Political Analysis Advance Access published August 11, 2005

doi:10.1093/pan/mpi030

Applying a Two-Step Strategy to the Analysis of Cross-National Public Opinion Data

Karen Long Jusko
Department of Political Science, University of Michigan,
5700 Haven Hall, 505 South State Street, Ann Arbor, MI 48109-1045
email: kjlong@umich.edu (corresponding author)

W. Phillips Shively
Department of Political Science, University of Minnesota,
1414 Social Sciences Bldg. 267 19th Ave. South, Minneapolis, MN 55455
email: shively@polisci.umn.edu.

In recent years, large sets of national surveys with shared content have increasingly been used for cross-national opinion research. But scholars have not yet settled on the most flexible and efficient models for utilizing such data. We present a two-step strategy for such analysis that takes advantage of the fact that in such datasets each "cluster" (i.e., country sample) is large enough to sustain separate analysis of its internal variances and covariances. We illustrate the method by examining a puzzle of comparative electoral behavior—why does turnout decline rather than increase with the number of parties competing in an election (Blais and Dobryzynska 1998, for example)? This discussion demonstrates the ease with which a two-step strategy incorporates confounding variables operating at different levels of analysis. Technical appendices demonstrate that the two-step strategy does not lose efficiency of estimation as compared with a pooling strategy.

1 Introduction

Which strategies are best suited to the analysis of cross-national public opinion data? By facilitating the study of political behavior across a variety of political contexts, cross-national public opinion projects enable researchers to link individual-level outcomes to institutional settings through increasingly "large-N" analyses. Researchers now routinely use the Comparative Study of Electoral Systems, World Values Survey, and the various Barometer project data in their studies of comparative political behavior. These datasets, and others like them, offer some special opportunities for analysis that current strategies for estimation do not appear to take full advantage of. Until recently, as we point out in connection with Table 1 below, the main ways such data were used were either to pull out

Authors' note: We are grateful for the comments and suggestions offered by the participants of the Princeton University conference on hierarchical models, especially Chris Achen and Larry Bartels.

Table 1 Traditional practice in the analysis of cross-national public opinion data

Type of strategy	*Frequency*
Partition strategies	
Comparison of means, proportions, etc.	29
Comparison of within-country multivariate analyses	22
Pooling strategies	
Aggregate-level analyses	18
Without corrections	10
With country indicators only	6
With country indicators and heteroscedastic corrections	7
With weighting factors to correct for differences in sample size	4
With country indicators and weighting factors to correct for differences in sample size	1
Multilevel strategies	
Varying intercepts	3
Varying coefficients	2

Note. The total number of articles using Comparative Study of Electoral Systems, Eurobarometer or European Community Surveys, the International Social Survey Program, or World Values Survey data, published in the *American Journal of Political Science* (1990–2003), *American Political Science Review* (1990–2001), *British Journal of Political Science* (1990–1999), *Comparative Politics* (1990–1999), *Political Behavior* (1990–1999), *Political Research Quarterly* (1993–2003), or *Public Opinion Quarterly* (1990–1999), is 74. As some authors use several strategies, this table reports the frequency of each strategy rather than the number of articles using each strategy. The periods chosen were those available through the JSTOR search engine.

countries for separate, independent analyses that were then compared fairly casually (thus losing the large-N data structure), or by pooling all of the data into one large dataset (with problematic results for estimation of standard errors).

Recently there has been much interest in using hierarchic linear models for these data, and this has been a huge step forward. But even current hierarchic linear models can fail to take full advantage of the special character of these datasets, especially the advantages offered by the fact that each cluster (country sample) in one of these sets is large enough to provide reasonable estimation of its internal variances and covariances. Usually, hierarchic models make no assumption that one can estimate directly individual-level variances and covariances; rather they borrow from the variances and covariances of all clusters to come up with common estimates.

We present in this article a two-step estimation strategy that draws heavily on meta-analysis, a variant of standard hierarchical linear models. Meta-analysis has traditionally been used to identify patterns in published results, generally in the medical sciences. This "study of studies" framework fits well with the collaborative research design of cross-national public opinion projects: The country studies that comprise the Comparative Study of Electoral Systems or World Values Survey are often stand-alone projects primarily, with the common survey module tacked on at the beginning or the end of an instrument designed explicitly for the analysis of each country's unique political dynamics. Although meta-analysis often aims to generate more reliable parameter estimates for the component studies (referred to as "shrinkage" estimates), this framework easily incorporates as well

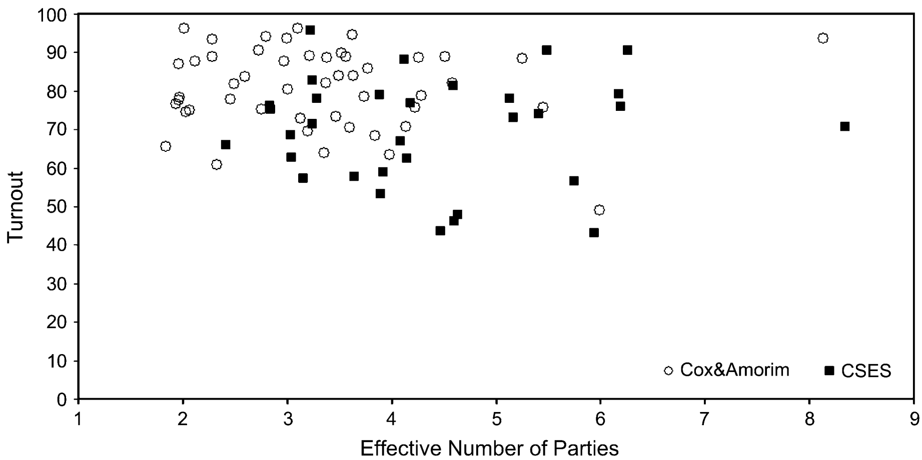

Fig. 1 Turnout and multiparty competition. The vertical axis reports the inverse normal transformation of p_j, the turnout proportion observed in each election. The horizontal axis reports the effective number of parties competing in the corresponding election, estimated using the Laakso-Taagepera index (Laakso and Taagepera 1979). *Sources:* Effective Number of Parties: Cox and Amorim (1997); CSES Modules 1 & 2. Voter turnout: IDEA (2004).

variables operating at and across different levels of analysis. This framework also offers more flexibility in modeling individual-level processes for each country; it is possible to use different sets of individual-level independent variables from one country to another, for instance. Most important, as we demonstrate in Appendices A and B, this added flexibility comes with no loss in efficiency.

Before presenting the model below, let us set it up with an example addressing a puzzle of comparative democratic politics: Advocates of multiparty elections often consider the number of parties competing in elections as an indication of whether or not the election presents voters with a meaningful choice set (Dahl 2002). We might reasonably expect turnout to increase with the number of parties competing. However, as seen in Fig. 1, voter turnout does not increase with the number of parties competing in an election.

Each of the data points in Fig. 1 is an election; the set is drawn from 60 countries over the period since 1980. The vertical axis reports the proportion of registered voters who cast ballots in each election, while the horizontal axis reports the effective number of parties for each election.[1] Square points are elections drawn from the Comparative Study of Electoral Systems database, either in its first module (1996–2001) or its second module (2002–2006). To demonstrate that the nonincreasing relationship seen here is not idiosyncratic to the elections included in the CSES database, additional cases from Cox and Amorim (1997) are included, signified by circular points; they display the same nonincreasing relationship as is seen in the CSES data.

One explanation for this pattern of association might be, as Downs (1957) and others (Blais and Dobrzynska 1998; Jackman 1987, for example) suggest, that the information costs of voting in multiparty systems discourage voters. Taking this as our hypothesis, we will examine the relationship between the information costs of voting and the number of

[1]The effective number of parties reports an estimate of the number of parties competing in an election, weighted by their share of the popular vote. In our example, we interpret this measure as the number of viable alternatives available to the average voter, i.e., the number of parties voters are likely to be aware of without much research (Laakso and Taagepera 1979).

parties competing in an election. This example will allow us to identify shortcomings in common practices for the analysis of cross-national public opinion data, and propose a strategy well-suited to comparative political research.

2 Traditional Practice

How can we best analyze the sources of a relationship like that of Fig. 1, using datasets consisting of collections of national surveys? To review traditional ways in which this problem has been addressed, let us consider first the basic individual-level model of interest, described in Eq. (1). Let

$$VOTE_{ij} = INFO_{ij}\beta_j + e_{ij}, \tag{1}$$

where by assumption, $E(e_{ij}) = 0$ and $E(e_j{}'e_j) = \sigma_j{}^2 I$ characterize the true relationship between individual-level variables *VOTE* and *INFO*, in each country $j = 1, 2 \ldots J$, for individuals $i = 1, 2 \ldots n_j$. Equation (1) implies that β_j varies across countries.[2] To evaluate this relationship, analysts often partition cross-national public opinion data, estimating the parameters separately for each country sample, perhaps using an ordinary least squares (OLS) estimator.[3]

$$\hat{\beta}_{\mathbf{j}} = (\mathbf{INFO_j' INFO_j})^{-1}\mathbf{INFO_j' VOTE_j} \tag{2}$$

$\mathbf{INFO_j}$ is the $n_j \times 1$ matrix containing the observed values of *INFO* for respondents in country *j*. $\mathbf{VOTE_j}$, also $n_j \times 1$, indicates whether or not each respondent in country *j* voted. Then, under the usual conditions $\hat{\beta}_j$ reports the unbiased least-squares parameter estimate for country *j*.

Conclusions drawn from partitioning strategies necessarily include proper names: "*INFO* is positively related to *VOTE* in countries Canada, Great Britain, Switzerland, and the United States, but not in Belgium or Australia" (see Appendix C). Although an analyst may identify patterns among clusters of countries, a partitioning strategy proceeds at only subnational levels of analysis. As a result, although coefficient estimates can be compared across models in a casual way (i.e., apparent patterns or clusters may be identified), analysts who use partition strategies lack the ability to draw general conclusions about how features of political systems structure political behavior.

[2]We acknowledge that a model with a binary-dependent variable is a more appropriate specification. However, to facilitate this discussion, we use the linear specification here. Jusko (2005) extends this discussion to binary response models.

The individual-level model Eq. (1) may also include other individual-level variables, as well as a constant. Our discussion is considerably simplified, although without loss of generality, by focusing on the role of political information in turnout. Moreover, although in this example the same individual-level model is estimated for each country, this is not necessary. For instance, the effect of information on turnout in Spain was different for Basques than for non-Basques. As it turned out, there were sufficiently few Basques in the sample that this did not affect the overall analysis significantly. But had Basques been more numerous, we could have estimated the model for Spain including a Basque/non–Basque term in that country only.

The inclusion of a constant would allow analysts to evaluate variance in the country fixed effects. Often, varying intercepts are of primary interest to those using hierarchical linear models. Although our discussion is simplified by their exclusion, the two-step strategy we propose applies to analyses using intercepts as the dependent variable in the second stage.

[3]This strategy is equivalent to the specification of a model in which an indicator for each country is included in an interaction with each individual-level variable, and variance parameters are permitted to vary across countries.

Alternatively, as seen in Table 1, analysts sometimes pool (or stack vertically) public opinion data across country samples and estimate a common parameter for the entire cross-national sample.[4]

$$VOTE_{ij} = INFO_{ij}\bar{\beta} + w_j + \epsilon_{ij}. \tag{3}$$

Note that pooling strategies effectively reduce cross-national variation to residual variance:

$$w_j = INFO_{ij}(\beta_j - \bar{\beta}) \tag{4}$$

where

$$\bar{\beta} = \frac{1}{m}\sum_{j=1}^{m}\beta_j. \tag{5}$$

Of course, if

$$E(\beta_j - \bar{\beta}) = 0 \tag{6}$$

there would be no cause for concern. In fact, a pooling strategy requires that there be no systematic deviations in the relationship between *VOTE* and *INFO*.

As suggested earlier, the complexity of decision making may increase with the number of parties competing in an election: Whether voters see elections as processes of government selection or as "expressions of preference," there should be an increasing relationship between the information costs of voter decision making (the information a voter must have available in order to choose between competing parties) and the number of parties competing in an election (Downs 1957 provides further discussion of this point). When these information costs exceed the benefits citizens expect to derive from an election, they may abstain from voting.

Some features of electoral systems, however, may act to focus attention on just a few of the competing parties, thereby moderating the effect of the number of parties on the information costs of voting. We suspect that concurrent presidential elections might work this way (see Golder forthcoming, for recent work on this topic). Building on the model presented in Eq. (1), the relationship between information costs of voting and the number of parties can incorporate concurrent presidential elections as:

$$\beta_j = \gamma_1 + PARTIES_j\gamma_2 + PRES_j\gamma_3 + (PARTIES_j \times PRES_j)\gamma_4 + u_j \tag{7}$$

where *PRES* is a binary variable, indicating whether there is a concurrent presidential election, $\times$ denotes scalar multiplication, and by assumption,

$$E(u_j) = 0 \text{ and } E(\mathbf{uu}') = \tau^2\mathbf{I} \equiv \mathbf{T}. \tag{8}$$

$\gamma_2 > 0$ would indicate that the information gap between voters and nonvoters, or the information costs of voting, increases with the number of parties competing in an election.

[4]Aggregate-level analyses are included as pooled analysis in Table 1 because the quality of their parameter estimates depends on the same assumptions as those underpinning pooling strategies.

$\gamma_4 < 0$ would suggest that concurrent presidential elections work to decrease the effect of the number of parties competing on the information costs of voting.

Suppose, however, that a pooling strategy is used, and

$$\hat{\beta}_{\mathbf{pool}} = \left(\sum_{\mathbf{j=1}}^{\mathbf{J}} \mathbf{INFO}_\mathbf{j}' \mathbf{INFO}_\mathbf{j} \right)^{\mathbf{-1}} \sum_{\mathbf{j=1}}^{\mathbf{J}} \mathbf{INFO}_\mathbf{j}' \mathbf{VOTE}_\mathbf{j} \tag{9}$$

is estimated for the entire cross-national sample. Then, using a general form of Bartels (1996) argument, the bias in Eq. (9) is calculated according to the following expression:[5]

$$\begin{aligned} E(\hat{\beta}_{\mathbf{pool}}) &= \left(\sum_{\mathbf{j=1}}^{\mathbf{J}} \mathbf{INFO}_\mathbf{j}' \mathbf{INFO}_\mathbf{j} \right)^{\mathbf{-1}} \left(\sum_{\mathbf{j=1}}^{\mathbf{J}} \mathbf{INFO}_\mathbf{j}' \mathbf{INFO}_\mathbf{j} \cdot \beta_\mathbf{j} \right) \\ &= \bar{\beta} + \left(\sum_{\mathbf{j=1}}^{\mathbf{J}} \mathbf{INFO}_\mathbf{j}' \mathbf{INFO}_\mathbf{j} \right)^{-1} \\ &\quad \times \left(\sum_{\mathbf{j=1}}^{\mathbf{J}} \mathbf{INFO}_\mathbf{j}' \mathbf{INFO}_\mathbf{j} \cdot (\gamma_1 + PRES_j\gamma_3 + PARTIES_j(\gamma_2 \right. \\ &\quad \left. + PRES_j\gamma_4) - \bar{\beta}) \right) \end{aligned} \tag{10}$$

with $\bar{\beta}$ defined as

$$\bar{\beta} = \gamma_1 + \overline{\mathbf{PRES}}\gamma_3 + \overline{\mathbf{PARTIES}}(\gamma_2 + \overline{\mathbf{PRES}}\gamma_4). \tag{11}$$

If the covariance of *PARTIES* and *PRES* is nonzero (systems with concurrent presidential elections are more likely to have fewer parties, for example), $\hat{\beta}_{pool}$ will be biased in the direction of γ_4. Assumption (6) can now be reformulated, neatly summarizing the key challenge of comparative research:

$$E(\beta_j | \mathbf{ELECTION_j}) = \bar{\beta} \tag{12}$$

where $\mathbf{ELECTION_j}$ includes all features of the election held in country *j*. Neither pooling strategies nor partitioning strategies allow us to deal fruitfully with the challenge implicit in this assumption. An ideal strategy for the analysis of cross-national research would promote broad comparisons (which partitioning fails to do), while allowing incorporation of confounding variables that operate across levels of analysis (where pooling may be misleading). We therefore apply a two-step strategy for analysis.

3 A Two-Step Strategy for Analysis

To estimate the β_js and γs, we use data from the Comparative Study of Electoral Systems, a collaborative public opinion project that binds post-election studies in over 40 countries through common survey modules. Both Module 1 (for elections held between 1996 and

[5]Here, to simplify the argument, we assume $\sigma_j^2 = \sigma^2$ for all *j*.

2001) and Module 2 (for elections held since 2002) include measures of political knowledge and turnout, yielding 32 legislative election cases (from 26 countries).[6]

Before proceeding to more complex analyses, some investigation of the individual-level relationships between political information and turnout is important, to ensure that the model accurately reflects the data generating process. The model

$$VOTE_{ij} = \alpha_j + INFO_{ij}\beta_j + e_{ij} \tag{13}$$

is specified, and the parameters α_j and β_j are estimated for each country sample, using the usual ordinary least squares (OLS) estimator, given by Eq. (2). Here, *INFO* is *i*'s within-country mean-centered score on a three-item political information index, which counts correct responses to a series of increasingly difficult questions about the election, elites, or *i*'s political system more generally. With few exceptions, the relationship between political information and turnout is positive and linear to a good approximation in each country.[7] Appendix C lists estimates of $\hat{\alpha}$, $\hat{\beta}$ and their standard errors, for each country, as well as the descriptive statistics for the individual-level variable, *INFO.* This completes the first step of the two-step strategy.

Figure 2 plots OLS estimates of $\hat{\beta}$ for each country and the effective number of parties competing in each election. As anticipated by Eq. (7), although with a few notable exceptions (discussed below), there is a positive and approximately linear relationship between the number of parties competing in the election and information costs: As the number of parties increases, so does the estimated effect of one's political knowledge on the likelihood that one will vote.

There are several features of the pattern of association reported in Fig. 2 worth noting. First, observe the location of parameters estimated for different elections held in the same country, but at different points in time (e.g., Poland). Their close proximity seems to confirm a more general relationship between the information costs of voting and turnout.[8] Moreover, this pattern suggests that if it is the complexity of decision making in a prior election that determines the difficulty of decision making at the current election (rather than characteristics of the current election itself), the substantive conclusions would not change.[9]

Second, there does indeed seem to be some correlation between concurrent presidential elections (indicated with solid points) and the number of parties competing in an election.

[6]Several elections were excluded from this analysis because they were exclusively presidential elections or because the required data are not available.

Usually, the "Level-2" unit of analysis is the country or system. However, in the analysis that follows, "election" is the more theoretically and empirically appropriate unit of analysis. Results are reported for both the entire sample of elections (labeled A in Table 2) and the set of unique country observations, including only the earlier election for those countries in which there are two (labeled B in Table 2). Note also that the linguistic regions of Belgium are treated as independent observations: These were independent studies, conducted by different research teams, and with different sample designs. Further, their parallel and exclusive party systems imply that the voter decision-making processes may vary across these regions.

[7]The following analysis was conducted excluding those countries in which there appears to be the largest deviations from linearity, with no difference in results.

[8]Any significant differences can reasonably be attributed to short-term shocks to the political system and are consistent with the expected direction of the relationship. The 1996 New Zealand election, for example, was the first election held under mixed member plurality rules. It seems to be the case that the high salience of this election, resulting from an intense publicity campaign, decreased the information costs associated with voting.

[9]Recall that the effective number of parties is calculated using electoral outcome data, resulting in a theoretical timing inconsistency. We reasonably assume that voters have access to accurate polls prior to the election and are able to estimate the number of viable choices with which they are confronted.

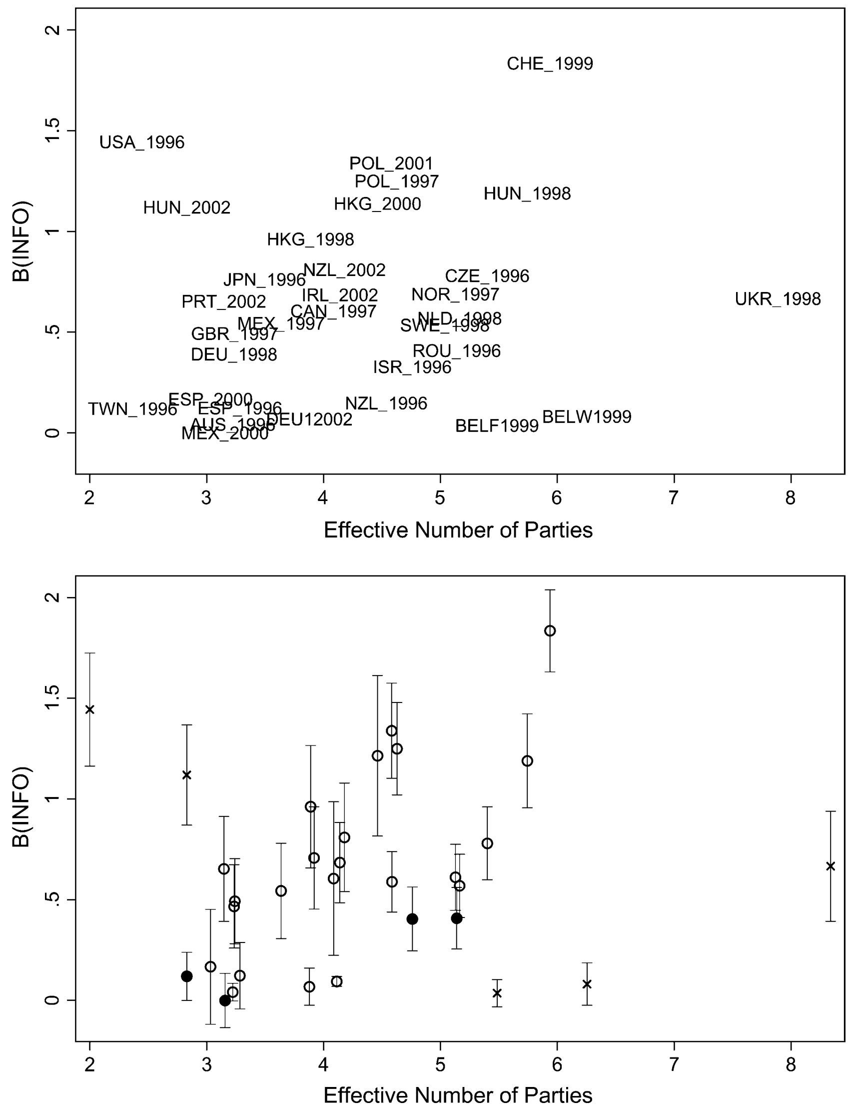

Fig. 2 Information costs of voting. The $\hat{\beta}_j$ are estimated by OLS. Bars denote 95% confidence intervals. The horizontal axis reports the effective number of parties, estimated using the Laakso-Taagepera index (Laakso and Taagepera 1979). Solid points denote those legislative elections that were held concurrently with presidential (or in the case of Israel's 1996 election, prime ministerial) elections. X denotes cases later excluded from this analysis. *Source:* CSES Modules 1 & 2.

In general, concurrent presidential elections tend to have fewer parties competing than strictly legislative elections.

Finally, the locations of outliers in the observed patterns of association are expected. The location of Belgium, for example, seen in the lower right corner of Fig. 2, likely results from Belgium's compulsory voting laws. As a result, information costs have little relevance to turnout in this system, even though the number of parties competing in Belgium is quite similar to those of other countries included in the analysis. Note that the effects of compulsory voting in other systems included in the study are less apparent because the numbers of parties competing in these systems are generally low (i.e., Australia is effectively a three-party system). Therefore, the effects of compulsory voting requirements on the relationship between the information costs of voting and turnout are likely to be quite different for Belgians.[10]

The locations of the other outlying systems—Hungary (2002), Urkaine, and the United States—are similarly expected and are also excluded from the analysis that follows. In the Hungarian 2002 election (seen in the upper left-hand corner of Fig. 2), for example, two of the main parties of the political system formed an electoral block (the governing FIDESZ-MPP and its coalition partner MDF) and launched a successful campaign, while maintaining their separate identities throughout and since the campaign period (MDF, for example, maintains its own parliamentary group). Therefore, it is not surprising that the information costs imposed by this election might resemble those of the multiparty system Hungary was and continues to be (note the location of the parameter estimated for the 1998 Hungarian election).

Although the effective number of parties competing in the 1998 Ukrainian election (seen at the far right margin of Fig. 2) is indeed high, the system is highly polarized, with parties composing five relatively coherent electoral blocks (Birch and Wilson 1999). As a result, the information costs imposed by this Ukrainian election are more in keeping with those expected of a four- or five-party system.

Finally, the unexpectedly high information costs imposed on voters in American elections, relative to the number of parties competing, are not surprising: Registration requirements, ballot questions, and the number of elections heighten the complexity of decision making beyond a comparatively simple vote choice.

These outlying cases are noted for two reasons: First, they provide further support for the claim that analysts must be sensitive to cross-national structural differences in the patterns of individual-level relationships. Second, recognition of the characteristics of the outlier countries identifies conditions under which institutions are likely to shape political behavior. Knowledge of the outlying systems, which may be obscured in a pooled analysis, and an understanding of their deviations can then be used to construct better-specified models.

Although the discussion in the previous paragraphs lends insight into the relationship between the information costs of voting and the number of parties competing in an election, the strategy employed—plotting coefficient estimates against estimates of *PARTIES*—does not provide an estimate of γ, or any estimate of γ's robustness. To generate these estimates, we exploit the linear structure in Eq. (7). First, note that estimates of $\hat{\beta}_j$ include a stochastic component:

$$\hat{\beta}_j = \beta_j + v_j \tag{14}$$

where $v_j = (\mathbf{INFO_j'INFO_j})^{-1}\mathbf{INFO'e_j}$ and

$$E(\mathbf{vv'}) = (\mathbf{INFO_j'INFO_j})^{-1}\mathbf{INFO'\Sigma INFO}(\mathbf{INFO_j'INFO_j})^{-1} \equiv \mathbf{V}. \tag{15}$$

When Eq. (7) is substituted in Eq. (14), the resulting expression,

[10]In an analysis that includes a larger number of compulsory voting systems with large numbers of parties competing, the appropriate model specification might include an interaction term, *PARTIES* $\times$ *COMP.*

$$\hat{\beta}_j = \gamma_1 + PARTIES_j\gamma_2 + PRES_j\gamma_3 + (PARTIES_j \times PRES_j)\gamma_4 + u_j + v_j, \qquad (16)$$

implies an estimator for the γs:

$$\hat{\gamma} = (\mathbf{ELECTION'\Omega^{-1}\ ELECTION})^{-1}\mathbf{ELECTION'\Omega^{-1}}\hat{\beta} \qquad (17)$$

where, as in the previous section, **ELECTION** is a matrix in which each row contains the election data for each country, and

$$\mathbf{\Omega} = \mathbf{E}((\mathbf{u}+\mathbf{v})(\mathbf{u}+\mathbf{v})') = (\mathbf{T}+\mathbf{V}) \qquad (18)$$

(see Appendix A for the complete derivation).[11] Equation (17) comprises the second step of the two-step estimation strategy: Parameter estimates for each country are regressed on system-level variables, in this case *PARTIES*, *PRES*, and *PARTIES* $\times$ *PRES*, and a constant, to generate $\hat{\gamma}$s. Estimates of the variance of the γs are similarly straightforward,

$$\mathbf{Var}(\hat{\gamma}) = (\mathbf{ELECTION'\Omega^{-1}\ ELECTION})^{\mathbf{-1}}. \qquad (19)$$

In practice, the observed residuals, estimated using the expression

$$\hat{\omega}_j = \hat{\beta}_j - \hat{\gamma}_1 - PARTIES_j\hat{\gamma}_2 - PRES_j\hat{\gamma}_3 - (PARTIES_j \times PRES_j)\hat{\gamma}_4 \qquad (20)$$

can be used to generate unbiased estimates of Ω, according to the strategies that Lewis and Linzer (2005; Hanushek 1974), for example, describe. (Our analysis uses the "FGLS with known variance" setup). Appendix B demonstrates that when the components of Ω are known, the two-step strategy yields coefficient estimates that are asymptotically equivalent to those estimated using a properly specified pooled model. Then, under the standard regularity conditions for FGLS estimates (see Fuller and Battese 1973), it can be shown that $\hat{\Omega}$ maintains the asymptotic unbiasedness and efficiency of Ω.

Table 2 reports the estimated parameters resulting from the second step of a two-step analysis, in which $\hat{\beta}$ is regressed on *PARTIES*. Model B includes only one election for each country; Model A includes additional elections for countries for which several elections were included in the CSES datasets. Briefly, the estimated coefficient $\hat{\gamma}_2 = 0.36$ is of the expected sign and is estimated with relatively little variance: Multiparty competition is associated with higher information costs of voting. Similarly, as $\hat{\gamma}_4 = -0.19$, there is evidence that concurrent presidential elections work to diminish the effect of *PARTIES* on the information costs of voting, although this parameter is estimated with substantial variance.

To interpret the substantive effect of these estimates—how *INFO* affects *VOTE*—it is helpful to recall the interactive structure implied by Eqs. (13) and (16):

$$\begin{aligned} VOTE_{ij} = \hat{\gamma}_0 + INFO_{ij}\big(&\hat{\gamma}_1 + PARTIES_j\hat{\gamma}_2 + PRES_j\hat{\gamma}_3 \\ &+(PARTIES_j \times PRES_j)\hat{\gamma}_4 + u_j + v_j\big) + a_j + e_{ij} \end{aligned} \qquad (21)$$

[11] An assumption made more explicit in Appendix A excludes cross-country covariance in the v_j. Specifically, we assume that variance in estimation is uncorrelated across countries. Within the context of our model, and as the country samples were independently drawn at different points in time and administered by different election study teams, this seems to us to be a plausible assumption.

Table 2 Information costs of voting: coefficient estimates (FGLS with known variance)

	A	*B*
$(\hat{\gamma}_0)$	0.84 (0.02)	0.84 (0.02)
INFO $(\hat{\gamma}_1)$	−0.84 (0.38)	−0.70 (0.41)
PARTIES × *INFO* $(\hat{\gamma}_2)$	0.36 (0.09)	0.32 (0.09)
PRES × *INFO*$(\hat{\gamma}_3)$	0.41 (0.82)	0.45 (0.95)
PARTIES × *PRES* × *INFO* $(\hat{\gamma}_4)$	−0.19 (0.19)	−0.19 (0.21)
J	27	21

Note. This table reports coefficients estimated in the second step of a two-step strategy, as described in the text. $\hat{\beta}_j$ is the dependent variable. Specification A includes all elections for which the data are available. Specification B includes only one election for each country for which several elections are included in the CSES data sets. *Source:* CSES Modules 1 & 2.

where $\alpha_j = \hat{\gamma}_0 + a_j$ and a_j includes the estimation and cross-national variance of the individual-level intercepts, α_j. Then, the predicted turnout can be estimated using the following expression:[12]

$$\begin{aligned} Pr(VOTE_{ij} = 1|INFO_{ij}, PARTIES_j, PRES_j) \\ = \hat{\gamma}_0 + INFO_{ij}\hat{\gamma}_1 + INFO_{ij}PARTIES_j\hat{\gamma}_2 + INFO_{ij}PRES_j\hat{\gamma}_3 \\ + INFO_{ij}(PARTIES_j \times PRES_j)\hat{\gamma}_4. \end{aligned} \tag{22}$$

Using elements of the matrix described by Eq. (19), the variance of the predicted effect of political information can be estimated with

$$\begin{aligned} \mathrm{Var}\left(\frac{\partial VOTE}{\partial INFO}\right) = {} & \mathrm{Var}(\gamma_1) + \mathrm{Var}(\gamma_2)PARTIES^2 + \mathrm{Var}(\gamma_3)PRES^2 \\ & + \mathrm{Var}(\gamma_4)(PARTIES \times PRES)^2 + 2\mathrm{Cov}(\gamma_1,\gamma_2)PARTIES \\ & + 2\mathrm{Cov}(\gamma_1,\gamma_3)PRES + 2\mathrm{Cov}(\gamma_1,\gamma_4)(PARTIES \times PRES) \\ & + 2\mathrm{Cov}(\gamma_2,\gamma_3)PARTIES{\cdot}PRES \\ & + 2\mathrm{Cov}(\gamma_2,\gamma_4)PARTIES{\cdot}(PARTIES \times PRES) \\ & + 2\mathrm{Cov}(\gamma_3,\gamma_4)PRES{\cdot}(PARTIES \times PRES). \end{aligned} \tag{23}$$

We must, therefore, take into account the number of political parties competing and whether a concurrent presidential election was held, when calculating predicted turnout and the variance of the effect of political information, our key individual-level variable.

In Fig. 3, we plot the predicted probabilities of voting, distinguishing between voters with low levels of information ($INFO = -0.3$, denoted with diamond-shaped points) and voters with high levels of information (denoted with square-shaped points, $INFO = 0.3$), and observing the effects of variation in the effective number of parties. The two lower lines track the predicted turnout of low-information voters in the presence (hollow points) and absence (solid points) of concurrent presidential elections. Note that turnout among low-information voters is negatively affected by increases in the effective number of

[12] If different individual-level variables are included for each country, predicted values may be estimated in a similar way, although with the effects of these variables incorporated.

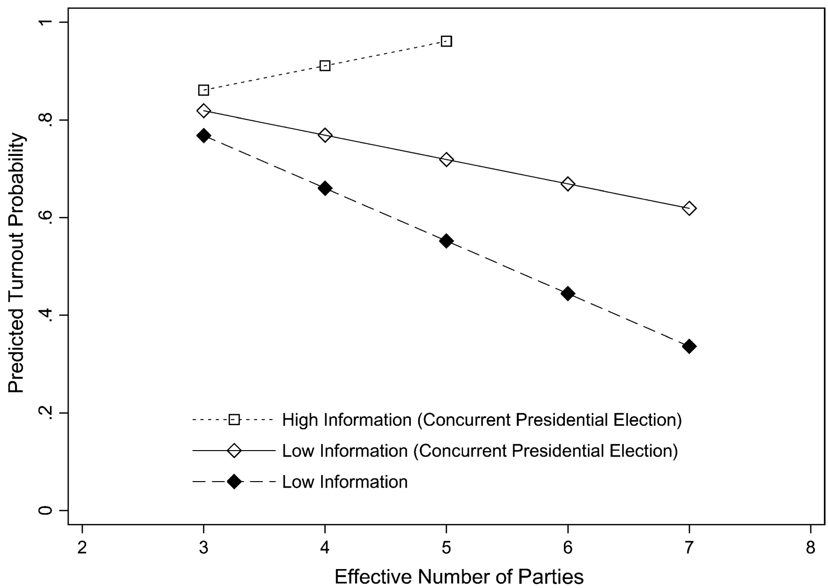

Fig. 3 Interpreting second-step coefficient estimates (predicted turnout). This figure reports predicted turnout probabilities among high- and low-information voters, in systems with and without concurrent presidential elections. Turnout probabilities are calculated according to Eq. (22). *Source:* CSES Modules 1 & 2.

parties, although the effect is moderated somewhat by the presence of concurrent presidential elections. High-information voters, however, may be expected to vote more frequently than low information voters in elections where a larger number of parties compete—a result that conforms to our initial intuition about how a larger number of parties influences incentives to vote. (High-information citizens in the absence of concurrent presidential elections vote with near certainty, regardless of the number of parties competing.)

This analysis, therefore, provides a nonintuitive insight into how multiparty competition structures voter behavior. In particular, notice how electoral institutions may produce unexpected results: To the extent that low-information citizens hold preferences that are distinct from those of high-information citizens, the presumably transparent and open system of proportional representation (PR), by encouraging multiparty competition, may actually inhibit the representation of these preferences. Further, to the extent that low-information voters outnumber high-information voters, the puzzling pattern observed in Fig. 1 may be accounted for by the information costs of voting.

4 Conclusion

Only recently have comparative political scientists begun to consider how best to analyze cross-national public opinion data. Most published research continues to apply partitioning and pooling strategies, but an emerging consensus appears to recognize the advantages of a hierarchical model. We have illustrated here a two-step strategy that draws heavily on the statistical foundations of hierarchical linear models (and meta-analysis especially) but maintains many of the advantages of both partitioning and pooling strategies. In particular, the two-step strategy, unlike standard hierarchic linear models, affords the investigator greater flexibility in model specification, including incorporation of confounding variables at different levels of analysis, identification of outliers, and the possibility of using different right-hand-side models for different clusters. And it accomplishes this without loss of efficiency as compared to pooling strategies.

A particularly attractive aspect of the two-step strategy is that it parallels closely the data generation process, which consists of aggregating large clusters (with their internal processes) into a larger multilevel data set. It is possible that the two-step strategy could represent a middle ground of sorts for those comparativists who are concerned that the complexities of individual cases must be understood, but still want to draw broad, large-N comparisons. The flexibility available at the first stage of the analysis strategy proposed here can allow the analyst to take into account variations in process from one country to another. The second step of the analysis facilitates cross-national generalization and the ability to incorporate system-level variables. The two-step strategy, therefore, seems to be especially well suited to the comparative study of political behavior.

Nevertheless, a word of caution is in order: Errors in the within-country specification are passed through to the cross-country model and may have unexpected results.[13] Further, although we suspect that the data structure that facilitates a two-step strategy—a large number of large and independent samples—may apply more generally to questions of comparative politics, further research will clarify the application of this strategy to cases in which the the assumptions underlying this analysis do not hold.

Appendix A. Two-Step Strategy

Let

$$\mathbf{y} = \mathbf{X}\beta + \epsilon \tag{A1}$$

represent the relationship between variables measured at the individual level of analysis, where

$$\mathbf{y} = \begin{bmatrix} \mathbf{y_1} \\ \mathbf{y_2} \\ \vdots \\ \mathbf{y_J} \end{bmatrix}, \quad \mathbf{X} = \begin{bmatrix} \mathbf{X_1} & 0 & 0 & \cdots & 0 \\ 0 & \mathbf{X_2} & 0 & & 0 \\ 0 & 0 & \ddots & & \\ & & & & \\ 0 & 0 & 0 & \cdots & \mathbf{X_J} \end{bmatrix}, \quad \beta = \begin{bmatrix} \beta_\mathbf{1} \\ \beta_\mathbf{2} \\ \vdots \\ \beta_\mathbf{J} \end{bmatrix}, \quad \text{and } \epsilon = \begin{bmatrix} \epsilon_\mathbf{1} \\ \epsilon_\mathbf{2} \\ \vdots \\ \epsilon_\mathbf{J} \end{bmatrix}$$

That is, for each of $j = 1 \ldots J$ countries, $\mathbf{y_j}$ contains n_j observations.[14] Similarly, each $\mathbf{X_j}$, and $\varepsilon_\mathbf{j}$, report n_j observations for k_{1_j} individual-level variables and residual terms,

[13]This is of particular concern if the bias is correlated with the system-level variables, as might be the case in an analysis of support for democracy, and level of economic development, for example. Suppose individual-level variables are more likely to be measured with error in countries with less resources to devote to survey research, and the parameters estimated in these countries are biased downwards. If these parameters are then regressed on a measure of aggregate wealth, second-step coefficients are likely to be estimated with bias and may lead to errors in the substantive conclusions reached. If, on the other hand, bias in the estimation of individual-level coefficients is not correlated with system-level covariates, this bias will likely result in inefficiency at the second step of the analysis. Note, however, that this bias and inefficiency does not result from the two-step estimation strategy and would be evident in either a pooled or two-step model. Fortunately, as Achen (2005) suggests, the effects of bias in the individual-level parameters on second step parameter estimates are likely to decrease as the number of country cases included in the analysis increases.

[14]Bryk and Raudenbush (1992) provide the jumping-off point for this section of the discussion.

respectively. Note that it is not necessary that either n_j or k_{1_j} be consistent across all J countries: The number of observations for each country and the set of relevant variables in comparative analysis often varies cross nationally. Finally, the stacked matrix β contains k_{1_j} parameters that correspond to each country included in the analysis.

Residual variances differ across countries, with

$$E(\epsilon\epsilon') = \begin{bmatrix} \sigma_1\mathbf{I} & 0 & 0 & \cdots & 0 \\ 0 & \sigma_2\mathbf{I} & 0 & & 0 \\ 0 & 0 & \ddots & & \\ & & & & \\ 0 & 0 & 0 & \cdots & \sigma_J\mathbf{I} \end{bmatrix} \equiv \mathbf{\Sigma}$$

The assumption of constant within-country variance is not necessary but facilitates the derivations of results presented in the next appendix.

Relationships between variables measured at the subnational level may vary in systematic ways. That is, suppose

$$\beta = \mathbf{Z}\gamma + \mathbf{u} \tag{A2}$$

where $\mathbf{Z}$ is a partitioned matrix of system-level variables, with $k_{1_j} \times k_2$ components for each country, and $\mathbf{u}$ is a J vector of country-level residuals. By assumption, $E(\mathbf{uu}') = \tau^2\mathbf{I}$.

To see how the $\mathbf{Z}$ matrix is constructed, suppose the subnational analysis for country j includes two covariates, with

$$\beta_j^1 = Z_j^{11}\cdot\gamma_1 + Z_j^{12}\cdot\gamma_2 + u_j^1 \tag{A3}$$

$$\beta_j^2 = Z_j^{21}\cdot\gamma_1 + Z_j^{22}\cdot\gamma_2 + u_j^2. \tag{A4}$$

In this case, $\mathbf{Z_j}$ is constructed

$$\mathbf{Z_j} = \begin{bmatrix} Z_j^{11} & Z_j^{12} & 0 & 0 \\ 0 & 0 & Z_j^{21} & Z_j^{22} \end{bmatrix}$$

The $\mathbf{Z_j}$ are stacked vertically in $\mathbf{Z}$.

The individual-level coefficients, $\beta_\mathbf{j}$, are unobserved and must be estimated with some uncertainty:

$$\hat{\beta}_j = \beta_j + \mathbf{v_j}. \tag{A5}$$

If β is estimated using the OLS estimator, $\mathbf{v_j} = (\mathbf{X_j'X_j})^{-1}\mathbf{X_j'e_j}$, and

$$E(\mathbf{vv}') = (\mathbf{X'X})^{-1}\mathbf{X'\Sigma X}(\mathbf{X'X})^{-1} = \mathbf{V}. \tag{A6}$$

Note that if $\sigma_j^2 = \sigma^2$ for all j, Eq. (A6) reduces to the more familiar $\mathbf{V} = \sigma^2(\mathbf{X'X})^{-1}$.

Combining Eqs. (A2) and (A5) yields the expression,

$$\hat{\beta} = \mathbf{Z}\gamma + \mathbf{u} + \mathbf{v}. \tag{A7}$$

Provided that the following assumptions hold,

$$E(\epsilon_{ij} \mid \mathbf{X}, \mathbf{Z}) = E(u_j \mid \mathbf{X}, \mathbf{Z}) = 0, \tag{A8a}$$

$$E(\epsilon_{ij} u_j \mid \mathbf{X}, \mathbf{Z}) = 0 \text{ for all i, j, and} \tag{A8b}$$

$$E(u_j v_j \mid \mathbf{X}, \mathbf{Z}) = 0, \tag{A8c}$$

this setup implies a LS estimator for the parameter of interest, γ (with the usual correction for heteroscedastic variance):

$$\hat{\gamma} = (\mathbf{Z}'(\mathbf{T} + \mathbf{V})^{-1}\mathbf{Z})^{-1}\mathbf{Z}'(\mathbf{T} + \mathbf{V})^{-1}\hat{\beta}. \tag{A9}$$

Clearly, this estimator has the capacity to incorporate context-specific variables and to provide evidence of patterns that hold across populations. The next section of this discussion demonstrates that Eq. (A9) maintains the desirable properties of a pooling strategy.

Appendix B. Properties of a Two-Step Estimator: Consistency and Efficiency

Note that Eqs. (A1) and (A2) imply

$$\mathbf{y} = \mathbf{XZ}\gamma + \mathbf{Xu} + \epsilon \tag{A10}$$

and that γ may be estimated with the GLS estimator,

$$\tilde{\gamma} = (\mathbf{Z}'\mathbf{X}'(\mathbf{XTX}' + \boldsymbol{\Sigma})^{-\mathbf{1}}\mathbf{XZ})^{-1}\mathbf{Z}'\mathbf{X}'(\mathbf{XTX}' + \boldsymbol{\Sigma})^{-\mathbf{1}}\mathbf{y}. \tag{A11}$$

Following Aitken (1935), $\tilde{\gamma}$ is an unbiased and minimum variance estimate of γ.[15]

Proposition 1. Provided the assumptions listed in Eqs. (A8a), (A8b), and (A8c) hold, Eq. (A9) is an unbiased estimate of γ.

It is sufficient to show that

$$\begin{aligned} \gamma = &\ (\mathbf{Z}'(\mathbf{T} + \mathbf{V})^{-\mathbf{1}}\mathbf{Z})^{-\mathbf{1}}\mathbf{Z}'(\mathbf{T} + \mathbf{V})^{-\mathbf{1}}\hat{\beta} \\ &= (\mathbf{Z}'\mathbf{X}'(\mathbf{XTX}' + \boldsymbol{\Sigma})^{-\mathbf{1}}\mathbf{XZ})^{-\mathbf{1}}\mathbf{Z}'\mathbf{X}'(\mathbf{XTX}' + \boldsymbol{\Sigma})^{-\mathbf{1}}\mathbf{y} = \tilde{\gamma} \end{aligned}$$

Note the structure of the matrix,

$$\boldsymbol{\Omega} \equiv (\mathbf{T} + \mathbf{V})^{-1}$$

$$= \begin{bmatrix} (\sigma_1^2(\mathbf{X_1'X_1})^{-1} + \tau^2\mathbf{I})^{-1} & 0 & 0 & \cdots & 0 \\ 0 & (\sigma_2^2(\mathbf{X_2'X_2})^{-1} + \tau^2\mathbf{I})^{-1} & 0 & \cdots & 0 \\ 0 & 0 & \ddots & & \\ 0 & 0 & 0 & & (\sigma_J^2(\mathbf{X_J'X_J})^{-1} + \tau^2\mathbf{I})^{-1} \end{bmatrix}$$

[15] Saxonhouse (1977) also uses this comparison to evaluate the properties of a two-step linear estimator, although more briefly and in a different context.

Applying Rao (1965, p. 29), the diagonal elements of $(\mathbf{T} + \mathbf{V})^{-1}$ can be re-expressed:

$$\mathbf{\Omega_j} = \frac{1}{\sigma_j^2}(\mathbf{X_j'X_j}) - \frac{1}{\sigma_j^2}(\mathbf{X_j'X_j})\left(\frac{1}{\tau^2}\mathbf{I} + \frac{1}{\sigma_j^2}(\mathbf{X_j'X_j})\right)^{-1}\frac{1}{\sigma_j^2}(\mathbf{X_j'X_j}). \tag{A12}$$

Similarly, the diagonal elements of the matrix $\mathbf{X}'(\mathbf{XTX}' + \Sigma)^{-1}\mathbf{X}$ can be expressed

$$\tilde{\mathbf{\Omega}}_{\mathbf{j}} = \frac{1}{\sigma_j^2}(\mathbf{X_j'X_j}) - \frac{1}{\sigma_j^2}(\mathbf{X_j'X_j})\left(\frac{1}{\tau^2}\mathbf{I} + \frac{1}{\sigma_j^2}(\mathbf{X_j'X_j})\right)^{-1}\frac{1}{\sigma_j^2}(\mathbf{X_j'X_j}). \tag{A13}$$

This implies that $(\mathbf{T} + \mathbf{V}) = \mathbf{X}'(\mathbf{XTX}' + \Sigma)^{-1}\mathbf{X}$ or that

$$(\mathbf{Z}'(\mathbf{T}+\mathbf{V})^{-1}\mathbf{Z})^{-1} = (\mathbf{Z}'\mathbf{X}'(\mathbf{XTX}' + \mathbf{\Sigma})^{-1}\mathbf{XZ})^{-1}. \tag{A14}$$

Using the same procedure, it can be shown that

$$(\mathbf{T}+\mathbf{V})^{-1}\hat{\beta} = \mathbf{X}'(\mathbf{XTX}' + \mathbf{\Sigma})^{-1}\mathbf{y} \tag{A15}$$

and the result directly follows.

Proposition 2. Provided the assumptions listed in Eqs. (A8a), (A8b), and (A8c) hold, Eq. (A9) is an efficient estimate of γ.

It is sufficient to show that

$$\Delta = \mathrm{Var}(\hat{\gamma}) - \mathrm{Var}(\tilde{\gamma}) = 0. \tag{A16}$$

Note that

$$\mathrm{Var}(\hat{\gamma}) = E[(\hat{\gamma} - \gamma)(\hat{\gamma} - \gamma)'], \tag{A17}$$

where

$$(\hat{\gamma} - \gamma) = (\mathbf{Z}'(\mathbf{T}+\mathbf{V})^{-1}\mathbf{Z})^{-1}\mathbf{Z}'(\mathbf{T}+\mathbf{V})^{-1}(\mathbf{u}+\mathbf{v}) \tag{A18}$$

with the result

$$\mathrm{Var}(\hat{\gamma}) = (\mathbf{Z}'(\mathbf{T}+\mathbf{V})^{-1}\mathbf{Z})^{-1}. \tag{A19}$$

Similarly, it can be shown that

$$\mathrm{Var}(\tilde{\gamma}) = (\mathbf{Z}'\mathbf{X}'(\mathbf{XTX}' + \mathbf{\Sigma})^{-1}\mathbf{XZ})^{-1}. \tag{A20}$$

By Eq. (A14), Eqs. (A19) and (A20) are asymptotically equivalent, and the result follows.

Appendix C. Data

Election	$\hat{\beta}$	$\hat{\sigma}_\beta$	$\hat{\alpha}$	$\hat{\sigma}_\alpha$	*PARTIES*	*PRES*	*INFO* (SD)	n_j
Australia 1996	0.04	0.02	0.99	0.00	3.22	0	0.10	1733
Belgium (Flanders) 1999†	0.04	0.03	0.97	0.00	5.49	0	0.10	2179
Belgium (Walloon) 1999†	0.08	0.05	0.94	0.01	6.25	0	0.10	1959
Canada 1997	0.60	0.19	0.86	0.02	4.09	0	0.10	1848
Czech Republic 1996	0.78	0.09	0.90	0.01	5.40	0	0.11	1228
Germany 1998	0.47	0.11	0.93	0.01	3.23	0	0.09	2011
Germany 2002*	0.07	0.05	0.99	0.00	3.88	0	0.09	2000
Great Britain 1997	0.49	0.11	0.83	0.01	3.24	0	0.10	2894
Hong Kong 1998	0.96	0.16	0.74	0.01	3.89	0	0.09	975
Hong Kong 2000*	1.21	0.20	0.60	0.02	4.46	0	0.09	659
Hungary 1998	1.19	0.12	0.75	0.01	5.74	0	0.10	1523
Hungary 2002*†	1.12	0.13	0.83	0.01	2.83	0	0.09	1200
Ireland 2002	0.68	0.10	0.87	0.01	4.14	0	0.09	2366
Israel 1996‡	0.40	0.08	0.91	0.01	4.76	1	0.10	998
Japan 1996	0.71	0.13	0.84	0.01	3.92	0	0.07	1325
Mexico 1997	0.54	0.12	0.75	0.01	3.64	0	0.10	2021
Mexico 2000*	0.00	0.07	0.94	0.01	3.16	1	0.11	1429
Netherlands 1998	0.57	0.08	0.91	0.01	5.16	0	0.09	1814
New Zealand 1996	0.09	0.01	0.99	0.00	4.11	0	0.10	4013
New Zealand 2002*	0.81	0.14	0.80	0.01	4.18	0	0.09	1674
Norway 1997	0.61	0.08	0.86	0.01	5.13	0	0.09	2055
Poland 1997	1.25	0.12	0.58	0.01	4.63	0	0.10	1993
Poland 2001*	1.34	0.12	0.59	0.01	4.58	0	0.10	1785
Portugal 2002	0.65	0.13	0.76	0.01	3.15	0	0.09	1301
Romania 1996	0.41	0.08	0.89	0.01	5.14	1	0.11	1160
Spain 1996	0.12	0.08	0.90	0.01	3.28	0	0.10	1208
Spain 2000*	0.17	0.15	0.81	0.01	3.03	0	0.08	1057
Sweden 1998	0.59	0.08	0.89	0.01	4.59	0	0.11	1157
Switzerland 1999	1.83	0.10	0.61	0.01	5.94	0	0.10	2047
Taiwan 1996	0.12	0.06	0.92	0.01	2.83	1	0.12	1191
Ukraine 1998†	0.67	0.14	0.77	0.01	8.34	0	0.09	1145
United States 1996†	1.44	0.14	0.78	0.01	2.00	1	0.08	1534

Note. $\hat{\beta}$ and $\hat{\sigma}_\beta$ are OLS estimates. *PARTIES* calculated using popular vote returns in lower house elections. *INFO* is mean-centered within each country and rescaled to the $[-1, 1]$ interval.
*Elections excluded in Model B (Table 2).
†Outlying cases that are excluded from all analyses.
‡Israel's prime ministerial election is coded as a concurrent presidential election.
Source: CSES Modules 1 & 2.

References

Achen, Christopher. 2005. "Two–Step Hierarchical Estimation: Beyond Regression Analysis." *Political Analysis* doi:10.1093/pan/mpi033.

Aitken, A. C. 1935. "On Least Squares and the Linear Combination of Observations." *Proceedings of the Royal Society of Edinburgh* 55:42–48.

Bartels, Larry. 1996. "Pooling Disparate Observations." *American Journal of Political Science* 40:905–942.

Birch, S., and A. Wilson. 1999. "The Ukrainian Parliamentary Elections of 1998." *Electoral Studies* 18:276–282.

Blais, Andre, and Agnieszka Dobrzynska. 1998. "Turnout in Electoral Democracies." *European Journal of Political Research* 33:239–261.

Bryk, A. S., and S. Raudenbush. 1992. *Hierarchical Linear Models.* Newbury Park, CA: Sage.

Cox, Gary, and Octavio Amorim. 1997. "Electoral Institutions, Cleavage Structures and the Number of Parties." *American Journal of Political Science* 41:149–174.

CSES. 2003a. "CSES Module 1, 1996–2001 (August 4, 2003 Version)." Center for Political Studies, University of Michigan [producer and distributor]. (Available from www.cses.org.)

CSES. 2003b. "CSES Module 2, 2001–2005 (May 1, 2003 Version)." Center for Political Studies, University of Michigan [producer and distributor]. (Available from www.cses.org.)

Dahl, Robert. 2002. *How Democratic Is the American Constitution?* New Haven: Yale University Press.

Downs, Anthnoy. 1957. *An Economic Theory of Democracy.* New York: Harper.

Fuller, Wayne, and George Battese. 1973. "Transformations for Estimation of Linear Models with Nest-Error Structure." *Journal of the American Statistical Association* 68:626–632.

Golder, Matthew. Forthcoming. "Presidential Elections and Legislative Fragmentation." *American Journal of Political Science* 50(1).

Hanushek, Eric. 1974. "Efficient Estimators for Regressing Regression Coefficients." *American Statistician* 28:66–67.

IDEA. 2004. "Voter Turnout from 1945 to Date." International IDEA Voter Turnout Project. (Available from www.idea.int/vt/index.cfm.)

Jackman, Robert. 1987. "Political Institutions and Voter Turnout in the Industrial Democracies." *American Political Science Review* 81:405–423.

Jusko, Karen Long. 2005. "Two-Step Binary Response Models for Cross-National Public Opinion Data: A Research Note." Presented at the Midwest Political Science Association National Conference, April 7–10, 2005. Chicago.

Laakso, M., and Rein Taagepera. 1979. '"Effective' Number of Parties: A Measure with Application to West Europe." *Comparative Political Studies* 12:3–27.

Lewis, Jeffrey, and Drew Linzer. 2005. "Estimating Regression Models in which the Dependent Variable Is Based on Estimates." *Political Analysis* doi:10.1093/pan/mpi026.

Rao, C. R. 1965. *Linear Statistical Inference and Its Applications.* New York: Wiley.

Saxonhouse, Gary. 1977. "Regressions from Samples Having Different Characteristics." *Review of Economics and Statistics* 59:234–237.