Political Analysis Advance Access published August 17, 2005

doi:10.1093/pan/mpi028

Context and the Economic Vote: A Multilevel Analysis

Raymond M. Duch
Department of Political Science,
University of Houston,
4800 Calhoun,
Houston, TX 77204–3011
e-mail: rduch@uh.edu (corresponding author)

Randy Stevenson
Department of Political Science,
Rice University,
P.O. Box 1892,
Houston, Texas 77251–1892
e-mail: stevenso@ruf.rice.edu

Voters use observed economic performance to infer the competence of incumbent politicians. These economic perceptions enter the voter's utility calculations modified by a weight that is minimized when the variance in exogenous shocks to the economy is very large relative to the variance in economic outcomes associated with the competence of politicians. Cross-national variations in the political and economic context systematically increase or undermine the voter's ability to ascertain the competency of incumbents. We test one hypothesis: As policy-making responsibility is shared more equally among parties, economic evaluations will be more important in the vote decision. We employ two multilevel modeling procedures for estimating the contextual variations in micro-level economic voting effects: a conventional pooled approach and a two-stage procedure. We compare the multivariate results of a pooled method with our two-stage estimation procedure and conclude that they are similar. Our empirical efforts use data from 163 national surveys from 18 countries over a 22-year period.

1 Introduction

The phenomenon of "economic voting" has been at the forefront of efforts to understand electoral accountability in democratic politics. In this essay we propose an explanation for why the economic vote varies across national contexts; we identify individual-level data from a large number of national contexts that can be used for testing this explanation; and we describe multilevel modeling procedures for analyzing the data. Our point of departure

Authors' note: We acknowledge the generous support of the National Science Foundation that funded this project with grant SBR-0215633. We thank the participants of the Princeton University Workshop on Multi-Level Modeling (Center for the Study of Democratic Politics, October 22–24, 2004) for their extremely helpful comments and suggestions.

is the insight, gained primarily from cross-national studies of aggregate election results, that variations in political and economic contexts explain national differences in the magnitude of the economic vote (Paldam 1991; Powell and Whitten 1993). For the most part, these explanations assume a sanctioning model of the economic vote. We adopt a different theoretical approach. The theory that we briefly summarize here, and develop much more fully elsewhere (Duch and Stevenson 2005), employs a competency model of the vote decision.

With the exception of analyses based on a relatively small number of countries (Lewis-Beck 1988; Anderson 2000; Alvarez et al. 2000; Duch 2001; Hellwig 2001; Blais et al. 2004), there is no systematic cross-national analysis of individual-level survey data that addresses how economic voting varies across different political contexts. This is surprising, because students of economic voting now have a wealth of individual-level data with which to work. Over the past five decades, political scientists have regularly surveyed people in the developed democracies about their perceptions of the economy and their vote choices, so there now is the potential for a large-scale individual-level study of economic voting that focuses on differences in political context.[1] Rather than relying on aggregate-level election results, we propose taking advantage of this large number of individual-level election surveys. We describe two multilevel modeling procedures for analyzing these diverse election studies: a conventional pooled approach and a two-stage approach. Finally, we test one of our hypotheses regarding cross-national variation in the economic vote using the large-N sample of individual-level elections studies and employing these two multilevel modeling techniques.

The essay is organized as follows: We first briefly outline our selection theory of the economic vote and motivate one specific hypothesis from this theory. We then suggest why the analysis of a large number of individual-level election surveys is a particularly effective strategy for testing our competency or selection model of the economic vote. This is followed by a description of two multilevel modeling procedures for estimating the contextual variations in micro-level economic voting effects: a conventional pooled approach and a two-stage procedure. We conclude with an empirical section that has three components designed to both test the hypothesis motivated by our competency theory and compare the results of a pooled method with our two-stage estimation procedure. First we generate a map of the magnitude of economic voting in each of the sampled countries. Second, we estimate the macro-level effects employing both pooled and two-stage procedures. The third section assesses whether the two-stage procedure generates results similar to those associated with a conventional pooled approach. Our empirical efforts use data from 163 different national surveys conducted in 18 countries over a 22-year period (1979–2001).

2 Theoretical Models of Contextual Effects

Students of economic voting have long suspected that features of the policy-making process should condition the magnitude of economic voting (Lewis-Beck 1988; Paldam 1991; Anderson 1995). These explanations have been loosely based on a sanctioning

[1]Of course, many of these surveys have been analyzed independently and the evidence for economic voting accessed (for an early example, see Eulau and Lewis-Beck 1985). However, our efforts (and indeed any effort that may have been tried by others) to turn these separate analyses into systematic conclusions about the impact of political context on economic voting have come to naught. This failure is due to the simple fact that the published empirical studies of economic voting in individual elections (or small numbers of similar elections) are so different in terms of the methods used that it is usually impossible even to construct a comparable metric from these studies, let alone to rule out differences in method as the chief cause of observed differences in economic voting across the studies.

model of economic voting in which voters are seen as principals who want economic policy makers (their agents) to provide future economic benefits but realize that these policy makers have incentives to shirk this responsibility (Barro 1973; Ferejohn 1986). Consequently, voters must provide disincentives for shirking via electoral punishments. However, since voters cannot actually observe the actions (or effort) of politicians, they cannot create the proper incentives by conditioning their votes on what politicians actually do, but instead must condition their votes on observable outcomes—i.e., the observed state of the economy.

Some have argued that when voters are unsure about which parties are responsible for economic policy making, their ability to use the vote to sanction politicians is compromised (Powell and Whitten 1993). Consequently, they do not use their vote as a sanction to induce economic voting, and the connection between the economic performance and vote choice is broken. The empirical hypotheses that flow from this depend on the additional observation that the voter's confusion of responsibility is more likely to occur when the cabinet is made up of coalitions of different parties or when nongovernmental parties are involved in policy making thorough consensual institutions like strong legislative committees and bicameral legislatures.

In Duch and Stevenson (2005) we suggest an alternative explanation for this relationship between power sharing and economic voting. That explanation builds on Alesina and Rosenthal's (1995) model of rational retrospective economic voting in which rational voters try to select competent economic managers from the set of parties competing in the election. As such, it is an example of a "selection" model of voting, rather than a "sanctioning" model (see Fearon [1999] for a discussion of the distinction between these two general classes of voting models).

Alesina and Rosenthal model economic growth using an expectations-augmented Phillips curve with an unobserved random shock. The shock includes a part that depends on the incumbent politician's competence and is persistent over time. The remainder of the shock is exogenous.[2] Voters form expectations about the actions of politicians and the state of the economy rationally and so anticipate economic policy, robbing it of its impact on growth. Thus the only difference in the ability of parties to deliver economic performance is their unknown level of competence, which is noisily reflected (given persistence) in current economic performance. Consequently, rational voters try to use observed economic performance to infer the competence of incumbent politicians. The solution to the voter's signal extraction problem results in economic performance (or perceptions of performance) entering the voter's utility calculations modified by a weight that varies between zero and one and that is minimized when the variance in exogenous shocks is very large relative to the variance in the competence of politicians. Under these conditions, observed fluctuations in economic performance carry little information or "signal" about the competence of politicians, so voters discount the economy in their voting calculus. Alternatively, when an economy is insulated against exogenous shocks (i.e., the variance in these shocks is quite small from period to period), observed movements in the economy will be relatively more informative about the competence of politicians and so more important in the vote choice of rational voters.[3]

[2]The competence shock should be understood as any action politicians can take that impacts growth but is not observable (and therefore predictable) by rational voters. Exogenous shocks are not a function of the actions of politicians but are similarly unobservable.

[3]This is true holding the distribution of competence shocks constant over the two hypothetical examples. Also, the mean of the exogenous shocks must be zero given rational voters and a definition of a shock as something unanticipated.

This logic immediately suggests at least one kind of context that should impact economic voting by rational retrospective voters—the extent to which the economy is subject to exogenous versus competence shocks.[4] Alesina and Rosenthal's model was developed for (and relies on assumptions specific to) the American case of two-party competition with a unified executive. When we generalize the logic of their model to a wider range of different institutional contexts (including those in which multiple parties compete, there are post-election negotiations over the makeup of the executive, and executive power is shared), the implications for the impact of power sharing on the magnitude of the economic vote are quite obvious (Duch and Stevenson 2005). Specifically, this generalized model implies that when executive power is shared, economic voting for every party is diminished—even if voters know the distribution of administrative responsibility among parties perfectly. Thus the reason for this smaller economic vote is not that voters cannot attribute responsibility over parties, but that a more equal distribution of responsibility weakens the signal that the previous economy provides about the competence of the incumbent parties. Hence, the following hypothesis:

> *As the distribution of policy-making responsibility is shared more equally among the parties in an election, retrospective economic evaluations will be more important in the vote choice.*

3 Empirical Models of Contextual Effects

Until very recently, much of what we know about cross-national variations in the economic vote has been based on the analysis of aggregate-level data. These efforts to identify a uniform economic vote function for a diverse set of countries employing aggregate-level data have generated often very different and sometimes contradictory results: Lewis-Beck and Mitchell (1990) analyze 27 elections and conclude that there is an economic vote. Paldam (1991), analyzing 197 election results, concludes that the economy has at best a very weak impact on the vote. Powell and Whitten (1993) examine approximately 100 and find an economic vote but conclude that its magnitude depends on the clarity of responsibility afforded by the institutional context. Some have challenged the robustness of this result claiming that there is little support for the clarity of responsibility argument, in fact claiming that economic voting is higher for coalition as opposed to single-party governments (Royed et al. 2000).[5]

We believe that these contradictions can be addressed by analyzing individual-level models that typically allow for a better specification of the economic vote model (Lewis-Beck and Stegmaier 2000). However, these analyses must include a relatively large number of diverse individual-level surveys to ensure consistent estimates of the economic vote and to ensure significant contextual variation. Until recently, studies of cross-national economic voting that employ individual-level data have been constrained, at least in terms of testing hypotheses regarding contextual effects, because of the lack of sufficiently similar survey studies conducted over diverse national contexts. Over the past five decades, though, political scientists have regularly surveyed people in the developed democracies about their perceptions of the economy and their vote choices. Examples of these efforts are well known to most students of comparative voting behavior: the British Election Studies since 1964; the Canadian National Election studies since 1965; the German election studies since 1969. In addition, the regularly administered Euro-Barometer studies have, until recently, included

[4] In Duch and Stevenson (2005) we pursue this implication as well as others.
[5] Although see the rigorous rebuttal by Palmer and Whitten (2003a and 2003b).

questions related to economic perceptions and self-reported vote (Anderson 2000).[6] As a result, there is now the potential for assessing the impact of context on the economic vote by analyzing a large number of individual-level surveys that cover diverse national contexts over a reasonably long time period.

Estimating cross-national economic voting with large numbers of individual-level studies can also improve our estimates of the relative importance of the economic vote compared to other factors shaping the vote decision. Again, some important advances have been made employing individual-level voter preferences studies (Alvarez et al. 2000; Blais et al. 2004), but they have been based on relatively small numbers of studies—eleven in the case of Blais et al. (2004) and seven in the case of Alvarez et al. (2000). Not surprisingly given these small numbers, the conclusions have been contradictory, with Blais et al. (2004), suggesting that issue voting outweighs economic voting and Alvarez et al. (2000) making the opposite claim. Basing these comparisons on a much larger sample of individual-level surveys will generate much more robust estimates of the relative importance of the economic vote.

Our principal concern in this article is to describe how we have used a large number of individual surveys from diverse national contexts to explore variation in the magnitude of economic voting and to test one hypothesis about the specific kinds of political of contexts that may explain this variation. We begin with the observation that our data—multiple individual-level surveys on vote choice and retrospective economic perceptions—can be represented as a hierarchical, multilevel structure with two levels, where level-1 units are individuals (i.e., the respondents in each of the election studies) and level-2 units are the election studies themselves. Since each survey corresponds to a particular country and time point (i.e., the date the survey was administered), this structure allows for variation in the individual behavioral relationship between economic perceptions and self-reported vote choice across countries and over time. In this article we will focus on the decision to vote for or against the party of the chief executive, so a general representation of this structure could be written using a logit link function as follows:

$$\begin{aligned} v_{ik} &\sim Bin(\pi_{ik}) \\ logit(\pi_{ik}) &= \beta_{0k} + \beta_{1k}\mathrm{X}_{ik} + \sum_{j=1}^{J_k} \phi_{jk} Z_{jik}. \end{aligned} \tag{1}$$

In this notation, v_{ik} indicates a vote for the chief executive party by voter i in each of k election surveys where $i = 1 \ldots . n_k$. Likewise, X_{ik} are retrospective economic evaluations measured at the individual level and Z_{jik} are other characteristics of individuals that shape self-reported vote choice where J_k indicates the number of these control variables associated with each k election study. The two coefficients that describe economic voting in any particular survey are β_{0k} and β_{1k}, and these are allowed to vary from survey to survey. If we think that economic voting is conditioned on known contextual variables that vary over the level-2 units, then we can account for this effect by including in the estimation C_k, which is a contextual variable measured at level 2.[7] We will use two different strategies for estimating its parameters. Both allow the coefficients describing the economic voting at the individual level to vary systematically as a function of C_k but do not assume this is the only source of variation in the coefficients. The first of these

[6]The EuroBarometer is a semiannual public opinion survey that has been conducted in each of the EU member states since 1973.

[7]To simplify exposition we assume only one contextual variable in the analysis here—in practice there can be multiple contextual variables specified in these models.

produces estimates of the individual-level relationships and the contextual effects in a single step with pooled data, while the second uses a two-step strategy. In this application, both methods produce similar results and identical substantive conclusions.

3.1 *Pooled Strategy*

The pooled strategy pools the data from all the level-2 units and assumes that, conditional on C_k, variation in the parameters across level-2 units can be described by a normal distribution. Consequently, estimating variation in the parameters reduces to estimating the parameters of this normal distribution. Specifically, the model assumes that:

$$logit(\pi_{ik}) = \beta_{0k} + \beta_{1k}X_{ik} + \sum_{j=1}^{J_k} \phi_{jk}Z_{jik} \tag{2}$$

$$\beta_{0k} = \gamma_{00} + \gamma_{01}C_k + \omega_{0k} \tag{3}$$

$$\beta_{1k} = \gamma_{10} + \gamma_{11}C_k + \omega_{1k} \tag{4}$$

$$\begin{bmatrix} \omega_{0k} \\ \omega_{1k} \end{bmatrix} \sim N(0, \Omega), \quad \Omega = \begin{bmatrix} \sigma^2_{\omega 0} & \sigma_{\omega 0,\omega 1} \\ \sigma_{\omega 0,\omega 1} & \sigma^2_{\omega 1} \end{bmatrix}. \tag{5}$$

This specification concentrates only on variation in the intercept and the coefficients on economic perceptions, but variation in the coefficients on Z_{jik} could also be included.[8]

Models such as this one are quickly becoming standard in political science (see Snijders and Bosker 1999 for a good introduction) and are usually estimated using either Bayesian simulation or a quasi-likelihood method developed by Goldstein (1995). The most important feature of the models for the purposes of this article is that the estimates of γ_{01} and γ_{11} and their associated standard errors provide direct tests of the impact of measured contextual effects on economic voting. Consequently, if we can measure how concentrated or dispersed the distribution of policy-making responsibility across parties is in each electoral context in our data, we can use this method to test our hypothesis directly. Further, the model does not assume that the distribution of policy-making responsibility is the only source of variation in economic voting. The model does assume, however, that any additional random variation arises from a normal distribution, the parameters of which can be estimated and used to describe that variation.

Estimation of this one-stage model can be rather complex if all the variables included in the model are not identical.[9] Hence the typical pooled estimation strategy restricts the model specification to variables that are identically measured. With respect to the core variables in the economic vote equation—v_{ik} and X_{ik}—this is not a constraint: we could not estimate the economic voting effect without these two core variables. The requirement that the control variables, Z_{jik}, should be measured in exactly the same fashion in all the individual-level studies is considerably more constraining.[10] In our case, and in many others in which the individual-level data come from diverse election surveys, this means the specification of the individual-level behavioral model will be sparse and may

[8] Over- or under-dispersion in the individual-level binomial variation can be included at the cost of estimating one additional parameter. In the models we estimate, we include this parameter and test for over- or under-dispersion.

[9] The complexity results from the host of dummy variables and dummy interaction terms needed to distinguish the variables that are distinct or measured quite differently from one survey to the next.

[10] By measured in exactly the same fashion we mean that the question wording and response sets in the different surveys are similar enough that they can be represented by a single variable in a model analyzing a stacked or pooled data set.

necessarily exclude some variables known to be important. Specification arising in this way should be a principal concern of researchers using multilevel modeling with diverse surveys with limited sets of identically measured control variables. In our results section, we demonstrate how careful exploratory data analysis of the individual cases can alert researchers to this possibility.

An alternative option is to restrict the pooled data set to those studies that have similarly measured v_{ik} and X_{ik} and have a full set of similarly measured control variables, Z_{jik}. This effectively restricts the estimation of contextual effects to those individual surveys that were designed explicitly to have the same set of variables. For example, Hellwig (2001) estimates a fully specified pooled economic vote model employing the CSES studies but at the cost of including only nine level-2 units.

3.2 *Two-Stage Strategy*

Rather than pooling the data and estimating the multilevel model in the last section, the two-stage strategy begins by estimating the parameters in Eq. (1) separately for each level-2 unit and then using these estimated parameters (or functions of them) in second-stage analyses that explore the sources of that variation.

While the first-stage estimation is straightforward, one specification issue requires some comment. Specifically, since the estimates are produced separately for each case, it is not necessary for Z_{jik} to be the same in each model, and so the question of which variables to include in the different models arises. In our view, one should include important control variables even if these differ across contexts or are measured differently in different surveys (e.g., religion, education, and class). Of course, the resulting estimates for these control variables should not be naively compared. The advantage of including important control variables is that they may improve the specification and so help us produce consistent estimates for the variables of real interest (measures of economic perceptions). Since these variables are measured the same across contexts, consistent estimates of their effects on self-reported vote will ensure that the patterns of variation in the effects are meaningful.

In some applications, including ours, it will be more useful to examine functions of the first-stage coefficients rather than the coefficients themselves. This will likely be true when the individual-level behavioral relationship is not fully described by a single coefficient, but instead by a set of coefficients.[11] Perhaps more important, transforming the coefficients of the model to predicted or expected values (or changes in these values) will often provide a more substantively meaningful comparison across cases, in which it is easier for a researcher to bring his or her knowledge of specific cases to the examination of the pattern of variation in the estimates.

In our case we use the estimated coefficients from each survey in the first stage to produce a predicted change in vote for the party of the chief executive due to a "typical" change in the distribution of economic perceptions over the surveyed population. We start with the predicted probability that a voter i, in survey k, will vote for the party of the chief executive:

$$\hat{\pi}_{ik} = \frac{e^{\hat{\beta}_{1k}(\mathrm{X}_{ik}) + \sum_{j=1}^{J_k} \hat{\phi}_{jk} Z_{jik}}}{1 + e^{\hat{\beta}_{1k}(\mathrm{X}_{ik}) + \sum_{j=1}^{J_k} \hat{\phi}_{jk} Z_{jik}}}. \tag{6}$$

[11] In the larger project of which this article is a part we use multinomial models rather than binomial models, so this is a larger issue. In the current article it is a more minor issue but nevertheless exists because we measure economic perceptions with a set of dummy variables.

Our measure of the magnitude of the economic vote for this individual is simply the change in $\hat{\pi}_{ik}$ produced by a given change in her economic perceptions (say, from X_{ik} to X'_{ik}):

$$EV_{ik} = \frac{e^{\hat{\beta}_{1k}(X_{ik}) + \sum_{j=1}^{J_k} \hat{\phi}_{jk} Z_{jik}}}{1 + e^{\hat{\beta}_{1k}(X_{ik}) + \sum_{j=1}^{J_k} \hat{\phi}_{jk} Z_{jik}}} - \frac{e^{\hat{\beta}_{1k}(\mathbf{X}'_{ik}) + \sum_{j=1}^{J_k} \hat{\phi}_{jk} Z_{jik}}}{1 + e^{\hat{\beta}_{1k}(\mathbf{X}'_{ik}) + \sum_{j=1}^{J_k} \hat{\phi}_{jk} Z_{jik}}}. \quad (7)$$

Based on our assessment of the distribution of economic perceptions over individuals in the 163 surveys we used in our empirical analysis, we defined a typical change in economic perceptions as a move of one category in our three-category measure. In all the empirical analysis we report, we chose the direction of this change to reflect a worsening economy. To obtain an estimate of the average magnitude of the economic vote in the sample, we calculated Eq. (7) for all individuals in the sample (i.e., using the measured values of Z_{jik} and X_{ik}), setting X'_{ik} to be one category worse than X_{ik}. If the voter's economic perceptions were already at the worst category, we did not change them.[12] The economic vote for each individual was then averaged to produce an estimate of the average economic vote in the sample, EV_k. Standard errors of the predicted changes were simulated using the procedures outlined in King et al. (2000).

Equation (1) provides our general model of the individual-level relationship between economic perceptions and self-reported vote, but since our goal is to explore the sources of variation in this relationship across contexts, we must engage in a second stage of analysis. For some purposes this second-stage analysis requires nothing more than looking at the first-stage estimates graphically. As Bowers and Drake (2005) argue (and we reiterate below), this kind of exploratory analysis should be an important part of any multilevel analysis. Of course, one can also go farther and specify a statistical model in the second stage with the goal of testing hypotheses about how known differences in context impact individual level behavior.[13] In our case we do this using a regression of the following general form:

$$EV_k = \alpha_0 + \alpha_1 C_k + \upsilon_k \quad (8)$$

The "data" that we use for the dependent variable in the second-stage model are the estimates of economic voting obtained in the first stage, and so this analysis must account for the uncertainty of these estimates. The common procedure for doing this is to weight the second-stage analysis by the inverse of the standard error of the estimates of the dependent variable (King 1997, p. 290). However, as Lewis and Linzer (2005) show, this procedure can produce incorrect estimates of uncertainly because it treats all of the second-stage error as if it came from estimation uncertainty, when in fact some of it is inherent to the stochastic process governing the second stage. Lewis and Linzer (2005) provide a weighting procedure that produces appropriate standard errors but they suggest (and their Monte Carlo results show) that using standard errors that are robust to unspecified forms of

[12] We also calculated the opposite change in which each voter's perceptions get one category better except those that are already the best. Comparing the results from the two measures reveals no asymmetry in economic voting—the size of the economic vote for chief executives is the same in both cases (though with opposite signs). This is in itself an important finding that corroborates previous failures to find asymmetry in individual voting models (Lewis-Beck 1988).

[13] Jusko and Shively (2005) show that in the case in which the first-stage estimates are coefficients from a linear regression, the coefficients on the contextual variables are consistent. While this has not been shown formally for the specific case described here, our comparison of the multilevel estimates (which have been shown to be consistent under the assumptions of that model) to the ones obtained from the two-stage method reveal few differences.

heteroscedasticity is almost always as good. Given this (and because we must combine this correction with others discussed below), we adopt the latter procedure.

Another challenge for our second-stage analysis is that the incumbent chief executive is often the same party in different surveys conducted in the same country. Since we might expect the size of the economic vote for the same party across nation studies to be correlated (even conditional on a well-specified set of covariates), we need to allow for this in the estimation. We do so by grouping the data (by party in this case) and calculating standard errors that allow for an unspecified pattern of correlation within parties, but assume that there is no correlation across parties. The resulting standard errors, when combined with the correction in the last paragraph, simply allow for unspecified forms of heteroscedasticty across surveys, correlation between surveys when the party of the chief executive is the same, and no correlation otherwise. These corrections are now common and can be implemented with standard statistical packages.

4 Data Considerations

The first data collection step in our two-stage estimation procedure involves identifying those surveys that ask respondents about their vote preferences (our dependent variable) and their perceptions of the economy (our primary independent variable). These questions are rarely asked in exactly the same way in each survey. Thus we need to make judgments about the degree of difference in question wording that is tolerable between surveys. Too high a tolerance will undermine the comparability of the resulting estimates, but too low a tolerance will unduly limit the number of surveys included in the analysis. Fortunately, however, there is a remarkable degree of consistency in the question wording for self-reported vote and economic perception questions across surveys, and where differences do exist, they are seldom idiosyncratic. Rather, there are usually only two or three ways a given question is asked and so there is almost always a group of surveys that ask questions in the same way. As a result, each type of question wording has a sufficient number of cases to allow for secondary analysis designed to identify any systematic impact of the differences on our estimates. For example, if there are two ways in which the vote preference question is asked and this difference matters for our analysis, then we should see some systematic difference in our estimates of economic voting for surveys in which the different questions were used.

We reviewed over 300 candidate surveys for inclusion in this study, and 163 were judged to have sufficiently comparable questions on self-reported vote and economic perceptions.

Comparable Vote Choice Questions. Surveys that asked respondents to indicate whom they would vote for in a single election were included in our study. The vote preference questions differed in one important respect: their relationship to the election for which the vote applied. Surveys conducted directly after elections ask respondents for their vote choices in the preceding election. Surveys that were conducted just before an election ask respondents for whom they intend to vote for in the upcoming election. Finally, surveys that were not proximate to an election (e.g., many of the EuroBarometer surveys) ask the voter about a hypothetical election: "If there were a general election tomorrow, which party would you support?" The key question for our analysis is whether these differences introduce systematic biases into our estimates of the strength of economic voting that will make them less comparable. In other work (Duch and Stevenson 2005), we have done an extensive secondary analysis looking for any systematic influence of this difference on our estimates of economic voting. No significant influence was found.

Comparable Economic Perception Questions. Each of the 163 surveys used in the analyses in this article asks respondents a question of the following general form:

> Looking back over the last year, would you say that the economic situation in [name of country] has gotten much better, somewhat better, stayed the same, somewhat worse, or much worse?

To be included in our study, we required that the question be retrospective, that it refer to the national economy, and that it concern change in the economy rather than the absolute level of the economy.[14] However, we tolerated less fundamental differences, like whether the retrospective evaluation was over a year or two years and whether the number of response categories was three rather than five. Again, our secondary analyses suggest that these different wordings of the evaluation question have no impact on our economic vote results.

Practical constraints dictate the adoption of the retrospective version of national economic evaluations because in the elections surveys conducted over the past 25 years, this question is included far more frequently than the question about prospective evaluations. This, of course, is not by chance but rather reflects the theoretical importance of the retrospective model of the voter's economic reasoning, beginning with Key's (1966) *The Responsible Electorate*; the micro-level findings in the United States of Fiorina (1978, 1981) and Kinder and Kiewiet (1979, 1981). Since the purpose of this article is not to determine the relative importance that voters accord retrospective versus prospective assessment in evaluating the macro economy, we focus exclusively on retrospective national evaluations.

Identifying Control Variables. The practical requirements of the one- and two-stage models call for different approaches to the inclusion of control variables. Specifically, in the two-stage approach we can include different sets of control variables in each first-stage specification if this is warranted by our knowledge of voting behavior for the case. In the pooled models this kind of tailoring for specific cases is possible on a limited basis (by using dummy variables to activate a particular variable for a single case), but practical limitations (the size of the resulting data matrices) limit this strategy as a general solution. Most scholars who use a pooled estimation strategy simply choose a set of controls that are measured the same across all cases. This is the strategy adopted here. But in our case, since we are analyzing such a diverse set of election surveys, we were constrained to include only a small number of control variables.

We are not restricted to including identically measured variables in the case of the two-stage procedure. The set of control variables can vary from one election study to the next. In choosing control variables for the two-stage models, we attempted to identify variables that are known to be important in voting in the particular country at the time of the survey. To do this, we mined the literature on comparative voting behavior and on the country-specific literatures on voting in each country. In theory, this strategy could lead to very different model specifications in different cases (at least for the two-stage models), but in practice differences in the list of controls across cases were fairly modest.[15]

[14]Questions concerning absolute levels of the economy would resemble the following: "Would you consider the current economic situation in the country to be very good, good, bad, or very bad?"

[15]This should not be surprising, since there is a great deal of agreement across countries about the basic factors that drive vote choice. This means that the literatures in different countries usually point to the same kinds of variables as important determinants of the vote. In addition, since the scholars who have written the voting literatures in each country are usually the same people who design the surveys, measures of these basic factors are usually included in election studies. This, of course, does not mean that the same concepts were *measured* the same way in different surveys, just that there tends to be agreement in the concepts included.

Given space constraints and the fact that the literature on vote choice in Western democracies is very well known, we do not provide details on the specification of each model in the two-stage procedure. Three categories of control variables are included in each election study model: socioeconomic cleavages (typically religion, age, gender, income, class, urban/rural residence, religion, region, language, and race); partisanship; and issue preferences (left-right identification and positions of voters and parties on different policies). We encourage readers to explore the "specification chart" that can be found at www.raymondduch.com/economicvoting in which we provide the exact specification for each of the 163 models estimated in the study.[16]

5 Exploratory Data Analysis

Bowers and Drake (2005) argue that hypothesis testing using multilevel methods should be preceded by careful exploratory analysis of the behavioral relationships in the different level-2 units. This should involve, at a minimum, graphing estimates of the behavioral relationship in each level-2 unit and comparing the extent and nature of the estimated variation as well as the precision with which the estimates are produced. Accordingly, we begin our empirical analysis by describing the variation in economic voting in each survey. Our graphical presentation of the data uses the estimates of EV_k we generated using the two-stage model. This is essentially the method recommended by Bowers and Drake (2005). It is also possible, however, to produce similar estimates of variations in the economic vote from the level-2 residuals in a pooled multilevel model without contextual variables. We end this section, then, by comparing the two estimation methods.

Figure 1 provides the estimates of EV_k from the one-stage model. This figure reveals first of all that there is significant variation in the magnitude of this causal relationship across national and temporal contexts. The dots in the plot are the estimated magnitude of the economic vote for the chief executive based on the two-stage method and the bars are their 95% confidence intervals. The plot illustrates quite nicely one of the chief advantages of the two-stage approach for practical data analysis—it lets you easily see the estimated cross-contextual variation in an individual behavioral relationship. Such plots are essential in understanding what is going on in the individual-level models, how they differ, and whether they produce plausible variation across context. They are essential in spotting outliers and potentially problematic surveys and may also reveal unexpected contextual variation that would be worth exploring theoretically.

Looking more closely at Fig. 1, we can see immediately that there are no significant outliers among our estimates and that the extent of variation in the estimates is plausible. The change in the chief executive's vote due to worsening economic perceptions is about −5% on average and ranges from −15% to 8%. This result is completely consistent with the historically modest electoral volatility that characterizes most elections in established democracies. In addition, this magnitude comports with other efforts to assess the magnitude of the economic vote. For example, this average level of economic voting is quite similar to the 6% that Blais et al. (2004) obtain employing a similar measure.

[16]This chart is simply too big to include in this article. For most countries, as we show below, the individual-level estimates are fairly robust to the richness (or sparseness) of the specification. More generally, for most countries we found stability in the economic voting relationship despite specification, as long as there was something in the model that captured partisan differences (e.g., left/right self placement, partisanship measures, or batteries of policy questions).

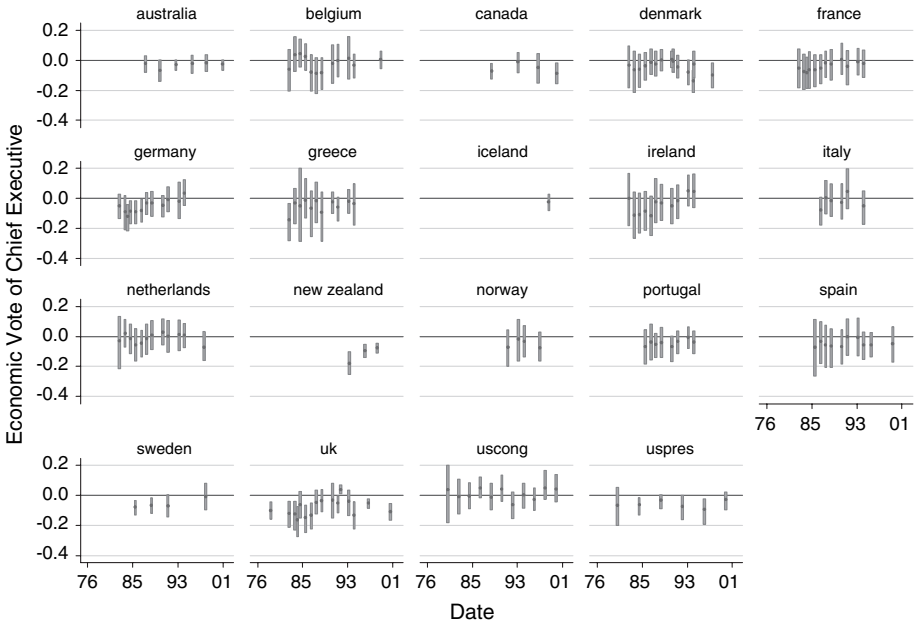

Fig. 1 A map of economic voting for the party of the chief executive. The upper bound of once confidence interval for Greece is truncated.

In addition to the plausibility of the range of the estimates, it is also clear from Fig. 1 that, in line with almost all theories of economic voting, ruling parties are hurt as economic perceptions worsen. Indeed, about 80% of the cases are negative and about 50% have values less than −5%. Further, if we examine the total area in the 95% confidence intervals across all the cases in Fig. 1, we find that over 70% of the density of these intervals falls below zero. The picture then leaves little doubt that there is a consistent cross-national economic vote that hurts incumbents who preside over a period of worsening economic perceptions. This finding is the first comprehensive evidence of economic voting at the individual level across a large number of countries and over an extended period of time. The preponderance of evidence here should ease any doubts about the generality of economic voting in the Western democracies (e.g., Paldam 1991; Cheibub and Przeworski 1999).

If we focus on the cross-national differences in economic voting that are revealed in Fig. 1, we find an identifiable group of countries in which the economic vote for the chief executive is usually close to zero: Italy, Belgium, and U.S. congressional elections, for example. We can also identify a cluster of countries in which the support of chief executives seems to depend heavily on economic perceptions: the U.K., U.S. presidential elections, Ireland, New Zealand, and Spain. This pattern of cross-national variation again suggests that the estimates are plausible since they are largely consistent with what we know from previous work. For example, one of the few efforts to identify the magnitude of cross-national differences in the economic vote, based on individual-level surveys, is Lewis-Beck (1988), and our rank ordering of countries in terms of the average magnitude of the economic vote is precisely the same as those of Lewis-Beck (from highest to lowest: Britain, Spain, Germany, France, and Italy).

In addition, the U.S. results line up exactly as the literature predicts. Specifically, the extensive literature on economic voting in the United States has consistently found

a strong economic vote in presidential elections (Kiewiet 1983; Erikson 1989; Erikson et al. 2002) while arguing for a much weaker impact of the economy in congressional elections (Erikson 1990; Alesina and Rosenthal 1995; Keech 1995). Our results are consistent with both of these expectations: U.S. presidential elections produce high levels of economic voting, compared to the other countries in our data, while we show almost no economic voting for the party of the president in congressional elections. It is also interesting (and encouraging) that the modest negative spikes that do occur in the congressional economic vote series (i.e., years of higher economic voting) tend to correspond with presidential elections, which is again what the U.S. economic voting literature would predict. In short, for those cases in which previous work provides some expectation about the likely pattern of cross-national variation in economic voting, our two-stage measures are consistent with these expectations.

One area in which we have little guidance from the existing literature is the temporal variation in the economic vote. We have argued, however, that patterns like those in Fig. 1 are useful in part because they can help us identify unexpected relationships in the data that are worth exploring in more depth. An inspection of the temporal variation of the economic vote in Fig. 1 provides one such insight and suggests at least one direction that our theoretical efforts to account for variation in economic voting might take. Specifically, many of the previous theoretical efforts to explain variation in the economic vote have focused on institutional characteristics that typically do not vary much over time (Powell and Whitten 1993; Anderson 1995). This is appropriate, of course, if most of the variation in the economic vote across contexts is confined to between-country differences rather than within-country variation. However, if there is considerable within-country variation in the extent of economic voting, we should search for explanatory variables that take on different values over time.

Figure 1 suggests that in fact there is considerable temporal variation within many of the 18 countries in our sample. The U.K. is one example: during much of the 1980s the economic vote for the party of the prime minister was extremely high, near $-.20\%$, while in the 1990s the economic vote was much closer to $-.05\%$. Belgium also exhibits an interesting pattern of temporal variation: During the early 1980s and much of the 1990s there is essentially no economic voting, but during the second part of the 1980s, the economic vote is actually quite large. This simple visual inspection of the second-stage point estimates thus illustrates the critical need of theorists to think about contextual influences on economic voting that change over time. What is it about the political and economic context that varies from election to election that might explain this temporal variation in economic voting? We take up this challenge elsewhere (Duch and Stevenson 2005), but our principal point here is that in many instances the simple visual inspection of second-stage point estimates can provide important insights into the most useful places to look for theoretical explanations of contextual variation.

6 Comparing the Two-Stage Procedure to the Pooled Multilevel Model

In this section we explore how the estimates of the economic vote obtained from each survey separately compare to those produced from the level-2 residuals of a pooled multilevel model. In order to produce the later estimates we combined individual-level data from our 163 electoral surveys to estimate a single logistic model of the individual-level vote for or against the chief executive party. This self-reported vote was modeled as a function of economic perceptions and a much reduced set of control variables that were available in most of our surveys. However, even with this reduced set of controls, the

requirement that we have a set of common control variables eliminated 17 studies, so 146 studies remain in the pooled data set. The hierarchical model we estimate is as follows:

$$\begin{aligned} logit(\pi_{ik}) &= \beta_{0k} + \beta_{1k} Worse_{ik} + \beta_{2k} Better_{ik} + \phi_{1k} \text{Ideology}_{ik} \\ &\quad + \phi_{2k} (\text{Ideology}_{ik} * \text{CE\_Ideology}_{k}) \qquad (9) \\ \beta_{0k} &= \gamma_0 + \omega_{0k} \qquad (10) \\ \beta_{1k} &= \gamma_1 + \omega_{1k} \qquad (11) \\ \beta_{2k} &= \gamma_2 + \omega_{2k}. \qquad (12) \end{aligned}$$

The usual assumption of multivariate normality of the second-level errors is also imposed. The variables included in the model are *Ideology*, which is the demeaned (employing the survey mean) left-right self-placement of the voter (coded so higher numbers indicate a more right-leaning voter) and its interaction with a dummy variable (*CE_Ideology*) indicating whether the chief executive's party is a leftist party.[17] We expect, then, that the coefficient on left-right self-placement (*Ideology*) will be positive and its interaction with the ideology of the chief executive (*Ideology* $*$ *CE_Ideology*) will be negative. The variable measuring retrospective perceptions of the national economy has been broken into three dummy variables: *Worse*, which equals one if the voter thought the economy had gotten worse in the last year; *Better*, which equals one if she thought the economy had gotten better; and *Same*, which equals one if she thought the economy had stayed the same. Two of these, *Better* and *Worse*, were included in the individual-level model. Clearly we expect these two dummy variables to have opposite signs, with worse being negative and better being positive. The pooled model is thus limited to one main control variable: the ideology of the voter relative to that of the chief executive. This sparseness in specification is almost always the cost one must pay for pooling data from disparate surveys. In our case, however, this particular variable appears to be the critical control for most of the models, so we get very sensible estimates out of the pooled models that largely match the estimates from the richer specifications.

Having pooled the data, we estimated the hierarchical model in Eq. (9), obtaining a single set of coefficients that describe the average relationships across all the studies in the sample. All these estimates were in the expected direction and were statistically significant.[18] Recall that in Eq. (9) the level-1 units are individuals while the level-2 units are election studies.[19] In this section we are not really interested in the estimates of the average level of economic voting across surveys (γ_1 or γ_2). Rather, we are interested in the survey-level residuals for each of the k election studies (estimates of ω_{1k} or ω_{2k}) that indicate how the impact of economic evaluations in different surveys deviates from its overall average effect. Specifically, we want to compare these to the estimates obtained from the separate, survey-by-survey estimates.

Differences in the estimates obtained from the pooled multilevel model and two-stage estimates can come from two sources. One is the estimation method itself and the other is

[17] In order to include as many studies as possible, we had to combine some left-right self-placement questions that gave the responded different response options. This was done by normalizing the various scales and then using the normalized scales.

[18] We do not show the coefficients from these models here since they are not needed for this discussion. However, they are provided later in the article as Model 1 in Table 2.

[19] One could designate a third level for country, but fixed country effects are easier to estimate simply by including country dummies and may make more sense, since we probably do not think of these countries as a sample out of a large set of similar units, which is really the motivation for treating the second-level coefficients as random and estimating features of the distribution of these coefficients.

the difference in the set of control variables included in the different models. The pooled multilevel model has a uniform specification of the control variables, so we were limited to a rather sparse specification of controls. In contrast, the model specifications in the first stage of the two-stage method were much richer and varied from one survey to the next. We will attempt to isolate these two potential sources of difference. First, we will compare the variation in estimated coefficients from specifications that differ only in the method of estimation; second, we will compare the extent of economic voting in estimations that differ only in the sets of controls included in the models.

Estimation effects. In order to isolate the effect of the different estimation procedures, we estimate separate logistic regressions for each survey that mirror the pooled, multilevel specification in Eq. (9)—the interaction term is dropped because *CE_Ideology* does not vary within election studies:

$$logit(\pi_{ik}) = \beta_0 + \beta_{1k} Worse_{ik} + \beta_{2k} Better_{ik} + \phi_{1k} \text{Ideology}_{ik} + \varepsilon_{ik} \tag{13}$$

We estimate this logistic model for each of the same 146 election studies included in the pooled model from Eq. (9). This generates 146 estimates of each of the β and ϕ coefficients in the model.[20] In order to compare these estimates to the estimates of the level-2 residuals from the pooled, multilevel model, we generate "survey-level residuals" for the estimated coefficient on the *worse* dummy variable:[21]

$$\psi_{1k} = \hat{\beta}_{1k} - \frac{1}{N} \sum_{k}^{K} \hat{\beta}_{1k} n_k, \tag{14}$$

where n_k is the number of observations in survey k and N is the total number of observations in all the surveys. Figure 2 plots these survey-level residuals (on the y-axis) against the level-2 residuals from the pooled multi-level model (ω_{1k} in Eq. [11]). Figure 2 reveals a tight fit between the two sets of estimates (the correlation is .81). Overall, we can conclude that there is not much systematic difference between the estimates that is due to the estimation method.[22]

Specification effects. Recall that in the separate estimations we are able to estimate models with a full set of control variables. In contrast, the control variables in the pooled multilevel estimates are restricted to those that are similarly measured in all election studies. In order to isolate the impact of the differences in the list of control variables included in the pooled and separate estimations, while holding the estimation method constant, we estimate two logistic models for each case. One logistic model is the restricted version described in Eq. (13). The other is more fully specified:

$$logit(\pi_{ik}) = \beta_0 + \beta_{1k} Worse_{ik} + \beta_{2k} Better_{ik} + \sum_{j}^{J_k} \phi_{jk} Z_{jik}, \tag{15}$$

[20]We focus on the coefficient on the worse dummy variable. An examination of the better variable generates the same conclusions.

[21]To be comparable to the multilevel result, this must be the weighted mean (where sample size provides the weight).

[22]One additional difference in the specification results from the fact that in the pooled model, we did not allow random coefficients for ideology and its interaction, while in the separate models these coefficients are clearly unconstrained. To be sure this was not critical, we estimated this model as well but the picture is essentially the same. Since the choice is inconsequential, we keep this presentation because it requires less notation.

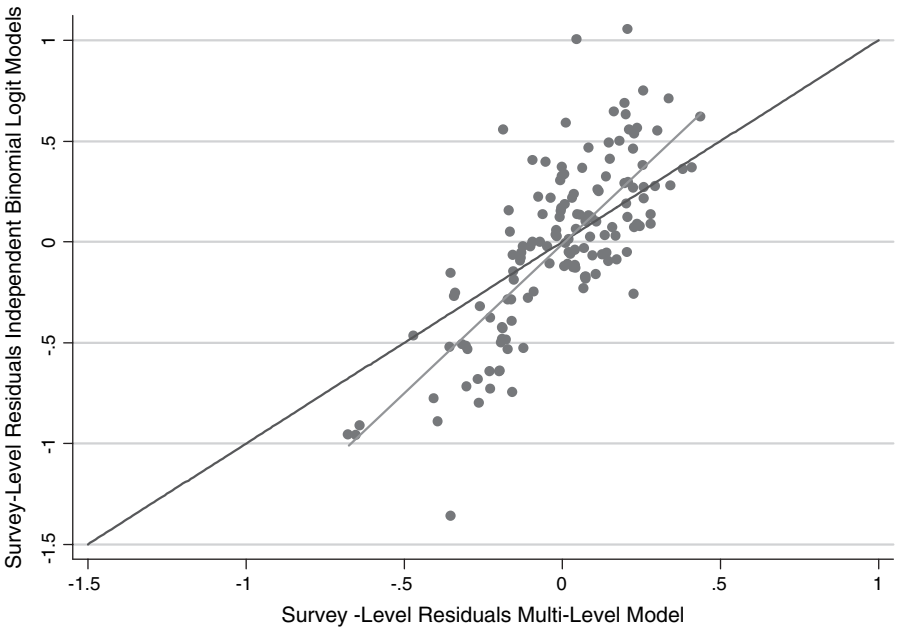

Fig. 2 Correspondence between estimated coefficients from separate logistic models and pooled multilevel model (with identical control variables in both models). Residuals from independent model are deviations of coefficient from mean across studies.

where Z_{ijk} is a potentially unique full set of J_k control variables for each of the k election studies.[23] For each survey, the full set of control variables used is the same one employed in separate estimates that produced Fig. 1. Since the estimation method is the same, the only source of difference in the estimates produced from Eqs. (13) and (15) is the set of control variables. Here we cannot simply compare coefficients from the two methods, as we did in Fig. 2, because the two equations have different specifications. Rather, for the assessment of specification effects we compare the differences in substantive effects resulting from the equations. We calculate EV_k for the fully specified and restricted specifications and compare these. Thus, Fig. 3 plots the estimated economic vote from the fully specified model (Eq. [15]) on the y axis and the estimated economic vote for the restricted specification (Eq. [13]) on the x axis.

This comparison makes it clear that for most cases the two-stage model produces estimates for economic voting that are quite similar to those produced by the more limited uniform specification of controls used in the multilevel case (the correlation is .61). However, estimates for Australia (and for a couple of Greek cases) are not similar across the different specifications (excluding Australia raises the correlation to .71). A close examination of the Australian case reveals why. Specifically, when we estimate individual-level vote models for Australia, we find that the estimates differ greatly depending on whether one includes a measure of respondents' feelings about the prime ministerial candidates and/or a battery of respondent policy positions (these are in addition to a left-right self-placement measure). If either of these variables is included, our estimates of the extent of economic voting are consistent with the general level of economic voting observed elsewhere—i.e., a change of 5% to 10% of the vote. However, as Fig. 3 shows,

[23]Ideology is included in the set of Z_{jik} variables.

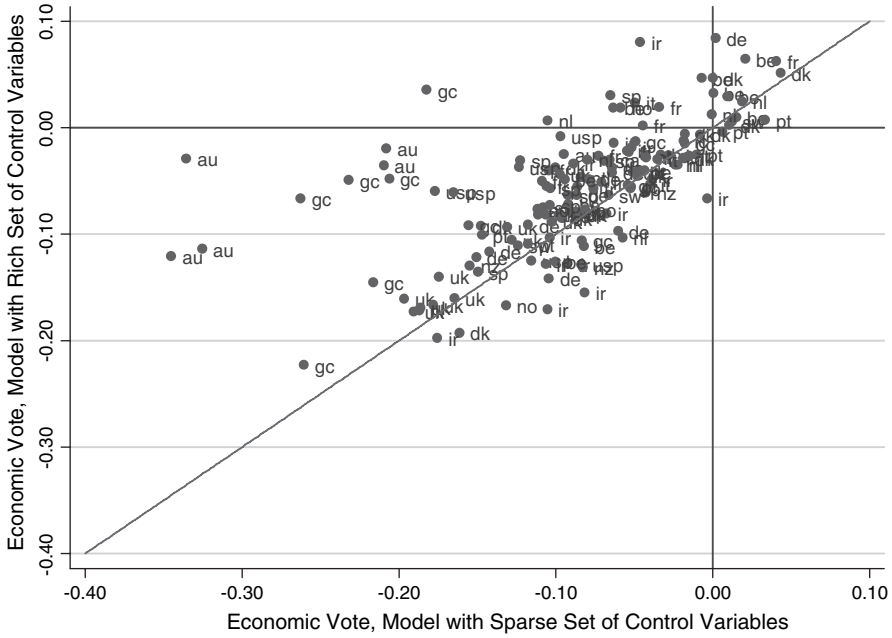

Fig. 3 Correspondence between levels of economic voting from independent logistic models with rich or sparse sets of control variables.

the sparse specification that does not include these variables results in dramatically higher estimates of economic voting: changes in support of between 30% and 40%. This suggests to us that in this case it is the sparse specification that is faulty—a conclusion that is in line with the literature on Australian voting behavior, which has emphasized the importance of PM evaluations in vote choice (McAllister 2001). The wider message of these results, however, is that the specification of the control variables included in the individual-level models must be appropriate for the country under consideration. This point is fairly obvious in the two-stage approach because one spends time fleshing out each individual-level model and can easily examine the sensitivity of the estimates to differences in the list of control variables. However, when a uniform specification is dictated because the data are pooled, it is easy to forget that the uniform specification may not be appropriate to each particular case. Exploratory data analysis of the kind pursued here should be part of any such analysis.

Finally, with the exception of the specification of the Australian vote models, our comparison of these two methods of inductively mapping the variation in economic voting using multiple election surveys does not appear to produce significantly different results. Indeed, there are almost no differences in the results that are due to the method of estimation itself. Of course, this fortunate result is very likely due to the fact that we have a large number of observations for each of our second-level units (i.e., surveys) and so *should not* be taken as evidence that the two methods would produce the same results when the number of observations for some second-level units is limited.

7 Testing the Impact of Power Sharing on the Economic Vote

The exploratory data analysis makes it clear that the economic vote in developed democracies varies significantly from one election to the next. However, the primary

concern of our empirical effort (here and in the larger project that this is part of) is to explain this variation in terms of specific aspects of the political and economic context in which elections take place. In this section we will test the contextual hypothesis described earlier using both the pooled and two-stage methods: *As the distribution of policy-making responsibility is shared more equally among the parties in an election, retrospective economic evaluations will be more important in the vote choice.*

Our measure of the distribution of policy-making responsibility takes into account both the coalition status of the cabinet and the extent to which power within coalitions is shared equally among cabinet partners. Specifically, given n parties in an election, we can define a vector δ as the n vector of equal shares of policy-making responsibly. δ has n equal elements that sums to one. Since this gives equal policy-making responsibility to all parties in the election, our hypothesis suggests that this vector of equal shares is the distribution of policy-making responsibility that should lead to the least amount of economic voting overall. Now, given some real-world n vector of shares of policy-making responsibility (call this vector λ), we can use the vector distance between λ and δ as a measure of the degree to which responsibility is concentrated in the system.[24] To construct this measure for each of our 163 cases, we use each party's share of cabinet portfolios for λ, construct a δ of appropriate length, and then calculate the resulting vector distance between the two.

If the elements of λ are equal across parties, the vector distance we calculate will be zero, no matter how many parties are in the election. However, the upper bound of this distance—which occurs when one party has all the administrative responsibility—will differ depending on the number of parties. Specifically, the more parties in the system, the bigger this upper bound will be. If one thinks that voters see power as more concentrated as the number of parties excluded from power is bigger, then this differing upper bound is appropriate. However, if voters perceive power as equally concentrated when, for example, one party holds all the power in a two-party system and when one party holds all the power in a five-party system, then we need to normalize this distance by dividing by the upper bound. Below, we report results using the second strategy, though the choice is not empirically consequential. We call this measure *concentration of authority*, and it can take values between zero (equality) and one (complete concentration of authority in one party).

With this measure in hand, we can now test the hypothesis using a pooled, multilevel model and doing a second-stage analysis in which we use the separate estimates of EV_k as a dependent variable. In the two-stage method, our model of the impact of *concentration of authority* on the magnitude of the economic vote is

$$EV_k = \alpha_0 + \alpha_1 A_k + \upsilon_k, \tag{16}$$

where A_k is the measure of *concentration of authority* defined earlier and EV_k is the first-stage estimate of the economic vote for the party of the chief executive. The various corrections for uncertainty in the measure of the dependent variable and for possible clustering on party are included. In addition, we have estimated the model with and without a full set of country and time dummy variables.[25]

We provide the results of our two-stage estimates of this model in Table 1. The expected sign on the coefficient is negative, indicating that the decline in the probability of voting for the incumbent (as a result of a deteriorating economy) is greater when there is less sharing of administrative responsibility.

[24]The vector distance is just the square root of the summed squared differences in the elements of the two vectors.
[25]These estimates are available from the authors. They are almost identical to those reported in Table 1.

Table 1 Concentration of authority and economic voting for the PM (two-stage method)

	All Cabinets	*Coalition Cabinets*
Concentration of authority	−.081 (0.023)	−.075 (.033)
Constant	0.02 (0.020)	.007 (.02)
Number of observations	152	76
Adjusted R-squared	0.05	0.04

Note. Numbers in cells are coefficients from OLS regression with robust standard errors clustered on party. Numbers in parentheses are standard errors. The effects are statistically significant at $p < .01$. U.S. congressional election surveys are excluded. The dependent variable is economic vote for chief executive (the party of the prime minister in all cases except the United States, where it is vote for the president).

The first model is for all cabinets including single-party prime ministers. The second model excludes single-party prime ministers. Both estimates show a strong relationship in the expected direction. That is, when authority is more concentrated, we observe more economic voting for chief executives than when it is shared more equally among parties. This result is consistent with our competency model of economic voting—contexts in which the competency signal is greater result in higher levels of economic voting. And, importantly for the purposes of this article, it demonstrates the usefulness of the two-stage method for investigating hypotheses about the specific ways political and economic context condition individual behavior.

We now turn to comparing these estimates with those from the pooled multilevel model. In the pooled estimation, we use the combined individual-level data from our 146 electoral surveys to estimate a single logistic model of the individual-level vote for the party of the chief executive. The specification of variables in the pooled model is similar to that described by Eq. (9) but adds *concentration of authority* to the model. In addition, for some of the models reported here we add a *size* variable, which is simply the average support for the chief executive party in the survey. This is obviously constant for each survey and essentially allows the constant-only model (which captures the probability that an ideologically average voter will vote for the prime minister's party) to vary based on the general level of support for the PM's party in the population. The model, with the *size* variable, is:

$$\begin{aligned} logit(\pi_{ik}) &= \beta_{0k} + \beta_{1k} Worse_{ik} + \beta_{2k} Better_{ik} \\ &\quad + \phi_{1k}\text{Ideology}_{ik} + \phi_{2k}(\text{Ideology}_{ik} * \text{CE\_Ideology}_{k}) \qquad (17) \\ \beta_{0k} &= \gamma_{00} + \gamma_{01} Size_k + \gamma_{01} A_k + \omega_{0k} \qquad (18) \\ \beta_{1k} &= \gamma_{10} + \gamma_{11} A_k + \omega_{1k} \qquad (19) \\ \beta_{2k} &= \gamma_{20} + \gamma_{21} A_k + \omega_{2k}. \qquad (20) \end{aligned}$$

Since we expect support for the chief executive to be negatively related to economic perceptions, we expect $\gamma_{10} < 0$ and $\gamma_{20} > 0$ (recall that the baseline category is voters who think the economy has stayed the *same*). Further, since greater concentration of authority is expected to increase the size of the economic vote, we expect $\gamma_{11} < 0$ and $\gamma_{21} > 0$ (since this will increase the difference between the overall impact of better versus worse perceptions).

Table 2 provides the estimates.[26] Model 4 is exactly Eqs. (17)–(20) and the coefficients listed next to the labels are for that model. This model clearly supports the hypothesis,

[26]The estimates were obtained using the PQL second-order linearization method outlined in Goldstein (1995). Diagnostics on estimated residuals at level 2 suggest that the assumption of normal variance in the level-two coefficients is not violated.

Table 2 Pooled, multilevel estimates

	Model 1	*Model 2*	*Model 3*	*Model 4*
Left-right self placement	0.459	0.479	0.473	0.482
(deviation from mean)	0.005	0.005	0.005	0.005
Left-right Self placement* Leftist CE	−0.691	−0.721	−0.725	−0.741
(Left chief executive is indicator variable)	0.007	0.007	0.007	0.007
Voter perceives economy got better	0.428	0.428	0.454	0.449
(indicator variable)	0.028	0.028	0.028	0.028
Voter perceives economy got worse	−0.56	−0.534	−0.590	−0.575
(indicator variable)	0.034	0.033	0.028	0.028
Concentration of authority	—	—	1.961	0.838
(deviation from mean)	—	—	0.185	0.144
Voter perceives economy got better*	—	—	0.568	0.565
Concentration of authority	—	—	0.144	0.148
Voter perceives economy got worse*	—	—	−0.976	−0.952
Concentration of authority	—	—	0.144	0.144
Size of chief executive party	—	5.252	—	4.552
	—	0.223	—	0.240
Constant	−0.920	−2.396	−0.883	−2.184
	0.046	0.067	0.036	0.072
$\sigma^2_{\omega 1}$	0.285	0.090	.163	.068
$\sigma^2_{\omega 2}$	0.121	0.119	.066	.065
$\sigma^2_{\omega 3}$	0.073	0.081	.066	.072
$\sigma_{\omega 1,\omega 2}$	−0.110	−0.048	−.038	−.020
$\sigma_{\omega 1,\omega 3}$	0.030	−0.023	$-.012^{ns}$	−.034
$\sigma_{u\ \omega 2,\omega 3}$	−0.053	−0.052	−.032	−.031
Survey-level variance for those who think economy is better (other variables at mean)	0.418	0.125	.205	.072
Survey-level variance for those who think economy is same	0.285	0.09	.163	.068
Survey-level variance for those who think economy is worse	0.186	0.113	.153	.093
N	147435	147435	147435	147435

Note. Standard errors listed below logit coefficients. All coefficients are statistically significant at $p < .001$ unless indicated by "ns." Congressional election studies are excluded.

since all the coefficients are statistically significant and in the expected direction. Thus, this pooled, multilevel model tells exactly the same story as the two stage estimates—shared power decreases the economic vote of the chief executive.[27]

The other models included in Table 2 provide estimates with and without the size variable and give estimates for the model without concentration of authority (Eq. [9]). Comparing the models with and without concentration of authority can tell us how much of the variation in the level-2 coefficients is accounted for by the concentration of authority variables. Specifically, the multilevel model does not assume that *concentration of authority* captures all the contextual variation in economic voting. Instead, it allows for random variation in the estimated coefficients for economic voting even after accounting

[27]We estimate a similar model that also includes a dummy variable for whether chief executive was a Left party and the results are virtually identical. These estimates are available from the authors.

for the impact of concentration of authority. These level-2 variances are estimated along with the coefficients and can be used to calculate the overall variation in the level-2 coefficients that remains once the systematic variation (captured by *concentration of authority* and, in some cases, *size*) is accounted for in the model. These estimates are provided at the bottom of Table 2.[28] In all cases, we find less random variation in the economic voting coefficients when differences in the *concentration of authority* across cases are explicitly included in the model. We can be sure, then, that some of the election-to-election variation in the importance of economic voting is due to differences in the *concentration of authority.*

8 Conclusion

Efforts to assess this relationship between retrospective economic perceptions and self-reported vote in different countries and at different times have resulted in a wide range of disparate results. Indeed, our analysis of data from 163 election studies confirms the basic variability of the economic voting relationship from election study to election study, while at the same time suggesting that the scope of this variability is within plausible limits (for example, the relationship is almost always in the expected direction).

In this article, we have proposed a competency model for explaining some of this variation and have examined two multilevel procedures (a pooled and a two-stage approach) for testing one hypothesis that follows from our theory and for exploring the nature of the variation in economic voting more generally. Both methods suggest that the concentration of policy-making authority enhances the importance of economic voting and so provides the first large-scale confirmation of the hypothesis using individual data from diverse national contexts.

Further, we hope our demonstration that substantively important information about the impact of context is available in the large number of existing election surveys will lead other students of comparative behavior to help open up this underexploited resource. We think our article demonstrates the feasibility of using these data, even when they include only a few variables that have been similarly measured across surveys.

References

Alesina, Alberto, and Howard Rosenthal. 1995. *Partisan Politics, Divided Government and the Economy.* Cambridge: Cambridge University Press.

Alvarez, R Michael, Jonathan Nagler, and Jennifer R. Willette. 2000. "Measuring the Relative Impact of Issues and the Economy in Democratic Elections." *Electoral Studies* 19:237–53.

Anderson, Christopher J. 1995. *Blaming the Government: Citizens and the Economy in Five European Democracies.* Armonk, NY: Sharpe.

Anderson, Christopher J. 2000. "Economic Voting and Political Context: A Comparative Perspective." *Electoral Studies* 19:151–170.

Barro, Robert. 1973. "The Control of Politicians: An Economic Model." *Public Choice* 14:19–42.

Blais, Andre, Mathieu Turgeon, Elisabeth Gidengil, Neil Nevitte, and Richard Nadeau. 2004. "Which Matters Most? Comparing the Impact of Issues and the Economy in American, British and Canadian Elections" *British Journal of Political Science* 34:555–564.

Bowers, Jake, and Katherine W. Drake. 2005. "EDA for HLM: Visualization When Probabilistic Inference Fails." *Political Analysis* doi:10.1093/pan/mpi031.

[28]Since the variance calculation depends on the values of the data, they are provided for several different cases (see Goldstein et al. 2002).

Cheibub, Jose Antonio, and Adam Przeworski. 1999. "Democracy, Elections, and Accountability for Economic Outcomes." In *Democracy, Accountability, and Representation*, eds. Adam Przeworski, Susan C. Stokes, and Bernard Manin. Cambridge: Cambridge University Press, pp. 222–250.

Duch, Raymond M. 2001. "A Developmental Model of Heterogeneous Economic Voting in New Democracies." *American Political Science Review* 98(4):895–910.

Duch, Raymond M., and Randy Stevenson. 2005. "Context, Competence and the Economic Vote." Unpublished manuscript.

Erikson, Robert S. 1989. "Economic Conditions and the Presidential Vote." *American Political Science Review* 83:568–573.

Erikson, Robert S. 1990. "Economic Conditions and the Congressional Vote: A Review of the Macrolevel Evidence." *American Journal of Political Science* 34:373–399.

Erikson, Robert S., Michael B. Mackuen, and James A. Stimson. 2002. *The Macro Polity.* Cambridge: Cambridge University Press.

Eulau, Heinz, and Michael S. Lewis-Beck, eds. 1985. *Economic Conditions and Electoral Outcomes: The United States and Western Europe.* New York: Agathon.

Fearon, James D. 1999. "Electoral Accountability and the Control of Politicians: Selecting Good Types versus Sanctioning Poor Performance." In *Democracy, Accountability, and Representation*, eds. Adam Przeworski, Susan C. Stokes, and Bernard Manin. Cambridge: Cambridge University Press, pp. 55–97.

Ferejohn, John. 1986. "Incumbent Performance and Electoral Control." *Public Choice* 50:5–25.

Fiorina, Morris. 1978. "Economic Retrospective Voting in American National Elections: A Micro-Analysis." *American Journal of Political Science* 22:426–443.

Fiorina, Morris P. 1981. *Retrospective Voting in American National Elections.* New Haven: Yale University Press.

Goldstein, H. 1995. *Multilevel Statistical Models*, 2nd ed. London: Edward Arnold.

Goldstein, H., W. J. Browne, and J. Rasbash. 2002. "Partitioning Variation in Multilevel Models." *Understanding Statistics* 1:223–232.

Hellwig, Timothy T. 2001. "Interdependence, Government Constraints, and Economic Voting." *Journal of Politics* 63(4):1141–1162.

Jusko, Karen Long, and W. Phillips Shively. 2005. "Applying Two-Step Strategy for the Analysis of Cross-National Public Opinion Data." *Political Analysis* doi:10.1093/pan/mpi030.

Keech, William. 1995. *Economic Politics and the Costs of Democracy.* Cambridge: Cambridge University Press.

Key, V. O. 1966. *The Responsible Electorate.* New York: Vintage Books.

Kiewiet, D. Roderick. 1983. *Macroeconomics and Micropolitics: The Electoral Effects of Economic Issues.* Chicago, IL: University of Chicago Press.

Kinder, Donald R., and D. Roderick Kiewiet. 1979. "Economic Grievances and Political Behavior: The Role of Personal Discontents and Collective Judgments in Congressional Voting." *American Journal of Political Science* 23:495–527.

Kinder, Donald R., and D. Roderick Kiewiet. 1981. "Sociotropic Politics: The American Case." *British Journal of Political Science* 11:129–161.

King, Gary. 1997. *A Solution to the Ecological Inference Problem: Reconstructing Individual Behavior from Aggregate Data.* Princeton, NJ: Princeton University Press.

King, Gary, Michael Tomz, and Jason Wittenberg. 2000. "Making the Most of Statistical Analyses: Improving Interpretation and Presentation." *American Journal of Political Science* 44(2):341–355.

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

Lewis-Beck, Michael S. 1988. *Economics and Elections: The Major Western Democracies.* Ann Arbor: University of Michigan Press.

Lewis-Beck, Michael S., and Glenn Mitchell. 1990. "Modelos Transnacionales de Voto Economico: Estudio de un Conjunto de Paises Europeos." *Revista del Instituto de Estudios Economicos* 4:65–81.

Lewis-Beck, Michael S., and Mary Stegmaier. 2000. "Economic Determinants of Electoral Outcomes." *Annual Review of Political Science* 3:183–219.

McAllister, Ian. 2001. "Prime Ministers, Opposition Leaders and Government Popularity in Australia." Presented at the Australian Politics stream, Australasian Political Studies Association conference, Brisbane.

Paldam, Martin. 1991. "How Robust is the Vote Function? A Study of Seventeen Nations over Four Decades." In *Economics and Politics: The Calculus of Support,* eds. Helmuth Northop, Michael Lewis-Beck, and Jean-Dominique Lafay. Ann Arbor: University of Michigan Press, pp. 9–31.

Palmer, Harvey D., and Guy D. Whitten. 2003a. "Ignorance Is No Excuse: Data Mining versus Theoretical Formulation of Hypotheses" *British Journal of Political Science* 33:159–160.

Palmer, Harvey D., and Guy D. Whitten. 2003b. "Questionable Analyses with No Theoretical Innovation: A Response to Royed, Leyden and Borrelli." *British Journal of Political Science* 33:139–149.

Powell, G. Bingham, and Guy D. Whitten. 1993. "A Cross-National Analysis of Economic Voting: Taking Account of the Political Context." *American Journal of Political Science* 37:391–414.

Royed, Terry J., Kevin M. Leyden, and Stephen A. Borrelli. 2000. "Is 'Clarity of Responsibility' Important for Economic Voting? Revisiting Powell and Whitten's Hypothesis." *British Journal of Politics* 30:669–698.

Snijders, Tom A. B., and Roel J. Bosker. 1999. *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling.* London: Sage.