Unpublished Supplement to

"Measuring Citizen and Government Ideology in the American States, 1960-93,"
***American Journal of Political Science*, vol. 42, Jan. 1998, pp. 327-48.**

William D. Berry
Florida State University

Evan J. Ringquist
Florida State University

Richard C. Fording
University of Kentucky

Russell L. Hanson
Indiana University Bloomington

This document contains:

(I) an expanded version of the description (pp. 330-31 in *AJPS* article) of our methodology for measuring state citizen ideology;

(II) a description of our procedure for estimating ideology scores for congressional party delegations with no members (cited in note 3 on p. 331 of *AJPS* article) [used on constructing our indicators of state citizen and government ideology];

(III) information on sources of data used in the published article; and

(IV) references cited in the document.

I. Our Methodology for Measuring State Citizen Ideology

To measure citizen ideology, we identify the ideological position of each member of Congress in each year, using interest group ratings (see Assumption 2 in *AJPS* article). Next, we estimate citizen ideology in each district (House and Senate) of a state using the ideology score for the district's incumbent, the estimated score for a challenger (or hypothetical challenger) to the incumbent, and election results that presumably reflect ideological divisions in the electorate (see Assumption 1 in *AJPS* article). Finally, citizen ideology scores for each district are used to compute an unweighted average for the state as a whole.[1]

Thus, the major challenge is estimating citizen ideology in a congressional district in a given year.[2] Our measure averages the ideology scores for major party candidates, using weights that are proportional to each candidate's share of support in the district:

$$CITIDEO_{d,t} = (INCSUPP_{d,t})(INCIDEO_{d,t}) + (CHALSUPP_{d,t})(CHALIDEO_{d,t}) \qquad [1]$$

where $CITIDEO_{d,t}$ denotes citizen ideology in district d in year t. $INCSUPP_{d,t}$ is the (estimated) proportion of the electorate in year t preferring district d's incumbent, and $CHALSUPP_{d,t}$ is the (estimated) proportion of the electorate in year t preferring the challenger (or a hypothetical challenger) to district d's incumbent. $INCIDEO_{d,t}$ is the ideology score for district d's incumbent in year t, and $CHALIDEO_{d,t}$ is the (estimated) ideology score for the challenger (or hypothetical challenger) to district d's incumbent in year t.

The ideology of the incumbent is observed and measured by a rating organization. The ideology of the challenger is not directly observable; to estimate $CHALIDEO_{d,t}$, we presume that the ideology score of the challenger is equal to the average ideology score of all incumbents in the state from the same party (relying on Assumption 4 in *AJPS* article).

Ideally, there would be annual elections allowing us to observe the levels of citizen support for both the incumbent and a challenger. In practice, the best we can do is to estimate support for both the incumbent and a "hypothetical challenger" in a particular year using election results for the previous and following elections. We give greater weight to the election that is closer in time to the year of analysis, and we assume that change in public support for an incumbent is gradual throughout his/her term (see Assumption 5). For the purposes of describing the methodology, we treat the November election in a calendar year as occurring *after* the year is completed.[3] For example, we would speak of the 1988 House election as being followed by a 1989 session, a 1990 session, and then the 1990 election. Thus, we conceive of two years (numbered 1 and 2) "sandwiched" between successive House elections, and six years (numbered 1 through 6) occurring between consecutive elections for each Senate "district."

---

[1] We use an unweighted average because all House districts have roughly equal populations, and all Senate districts are statewide.

[2] Each Senator in a state is assumed to be elected in a separate (state-wide) district.

[3] We believe this is reasonable, since Congress rarely convenes in a lame duck post-election session.

In a given year t, assumed to be the T'th year after the most recent *Senate* election in district d (e.g., if the year to be estimated is 1984, and the last election was in 1982, then t=1984 and T=2), we assume that:

$$INCSUPP_{d,t} = [T/7][INCVOTE_{d,+}] + [(7-T)/7][INCVOTE_{d,-}] \quad [2]$$

where $INCVOTE_{d,-}$ denotes the proportion of the two-party vote received by district d's incumbent in the election preceding year t, and $INCVOTE_{d,+}$ is the vote share received by either the incumbent (if he/she is running) or his/her successor candidate from the same party in the election following year t. For cases in which a district's incumbent does not compete in the following general election, we assume that voters perceive the ideological position of the candidate from the incumbent's party to be similar to that of the incumbent (see Assumption 3 in *AJPS* article) so that the election can still be viewed a referendum on the incumbent's ideology.

Similarly, for the hypothetical challenger:

$$CHALSUPP_{d,t} = [T/7][CHALVOTE_{d,+}] + [(7-T)/7][CHALVOTE_{d,-}] \quad [3]$$

where $CHALVOTE_{d,-}$ denotes the proportion of the vote received by the challenger to district d's incumbent in the election preceding year t, and $CHALVOTE_{d,+}$ is the proportion received by the challenger to the incumbent party's candidate in the election after year t. For *House* districts, the formulae are similar, but reflect two years between elections rather than six:[4]

$$INCSUPP_{d,t} = [T/3][INCVOTE_{d,+}] + [(3-T)/3][INCVOTE_{d,-}] \quad [4]$$

$$CHALSUPP_{d,t} = [T/3][CHALVOTE_{d,+}] + [(3-T)/3][CHALVOTE_{d,-}] \quad [5]$$

When measuring INCVOTE and CHALVOTE for equations 2 through 5, if an incumbent is uncontested, we do not score INCVOTE as 1 and CHALVOTE as 0, as this choice would overstate the estimated level of support for the incumbent and thus misrepresent citizen ideology in the district. Indeed, when an uncontested election is held, our goal is to estimate the *contested vote share* -- the vote share that would have been received by an incumbent had a challenger run. We believe that a reasonable estimation strategy is to linearly interpolate the vote share of the incumbent between the most recent previous contested election and the first following contested election, and then add a "bonus" to the incumbent's interpolated share equal to 10% of the share estimated to be won by the challenger. For example, if in a district, the incumbent Republican representative ran unopposed in 1988 but was opposed in both 1986 (winning 69% of the vote)

[4] There are two situations in which equations 2 through 5 are inapplicable: (i) in House districts in the first two years following a decennial census (i.e., years ending in 1 or 2), where redistricting makes it so that $INCVOTE_{d,+}$ and $CHALVOTE_{d,+}$ cannot be measured for the same districts for which $INCVOTE_{d,-}$ and $CHALVOTE_{d,-}$ are observed, and (ii) for a district in a year in which the following election has not yet occurred (e.g., a Senate district for a term beginning in 1991). In these situations -- where information about the following election is absent -- we assume stability in the relative support for the incumbent and the challenger as revealed in the previous election.

and 1990 (garnering 59% of the vote), the incumbent's contested vote share in 1988 would be estimated as .64 (the average of .69 and .59) plus .036 [=(.10)(1.00-.64)], for a total estimated share of .676.[5]

There are some exceptions to our "linear interpolation with bonus" strategy for estimating the contested vote shares for an *un*contested election. First, we assume that an incumbent in an uncontested election would have won at least 60% of the vote had he or she been challenged. Second, we cannot interpolate vote shares if we do not have an observation *during the period of analysis* for both a previous contested election and a following contested election in the same district. If either a "post-" or "pre-" observation (but not both) is missing, we generally assume that the incumbent's contested vote share is equal to the incumbent's share in the one observed election plus a bonus of 10% of the challenger's share in that same election. However, ten years is the maximum length of time over which we are prepared to assume this kind of stability in vote shares. If either a "post- " or "pre-" observation, or both, are missing, and there is no observation within ten years of the election result we seek to estimate, we assume that the incumbent's contested vote share is equal to the floor value of .60.[6]

Through the above procedure, we can estimate $CITIDEO_{d,t}$ -- citizen ideology in year t in a congressional district, d. We conclude by aggregating to the state level using the formula:

$$CITIDEO_{s,t} = \left(\sum_{\substack{\text{all districts d} \\ \text{in state s}}} CITIDEO_{d,t} \right) / D_s \qquad [6]$$

where $CITIDEO_{s,t}$ represents citizen ideology in state s in year t, and D_s is the number of districts in state s.

[5] We give the incumbent a bonus in estimating the contested vote share of an uncontested election to reflect (i) a recognition that the interpolation is based purely on the observation of the outcomes of *contested* elections and (ii) an assumption that the absence of a challenger likely reflects a perception by potential challengers that the incumbent has an especially high level of public support in the district. The specific rule for calculating the bonus is designed so that its magnitude becomes smaller as the pre-bonus estimated vote share for the incumbent gets larger, consistent with the knowledge that a vote share is mathematically constrained not to exceed 1.00.

[6] Fortunately, during the years for which we implement our strategy for measuring citizen ideology, this 60% rule is operative for only a very small percentage of elections: 2.2%.

II. Procedure for Estimating Ideology Scores for Congressional Party Delegations with No Members

Our methodologies for measuring state ideology require knowledge of the average ideology score for members in both parties' congressional delegations for all years in the period of analysis. Let $DEMCONID_{s,t}$ denote the average ideology score for congressional Democrats (House and Senate) in state s in year t, and let $REPCONID_{s,t}$ be the average score for congressional Republicans in state s in year t. For the *AJPS* article (and companion projects), we estimated mean congressional party delegation ideology using scores provided by several different rating organizations (and for two different periods of analysis):

For 1960-93:

-- COPE scores: 1960-93,
-- ADA scores: 1960-93,
-- average of COPE and ADA scores (denoted COAD): 1960-93[7]

For 1977-92:

-- ACLU scores: 1977-92
-- average of COPE, ADA and ACLU scores (denoted ALL): 1977-92

The scores assigned by all these rating organizations range between 0 and 100, with greater values associated with greater liberalism.

In this supplement, we will designate the mean party delegation ideology for a particular state by indicating the rating organization(s) on which the measure is based (ADA, COPE, ACLU, COAD, or ALL), followed by a party code (D or R), and then followed by a two digit state code, as follows:

1 AL	11 HI	21 MA	31 NM	41 SD
2 AK	12 ID	22 MI	32 NY	42 TN
3 AZ	13 IL	23 MN	33 NC	43 TX
4 AR	14 IN	24 MS	34 ND	44 UT
5 CA	15 IA	25 MO	35 OH	45 VT
6 CO	16 KS	26 MT	36 OK	46 VA
7 CT	17 KY	27 NE	37 OR	47 WA
8 DE	18 LA	28 NV	38 PA	48 WV
9 FL	19 ME	29 NH	39 RI	49 WI
10 GA	20 MD	30 NJ	40 SC	50 WY

[7] COAD is the delegation score used to construct the 1960-93 ideology measures discussed in the *AJPS* article.

For example, COAD_R(15) denotes mean congressional delegation ideology based on an average of COPE and ADA scores for IA (state 15) Republicans.

In some states, in some years, however, there are no members of Congress from one of the parties. (For the period 1960-93, 30 of the 100 state party delegations did not exist in at least one year; of the 3400 state-party-year observations during this period, 244 -- or 7% -- were missing. For the period 1977-92, 12 of the 100 state party delegations did not exist in at least one year; of the 1600 state-party-year observations during this period, 84 -- or 5% -- were missing.) We need to estimate what the "mean ideology" score for these delegations would have been had members from these parties been elected in these years.

For example, assume that state **s** had no Democratic congressional delegation in at least one year during the period of analysis. This means that one or more values of $DEMCONID_{s,t}$ needs to be estimated. Our strategy was to estimate a time-series regression (where the units are the observations of state **s** for each year during the period) of $DEMCONID_{s,t}$ on at least three independent variables. One is the ideology score for the other party in the same state (in the example, $REPCONID_{s,t}$). The other regressors are the ideology scores for the same party in *other* states. We included variables for each *neighboring* state for which there are at most four missing DEMCONID observations during the period of analysis, and for which there are no missing observations in any year for which $DEMCONID_{s,t}$ is missing.[8]

If for state **s**, there were not at least two variables based on neighboring states that satisfied the above conditions, we supplemented the regressors to a level of three by choosing states in the following order (if they exist): (i) the nearest state for which data are available for a sufficient number of cases (distance measured from capital to capital) having the same political culture as state **s** [where equality of political culture is judged based on the eight categories on Elazar's (1984, 135) map], (ii) the second nearest state for which sufficient data are available with an identical political culture, and (iii) the state having the closest score on Wright, Erikson, and McIver's [hereafter WE&M] (1985, Table 3) weighted state ideology scale (for which sufficient data are available).[9] [If there is a tie involving two or more states with respect to similarity on the WE&M scale, we break the tie by using the state geographically closest.]

---

[8] In determining which states are neighbors, we use Berry and Berry's (1990) rules [which treats some pairs of states that are very close as neighbors even if they do not share a common border] (see their appendix).

[9] When state **s** is AK, we modified slightly the distance rule for choosing among states with an identical political culture, since the differences in distance among the like-cultured states east of the Mississippi River seemed trivial in a substantive sense. The states with cultures identical to AK are NV, IL, IN, OH, PA, NJ, MD and DE. Since NV stands out from the rest in being appreciably closer to AK, we use NV as a regressor state if sufficient data are available for it. Then among the states east of the Mississippi, we take states in order of their similarity to AK's score on the WE&M scale [.166, as estimated by Ringquist (1993)]. IN is next, since it is appreciably closer on the WE&M scale than the rest (.159). Next is DE (.127). These three states were sufficient to apply the methodology to all AK cases.

Thus, in the illustration, the regression conducted is

$$DEMCONID_{s,t} = a_0 + b_0 REPCONID_{s,t} + b_1 DEMCONID_{1,t} + b_2 DEMCONID_{2,t} + ... + b_k DEMCONID_{k,t} \quad [7]$$

where 1,2, ..., k represent a set of k (>=2) other states. Missing values of $DEMCONID_{s,t}$ are then estimated by using the coefficient estimates to predict values of the dependent variable given the values of the independent variables during the relevant years.

There were two types of cases where estimating this regression was either impossible or inappropriate:

> (i) For ACLU_D(29) and ACLU_D(14) [for the 1977-92 data set], there was a high enough level of collinearity that coefficient estimates could not be calculated. Our strategy for these cases was to remove the independent variable for the state most dissimilar on the WE&M scale and attempt to reestimate. If coefficient estimates still could not be computed, we reinserted the excluded variable, removed the independent variable for the state next most dissimilar on the WE&M scale, and tried to reestimate. We stopped this procedure the first time we found a regression equation that could be estimated. For ACLU_D(29), we stopped when the variable for MA was removed, and for ACLU_D(44), we stopped when the variable for CO was deleted.
>
> (ii) If estimating the regression equation for a particular state party would require estimation of more missing values than the number of cases on which the regression would be conducted, we assumed that there were too many years in which a party delegation is missing to base our estimation on information about that state (and its neighbors) alone. In such cases [four of them], we pooled data across states, and used an estimation procedure based on the pooled data. [We describe this alternative procedure in the last paragraph of section II.]

We accept as reasonable estimates of missing ideology scores for state party delegations for which all three of the following conditions regarding regression equation 1 hold:

(condition a) There are no estimated ideology scores greater than 110 or less than -10[10].

---

[10] Despite condition a, any predicted ideology scores outside the range between 0 and 100 are adjusted to 0 or 100 for the final estimated scores. Final scores for the following state parties include one or more estimated values that are set to 0 or 100:

For the 1960-93 data set: ADA_R(4), ADA_R(10), ADA_R(18), ADA_R(31), ADA_R(40), ADA_R(48), COPE_R(10), COPE_R(28), COPE_R(40), COPE_R(48), COAD_R(28), COAD_R(40).

For the 1977-92 data set: ACLU_D(12), ACLU_D(44), ACLU_R(21), ALL_D(31).

(condition b) either (i) the largest coefficient estimate from the set of partial slope coefficient estimates for those variables that are ideology scores for the *same* party in another state is positive, or (ii) there is no coefficient estimate in this set of partial slope coefficient stimates that is more than weakly negative [defined operationally as no negative coefficient estimate with a t-ratio magnitude in excess of .40].[11]

(condition c) The R^2 value is greater than .33.

The majority of state parties for which estimation of missing delegation scores was necessary produced regression results which satisfied the above three conditions: in particular, in the case of COAD for 1960-93, 19 of 30 state parties, and in the case of ALL for 1977-92, 5 of 8 state parties.

If conditions a and b are met for a state party, but the R^2 value for the estimation regression is less than .33, we also accepted the associated estimates of ideology scores, since although we recognize that the quality of the estimates may be weaker than for those state parties satisfying condition c, we did not believe that we could do any better with a different estimation strategy.[12] However, see below for a strategy of exposing estimates based on a regression with an R^2 less than .33 to a further test.

We supplemented the data set consisting of *observed* party delegation ideology scores with all estimated party delegation scores that we deemed reasonable by the above standards.

This left us cases of two types: (i) the estimation regression produces predicted ideology scores in at least one year that are less than -10[13], and (ii) the largest coefficient estimate among the set of partial slope coefficient estimates for those variables that are ideology scores for the

---

[11] While theory leads to an unambiguous prediction that the slope coefficient for variables representing ideology scores for the *same* party in a neighboring (or like-cultured) state should be positive, there are reasonable competing predictions about the sign of the slope coefficient for the ideology score for the *other* party in the same state. A theory which suggests that both parties will "chase" a moving median voter over time would predict a positive coefficient, while a theory which suggests that both parties tend to move closer to the median voter (each from a different side) in election years and then toward opposite ideological extremes during nonelection years would predict a negative relationship. For this reason, we use the sign of the coefficient estimates for variables measuring the *same* party in other states to test the quality of the estimation strategy, but do not use the sign of the coefficient estimate for the variable measuring the *other* party in the same state for this purpose.

[12] In effect, we reject the strategy of trying (on an ad hoc basis) ideology scores for parties from more distant and less culturally similar states as there is no reasonable theory that would suggest that such states are more appropriate for estimation purposes. We could try many combinations of other states, and rely on the combination of states that maximizes R^2, but we would be capitalizing too much on chance factors.

[13] For the construction of 1960-93 data, ADA_R(48), COPE_R(11), COPE_R(48), and COAD_R(48). For 1977-92 data, ALL_D(31) and ALL_R(34).

same party in another state is *negative*.[14] We used the following rules to modify the regression estimation procedure for these cases.

Rule 1

If any variable for a neighboring state was excluded from an original regression because of too many missing observations, we added that variable (now supplemented with the estimates for the missing observations) to the regression. If one or more states were included as regressors in the original regression based on an identical political culture, we deleted one identical-culture state (beginning with the one farthest away) for each new neighboring state that was added.[15]

Rule 2

In a few cases, Rule 1 leads to no change in the estimating regression equation. We deal with different state party delegations separately:

(i) Hawaii, ADA_R(11), 1960-93: Here, MO and CO were the original other-state regressors. MO was the first state included because it was the only state with an identical political culture to HI; CO was included because it was tied with Iowa in having the WE&M score most similar to HI's estimated WE&M score [.092, as estimated by Ringquist (1993)]. In the regression, the coefficient for CO was negative. The change

[14] For the construction of 1960-93 data, ADA_D(27), COPE_D(27), COAD_D(27), ADA_R(7), ADA_R(11), ADA_R(12), COAD_R(4), COAD_R(7), and COAD_R(48). For the 1977-92 data, ALL_D(12), ACLU_R(21), ACLU_R(34), and ALL_R(34).

[15] This choice reflects our belief that, a priori, neighbors provide better information.

Implementing Rule 1 led to the following changes:

For construction of 1960-93 data:

ADA_R(7): add RI; ADA_R(12): add MT, NV; COAD_R(4): add MS, TX; COAD_R(7): add RI; ADA_D(27), COPE_D(27) & COAD_D(27): add KS, SD, WY.

For construction of 1977-92 data:

ALL_D(12): add UT; ALL_D(31): add UT.

These changes resulted in estimates for the following state parties that satisfy conditions a and b:

For construction of 1960-93 data: COAD_R(4)

For construction of 1977-92 data: None

But the changes produced results that failed to satisfy both conditions for the remaining state parties.

most consistent with the spirit of our other rules was to keep MO in the equation, but substitute Iowa for CO. With this change, conditions a and b were satisfied.

(ii) Hawaii, COPE_R(11), 1960-93: Again, MO and CO were the original other-state regressors. The initial estimation regression produced positive estimated partial slope coefficients for all other-state variables, but some very large negative predicted ideology scores. We tried the same procedure as in (ii) -- replacing CO with IA -- with the same basic result: positive partial slope coefficients, but some large negative predicted values. We continued this process of replacing the state included based on WE&M scores with the next closest state on the WE&M scale. We decided to stop only if we tried the five states closest on the WE&M scale and always failed to satisfy conditions a and b, or if we tried all states within a distance of .10 on the WE&M scale without satisfying these conditions.[16] After replacing IA with OH, NH and then MD, we had tried regressions with the five most similar "WE&M states." In every one, both partial slope coefficient estimates were positive, yet there were some large negative predicted ideology values in some years. Seeing no better option, we decided to return to the original estimates (based on MO and CO) and rely on these (yet modified to set the four "below-zero" predictions equal to zero).

(iii) In the case of WV COPE_R and COAD_R, all neighbors had been included in the original equation. We decided to treat this case according to Rule 3 below.

Rule 3

After implementing Rules 1 and 2, several problem cases still remained.[17] At this point, we had added all the states that we were comfortable adding, and decided that we needed to exclude regressors that are producing inconsistent sign estimates. We excluded the variable for another state producing the largest negative unstandardized slope coefficient estimate, and reestimated. If the resulting estimates satisfied conditions a and b, we stopped. If not, we continued the process of excluding the other-state variable with the largest negative coefficient until conditions a and b were satisfied. If this step at any point led to a reduction in the number of other-state regressors to less than two, we added the state with the closest WE&M score that was not already included. Similarly to our decision in Rule 2 (iii), we decided to abandon Rule 3 only if we tried the five states closest on the WE&M scale and always failed to satisfy conditions a and b, or if we tried all states within a distance of .10 on the WE&M scale without satisfying these conditions.

Rule 4

Rule 3 failed in the estimation of ALL_R(34) for the 1977-92 data set after including variables for the five states closest to ND on the WE&M scale (MS, OK, UT, SC and AL).

[16] The WE&M scale ranges from .008 to .396, if the single outlier -- NV -- is excluded.

[17] For 1960-93 data construction: ADA_D(27), COPE_D(27), COAD_D(27), ADA_R(7), ADA_R(12), COPE_R(48), COAD_R(7), COAD_R(48). For 1977-92: ACLU_D(29), ACLU_D(44), ALL_D(12), ALL_D(31), ACLU_R(34), ALL_R(34).

However, the above rules permitted satisfactory estimation of all missing values for ADA_R(34) and COPE_R(34) for 1960-93, and the estimation of all missing values for ACLU_R(34) for 1977-92. We computed missing values for ALL_R(34) by averaging values of ADA_R(34), COPE_R(34) and ACLU_R(34).

There were numerous state parties for which the final estimation regression satisfied conditions a and b, but not c (i.e., the R^2 value was below .33) [see list below].[18] To test for the sensitivity of the error we believe almost certainly exists in these estimates, we also used an entirely different estimation strategy (based purely on information about the history of the "same party in the same state"): For those state parties where all missing observations are "bounded" on both ends, the "different" estimation strategy is linear interpolation. For those state parties where there are missing cases that are "bounded" on only one side (e.g., in constructing data for 1960-93, and there are data for all years 1960- 82, but no observations after 1982), the "different" estimation strategy is assuming that all missing ideology scores are identical to the nearest observed "boundary score" (in the example, values for 1983 through 1992 would be set to the [boundary] value in 1982).[19] Then, for each of these state parties, we compared the estimates produced by the two strategies (our regression-based estimates and our internal-history based estimates) by calculating the time-series correlation between the two series.

For 1960-93 data, the correlations (over the 34 cases from 1960-93) are as follows:

COAD_R(12):	.99
ADA_R(41):	.99
COPE_D(12):	.99
COAD_R(41):	.99
ADA_R(2):	.98
COPE_R(41):	.98
ADA_D(41):	.97
ADA_R(31):	.97
COPE_R(26):	.96

[18] For 1960-93 data, there were 25 such cases. The distribution of the R^2 for the final regression used for estimation is as follows:

R^2	Freq.
>= .30	6
>= .25 & < .30	9
>= .20 & < .25	5
>= .15 & < .20	4
= .14	1

Thus, even among cases in which the estimating regression has an R^2 less than .33, the distribution is dominated by values in excess of .25.

[19] For most state parties, the over-time series for average ideology score is sufficiently "volatile" that such an interpolation/extrapolation strategy is not really substantively plausible.

ADA_R(12):	.95
COPE_R(18):	.92
COAD_D(50):	.92
COAD_R(10):	.91
COPE_D(50):	.90
COAD_R(1):	.90

ADA_D(29):	.89
ADA_R(1):	.88
ADA_R(4):	.88
ADA_D(16):	.87
COPE_D(45):	.87
ADA_D(27):	.86
COAD_R(18):	.81

COAD_R(39):	.75
ADA_R(34):	.74

COAD_D(27):	.65

Fifteen of these correlations are at .90 or greater; an additional seven are between .80 and .90. Only three correlations are below .75, and only one is below .70.

For 1977-92 data, only in one case -- ALL_D(12) -- was the final data constructed based on a regression estimation equation with an R^2 less than .33. The correlations (over the 16 cases from 1977-92) between the ALL_D(12) series based on regression estimates and the same series based on internal-history estimates is .97.

Finally, if estimating the above regression procedure for a particular state party would require estimation of more missing values than the number of cases on which the regression would be conducted, we assumed that there were too many years in which a party delegation is missing to base our estimation on information about that state (and its neighbors) alone. In such cases, we pooled data across states, and used an estimation procedure based on the pooled data. For the 1960-93 data set, this alternative procedure was not necessary in any case. However, for the 1977-92 data set, the individual-state procedure would yield a regression equation estimated with four observations for HI Republicans and AK Democrats, and with only two observations for WY Democrats and WV Republicans. To estimate missing observations for these four delegations, for each party, we pooled all state-years between 1960 and 1993 for which congressional delegations existed, and regressed a delegation's ACLU or ALL score on both its COPE and ADA scores.[20] The regression coefficient estimates were then used along with estimated or observed values of COPE and ADA from the 1960-93 data set to predict missing values for the dependent variable.

[20] The R^2 values for these four regressions were: ACLU_D, .75; ALL_D, .97; ACLU_R, .76; ALL_R: .96.

III. Sources of Data in *AJPS* Article

Congressional Ratings. For 1960-82, ADA and COPE scores for individual representatives and senators were drawn from ICPSR data set #7645 (originally collected by Congressional Quarterly). For 1983-91, ADA and COPE data were taken from various issues of *Congressional Quarterly Weekly Report*. CQ stopped publishing all ratings data after 1991 (and has never published ACLU ratings). Therefore, 1992-93 ADA and COPE scores, along with 1977-92 ACLU ratings, were obtained directly from these organizations.

Party of Governor. For years through 1983, these data were taken from ICPSR data set #16 (originally collected by Walter Dean Burnham). After 1984, data on party control of the executive were drawn from selected issues of *Statistical Abstract*.

Party Composition of State Legislatures. For years through 1975, biennial observations were drawn from ICPSR data set #16; the same data set includes annual observations from 1976-83. We were able to account for most observations not available from ICPSR by relying on published reports from various sources, including *Book of the States*, *Statistical Abstract*, and *State Yellow Book*. This left us with two years, 1986 and 1988, which could not be obtained from published reports. Data for these years were obtained directly from the states. The final data set thus contains yearly observations for each state for all years 1960-93.

Congressional Election Results. Election results for the House were provided by Gary Jacobson for all years between 1960 and 1992. Data for the Senate were collected by the authors from CQ's *America Votes*.

IV. References

Berry, Frances S. and William D. Berry. 1990. "State Lottery Adoptions as Policy Innovations: An Event History Analysis." *American Political Science Review* 84: 395- 415.

Elazar, Daniel. 1984. *American Federalism: A View from the States*, 3rd Edition. New York: Harper.

Ringquist, Evan J. 1993. *Environmental Protection at the State Level: Politics and Progress in Pollution*. Armonk, NY: M.E. Sharpe.

Wright, Gerald C., Jr., Robert S. Erikson, and John McIver. 1985. "Measuring State Political Ideology with Survey Data." *Journal of Politics* 47:469-89.