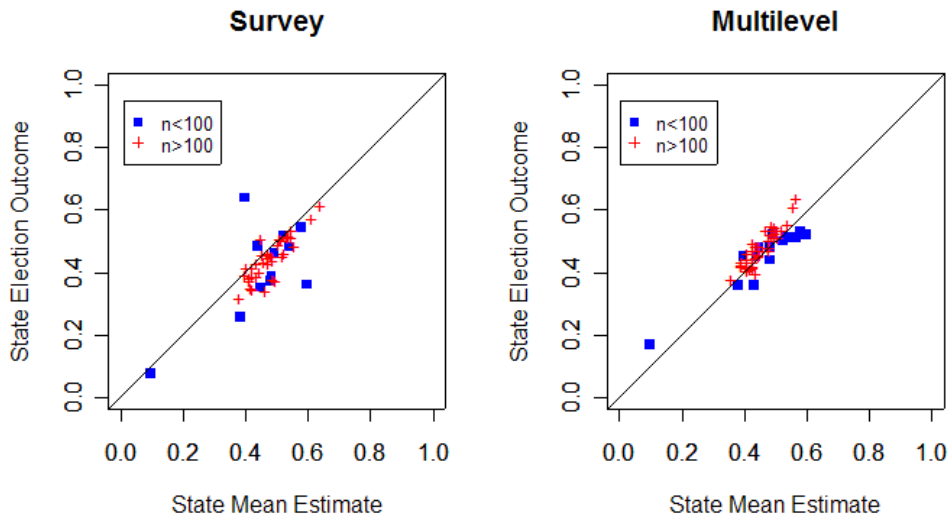
Figure 3. Partisanship self identification in the United States: Erikson, Wright, and McIver and vs. Multilevel
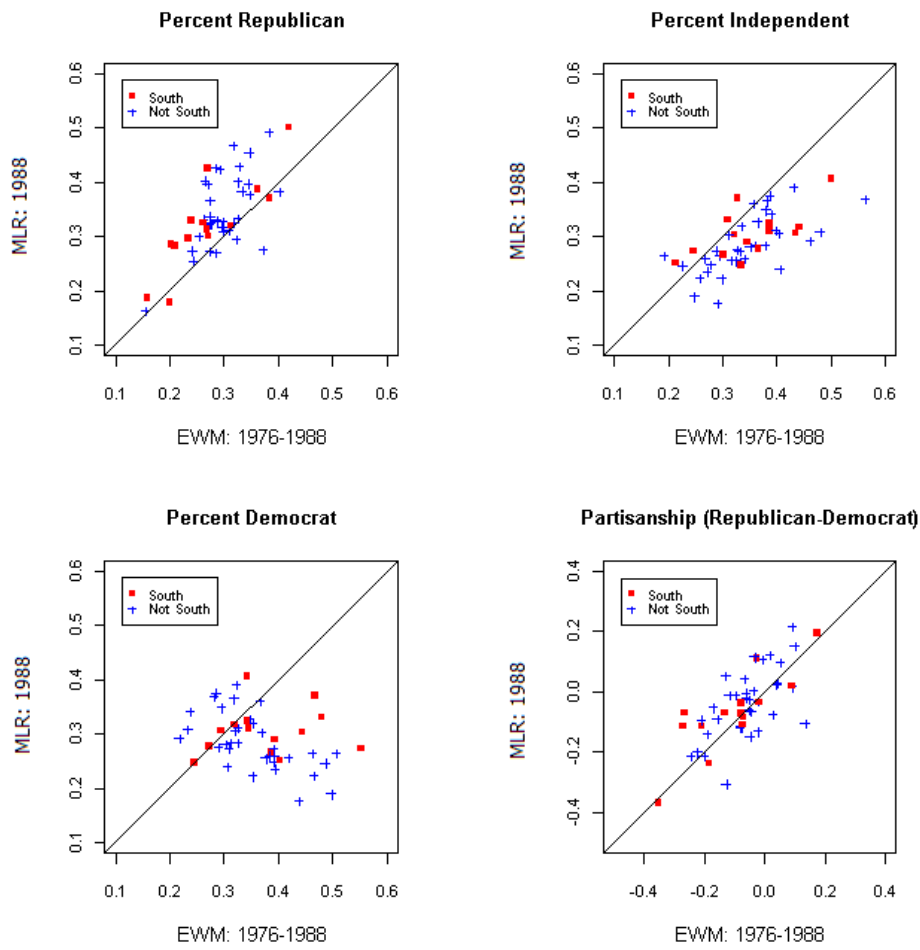(a) Republican (b) Independent (c) Democrat (d) Republican-Democrat Difference

Figure 4. Ideology self identification in the United States: Erikson, Wright, and McIver vs. Multilevel model (a) conservative (b) moderate (c) liberal (d) conservative-liberal difference.