Analysis of Voteshares Across Party Systems*

James Honaker† Drew Linzer‡

Draft prepared for the meetings of
the American Political Science Association,
August 2006.

ABSTRACT

Three problems complicate the analysis of voteshares; the same parties do not contest elections 1) across districts, 2) across time, or 3) across countries. Adding a model of the policy space to estimate substitution effects presents a solution.

Models of compositional variables are increasingly common through a number of disciplines, and within political science are commonly applied to vote shares, particularly to model the flow of votes between parties in a many-party system. However, broad, encompassing models of party vote shares are hamstrung whenever the set of parties contesting the elections differs across observations. The transformations used in the compositional literature assume the elements of the composition remain constant, and modifications that relax this requirement can only cope with small differences between observations.

We examine the use of log cumulative transformations to model the cutpoints between parties in a policy space, and show how this model is robust to the an ever-changing series of parties across observations. We demonstrate our model with an analysis to determine the effect of incumbency on party vote share in parliamentary democracies.

*Our thanks for helpful comments to Jonathan Wand, for research assistance to Catherine Brewer, and for providing their data to Marty Cohen, Hans Noel and John Zaller.

†Department of Political Science (Bunche Hall, Box 951472 Los Angeles, CA 90095-1472; tercer@ucla.edu).

‡Department of Political Science (Bunche Hall, Box 951472 Los Angeles, CA 90095-1472; dlinzer@ucla.edu; http://dlinzer.bol.ucla.edu).

1. INTRODUCTION

The share of the vote that each party receives in an election is an example of a *compositional variable*. In any election, each party's voteshare is positive and the sum of the voteshares across all parties must be one hundred percent. Compositional variables are common in the natural and social sciences, in questions such as the proportions of various minerals in rocks, the distribution of species in ecologies, and the allocation of workers' time to different tasks. The literature for models of such variables is grounded in the work of Aitchison (1986).

Compositional variables are also common in Political Science. The allocation of government budgets between different categories of expenditure, the distribution of campaign resources, the uses of bureaucratic time, and the ethnic composition of the voting electorate are all examples of issues that revolve around compositions. Voteshares, however, are the variables in Political Science most commonly, explicitly modeled as compositional data, and are possibly the most challenging (Katz and King (1999), Honaker Katz and King (2002), Jackson (2002), Tomz, Tucker and Wittenberg (2002)).

Voteshares, that is the fraction of the vote that each party receives in an election, bring additional complications rarely present in the other examples mentioned in the natural, social and political sciences, because the categories of the composition may change between observations. First, within an election, not all parties may run on the ballot in all districts. Second, within a country, but across time, which parties exist may change, as may the total number of parties. Third, across countries the variety, platforms and number of parties certainly varies[1].

In summary, these three problems are that the same parties do not contest elections across districts, across time, or across countries. Katz and King (1999) (hereafter KK) and in expansion, Honaker Katz and King (2002)[2] provide a model that is well suited to the first of these problems, and sometimes the second, but never the third.

2. COMPOSITIONAL DEPENDENT VARIABLES

Consider the voteshare, V, in election i for party j. The compositional nature of the variable is expressed by the constraints that the fraction of the vote any party might receive is doubly bounded,

$$V_{ij} \in [0,1] \qquad \forall\; i,j \tag{1}$$

and the set of votes in a district sums to unity,

$$\sum_{j=1}^{J} V_{ij} = 1 \qquad \forall\; i \tag{2}$$

where J is the total number of parties. The space of the raw voteshare, then, is the J-dimensional simplex. The transformation of Aitchison creates a set of $J-1$ *log vote ratios*

[1] It is extremely rare for the same party to contest elections across different countries. One possible example would be the parties in West Germany that then exist in the unified Germany, and in the other direction, the parties of Czechoslovakia that continue after the partition. Another example is *Shin Féin* which contests elections in both Northern Ireland and the Irish Republic.

[2] The Katz and King model treats districts where not all parties contest the election as censored observations of the effective vote, and the model is estimated via Full Information Maximum Likelihood requiring numerical integration. The Honaker, Katz and King model makes computation numerically feasible for larger numbers of parties by treating the censored observations of the effective vote as a missing data problem.

each of which compare the vote of some particular party to that of a baseline or reference party (here, party J by convention):

$$Y_{ij} = \ln\left(\frac{V_{ij}}{V_{iJ}}\right) \qquad \text{for } j = 1, \ldots, J-1. \tag{3}$$

The set of log vote ratios, Y_{ij}, are now individually and collectively unconstrained, thus all i observations (elections) for party j can be modeled in one equation, giving us $J-1$ total sets of equations. If we assume that these transformed Y's are distributed multivariate normal (or multivariate t), then this corresponds to the original votes, V, having an additive multivariate normal distribution (or additive multivariate t). After modeling, estimates are mapped back onto the simplex and the results are recovered in their original scale of interest. The reverse transformation from log vote shares to votes is:

$$V_{ij} = \frac{e^{Y_{ij}}}{1 + \sum_{k=1}^{J-1} e^{Y_{ik}}} \qquad \forall j, \text{ where } Y_{iJ} \equiv 0. \tag{4}$$

Although this transformation facilitates a simple estimation procedure with a system of normal or t distributions, and a linear functional form of the independent variables, the coefficients that are created are of an unintuitive dependent variable, and thus not readily interpretable. Each coefficient describes how the log ratio of some party changes with regards the reference party. Interpretation of parameters requires the set of techniques familiar to users of nonlinear models, such as predicted and estimated values, first-differences, and relative risks.

However, this transformation does simplify quite nicely in the two-party case, which many students of elections are familiar with dealing with. Often researchers are investigating two-party systems, or fabricate a dichotomy out of a multiparty system (perhaps all post-communist parties versus all reform parties). Then they choose one of these two vote shares and simply run linear regression. Any coefficient that is positive means that the dependent variable chosen gains, and the variable not explicitly modeled loses, as that variable increases, in the linear fashion we are familiar with in the OLS model[3] The model of Aitchison, although it is unintuitive, does mimic this common linear model in the two-party case. Figure 1 shows the log ratio transformation when $J = 2$, and thus:

$$Y_{i1} = \ln(V_{i1}/V_{i2}) = \ln(V_{i1}/(1 - V_{i1})) \tag{5}$$

While this transformation has the shape of a rotated S over the possible values of V_{i1}, contested elections rarely fall outside a 40-60 split. Inside this range, as graphed in bold on the left of figure 1 and magnified on the right, the log vote share transformation is very close to linear. Therefore, the common model used by applied practitioners in two-party cases, or constructed dichotomies, does not disagree with researchers who are using Aitchison's transformation.

3. "STRUCTURAL" OR "ESSENTIAL" ZEROS

The possibility exists in any model that one of the categories in the composition (in our case, one of the voteshares) is at the lower bound of zero. This would happen whenever a party receives exactly zero votes. This might happen because a party chose not to run in that district, or the party does not exist at that point in time, or the party does not belong

[3]The choice of which party to model is irrelevant, as this will only serve to change the signs of all the coefficients.

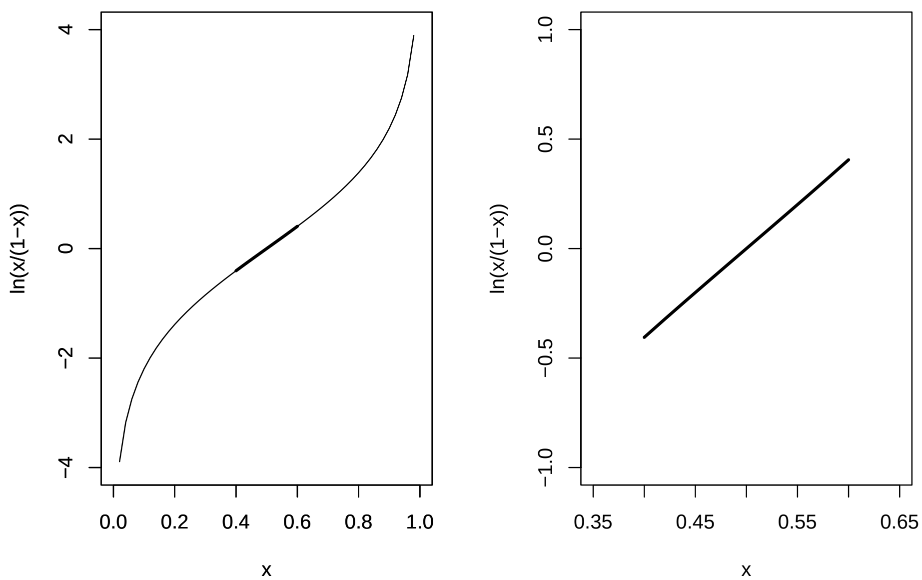

Figure 1: *In two party systems, the log vote ratio reduces to the familiar linear model within the typical range of vote shares. The graph on the left shows the value of the log vote ratio transformation,* $ln(x/(1-x))$*, across the ranges of possible votes for one party,* x*. While this transformation has the shape of a rotated S over the possible values of* x*, in contested elections, even in the heaviest defeats, two party systems rarely return results greater than a 40-60 loss. Within this 40-60 range, as graphed in bold on the left and magnified on the right, the log vote share transformation is very close to linear.*

in that country or cross-sectional unit. These zeroes pose a problem to the additive log ratio (ALR) transformation in 3 as either the numerator or denominator becomes zero, and we can not take the log of zero or infinity respectively. In different settings and disciplines these zeros may have reasonably different interpretations. They may represent a small number that has been rounded down, or a small measure beneath the detection of the data instrument. They may represent an experimental design, such as a response that is not offered to a subsample of a survey, or an treatment path that can not be chosen by a subsample of a clinical trial. Or they may represent a choice on the part of the unit of observation not to engage in a category of conduct, such as a teetotaler who expends none of their budget on alcohol and cigarettes, or, importantly in our focus, a party that receives none of the vote in a district because it was not on the ballot.

Because zeros in the composition occur for different substantive reasons and at different frequencies, a number of solutions and fixes have been proposed across the fields using compositional techniques. These solutions, and their relevance to the analysis of vote shares, are worth considering at this point.

3.1. *Multiple Analyses*

One conceptually simple, but parametrically demanding, approach is to estimate one separate set of equations for *each* compositional pattern. This is the approach to voteshares advocated by Tomz, Tucker and Wittenberg (2002) and implemented in Tucker's (2001) analysis of post-communist elections and Cohen, Noel and Zaller's (2005) study of American presidential primaries. Clearly the complication here is that within any country there may be multiple patterns of contestation. For example in Cohen, Noel and Zaller, every election year has a new set of candidates running in each primary. Thus elections across years (and across parties) are not pooled, and while the analysis is exhaustive and thorough (spanning all primaries from 1972 to 2004) this results in 13 models, each of which is itself a system of equations due to the nature of the Aitchison transformation. Similarly, cross-national studies in the style of Tucker require multiple analyses both *within* countries, if the set of parties changes over time or district, and necessarily *across* countries. The multiple analysis approach is reasonable when the set of patterns of party contestation can be streamlined. However, when the data consists of many different types of elections, all of which we believe are appropriately explained by the same theory, then the point-of-view of this paper is that the necessity of multiple analyses are the problem to be overcome, and not the solution.

3.2. *Zero Replacement*

The original work of Aitchison (1986) briefly suggested that small positive values be added to all zeroes in the composition, so that the transformation is viable. This is realistic when the researcher believes that zeros are simply caused by the variable in the composition being below the detection threshold of the instrument of measure. That is, the researcher believes that all elements of the composition are always present, however, trace quantities are not detectable (such as geology where there are traces of most rock types in any sample)[4]. However, in vote share, almost all zeros are the result of structural rather than measurement reasons. Parties either do not choose to run, or are physically not present

[4]Recent work suggests results may depend on whether the values of the replaced observations retain the proportions of the observed data, and the magnitude of the replaced value can sharply increase or attenuate the predicted flows in the composition.

in that country-election-year. However, both Katz and King (1999) and Tomz Tucker and Wittenberg (2002) suggest some form of zero replacement as a possible simple solution when zeros are rare in the dataset. Fry and Chong (2005) adapt this approach in their study of vote shares in the 2001 Federal election in Australia.

3.3. *Conditional and Zero-added approaches*

In recent work, Aitchison and Kay (2003), have advocated a two stage approach, where the first stage estimates whether an element of the composition is zero or non-zero as a binary variable, and then the second stage estimates the size of the log ratio, conditional on the results of the first stage model. In political science, Adolph (2004) uses this model to analyze the career background of central bank appointees. This first stage requires that there be J sets of equations to estimate the existence of each party in the composition, and then an additional $J-1$ sets of equations to estimate the log vote ratios. Adolph suggests that all parameters be included in all models, thus the total number of parameters estimated can be large if the set of possible parties J is large. Moreover if we were attempting to combine models across countries, the number of parameters to estimate would continue to grow as the dataset expands by increasing the cross-section. Simple computation of this model also requires that there is a reference party that contests in all observations to form the denominator of the log ratio. This approach seems best suited to low numbers of possible elements in the composition, even if each of these may be frequently missing, and there are many different patterns of contestation across observations.

3.4. *Effective Votes*

The model of Katz and King (1999) attempts to address the problem of structural zeros created by parties that do not contest elections, by creating a counterfactual estimate of what the vote share *would have been* if all parties *had contested*. These counter factual voteshares are labeled the "effective votes" and they show how this can be estimated as full-information maximum likelihood (FIML) model. Their model works across time in their example of the series of British elections from 1950 to 1992. Here the Labour and Conservative parties have a strong presence in every election, and the Alliance party more consistently contests elections over time. If, however, we were interested in the voteshares of candidates in Democratic primaries in the United States, the model of KK would work in a given election where not all primary candidates contest all States, but would fail across time since the variety of primary candidates would vary greatly from election to election. Additionally, the Katz and King model requires numerical integration that is not feasible beyond elections with four or five parties. Honaker, Katz and King (2002) expand the "effective vote" approach with EM and missing data methods, however, although this model allows more parties in the composition, it is limited in the number of patterns of contestation it can estimate. For this reason, the "effective vote" model of the Polish elections in 1993 presented in Honaker, Katz and King (2002) could not be expanded *over time* as the number and character of parties has changed dramatically in just the two elections since the 1993 election due to extinction and merging of parties for strategic and institutional reasons.[5,6] An MCMC missing data approach to structural zeros is similarly explored in Martin, Palarea and Gomez (2005), however, it has the same limitations.

4. MODEL

Rather than expanding or developing one of these generic techniques for structural zeroes to allow Aitchison's transformation to be viable, the approach taken here is to consider alternate transformations that exploit our theoretical understanding of voteshares. A simple unidimensional Downsian-styled model might be expressed as follows: assume there is a single policy dimension, x, and some distribution of voter ideal points, $f(x)$, and a set of parties, P. Each party has a position, p_i in this dimension. If there are K parties, let us label them in an ordered fashion for ease of notation, such that:

$$p_0 < p_1 < p_2 < \ldots < p_K < p_{K+1}, \tag{6}$$

where p_0 and p_{K+1} are not party positions, but the bounds of the policy space[7]. If voters quite simply cast their ballot for the party closest to their individual ideal point, then the voteshare for any party can be determined as:

$$v_i = \int_{\frac{p_{(i-1)}+p_i}{2}}^{\frac{p_i+p_{(i+1)}}{2}} f(x)\ \delta x \tag{7}$$

Where again, $f(x)$ is some distribution that describes the density of voters in the space.

In this decision calculus, the locations:

$$c_i = \frac{p_i + p_{(i+1)}}{2} \tag{8}$$

represent cutpoints[8]. Voters who are only some small ϵ apart from each other, but on different sides of these cut-points vote for different candidates. Cutpoint c_i divides voters who choose party i from those who vote for $(i+1)$. Models by Snyder (1994) and Ansolabehere and Snyder (1996) allow these cut points to move in response to unforeseen valance issues that occur during the course of the election, such as scandals, policy failures, or crises.

Following this logic, let there be some number of valance concerns, which have an aggregate, additive weight for each party. The cutpoints are modified by these weights as:

$$c_i(P, V) = \frac{p_i + p_{(i+1)}}{2} + v(i) - v(i+1) \tag{9}$$

Where $v(i)$ is some function summing together the different valance effects concerning party i. Positive valance issues about party i (or $i+1$) move the cut point to the right

[5] We could use the effective vote approach, but we would require a separate analysis of every election year. If we believe that primaries are not fundamentally different from year to year, and each election is contributing to some common understanding, then we need a method to combine these models, or a single model capable of overcoming this problem.

[6] This should not be misconstrued to say that the counterfactual nature of the effective vote is problematic. The effective vote is a useful and pleasing abstraction much like the heaven, or the "effective afterlife". The argument here is that there is a broader class of compositional problems, of importance and interest, to which this model can not extend.

[7] Or more precisely, if there is a lower bound on the policy space, l, then $p_0 \leq (2l - p_1)$, and $p_{(K+1)} \geq (2u - p_K)$ for upperbound u, simply so that all of the policy space is contained within the cutpoints about to be defined.

[8] Excluding those at the edges, that is $\forall\ i : K > i > 0$.

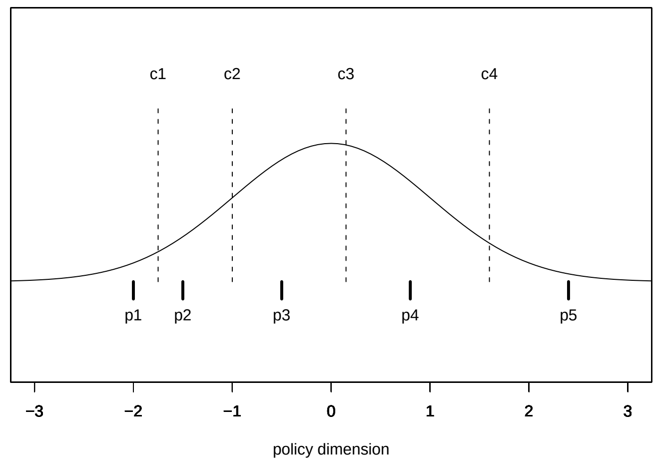

Figure 2: *Cut points in the simple case where voters choose the party closest to their position in the policy dimension. Here the cut points, c, represented here by dashed lines, are simply the points equidistant between the two adjacent parties, p, as in equation 8. The distribution of voters here is a standardized normal.*

(left), thus some voters near the cutpoint switch their vote to party i increasing (reducing) i's vote share. The total magnitude of change to the vote share depends both on the size of $v(i) - v(i+1)$ and the local density of voters, $f(x)$.

Notice, that valance issues for party i will move the location of two cut points, that dividing parties $i-1$ and i and that dividing i and $i+1$. However, all other cut points will be unchanged. In some cases, $v(i)$ may be of sufficient magnitude that a party's location is not within it's own cut points. Thus a voter who perfectly shares that party's policy position would not vote for that party, although there exist more distant voters that would.

Figure 3 demonstrates some intuitions of valance issues in this model. Assume party 2 and party 3 form an incumbent coalition, and incumbency is an issue that contributes to the valance shift of the cut point. The center graph repeats figure 2 and represents the reference case where incumbency has no effect, and voters simply cast their ballot for the closest party. In the top graph, there is an incumbency advantage. *Ceteris paribus*, voters prefer incumbents, and in this example the valance shift created by incumbency, $v(i)$ is 0.5. Thus the option of voting for a party one half unit closer to their ideal point, and voting for an incumbent party, are equally weighted in the decision calculus of voters. Notice that the cutpoint, c_2, between the two incumbent parties does not move. Although incumbency has an advantage, the advantage is equally balanced when voters near this cut point are deciding who to vote for. However, the cutpoints c_1 and c_3 move and increase the voteshares of the incumbent parties. It is important to see the compositional nature of the incumbency effect. Although the cutpoints move the same distance, the voteshare of party 3 is increased more than the voteshare of party 2, because the cutpoint moves over a denser portion of the distribution of voters. The voteshares of parties 2 and 3 increase at the expense of parties 1 and 4, but party 5 does not change its voteshare. If there had been a party located *between* parties 2 and 3, and this party was similarly a member of the incumbent coalition, this hypothetical party's voteshare would not be changed by the existence of an incumbency contribution to the valance term, because both of its neighbouring parties would also have that valance term and the effects would cancel. Thus, although the model assumes the very simple proposition that incumbency has a constant value in the decision calculus of all voters, the effect of incumbency on parties can vary enormously depending on the location of parties, the incumbency status of adjacent parties, and the local density of voters. A constant incumbency advantage in the voter calculus might result in some incumbent parties seeing no increase in their voteshare, and some parties out of the coalition receiving no decrease in their election results, and a whole spectrum of intermediate changes across other parties.

The bottom graph in figure 3 shows the opposite possibility, that incumbency decreases voteshare as voters penalize incumbents. Many of the previously mentioned patterns continue but are reversed. One additional point to note is that the region where voters will vote for party 2 (that is the area between cutpoints c_1 and c_2), does not itself contain the location of party 2. Thus in this model, voters who exactly share the ideal point of a party may not vote for them, although more distant voters may see that party as their best option.

5. ESTIMATION OVER DIFFERENT PARTY SYSTEMS

In the model of the previous section, the compositional interdependence of vote shares is at the forefront. Some parties benefit or lose due to valance issues, and when they win they transfer votes from specific parties. Some parties suffer no gains or losses when adjacent parties that voters might turn to have the same valance values and these variables balance out in the voter calculus.

The driving forces of vote shares are 1) the party positions, 2) the valance attributes of the parties, and 3) the voter density in the policy space. The first two of these pieces drive the bounds of the integral, and the third piece determines the distribution we integrate over. Total voteshare, is expressed as:

$$v_i = \int_{c_{i-1}}^{c_i} f(x)\ \delta x \qquad (10)$$

where c_i represents the cutpoint between party i and $i+1$ and $f(x)$ is the density of voters in the policy space.

5.1.

Consider the following extreme simplification of world. If the following circumstances were met:

1. The data contains elections for only one country.
2. The set of parties is fixed and all parties contest all elections.
3. The locations of the cutpoints, c_i, remained fixed in all elections.
4. The distribution of voters, $f(x)$ was a standardized normal distribution.

then the foresighted reader may have seen that the model in 10 is appropriately reduced to (Grouped) Ordered Probit. The likelihood of observation i is:

$$L(\beta_0,\ \tau|V_i) = \prod_{j=1}^{J} \Big(\Phi(\tau_j - \beta_0) - \Phi(\tau_{j-1} - \beta_0)\Big)^{v_{ij}} \qquad (11)$$

and the maximum likelihood estimates of β_0 estimate the position of the mean voter in the national electorate, while the vector of τ's estimate the cutpoints between each party.

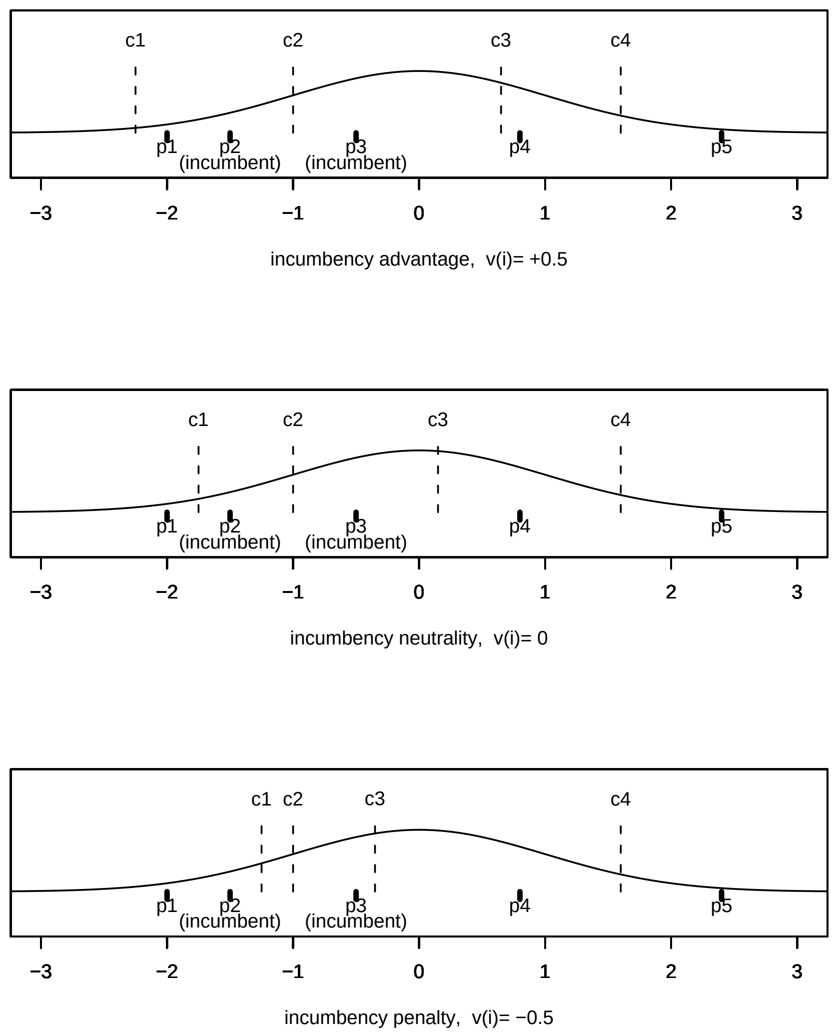

Figure 3: *Here party 2 and party 3 form an incumbent coalition. If incumbency has a positive valance effect, then voters who were just beyond the cutpoint for these two parties now prefer these parties to the party that was closer to them in policy dimension, and the cutpoints move outwards. Notice the cutpoint between party 2 and 3 is unaffected, since the valance issue cancels. Similarly, the cutpoint between party 4 and 5 does not change in any of these example cases.*

Instead of this abstraction, we would certainly like a more versatile model. Let us examine how to overcome each of these four simplifications in turn. First, when we have multiple countries, obviously we will need to estimate separate means for each country, thus we replace β_0 with a vector β of country fixed effects.

One key difference between our model and the standard implementation of ordered probit is that these β's, although necessary to model elections across countries (which have unique voter distributions), are not the coefficients of key interest. Instead, generally all the β's will be regarded as nuisance parameters. Instead, parameters of interest will modify the locations of the cutpoints, thus we need to parameterize the τ's as a function of relevant independent variables.

One method to parameterize the τ's is to decompose the J categories of the dependent variable into $J-1$ *proportional odds* that give the ratio of the dependent variable realized above and below some (interior[9]) τ value (Snell 1964, McCullagh 1980). The logarithm of these odds ratios can be modeled linearly with normal disturbance (or the denominator of this transformation can be modeled with a Logit or Probit).

$$Y^*_{ij} = \ln\left(\frac{\sum_{p=1}^{j} V_{ip}}{\sum_{q=j+1}^{J} V_{iq}}\right) \qquad \text{for } j = 1, \ldots, J-1 \tag{12}$$

Thus for every election, if there are J parties, there are $J-1$ *log cumulative votes*, and we estimate the model:

$$\begin{aligned} Y^*_{ij} &\sim f_{normal}(\theta_{ij}|\sigma_j) \qquad (13)\\ \theta_{ij} &= \beta_i + \gamma(x_j) - \gamma(x_{j+1}) \\ &= \beta_i + \gamma\Big(x_j - x_{j+1}\Big) \qquad (14) \end{aligned}$$

The reverse transformation from log proportional votes back to votes is:

$$V_{ij} = (1 + e^{-Y^*_{ij}})^{-1} - (1 + e^{-Y^*_{i(j-1)}})^{-1} \qquad \forall j, \text{where } Y^*_{i0} \equiv -\infty \tag{15}$$

A graphical heuristic is shown in figure 4 for the difference between these two transformations, the log vote ratio, Y in equation 3, and the log cumulative vote, set out above in equation 16 as Y^*. The top row of boxes represents how one observation of the original compositional dependent variable, V_i, might appear in row form within a dataset. Here there are five parties and thus five boxes in the top row, and five columns of boxes. The elements of V_i that are used in the various log vote ratio transformations, Y_{ij} are shown on the left. The log vote ratio creates four ratios between the values of parties 1 through 4, compared to the reference or baseline category, the vote share of party 5. These four ratios add one observation to each of four different sets of simultaneous equations. The right side of figure 4 shows the elements of the dataset V_i that are used in each log cumulative vote, Y^*_{ij}. Again, there are four transformations created from the original five values of V_i. However, each element of the log proportional vote uses all elements of V_i, and progressively moves elements from the denominator to the numerator, to create the ratio of voters on either side of c_i as i moves here from 1 to 4.

Let us consider now removing the third of the four abstractions in section 5.1. We want to parameterize the location of the cutpoints, across observations. The standard Aitchison

[9]By interior, we mean to exclude the boundary values τ_0 and τ_J which exist simply as a notational convenience, and do not serve as cut points between parties.

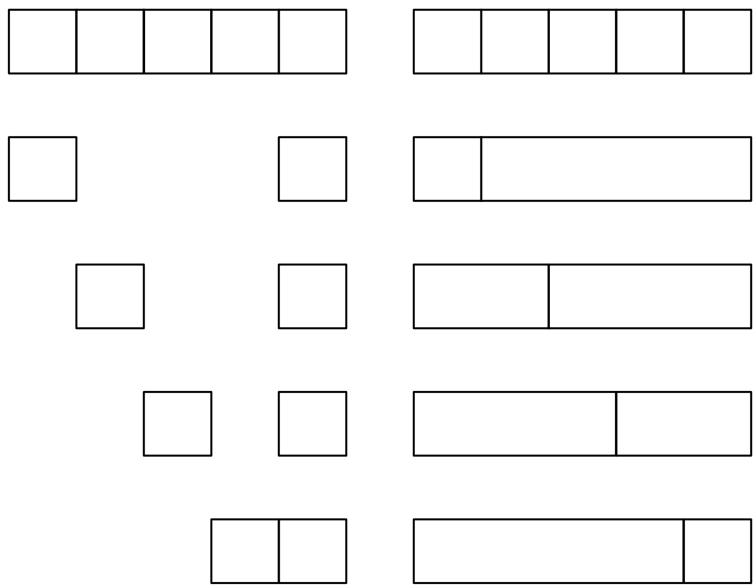

Figure 4: *A graphical representation of the observations used in the Aitchison styled log vote ratio and the log cumulative vote. The top row of boxes represents how one observation with five parties and voteshares might appear in row form within a dataset. On the left, the log vote ratio is represented. It creates four ratios between the values of parties 1 through 4, compared to the reference or baseline category, the vote share of party 5. On the right, the log cumulative vote creates four different ratios of the voteshare to the left and to the right of some cutpoint dividing parties.*

transformation that gives us the log vote ratio allows an event to have different, or indeed opposite effects on different terms of the transformation. Imagine event Q (perhaps a scandal) occurs before an election, and because of Q we expect certain voters to switch allegiance from party 2 to party 3 (in a many party system). For this to be true, the coefficient on the measure of Q must be *negative* for Y_2, that is party 2 decreases its vote share compared to the reference party J, which is unaffected by the scandal. Also, the coefficient on the measure of Q in the equation of Y_3 must be *positive* as party 3 gains, and the ratio of V_3 to V_J increases. As we see, effects in compositional variables transfer votes from one category to another, and this is usually expressed in the Aitchison transformation as opposite coefficients in separate sets of regressions. Clearly then, we can see part of the reason we estimate $J-1$ systems of simultaneous equations in Aitchison's model. If variables transfer votes between parties, the related coefficients need to have opposing signs across equations[10]. We can not pool equations, instead we need one equation for every ratio, and any structural zeros in the dataset will cause problems as previously discussed in section 3.

However, let us consider again event Q that transfers votes from party 2 to party 3, in the context of the log cumulative vote transformation. Here we expect the cutpoint between party 2 and party 3 to shift, while all other cutpoints remain constant. If we expect all Q-events to adjust cutpoints in the same fashion then we expect the coefficients of to be the same in each pool.

The most important reason why we can not pool equations in the Aitchison model is that we do not want to restrict the possible flows over the composition. Without the leverage of theory, the Aitchison model is exceptionally flexible and will allow any variables to move weight in the composition from any element to any other element, or distribution of elements. If we consider the elements of the composition to be nominal categories, and have no prior restrictions on feasible compositional flows, this flexibility suitably allows us to explore how the composition changes in response to our covariates. However, if we are willing to invoke the assumption of the spatial Downsian model, that parties have an ordering along a policy dimension, and the voters we expect may switch votes in response to covariates are those that are near cutpoints, then we can restrict the set of feasible flows. This restriction allows us to use the log cumulative transformation, as the only flows that need to be modeled are the flows between adjacent parties in the ordered policy space. This will require us to have additional information that orders the parties in the policy space, which the Aitchison model would not require us to collect. However, to the extent we believe in policy space representations of vote choice, the very fact that the Aitchison model makes no use of this information should alert us to the idea that we can tailor a model more specific and useful to our dependent variable of vote choice.

While the leverage of incorporating Downsian notions of cutpoints will allow us to restrict the set of flows we need to consider, most important of all we should now realize that we expect the effects of all variables across all cutpoints to be the same. If there are multiple Q-events between different sets of parties, we should expect the cutpoints to move in the same fashion in each party-pair. What is key from this is that all analyses of the cutpoints can be pooled as we expect the location of the cutpoint to move in the same way. Therefore, if we are analyzing the location of cutpoints, which is what the log cumulative vote transformation measures, we can pool all observations, since all cutpoint locations respond to the same variables under the same theory, and all coefficients on the same variables can be restricted to be equal, regardless of which cutpoint we are considering. Since we can pool all observations, structural zeros, and the set of parties that make up the

[10]Unless all variables always transfer from parties $1 \ldots (J-1)$ to party J or vice-versa.

composition do not impede our analysis. If the set of parties contesting elections changes, across districts, across time, or even across countries, the cutpoint that exist within any particular election are still valid observations of the behavior of cutpoints between adjacent parties. Shifting the analysis to this unit of observation, if a party drops out, the cutpoints that result will be between new parties, which will have different independent variables (in our parlance, valance characteristics $v(i)$ and $v(i+1)$) but the cutpoint is still equally valid in this population of observations. A model that predicts the location of cutpoints between adjacent parties will not depend structurally on the parties remaining constant.

The last of the four abstractions in section 5.1 to overcome is that the voter distributions are stylized normal. Clearly, if every country can have different means to their voter distribution, the variances may also differ cross-nationally. To parameterize this we would need to estimate individual variances for every country or cross-sectional unit. It may often be that we have additional information, perhaps from surveys of voter attitudes, from which we can construct an empirical distribution of the voters across the policy space. The previously mentioned paper of Cohen, Noel and Zaller (2005) do precisely this from ANES data in the construction of some of their independent variables. We might use empirical measures of the voter distribution to give us an alternate function to integrate in our model, or to fix the relative variances between cross-sectional units.

5.3.

Incorporating all these changes together gives us the following. For a dataset of voteshares, V, in country h, election i, for party j, we can create a log cumulative vote:

$$Y^*_{hij} = \ln\left(\frac{\sum_{p=1}^{j} V_{ip}}{\sum_{q=j+1}^{J} V_{iq}}\right) \qquad \text{for } j = 1, \ldots, J-1 \tag{16}$$

which is transformation that takes the logarithm of the ratio of votes on either side of cutpoint i. We estimate this as:

$$Y^*_{hij} \sim f_{\text{Normal}}(\theta_{hij}, \sigma) \tag{17}$$

$$\theta_{hij} = X\beta + \Big\{C_{hij}\delta + \gamma\Big(V(j) - V(j+1)\Big)\Big\}\Big(X\alpha\Big) \tag{18}$$

Where X is a matrix of country dummies, $V(j)$ is a matrix of characteristics of party j, and C is an (optional) additional set of variables that estimate the position of the cutpoint between parties i and $(i+1)$, such as an estimate of the midpoint between the party locations, or the lagged value of Y^*. If we are willing to assume that all cross-sections have voter distributions with the same variance, we can remove the term, $X\alpha$.

This transformation and estimation is similar in style to Aitchison's log vote ratio. We transform J voteshares into $J-1$ terms, each of which is the logarithm of a ratio, and the transformed variables may be estimated with linear regression. However, this transformation does not have a common baseline party, and restricts flow in composition to be between adjacent parties by movement of the cutpoint over the distribution of voters. The most important difference, though, is that all observations of Y^* are pooled into one model.

One last side note to observe is that in the two-party case, with parties receiving votes shares x and $(1-x)$, equation 16 reduces to:

$$Y^*_{ij} = \ln\left(\frac{x}{1-x}\right) \tag{19}$$

Which has near-linear properties within the range of vote shares typically seen in two party systems, as previously detailed in figure 1. Thus this model reduces to the linear link typically used in two-party systems such as researchers in American politics.

6. THE ELECTORAL LIABILITY OF INCUMBENCY IN PARLIAMENTARY DEMOCRACIES

The objective of the following application is to examine whether or not, in parliamentary democracies, there exists an electoral cost to governing for political parties that constitute the incumbent government. Should political parties count future electoral losses among the potential costs of participating in the government? Are parties who remain in opposition "rewarded" at the polls in an ensuing election? And if there is a systematic difference between government and opposition parties in the change in vote shares received from one election to the next, then what is its magnitude, and under what conditions should it be expected to obtain?

Resolving these questions has important implications for the study of parliamentary government formation; in particular what Strøm (1984; 1990) terms the "cost-benefit" electoral calculus made by political parties when deciding whether or not to pursue joining the government. Existing cross-national studies have consistently shown that incumbent governments *in aggregate* tend to lose vote shares from one election to the next (Rose and Mackie 1983; Strøm 1990; Powell and Whitten 1993; Paldam and Skott 1995; Paldam and Nannestad 1999; Powell 2000; Stevenson 2002), but no study has ever examined the effect of incumbency on the vote shares of *individual* parties that constitute the government. As the decision to join the government is made on a by-party basis, this seems a fairly significant omission.[11] Observing aggregate vote losses by incumbent parties may inform our understanding of voter behavior and mechanisms of voter accountability (Powell 2000), but it tells us almost nothing about factors influencing the decision-making process of individual parties.

Consider, for example, Strøm's (1984: 212) well-known claim that because "government incumbency tends to result in subsequent electoral losses," there are oftentimes electoral incentives for parties to "remain in opposition temporarily... [and] wait for more favorable circumstances."[12] But to support this theory requires evidence *not* that incumbency has a detrimental effect on *aggregate* incumbent vote shares, but rather that *ceteris paribus*, joining the government depresses *individual party* vote shares at the ensuing election.

It is certainly the case that if on average, incumbent coalitions lose votes in the election following incumbency, then on average incumbent parties lose votes as well.[13] However, in practice, changes in party vote shares tend to vary widely around the (negative) mean value, and simply averaging across incumbent parties' vote shares hides patterns of vote gains and losses among what may be multiple parties *within* the same governing coalition. Indeed, it is not the case that because incumbent governments tend to lose votes in aggregate, all incumbent parties lose votes individually. Rose and Mackie (1983) have demonstrated that in only approximately one-third of coalition governments do all incum-

[11] Laver and Schofield (1990), in their exhaustive review of coalition formation theories and empirical evidence, without fail describe this choice as belonging to individual political parties. See chapter 5.

[12] In his 1990 book-length treatment of the subject, Strøm elaborates upon these ideas in greater detail, noting that "government incumbency typically represents an electoral disadvantage." See in particular pp. 45-52.

[13] To see this, consider any set of multi-party governments comprised of a mean number of parties, $\bar{p}$. Then simply divide the average coalition's change in vote share—which we know to be negative—by $\bar{p}$, to find that the average party's change in vote share will also be negative.

bent parties lose votes in a subsequent election. In nearly two-thirds of coalitions, some incumbent parties lose votes while others actually *increase* their vote share. Analysis at the party level enables us to investigate various factors that may explain this apparent variation in the effect of incumbency on changes in party vote share. Most importantly, it allows us to see if, once these factors have been controlled for, incumbency retains any systematic effect on vote shares at all.

If indeed there is *no* systematic difference between electoral prospects for incumbent and opposition parties, then it would be erroneous to infer that parties factor potential electoral "costs" of governing into their decision to join the government. Only if parties expected this incumbency "penalty" to obtain might they decide that their long-term policy-making goals were better served by waiting in opposition for a larger electoral vote share—and, presumably, a larger legislative seat share[14]—in a future election.

The next sections reviews two theories of why incumbent parties should be expected to lose votes in elections following their terms in office, and discuss the possible limitations and complementarities of these two theories. We then briefly explore some quantitative differences between patterns of vote losses for incumbents and opposition parties, and show how previous studies differ from the present work by failing to disaggregate incumbent and opposition vote shares into their constituent party units. Next, we model patterns of changes in incumbent and opposition party vote shares, to test the theories of vote losses and specify the conditions under which incumbent parties should expect to lose votes. These conditions are then shown, in addition, to predict incumbent party seat losses as well as vote losses. A discussion of the implications of these findings for the decision-making process of political parties concludes.

6.1. *Theories of Incumbent Party Vote Loss*

Why should incumbent parties expect to lose votes in a future election? Are these costs of governing avoidable under some conditions, or are they, on average, inevitable? There are two broad families of hypotheses that attempt to explain incumbent party vote losses that may be loosely categorized as *incumbency-based* and *size-based*.

In the first group are theories that imply incumbency *itself* leads to vote losses in a subsequent election, via a mechanism of retrospective voter accountability of poorly performing incumbents. Strøm (1990: 45) notes broadly that "there are several reasons why incumbency should be an electoral liability, all of them relying on some assumption of retrospective voting".[15] More strongly, Paldam and Skott (1995) and Stevenson (2002) have advanced a "median gap" model of incumbent vote loss whereby "voters who prefer policies that lie between the positions of the parties will want an alternation in power so that the policies that obtain will, on average, be closer to their positions than if a there were no alternation in government" (Stevenson 2002: 157). Hypotheses in the economic voting literature argue that voters condition future electoral support for incumbents on national-level economic conditions (Lewis-Beck and Eulau 1985; Lewis-Beck 1988; Lewis-Beck and Stegmaier 2000; Anderson 2000), such that incumbents are held accountable (i.e., punished) for poor economic outcomes—but may also be positively rewarded when the economy performs well.[16] Finally, the idea that voters retrospectively differentiate

[14]The correlation between party vote share and seat share has varying degrees, and may not be linear depending upon the election rule. See Taagepera and Shugart (1991) on the translation of votes into seats in the legislature.

[15]See also Rose and Mackie (1983) for an excellent, succinct review of the many reasons why retrospective accountability should produce incumbent party vote losses.

[16]See especially Lewis-Beck (1988: 77-79) regarding the "symmetry" of economic effects: "electorates

politician performance on any number of economic or policy factors, and retain an implied electoral threat to sanction representatives when they act in a manner contrary to constituent interests, also underlies a diverse array of formal models of democracy.[17]

In the second group, attention is focused on short-term exogenous factors that influence election outcomes but have nothing to do with incumbency or accountability *per se*. These theories view parties' popular vote shares as the sum of a long-term underlying "baseline" level of support in the electorate, plus or minus a short-term "surge" component that lasts only through the current election. For each party, this short-term "surge" in vote fluctuates randomly between elections, with mean and mode zero (Converse 1966).[18] To explain incumbent party vote loss, then, short-term effects theories note correctly that incumbent governments are comprised of unusually large parties that enjoyed particularly large "surges" in the election preceding their ascension to government, and speculate that the vote losses that befall government parties arise because of the contraction or "backswing" of the short-term "surge" in a subsequent election (Paldam 1986). Similarly, parties that do not make it into government because of their small size appear to "earn" votes because their short-term *negative* "surge" also recedes in the following election. The key to this up-and-down pattern of fluctuations is the improbability of obtaining consecutive large or small random vote "surges".[19] Size effects should obtain for all parties, regardless of their government status. The implication is that large *opposition* parties should also tend to lose votes, and small *incumbent* parties should tend to gain votes. So, incumbency itself may not be responsible for incumbent vote loss; rather, by selectively observing incumbent parties—which tend to be large and fortunate in a previous election—it appears that incumbency possesses an electoral cost when in fact short-term size effects are doing the work.

Incumbency and size effects are by no means necessarily mutually exclusive. Rather, how well incumbents respond to the exogenous "random" factors that arise during their tenure shapes the degree to which they will be held accountable in a future election. Consider that for example, in Germany's September 2002 parliamentary election, the incumbent Social Democratic Party (SDP) lost 2.4% of its vote from the previous 1998 election (down to 38.5% from 40.9%)—but would have likely fared even worse if not for Chancellor Gerhard Schröder's effective response to floods in the east and his popular stance against war in Iraq (*Economist* 2002: 45). In the event, the Social Democrats retained office, but only because of the success of their coalition partners, the much smaller Green Party, which—as it happened—increased its vote share by 1.9%, from 6.7% in 1998 to 8.6% in 2002. Perhaps there was no size-effect "swingback" and the SDP simply lost votes because of its unpopularity prior to the floods and the Iraq issue. Or perhaps the SDP was regressing to its baseline level of support, following the cumulative set of events that amounted to a smaller vote "surge". But it could also be that the SDP had overachieved in the 1998 election *and* was due for an electoral "swingback", but was able to mitigate its losses by governing effectively. The essential unanswered question is the extent to which

are 'even-handed' in their economic judgments, voting for governments that are liked, against governments that are disliked."

[17] These include Downs (1957), Persson, Roland and Tabellini (1997), and McGillivray and Smith (2000), to mention only a few.

[18] To be precise, the claim is not that any particular election outcome is "random". Rather, it is that elections measure "true" voter preferences on election day, with the "surge" component of the vote being the sum of deliberate, "rational" voter decisions based on "random" exogenous factors that vary between elections and are not, over time, systematically biased in favor of any particular party or parties.

[19] Formally, this pattern is known in statistics as regression to the mean, and arises whenever two variables are imperfectly correlated (as are consecutive party vote shares), and when sampling is biased (as by only observing large incumbent parties).

incumbency and size effects *each independently* help (or do not help) explain long-term patterns of gains and losses in party vote shares between incumbent and opposition parties.

There are a number of reasons, before continuing further, why we might reasonably be skeptical that either or both of these two theories are incorrect. Theories of short-term effects have not been widely criticized in the literature, and, if anything, have been more neglected than subjected to rigorous tests.[20] Two cross-national comparative studies stand out. Powell and Whitten (1993) demonstrate that both the "previous government vote percentage" and the "previous government vote swing" are statistically significant predictors of future change in vote for governing coalitions in aggregate, in the predicted negative direction. Paldam (1986: 14) also argues empirically in favor of a "small swing-back effect," but cautions that the correlation between the change in aggregate vote share for incumbent governments prior to and following incumbency is extremely weak compared to "a genuine cost of ruling" that is independent of the previous change. Paldam does not analyze the determinants of this "cost of ruing" at the party level, nor does he test whether the size of the incumbent bloc is a significant predictor of change in vote, as the short-term effects hypothesis predicts.

Criticisms of theories underlying incumbency effects are both more widespread and more severe. As described above, the supposed mechanism behind incumbent party vote loss is one of retrospective voter accountability. Yet the current literature on electoral accountability is increasingly questioning the viability of the electoral link to sustain politician compliance with voters' interests.[21] Even Strøm (1990: 46), who believes that retrospective accountability explains incumbent party vote loss, admits to no "definitive conclusion as to precisely why incumbency is an electoral liability."

The expectation that voters use elections to monitor and sanction legislators assumes that voters both have sufficient knowledge of their representatives' actions, and furthermore exercise their dissatisfaction at the polls in a manner that affects the composition of the legislature. Yet detection breaks down when voters are unable to discern unfavorable behavior on the part of elected representatives; Powell and Whitten (1993) and Powell (2000) both argue convincingly that government vote losses decline as responsibility for policymaking is diffused in voters' minds through lower levels of political party cohesion, shared committee chairs, and strong opposition parties. With respect to economic voting, Anderson (2000: 168) writes, "Political context interacts with economic perceptions to affect voting behavior.... Voters' ability to express discontent with economic performance is enhanced when accountability is simple." Poor economic conditions may provoke a desire by voters to vote against incumbents, but unless voters can "identify who is in charge, how much responsibility they have, and what the alternatives are," then accountability will be severely constrained. And as Maravall (1999: 192) makes clear, politicians possess a wide range of rhetorical and procedural options to "manipulate to their advantage the problems of information, monitoring, and commitment of citizens when assessing whether the incumbent is pursuing their interest."

The sanctioning component may also break down if voters desire but are effectively unable to punish legislators for poor performance in office. Institutional features such as party rules for candidate nomination and formal election laws—including voting rules,

[20] Indeed, most short-term effects theories are found in the American politics literature on the pattern of congressional midterm vote losses by the president's party, and include, most prominently, the surge-and-decline hypothesis of Campbell (1966) and the political-business cycle hypothesis of Alesina and Rosenthal (1995). But see also Campbell (1985), Erikson (1988), Hinckley (1967), and Scheve and Tomz (1999).

[21] Certainly, not all new research is "against" accountability. See Canes-Wrone, Brady, and Cogan (2002) and Erikson, Mackuen, and Stimson (2002) for compelling evidence of accountability at work in elections in the United States.

formulas for translating vote shares into legislative seat shares, time between elections, district magnitudes, etc.—can diminish voters' power by placing hurdles along the electoral link to representatives. Mitchell (2000: 346) concludes that "there exists a trade-off that is crucial to voters' abilities to sanction agents between the directness of the link between voter and agent and the choice of agents." Again, even if parties "deserve" to lose votes because of their (poor) performance in office, there is no guarantee that they will actually be held accountable at the polls for their actions.

In summary, to the extent that incumbency effects fall short, it is quite theoretically possible to explain observed patterns of incumbent party vote losses through size effects—and vice-versa. It remains to test empirically whether either or both of the hypothesized effects exist, and quantify the magnitude of each of the effects, once the other has been systematically controlled for. The incumbency hypothesis predicts that incumbency itself will carry an electoral cost. If it is correct, we should expect our model to predict that regardless of previous vote share, a party will lose votes in a future election if it is an incumbent and gain votes if it is in opposition. The short-term effects hypothesis predicts that party size will carry a penalty, so that large parties lose votes and small parties gain votes, regardless of incumbency status. Short-term effects also imply that "surges" in the share of votes received in a previous election should dissipate in a latter election. If this is correct, our model should show that the posited negative effects of incumbency are mitigated for small parties or parties experiencing negative surges in a previous election, and exacerbated for large parties or parties experiencing positive surges in a previous election.

6.2. *Do Incumbents Lose Votes?*

On average, the *aggregate* vote share of *all parties* in incumbent governments declines from one election to the next over the course of their administration. Paldam and Nannestad (1999: 21) observe that indeed, "the cost of ruling is almost *constant* in all stable developed democracies. The constancy applies both across countries and over time." The empirical evidence is fairly straightforward. Strøm (1990: 124) calculates the mean aggregate vote share loss among all parties in government to be 3.15%, with standard deviation 6.39%, in 327 governments. Assuming that the aggregate change in vote share for governing coalitions is distributed normally, then we may estimate from these figures that approximately 69% of incumbent governments lose votes in the election after their incumbency.[22] Using a more current dataset, Powell (2000: 53) observes that: "incumbent governments generally lose votes; it seems easier to be blamed for failures in office than to exploit successes." By his count, approximately 31% of incumbent government parties lose, in aggregate, over 5% of their vote from the previous election. Another 44% lose between 1% and 5% of their vote, and only 25% experience either no change or an increase in their collective vote shares. Paldam (1986) and Rose and Mackie (1983) arrive at similar results. Considering these studies, then, approximately two-thirds to three-quarters of all incumbent governments lose votes in subsequent elections.

These analyses are basic enough—but they do not permit us to analyze the electoral fortunes of the individual parties of which incumbent governments are composed. Because previous researchers have been only interested in patterns in the aggregate vote shares of incumbent parties, and because votes gained and lost by incumbent and opposition blocs sum to zero, it was sufficient for these studies to confine their analyses to changes in the total share of votes received by all incumbent parties. However, in examining *individual*

[22] Calculated using standard normal probabilities: $z_0 = \frac{3.15}{6.39} = 0.49$. $Pr(z < z_0) = 0.69$.

parties as the unit of analysis, it is no longer the case that incumbents' lost votes are necessarily transferred to opposition parties, and vice versa. It is also possible for voters to switch their vote from one incumbent party to another, or from one opposition party to another (Rose and Mackie 1983). Anderson (1995) has shown that this is precisely what happens, for example, in elections in Denmark and the Netherlands. When we decompose government and opposition vote shares into their constituent party shares, we must also broaden our sample universe. Parties in opposition cannot be excluded from an unbiased sample; failure to consider the patterns of vote gains and losses among parties that are *not* already in government would lead to selection bias. Most simply put, it is impossible to compare patterns of votes gains and losses between incumbent and opposition parties if the sample universe contains only incumbent parties.

6.3. *Results*

Please see the most recent version of this paper for the results of the incumbency application.

REFERENCES

Adolph, Christopher. 2004. "Succession in the Temple: Central Banker Careers and the Politics of Appointment." *Working Paper.*

Aitchison, J. 1986. *The Statistical Analysis of Compositional Data.* London: Chapman and Hall.

Aitchison, J. and J. W. Kay. 2003. "Possible Solutions in Some Essential Zero Problems in Compositional Data Analysis." Working Paper, presented at CoDaWorks03.

Ansolabehere, Stephen and James M. Snyder, Jr. 1996. "Party Platform Choice in Single-Member-District and Party-List Systems." *Unpublished Working Paper.*

Campbell, Angus. 1966. "Surge and Decline: A Study of Electoral Change," in Angus Campbell, Philip E. Converse, Warren E. Miller, and Donald E. Stokes, eds. *Elections and the Political Order.* New York: John Wiley and Sons, Inc.

Campbell, James E. 1985. "Explaining Presidential Losses in Midterm Congressional Elections." *Journal of Politics.* 47(4): 1140-1157.

Canes-Wrone, Brandice, David W. Brady, and John F. Cogan. 2002. "Out of Step, Out of Office: Electoral Accountability and House Members' Voting." *American Political Science Review.* 96(1): 127-140.

Cheibub, José Antonio, and Adam Przeworski. 1999. "Democracy, Elections, and Accountability for Economic Outcomes," in Adam Przeworski, Susan C. Stokes, and Bernard Manin, eds. *Democracy, Accountability, and Representation.* Cambridge: Cambridge University Press.

Converse, Philip E. 1966. "The Concept of a Normal Vote," in Angus Campbell, Philip E. Converse, Warren E. Miller, and Donald E. Stokes, eds. *Elections and the Political Order.* New York: John Wiley and Sons, Inc.

Cox, Gary W. 1997. *Making Votes Count; Strategic Coordination in the World's Electoral Systems.* Cambridge: Cambridge University Press.

Downs, Anthony. 1957. *An Economic Theory of Democracy.* New York: Harper and Row.

Economist. 2002. "Germany's general election: Gerhard Schröder clings on." 28 September. Pp. 45-46.

Honaker, James, Jonathan Katz and Gary King. 2002. "A Fast, Easy, and Efficient Estimator for Multiparty Electoral Data with Improved Lemon Scent." *Political Analysis* 10(1): 84-100.

Jackson, John E. 2002. "A Seemingly Unrelated Regression Model for Analyzing Multiparty Elections." *Political Analysis* 10(1): 49-65.

Katz, Jonathan and Gary King. 1999. "A Statistical Model for Multiparty Electoral Data." *American Political Science Review* 93(1)(March): 15-32.

King, Gary, James Honaker, Anne Joseph, and Kenneth Scheve. 2001. "Analyzing Incomplete Political Science Data: An Alternative Algorithm for Multiple Imputation." *American Political Science Review* 95(1)(March): 49-69.

Lange, Kenneth L., Roderick J. A. Little, and Jeremy M. G. Taylor. 1989. "Robust Statistical Modeling Using the t Distribution." *Journal of the American Statistical Association* 84, 408: 881-896.

McCullagh, P. 1980. "Regression models for Ordinal Data (with Discussion)." *Journal of the Royal Statistical Society* B,42,:109-149.

McLachlan, Geoffrey and Thiriyambakam Krishnan. 1996. *The EM Algorithm and Extensions.* New York: Wiley.

Mackie, Thomas T. and Richard Rose. 1991. *The International Almanac of Electoral History.* London: Palgrave Macmillan.

Paldam, Martin. 1986. "The Distribution of Election Results and the Two Explanations of the Cost of Ruling." *European Journal of Political Economy.* 2(1): 5-24.

Paldam, Martin and Peter Nannestad. 1999. "The cost of ruling: A foundation stone for two theories." Presented at the workshop on "Institutions as intermediaries between economics and politics?" Panel 12. European Consortium for Political Research Joint Session of Workshops. Mannheim, 1631 March.

Paldam, Martin and Peter Skott. 1995. "A rational voter explanation of the cost of ruling." *Public Choice.* 83: 159172.

Persson, Torsten, Gerard Roland, and Guido Tabellini. 1997. "Separation of Powers and Political Accountability." *Quarterly Journal of Economics.* 112(November): 1163-1202.

Powell, G. Bingham, Jr. 2000. *Elections as Instruments of Democracy.* New Haven: Yale University Press.

Powell, G. Bingham, Jr. 1982. *Contemporary Democracies; Participation, Stability, and Violence.* Cambridge: Harvard University Press.

Powell, G. Bingham, Jr., and Guy D. Whitten. 1993. "A Cross-National Analysis of Economic Voting: Taking Account of the Political Context." *American Journal of Political Science.* 37(2): 391-414.

Przeworski, Adam, Michael E. Alvarez, Jose Antonio Cheibub, and Fernando Limongi. 2000. *Democracy and Development; Political Institutions and Well-Being in the World, 1950-1990.* Cambridge: Cambridge University Press.

Przeworski, Adam, Susan C. Stokes, and Bernard Manin, eds. 1999. *Democracy, Accountability, and Representation.* Cambridge: Cambridge University Press.

Rose, Richard and Thomas Mackie. 1983. "Incumbency in Government: Asset or Liability?" in Hans Daalder and Peter Mair, eds. *Western European Party Systems:*

Continuity and Change. Beverly Hills: Sage.

Schafer, Joseph L. 1997. *Analysis of Incomplete Multivariate Data.* New York: Chapman and Hall.

Scheve, Kenneth and Michael Tomz. 1999. "Electoral Surprise and the Midterm Loss in US Congressional Elections." *British Journal of Political Science*. 29: 507-521.

Snell, E. J. 1964. "A scaling procedure for ordered categorical data. *Biometrics*, 20:592-607.

Snyder, James M., Jr. 1994. "Safe Seats, Marginal Seats, and the Logic of Party Platform Differentiation." *Economics and Politics* 6:201-213.

Stevenson, Randolph T. 2002. "The cost of ruling, cabinet duration, and the 'median-gap' model." *Public Choice*. 113: 157178.

Strøm, Kaare. 1990. *Minority Government and Majority Rule*. Cambridge: Cambridge University Press.

Strøm, Kaare. 1984. "Minority Governments in Parliamentary Democracies; The Rationality of Nonwinning Cabinet Solutions." *Comparative Political Studies*. 17(2): 199-227.

Tomz, Michael, Joshua A. Tucker and Jason Wittenberg. 2002. "An Easy and Accurate Regression Model for Multiparty Electoral Data." *Political Analysis* 10(1): 66-83.

Tucker, Joshua A. 2001. "Economic conditions and the vote for incumbent parties in Russia, Poland, Hungary, Slovakia, and the Czech Republic from 1990 to 1996." *Post-Soviet Affairs* 17(4): 309-331.