Bias and Responsiveness in
Multiparty and Multigroup Representation*

Paper prepared for presentation to the
Political Methodology Summer Meeting
July 23-26, 1998
University of California - San Diego

Burt L. Monroe
Indiana University, Department of Political Science
Woodburn Hall 210, Bloomington, IN 47405-6001
812/855-9326
bmonroe@indiana.edu

Abstract

There is an extensive and expanding literature that examines methods for estimating the responsiveness and partisan bias of two-party electoral systems. Attempts to extend these methods into the multiparty domain appropriate for the vast majority of electoral systems, or to the analysis of the representation of other types of groups (e.g., regions, ethnic groups), have been limited. I describe index, multiyear, uniform swing, and variable swing methods -- along with novel graphical displays -- for analyzing seats-votes curves, bias, and responsiveness in multiparty systems. The variable swing method is a multiparty generalization of Gelman and King's "JudgeIt" model. Examples discussed include elections in the UK, Mauritius, and Costa Rica, and geographic representation worldwide. In comparing the various methods it is argued that variable swing is ideal for most applications, that uniform swing and index methods provide useful answers to a limited set of questions despite faulty assumptions, and that multiyear methods are generally not useful.

I would like to thank Patrick Brandt, Gary Cox, and John Williams for comments and guidance in this research, Gary Hetland and Anne Ruach for their assistance in preparing data sets, and Dick Katz, Gary King, and Jonathan Katz for providing data from their own research. This project is in many ways intermingled with work being pursued, both independently and jointly, as well as both by and with, Jonathan Katz and Gary King; their assistance and comments are appreciated. For research support, I would finally like to thank the National Science Foundation (Grant #SBR 91-13328), Harvard University, the International Foundation for Electoral Systems, as well as the Department of Political Science and the Workshop in Political Theory and Policy Analysis at Indiana University.

Bias and Responsiveness in Multiparty and Multigroup Representation

Burt L. Monroe, Indiana University

There is an extensive and expanding literature that examines methods for estimating the *responsiveness* (or *majoritarianism* or *subproportionality*) and *partisan bias* of two-party electoral systems. "Responsiveness" is the tendency for shifts in votes to be translated into shifts in legislative seats; when an electoral system is highly responsive (as are most plurality electoral systems, for instance), a small shift in vote shares leads to a large shift in seats shares.[1] Partisan bias is simply the party-specific deviations from the general pattern of responsiveness. Responsiveness is arguably *the* defining feature that distinguishes equally "fair" electoral systems, and a reasonable method of measuring it is of great importance to comparative electoral scholars. Partisan bias, which can emerge from several sources including districting and apportionment processes, is arguably *the* measure of electoral fairness; as such, its measurement is also of great importance.

Attempts to extend two-party methods into the multiparty domain appropriate for the vast majority of electoral systems, or to the analysis of the representation of other types of groups (e.g., regions, ethnic

---

[1] "Responsiveness" has an unfortunate positive normative connotation, as no one would defend an "unresponsive" electoral system. A more descriptive set of terms is that originated by Dixon (1968), popularized by Lijphart (1984), and used partly in this context by King and Browning (1987): "majoritarianism", "proportionalism", and "consensualism". "Majoritarianism" captures the idea of the large being overrepresented relative to the small, especially in the limit where the largest party takes all seats, the key feature of a highly "responsive" electoral system. "Consensualism", on the other hand, describes the opposite situation where the small are overrepresented relative to the large; in the limit of pure consensualism all parties receive equal seats regardless of vote share. These also provide a central point, "proportionalism", at which seat shares exactly match vote shares. Of course, these also have normative connotations of their own.

The most neutral language offered to date is probably that of "subproportionality" (advantage to large) and "superproportionality" (advantage to small), appearing in Christenson and Johnson (1995) and Cox (1996), but given the potential problems there as well ("sub" sounds bad; "super" sounds good), these confusing terms are less appealing to me. The term "disproportionality" is often used as well, but since majoritarianism and consensualism are different types of deviation from proportionality, and we would like to distinguish them, that term is not useful here. Finally, many authors have used the term "swing ratio", the change in seats for a unit change in votes, a quantity that is equivalent to responsiveness as measured in many two-party methods. I do not use the term here, but the basic concept — responsiveness as the slope of the seats-votes curve — remains the same.

groups, gender groups), range from the unsatisfactory to the unacceptable. Most problems stem from the drawing of inappropriate analogies between the canonical Democratic-Republican electoral system analyses and the problem at hand. While some analogies from the two-party literature are quite useful, many others are not. The inherent multidimensionality of this set of problems has limited our ability not only to develop methods, but to understand what the methods we have are telling us and to recognize when they are making no sense. I introduce a new graphical technique here for visualizing multiparty seats-votes curves, which I hope will help eliminate such problems.

The methods to be discussed in this paper can be classified, in rough order of sophistication, as *index*, *multiyear, uniform swing,* and *variable swing* methods. I treat each of these in turn, in each case briefly describing methods for two-party systems and then multiparty extensions. The variable swing discussion is an extended one, as that is by far the most important of these methods and its description here is new. Throughout these discussions, application of the methods to British general elections serves as a running example. Further examples (from elections in Mauritius and Costa Rica[2]) illustrate the range of applicability of the variable swing method and the new seats-votes graphics. Finally, examples from the measurement of geographic representation illustrate one area in which simple index methods remain the most useful. Multiyear methods (including the method with the greatest amount of scholarly acceptance to date, that based on King's (1990) "Multinomial Bilogit" model) are shown to provide highly misleading results for all of these applications.

Index Methods

Most of the early entries in the bias and responsiveness literature can be classified as discussing index methods. The first of these asserted, or examined qualitatively the assertion, that a *cube law* was an appropriate descriptor of the seats-votes relationship for the two major parties in Britain, or elsewhere (e.g., Smith 1909[3], crediting McMahon; *The Economist* 1950; Kendall and Stuart 1950, 1952; McHenry 1955; Rydon 1957; Butler 1963; Qualter 1968; March 1957-8; Brookes 1959; Taagepera 1973; Theil 1970). The cube law is the relationship

$$\frac{s_{2t}}{s_{1t}} = \left(\frac{v_{2t}}{v_{1t}}\right)^3$$

---

[2] This version of the paper does not yet include the Costa Rica example.

[3] An occasional first cite is to Edgeworth 1898, due to a citation in Kendall and Stuart 1950. Kendall and Stuart were only pointing out Edgeworth's comment that vote outcomes were distributed approximately normally, the basis, they argue, for the "cubic" law. Edgeworth did not discuss the cube law or anything like it.

$$s_{2t} = \frac{v_{2t}^{\ 3}}{v_{2t}^{\ 3} + v_{1t}^{\ 3}} = \frac{v_{2t}^{\ 3}}{(1 - v_{1t})^3 + v_{1t}^{\ 3}} \tag{1}$$

where s_{1t} is the (national / jurisdictional) share of seats for party 1 in election t and v_{1t} is its vote share.

This is easily extended to a more general power law, where the exponent is an unknown to be estimated, or even further to a system with J parties. Equivalent formulations of the power law include:

$$s_{jt} = \frac{v_{jt}^{\ \rho}}{\sum_{j=1}^{J} v_{1t}^{\ \rho}}, \tag{2}$$

$$s_{jt} \propto v_{jt}^{\ \rho}, \quad \text{and} \tag{3}$$

$$\rho = \frac{\ln(s_{it}/s_{jt})}{\ln(v_{it}/v_{jt})} = \frac{\psi_{ijt}}{y_{ijt}}, \tag{4}$$

where $y_{ijt} \equiv \ln(v_{it}/v_{jt})$ and $\psi_{ijt} \equiv \ln(s_{it}/s_{jt})$ for any two parties i and j. The logistic transformation is so useful that I introduce here the notation for the logratios of votes and seats, y and ψ, which I use throughout this discussion. When the "j" subscript is suppressed, party one is the reference party (j = 1).

Estimates of ρ can then be found by a variety of methods. A natural two-party index, for instance, is

$$\hat{\rho}_t = \frac{\ln(s_{2t}/(1 - s_{2t}))}{\ln(v_{2t}/(1 - v_{2t}))} = \frac{\psi_{2t}}{y_{2t}}, \tag{5}$$

which can then be charted over time or averaged for a pooled estimate (Eldersveld 1951; Brookes 1954; March 1957-8; Qualter 1968; Theil 1969; Casstevens and Morris 1972; Taagepera 1973, 1986; Laakso 1979).[4]

In a multiparty system, one party can be designated as a referent (party one) and then averages, perhaps weighted by vote share, can be taken across parties. This vote-weighted index would be

---

[4] In fairness, not all of these studies use exactly the method described. In particular, several of them do not estimate ρ, but argue only that the cube law does or does not apply in a particular polity (Brookes; Eldersveld; Qualter). Taagepera (1986) does not measure ρ, but predicts it as a function of the number of voters, V, districts, D, and district magnitude, M $\left(\rho = [\ln(V)/\ln(DM)]^{1/M}\right)$, calculates predicted seats using the power law, and compares these to actual seats. Differences in index methods are occasionally exaggerated, however; at least one describes equation 2 and equation 4 as distinct relationships (Laakso 1979).

$$\hat{\rho}_t = \frac{1}{1-v_{1t}} \sum_{j=2}^{J} \left[v_{jt} \frac{\ln(s_{jt}/s_{1t})}{\ln(v_{jt}/v_{1t})} \right] = \frac{1}{1-v_{1t}} \sum_{j=2}^{J} \left[v_{jt} \frac{\psi_{jt}}{y_{jt}} \right], \tag{6}$$

These can also be averaged across elections for a pooled estimate (Taagepera 1973, 1986 is close to this approach). Parties with zero seats or zero votes (not usually a problem) have their votes lumped with other parties. Note that there is no stochastic model specified here and such methods do not provide any measures of confidence in the estimates.

Of course, ρ is our measure of the electoral system's "responsiveness". For party representation, we expect estimates of ρ to be greater than or equal to one.[5] If equal to one, all parties receive votes in exact proportion to seats (*proportionalism*). If infinite, only the largest party receives seats (pure *majoritarianism*). Some studies have produced estimates of ρ less than one (*consensualism*), however, as I discuss below. At the pure consensualism extreme, where ρ is equal to 0, all parties receive the same number of seats regardless of votes. If we wish to extend methods to other types of representation, such as that of geographic regions or ethnic groups (where "vote" is replaced by "population" in all equations) we can expect consensual estimates of ρ.[6] It would be nice if methods were flexible enough to deal with such situations appropriately.

Particularly if we are going to do the type of pooling discussed above, it is useful to transform ρ (before averaging) so that it is symmetric around the proportionality point. That is, we want the "distance" between proportionality and pure majoritarianism to be the same as that between proportionality and pure consensualism. One such transformation is

$$\rho^* \equiv 1 - 2^{1-\rho}, \tag{7}$$

where ρ^* is -1 at pure consensualism ($\rho = 0$), 0 at pure proportionalism ($\rho = 1$), and 1 at pure majoritarianism ($\rho \to \infty$). The symmetry provides an improvement over ρ in the quality of averages; it

---

[5] All else equal, if ρ is less than one, a party can split in two (or clone itself) and increase its representation. For a deeper explanation, with contrasts to the representation of other types of groups where such splitting and cloning is difficult (e.g., geographic regions), see Monroe (1997b).

[6] The United States Senate, for example, has purely consensual geographic representation (all states represented equally). Ethnic representation in many places (e.g., Belgium, Switzerland, South Africa, Bosnia & Herzegovina) is relatively consensual, with smaller ethnic groups holding more than their population share of legislative seats.

also allows groups with zero seats (but some population) or zero population (but some seats)[7] to be included with appropriate values of ρ^* = 1 or -1, respectively.

Throughout this discussion, I will use British general elections as a running example. For ease of exposition and comparison, I examine only the votes and seats of the three major parties (Conservatives, Labour, and Liberals[8]) in England only. This excludes not only Northern Ireland, where the party system is completely different, but Wales and Scotland where nationalist parties compete with the other three and both Labour and the Liberals receive a great deal of their support. Figure 1 shows the index from equation 6, over the 1950-1997 period. For reference, lines have been drawn at $\rho = 0$ (consensualism), $\rho = 1$ (proportionalism), and $\rho = 3$ (the cube law). Note that the estimate provided for 1951 is nonsensical ($\rho = -5.45$). This is because Labour received more votes than the Conservatives, but fewer seats. This is clear evidence of partisan bias, but the index method cannot adjust for that. The other unusually low estimate, from the February 1974 election ($\rho = 1.10$) is due to the reverse occurrence, Labour winning more seats on fewer votes. The pooled estimate of the index is $\hat{\rho} = 1.74$ $(\hat{\rho}^* = 0.21)$; if we throw out the 1951 and Feburary 1974 elections, which is reasonable, the pooled index is $\hat{\rho} = 2.40$ $(\hat{\rho}^* = 0.59)$[9], a value that is highly majoritarian, if not quite to the extent indicated by the cube law. This is similar to the "2.5-law" asserted by Laakso (1979).

Such index methods have several advantages. Most important, they are simple and they can be estimated with observations from a single period *t*. I will argue below that they remain the only option for analyzing geographic and ethnic group representation in many systems.

The disadvantages are substantial, however. Indices imply a deterministic relationship between group votes / population and seats that is obviously false; this in turn also implies no measures of confidence (or statements of extreme confidence, such as standard errors of zero). Also of fundamental importance is the assumption of the functional form in the vote-seat relationship. The power law asserts a

---

[7] This is not as ridiculous as it seems at first glance. Many systems have appointed seats, or reserved seats for particular parties, ethnic minorities, or women, groups that may then receive these seats without competing in the election (and thus appear to have zero votes in electoral statistics). Furthermore, electoral data are often incomplete, reporting no votes for parties that clearly received some.

[8] In this paper I use "Liberals" to refer generically to the Liberals in pre-1983 elections, the Liberal / Social Democratic Party Alliance in the 1983 and 1987 elections, and the Liberal Democratic Party for elections from 1992 to the present. I use "Alliance" or "Liberal Democrats" (and abbreviations) where context makes it appropriate.

[9] This is an example of the usefulness of the transformed ρ^*, showing that ρ of 2.40 (and even more so, 3) is closer to pure majoritarianism than to proportionality, even though it appears closer to one than to infinity.

specific type of vote-seat relationship in which seats depend (in a specific way) only on the relative sizes of groups, not on any party-specific features. That is, the smaller of two parties may be underrepresented relative to its votes in a majoritarian electoral system, but the implication is that it would be overrepresented once it became the larger. The absence of such symmetry, or *partisan bias*, is one of the key measure of the fairness of a representative system, and extremely important to most analysts.[10] It cannot be measured with such indices. Indeed, when it is present (as it must be when the votes-seats relationship appears amonotonic), the index provides nonsense answers. King (1990) argues further that, if the appropriate model is one that includes partisan bias, ignoring it leads to poor estimates of responsiveness.

Multiyear Methods

In multiyear methods, an observation consists of a jurisdictional (e.g., national) seat-vote distribution. Typically the power law, with the addition of party-specific bias parameters, is still taken to capture the basic relationship between seats and votes. In all cases, identification of such models requires an assumption of at least some persistence in the responsiveness and bias parameters across time. Given some specific stochastic structure in which to place this relationship, along with a sufficient number of observations and identifying assumptions, the bias and responsiveness parameters can then be estimated.

A natural form for introducing bias into the two-party power law is

$$\frac{s_{2t}}{1-s_{2t}} = e^{b}\left(\frac{v_{2t}}{1-v_{2t}}\right)^{\rho}, \tag{8}$$

where b is bias in favor of party 1. Taking logs again produces a linear relationship. Tufte (1973), Linehan and Schrodt (1978), Shrodt (1981), Niemi and Jackman (1991), and Jackman (1994) are among those who have examined (two-party) models of forms similar to

---

[10] Tufte (1973), Niemi and Deegan (1978), Grofman (1983) and King and Browning (1987) all discuss, in various ways, bias as deviation from partisan symmetry. Majoritarian, proportional, and consensual system can all be unbiased or fair, in that if parties were to switch vote shares, they would also switch seat shares. This is reasonable in party representation context, although it should be noted that if group populations are fairly stable (as with ethnic groups), different forms of responsiveness in group representation will then also appear more or less "fair" over time even absent any group-specific bias. For example, the perception of majoritarianism in racial representation in the United States is one source of the "tyrrany of the majority" argument for greater minority representation, espoused most famously by Guinier (1994); majoritarianism is biased, in a very meaningful way, when majorities and minorities never change.

$$\ln\left(s_{2t}/\left(1-s_{2t}\right)\right)=b+\rho\ln\left(v_{2t}/\left(1-v_{2t}\right)\right)+\varepsilon_t, \text{ or}$$
$$\psi_{2t}=b+\rho y_{2t}+\varepsilon_t, \tag{9}$$

making the relationship an explicitly stochastic one.[11] Bias is now just the intercept and represents a persistent extra boost in party 2's seat share.[12] With the usual assumptions about the distribution of errors, this can simply be estimated using OLS.

There have been many twists on this basic idea. For example, Mooney (1996) examines such a model using bootstrapping methods and allowing explicitly for temporal (AR1) dependence. King and Browning (1987) describe a maximum likelihood model, the *binomial bilogit* (BNBL) which accounts for the discreteness of seat distribution. The BNBL model assumes a binomial stochastic process producing the seats. Solving equation 8 for *s*, the expectation of the binomial variable is

$$E\left(s_{1t}\right)=\left\{1+\exp\left[-b-\rho\ln\left(\frac{v_{2t}}{1-v_{2t}}\right)\right]\right\}^{-1}. \tag{10}$$

From this, the likelihood function is derived, and maximum likelihood is used to estimate b and ρ, along with their standard errors.

The natural multiparty extension of the basic two-party OLS model [referred to herein as multiparty OLS or MPOLS] is:

$$\ln\left(s_{jt}/s_{1t}\right)=b_2+b_3D_3+b_4D_4+\cdots+b_JD_J+\rho\ln\left(v_{jt}/v_{1t}\right)+\varepsilon_{jt},$$
$$\psi_{jt}=b_2+b_3D_3+b_4D_4+\cdots+b_JD_J+\rho y_{jt}+\varepsilon_{jt} \tag{11}$$

where party 1 is used as a reference party, D_j is a dummy for party *j*, b_j is the partisan bias for party *j* (relative to party 1), and $b_1=0$. Using the OLS estimator requires, of course, that we assume all of the errors are independently and identically distributed according to a univariate normal.

Having multiple parties introduces the possibility of correlation in the errors across seat distributions (given that draws in any election *t* are simultaneous). This can be taken into account with a simultaneous equation model of similar form

[11] There is considerable debate concerning the nature of the error term. This continues to be of interest (see, e.g., Mooney 1996).

[12] Note that bias *can* be a linear term when we use the logistic transformation, but could not be in the power law. The extra seat share value of a constant bias b does shift through the vote share domain.

$$\begin{aligned} \psi_{2t} &= b_2 + \rho y_{2t} + \varepsilon_{2t} \\ &\vdots \\ \psi_{Jt} &= b_J + \rho y_{Jt} + \varepsilon_{Jt} \end{aligned} \tag{12}$$

where the errors are independent across time periods, but, within any period *t*, assumed to be distributed according to some multivariate distribution. If the error distribution is multivariate normal, this is a seemingly-unrelated regression model with ρ constrained to be equal across equations, and referred to herein as the CSURE model. Note that if all correlations are to be estimated, identification of the CSURE model quickly becomes a problem as the number of parties grows.[13] MPOLS or CSURE could also be easily adapted (subject again to identification) to allow for temporal correlations, as with the AR1 model suggested by Mooney.

King (1990) expanded the BNBL model to multiparty systems in the form of the *multinomial bilogit* (MNBL) model, a model that has been used to evaluate 93 electoral systems comparatively (R. Katz 1997), as well as multicandidate U.S. primaries (Ansolabehere and King 1990). In this model, seats are distributed through a multinomial process, with the expected number of seats following the basic power law structure, with biases, as above. Conforming notation, the MLE estimates are those that maximize the log-likelihood

$$\ln L(\rho, b_2, \ldots, b_J \mid s_{1t}, \ldots, s_{Jt}) = \sum_{t=1}^{T} \left\{ \sum_{j=1}^{J} s_{jt} \left[b_j + \rho \ln(v_{jt}) - \ln\left(\sum_{k=1}^{J} e^{b_k + \rho \ln v_{jk}} \right) \right] \right\}. \tag{13}$$

Derivation of this likelihood from the assumptions given is straightforward; the details are available in King (1990). Note that, despite the apparent complexity, interpretation of the *b* and ρ estimates is exactly as before. Choosing among these methods is primarily a matter of deciding which implicit data-generating process is most plausible.

Returning to the British example and using data from 1955-1992, the various multiyear methods provide estimates as summarized in Table 1. These models agree on what appears to be a substantial anti-Liberal bias and two show a minimal anti-Labour bias. They disagree completely on the estimate of responsiveness, with MPOLS asserting mild majoritarianism, MNBL asserting near proportionality, and CSURE actually asserting consensualism (an advantage to small parties). These results -- asserting that elections in Britain under the plurality electoral system do not systematically advantage large parties over small, but simply the Conservatives over the Liberals -- should not sit well with our intuition.

Along these lines, it is worth commenting on the major published results using MNBL. King (1990) analyzes nine democracies. None of these systems produces an estimate of $\rho > 2$, and one

[13] The natural solution is to estimate correlations for the largest parties and set the remainder to zero.

(Austria) has a consensual estimate of $\rho < 1$ (0.87). Richard Katz uses the MNBL model to analyze 93 countries, and estimates of $\rho < 1$ are found in fully a fourth of them, several of them (including the UK) plurality systems where we would expect to see clearly majoritarian estimates.[14] It is possible, of course, that these results are "good counterintuitive" — they serve to correct our faulty intuition. It is also possible that these results are "bad counterintuitive" — they are wrong. I argue below that the latter is true of results from all of the multiyear methods.

These methods all have the advantage, over indices, of asserting a specific stochastic process (which allows estimates of standard errors) and of providing estimates of partisan bias. They do, however, provide estimates that are pooled across time and, therefore, unhelpful for many purposes. We often, for instance, want to know how partisan bias has shifted over time, as electoral systems have been changed or districts redrawn. Multiyear methods cannot provide answers to such questions.

The disadvantages are substantial. Some statistical problems have already been noted. Many of these models are not identified unless the number of parties is small relative to the number of elections in the sample. Also important is that, where MLE estimators are required, the small sample properties (typical samples consist of five to twenty elections) are unknown.

A more fundamental problem is that the answers provided by these methods appear often to be nonsensical (although offered with confidence). Like indices, these methods continue to assume the power-law functional form, a form that does not allow for nonlinearity in the relationship or for partisan differences in the responsiveness estimates.[15] The overconfidence of estimates is also probably due in part to unmodeled temporal dependencies, resulting in overfitting of the data, unreasonably small confidence bounds, and biasing of the responsiveness estimates down. Consider, for instance, an electoral system where incumbency advantage is strong. From one election to the next, the seat distribution is stickier than the vote distribution. This tendency of parties to receive the same number of seats from election to election, regardless of votes, drives all of the variance into the bias terms (dummy coefficients / intercepts). More important, however, is that, since multiyear models assume a persistence across time, they may provide estimates that are politically meaningless due to electoral system and other changes. Jackman (1994) noted this in two-party systems, with considerable evidence of same from Australia, and it is also

---

[14] Furthermore, Katz found estimates ranging dramatically, depending on the number of parties included in the analysis. For example, Canada estimates ranged from 1.99 to 3.17, Turkey estimates ranged from 1.44 to 5.86, and UK estimates ranged from 0.58 to 1.39. These ranges are well beyond the standard errors reported by the model in each case.

[15] We could, of course, generalize MPOLS, CSURE, or MNBL further to include, say, interaction terms with the dummies and vote logratios, but this increases substantially the number of parameters to be estimated.

true in multiparty systems. I provide some evidence that multiyear estimators are indeed typically "nonsense" / "bad counterintuitive" below.

Uniform Swing Methods

Uniform swing has a long history, much of it as an object of derision, in the seats-votes literature. Butler (1947, 1951) appears to have been the first to use it; classic examples include Gudgin and Taylor (1979) and Niemi and Fett (1986); recent applications include Garand and Parent (1991) and Grofman, et al. (1997). Criticisms of uniform swing can be found in King and Browning (1987), King and Gelman (1991), Gelman and King (1994), Jackman (1994)[16]. With uniform swing, we need information about election results *at the district level*. From an observed election result, which gives us one point on the seats-votes curve, we assume that the vote changes ("swings") in exactly the same way across all districts and then calculate the new seat distribution for this distribution of votes. We can then simply define the seats-votes curve by allowing the uniform swing to vary across some vote domain we consider relevant. This dispenses entirely with the power law assumption, although one might then check calculated seat-vote relationships at particular points (at or near the current outcome or one of equal votes) for the closest power-law fit. In the two-party literature, it is common practice to calculate the swing ratio / responsiveness of the seats-votes curve as its slope at, or in the region of, 50% vote share.

As with bias, although we think of swing in terms of vote or seat share, the natural mathematical formulation is in terms of the logratios, y and ψ. This allows swing, like bias, to enter as an additive linear component, without concern for where we are in the vote share domain. This effectively renders the arguments about the "internal inconsistency" of swing (e.g., 1% swing to Party A is not equal to a 1% swing against party A) moot. (McLean 197?; King and Gelman 1991)

The actual election result provides us with one point on the uniform swing seats-votes curve. This point is identified by the mapping from the logratios of (total) votes for parties 2 through J (relative to party 1) — suppressing time subscripts, call this (J-1)-tuple, $\bar{\mathbf{y}}$ — to the logratios of seats for parties 2 through J (relative to party 1) — call this (J-1)-tuple, $\bar{\boldsymbol{\psi}}$. To calculate the curve for a different point in the vote domain, $\bar{\mathbf{y}}'$, we simply add a vector of constant swings, $\delta_j = \bar{y}'_j - \bar{y}_j$, to the vote logratios, y_{ji}, in each district i. From there we calculate the seats won in each district, and then reaggregate to a jurisdictional seat logratio, $\bar{\boldsymbol{\psi}}'$, for a new point on the seats-votes curve. This can then be done over a large domain of possible $\bar{\mathbf{y}}'$.

One potential dilemma in this extension to multiparty systems is this calculation of seats. This remains simple for plurality electoral systems, especially where parties generally have only one candidate. Overall swing for or against a particular party is attributed to only one candidate; the candidate / party with

[16] See also Jackman's long list of citations on the subject (Jackman 1994, n. 41, 42).

the most votes wins. It is straightforward, if not simple, for most districted proportional representation (PR) systems -- votes merely shift across the party lists, and seats are calculated by the appropriate formula (largest remainders with Hare quota, d'Hondt with 5% threshold, etc.). Additional tiers of aggregating excess votes (e.g., as used in Greece) do not present any additional conceptual difficulty.

The calculation becomes involved and idiosyncratic for many other electoral systems, however. Where there is more than one type of vote or districts (e.g., additional member systems as used in Germany and New Zealand), one must decide how to handle these different types of districts. Under single nontransferable vote systems (as used in Taiwan), one must decide how party swings are to affect individual candidates. In preferential electoral systems, where more than first preferences are expressed (e.g., alternative vote as used from some elections in Australia, double ballots as used in France, or single transferable vote as used in Ireland), even more severe assumptions must be made. For the purposes of this paper, I avoid such complications, but it should be said that barriers remain to applying uniform (and variable) swing methods to such electoral systems.[17]

Defining a multiparty seats-votes curve is one thing; *visualizing* it is quite another. In a two-party system, we can graph seat share, s_2, as a function of vote share, v_2 (a mapping from the unit interval to the unit interval) or seat logratios, ψ_2, as a function of vote logratios, y_2 (a mapping from the real line to itself). A cube law curve, for example, is an S-shaped curve on the vote share domain (see, e.g., King and Browning 1987), and a line with slope three on the logistic domain. In a *J* party system, the logratio curve is a mapping from $\Re^{J-1}$ to $\Re^{J-1}$. That is, even a three-party seats-votes curve would require four dimensions to be displayed as a simple graph. Figures in the literature to date are typically either plots of the function $s = v^{\rho}$ (e.g., King 1990) or "proportionality profiles" (Taagepera and Laakso 1980; Taagepera and Shugart 1989). Neither of these appropriately captures the multidimensional nature of seats-votes data and curves.

We can, however, effectively halve the dimensionality requirement (to *J*-1) by exploiting the identity between the graph's range and domain. For a three-party system, we can illustrate selected points in the $\Re^2 \rightarrow \Re^2$ mapping as arrows in $\Re^2$ from a pair of vote logratios $(\bar{y}_2, \bar{y}_3)$ to a pair of seat logratios $(\bar{\psi}_2, \bar{\psi}_3)$. Figure 2 illustrates a seats-votes curve with a typical power law relationship; it is, in fact, the relationship ($\rho = 2.40$) previously estimated with the index method.

The reader should probably take a moment to become oriented with this graph. The *x*-axis indicates Labour performance relative to the Conservatives; the *y*-axis indicates Liberal Democrat performance relative to the Conservatives. Strong Labour performance is indicated by points to the right of

---

[17] Extending the methods here to any particular electoral system is a matter of finding an appropriate model of votes and seats at the district level. For a small minority of these systems (e.g., SNTV or runoff), these models will almost definitely have to be stochastic. If one is willing to incorporate such a model, then it should almost definitely be combined with the variable swing model, rather than with uniform swing.

the graph. Strong Liberal performance is indicated by points at the top of the graph. Strong Conservative performance is indicated by points in the lower left quadrant. The origin is the point at which all parties do equally well. Again, both vote and seat performance are displayed on the same graph, with votes at the base of an arrow, and seats at the tip.

Note first the linearity of the power law curve in the logistic domain; the tip of each arrow is approximately a multiple (ρ) further away from the origin than its base, and in approximately the same direction. The curve is majoritarian, since ρ>1. Visually, this means that the arrows point away fixed point in the mapping where seats exactly equal votes, $\overline{\mathbf{y}}' = \overline{\boldsymbol{\psi}}'$. If the curve were consensual (ρ<1), the arrows would point toward the fixed point. If it were perfectly proportional, there would be no fixed point. Note also that in this case the origin is the fixed point, because the index method assumes no partisan bias. I refer to the fixed point throughout as the *proportionality point*, and denote it $\overline{\mathbf{y}}^{(\mathrm{p})}$.

Finally, note the arrowhead shape outlining the arrows. This is a function of discreteness in the seat distribution. Logratios are undefined / infinite where one of the parties receives zero seats; the arrowhead outline is defined by seat distributions where each party, in turn, receives one seat.[18] It is important that this apparently asymmetric shape in the seats-votes curve not be mistaken for some partisan distortion in the curve. Note also that discreteness is an important consideration as seat distributions approach this boundary, creating distortions in the strict power law relationship. This becomes important where *D* is small, or even with large *D* where some party receives, say, three or fewer seats on a national basis, and may affect our inferences if we do not take it into consideration.

Figure 3 shows the curve implied by the MPOLS estimates from Table 1. This is still a majoritarian power law curve with arrows pointing out from a central point. The arrows are shorter (and many more fit in the arrowhead) than in Figure 2, since the estimate of ρ (1.31) is closer to proportionality. The dramatic partisan bias against the Liberals is reflected in a translation of the central point well *up* from the origin. For a power law curve, the vector of partisan biases is indicated exactly by the arrow with its base at the origin ($\boldsymbol{\psi} = \mathbf{b}$ if $\mathbf{y} = \mathbf{0}$ in a power law curve). In a power law curve, the proportionality point and the bias vector are related by

$$\mathbf{y}^{(\mathrm{p})} = \left(\frac{1}{1-\rho}\right)\mathbf{b} \tag{14}$$

[18] In a two-dimensional logratio space, the outline is a set of three parametric curves. They are all contained within a square centered at the origin, with sides of length 2ln(*D*), where D is the number of districts (in this case *D*=523).

Figures 4 and 5 show the curve implied by the MBNL and CSURE methods, respectively. We can see exactly what inference we are supposed to draw from the estimates, and each looks bizarre. The MBNL estimate of ρ is very close to one, indicating almost perfect proportionality. That means that seats are determined almost entirely by partisan bias, a severe anti-Liberal bias in this case. The CSURE estimate of ρ is less than one, along with substantial anti-Liberal bias, implying that seat distributions are quite similar regardless of vote distribution. The inferences from the MNBL and CSURE models, that the British electoral system is proportional or even consensual, should sit not sit well.

Figure 6 shows the seats-votes curve calculated with uniform swing. Several features are worth noting. The first is that the general pattern is majoritarian, with arrows pointing away from the proportionality point. The second is that that central point is not the origin, indicating partisan bias. The third, and generally most important, is that the curve is not exactly a power law distribution, although it does resemble one very closely. There are distortions away from a pattern like that of Figure 2 that are more than a simple translation away from the origin. That said, Figure 6 looks much more like Figure 2 than any of the curves calculated by multiyear methods. Note that all of the graphs in Figures 2-6 offer similar vote-seat relationships in the region where most postwar British results have occurred, along the *y*-axis below the origin (Conservatives and Labour close, with the Liberals in third). The implications for seats if votes shift even slightly away from this general pattern (such as they did in 1997 are quite different), are quite different for the different models.

For comparison's sake, if nothing else, we need some way of extracting summary measures of partisan bias and responsiveness from the uniform swing seats-votes curve. There are many ways to do this, but the following has proved to provide the most robust estimates. We first locate the proportionality point, $\overline{\mathbf{y}}^{(\mathrm{p})}$ (which is straightforward).[19] For a given vector pair $(\mathbf{y}', \boldsymbol{\psi}')$ on the uniform swing seats-votes curve, we then find the bias and responsiveness values of the power law that locally approximates the curve:

$$\begin{aligned} \rho' &= \frac{1}{J-1}\sum_{j=2}^{J}\frac{y_j^{(\mathrm{p})}-\psi_j'}{y_j^{(\mathrm{p})}-y_j'} \\ b_j' &= \frac{y_j^{(\mathrm{p})}\left(\psi_j'-y_j'\right)}{y_j^{(\mathrm{p})}-y'} \end{aligned} \tag{15}$$

As with index methods, these values tend to cluster around a particular point, with some wild outliers (e.g., $\rho' = -1000$). The median values calculated on the grid serve as stable estimates. Using this method in our example, we find $\hat{\rho} = 2.68$ $(\hat{\rho}^* = 0.69)$ and $\hat{\mathbf{b}} = (.897, -.227)$. The estimate is majoritarian,

[19] Unless the system is perfectly proportional, which it never is. If it were, however, we simply set $\hat{\rho} = 1$ and calculate bias as the average distance / direction between vote and seat vectors.

as is clearly implied by Figure 2. Indeed, we are not too far either from the cube law or the index estimates. Contra the multiyear methods, we also find a smaller anti-Liberal bias and a fairly large pro-Labour bias. The latter is somewhat contradictory to conventional wisdom, especially given that Scottish and Welsh seats, where Labour does much better than the Conservatives, have been excluded from the analysis. Remember, however, that this does not suggest that Labour will win more votes than the Conservatives, only that their vote share will be more efficiently translated into seat shares. Note again that, despite the machinations involved, these are point estimates with no implied standard errors.

The primary advantage of the uniform swing method is that, from a single election result, it provides a great deal of information about vote-seat relationships. In particular, it provides estimates of bias and responsiveness, allows us to recognize deviations from the power law functional form, and it provides evidence that multiyear methods are misleading. The advantages over index and multiyear methods appear because of the theoretical advantage they have. Uniform swing takes into account how seats are actually distributed through the electoral system; index and multiyear methods treat the electoral system as a black box transferring aggregate votes into aggregate seats.

The primary disadvantages, of course, are the false assumption of uniform swing and the implication of a deterministic relationship. Uniform swing has been shown to overestimate responsiveness in two-party systems (Jackman 1994; King and Gelman 1994), a tendency which we do observe in our examples here as well. More difficult cases — involving complicated electoral systems, shifting party systems, regional difference in party systems — require additional nuance which as been put aside here. And while figures for a four-party system could also be drawn, beyond that we would be limited to two, three, and four-party cross-sections and projections.

Variable Swing Methods

Methods that model swing, but specifically allow for variation, comprise the state of the art in estimation of two-party seat-vote relationships. Examples include King (1989), King and Gelman (1991), Jackman (1994), and Gelman and King (1994). In particular the Gelman and King "Judgeit" approach, generalized to the multiparty environment, provides us with a variable swing method. This generalization is nontrivial, and has applications well beyond the present goal of analyzing multiparty seat-vote curves. This model is outlined in greater detail in a companion article (Monroe 1998).

In overview, the approach here is to take any given election result as one of many possible hypothetical election results produced by the same underlying stochastic process. The goal is then to determine this process so that election results can be simulated under a variety of counterfactual conditions, from which we can calculate certain quantities of interest. In the present case, we wish to model district vote logratios as a function of various covariates and counterfactual national swings in logratios. From this model, we can simulate elections, calculate seat distributions and, in turn, calculate seats-votes curves, responsiveness, and biases.[20]

The Model[21]

As before, I assume that there are D electoral districts at the lowest level of aggregation, and that there are J parties that compete in every district in every election under consideration. We again choose one party (party 1) to be a reference party. It typically does not matter which party is chosen, but if there is a party that is consistently the most successful across the elections considered, it aids interpretation to choose it as the reference.

We must now keep very close track of districts, so where possible I suppress time subscripts as well as subscripts indicating the reference party. So, for a given election, I use v_{ji} to denote the proportion of votes for party *j* in district *i*. I use y_{ji} to denote the vote logratio for party *j* (relative to party 1) in district *i*: $y_{ji} = \ln(v_{ji}/v_{1i})$. We will denote the $D \times 1$ vector of vote logratios for party *j* as $\mathbf{y}_j$, and define $\mathbf{y} = [\mathbf{y}'_1, \mathbf{y}'_2, \ldots, \mathbf{y}'_{J-1}]'$, the $D(J-1) \times 1$ vector containing the stacked vote logratios for all parties in all districts. Logratios have the advantage over vote shares already mentioned — consistency of swing, linearity of swing, bias, and responsiveness in power law curves — plus one more. Vote shares (and seat shares) constitute *compositional data*, data vectors in which the elements must be positive and sum to unity. Modeling compositional data processes can be complex, but one simple approach is to make the "additive logistic" transformation (i.e., use logratios) and proceed using well-known multivariate techniques. The logratios being the natural format for vote-seats data in any case, that is how I proceed here. (On compositional data see Aitchison 1986; Katz and King 1997; Monroe 1997a).[22]

We model the process producing $\mathbf{y}$ as

---

[20] The model is considerably more general than this and can be used to calculate a variety of other types of quantities of interest, such as incumbency advantage (Gelman and King 1994; Monroe 1998).

[21] I should note that, except for the changes required to generalize the model to multiparty systems, this is quite similar to the Gelman and King model. In explication, I have in several cases simply used their language, rather than quote excessively or go through fifth-grade efforts at plagiarism avoidance (World Book Encyclopedia: "Kentucky's primary exports are coal and tobacco." Me: "Tobacco and coal are Kentucky's main exports."). I presume this indulgence can be granted in the interest of readability, without any suggestion that I am presenting their work as my own.

[22] This problem is ignored in typical two-party applications, because vote shares remain near 50%. Ignoring the 0%-100% bounds has little consequence. We do not have that luxury in settings of larger dimensionality.

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\gamma} + \boldsymbol{\varepsilon} \tag{16}$$

Party *j*'s logratios, $\mathbf{y}_j$, are predicted by K_j explanatory variables, which we gather in a $D \times K_j$ vector $\mathbf{X}_j$. These are also stacked in a $D(J-1) \times \sum_{j=2}^{J} K_j$ matrix, $\mathbf{X}$, such that:

$$\mathbf{X} \equiv \begin{bmatrix} \mathbf{X}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{X}_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{X}_{J-1} \end{bmatrix}$$

The linear coefficients are summarized in the $\sum_{j=2}^{J} K_j \times 1$ vector $\boldsymbol{\beta} = \left[\boldsymbol{\beta}_1', \boldsymbol{\beta}_2', \ldots, \boldsymbol{\beta}_{J-1}'\right]'$, where $\boldsymbol{\beta}_j'$ is the $K_j \times 1$ vector of coefficients for the party *j* equation.

Finally, there are two $D(J-1) \times 1$ vectors of error terms, $\boldsymbol{\gamma}$ and $\boldsymbol{\varepsilon}$. The errors $\boldsymbol{\varepsilon}$ are the traditional error terms; the $\boldsymbol{\gamma}$ are "random components" error terms, used to correct for omitted variables and measurement error in $\mathbf{X}$. Within each district *i*, these error terms are distributed according to independent multivariate normal distributions,

$$\begin{aligned} \boldsymbol{\gamma}_i &\sim N(\mathbf{0}, \boldsymbol{\Sigma}_\gamma) \\ \boldsymbol{\varepsilon}_i &\sim N(\mathbf{0}, \boldsymbol{\Sigma}_\varepsilon) \end{aligned} \tag{17}$$

with covariance matrices, $\boldsymbol{\Sigma}_\gamma$ and $\boldsymbol{\Sigma}_\varepsilon$ that must be estimated. Since $\boldsymbol{\gamma}$ and $\boldsymbol{\varepsilon}$ are independent from one another, we can think of the model as a seemingly unrelated regression with a single vector of error terms, $\mathbf{u}$:

$$\mathbf{y} = \mathbf{X}\mathbf{B} + \mathbf{u}, \text{ where } \mathbf{u}_i \sim N(\mathbf{0}, \boldsymbol{\Sigma}) \text{ and } \boldsymbol{\Sigma} = \boldsymbol{\Sigma}_\gamma + \boldsymbol{\Sigma}_\varepsilon. \tag{18}$$

It is useful to reparameterize this with a new matrix $\boldsymbol{\Lambda}$, such that

$$\boldsymbol{\Lambda} = \boldsymbol{\Sigma}_\gamma \boldsymbol{\Sigma}^{-1} \tag{19}$$

The covariance matrix, $\boldsymbol{\Sigma}$, of $\mathbf{u}$ is then a matrix-weighted sum of $\boldsymbol{\Sigma}_\gamma$ and $\boldsymbol{\Sigma}_\varepsilon$, with $\boldsymbol{\Lambda}$ determining the weights:

$$\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_{\gamma} + \boldsymbol{\Sigma}_{\varepsilon} = \boldsymbol{\Lambda\Sigma} + (\mathbf{I} - \boldsymbol{\Lambda})\boldsymbol{\Sigma} = \boldsymbol{\Sigma} \tag{20}$$

Recall that we must also model *hypothetical* election outcomes. The probability model for hypothetical vote logratios, $\mathbf{y}^{(\mathrm{hyp})}$, is then defined analogously as

$$\mathbf{y}^{(\mathrm{hyp})} = \mathbf{X}^{(\mathrm{hyp})}\boldsymbol{\beta} + \boldsymbol{\delta}^{(\mathrm{hyp})} + \boldsymbol{\gamma} + \boldsymbol{\varepsilon}^{(\mathrm{hyp})}, \tag{21}$$

where $\boldsymbol{\varepsilon}^{(\mathrm{hyp})}$ is a new $D(J-1)\times 1$ vector of error terms, with the $(J-1)\times 1$ vector of errors in district *i*, $\boldsymbol{\varepsilon}_i^{(\mathrm{hyp})}$, being distributed independently with covariance matrix $\boldsymbol{\Sigma}_{\varepsilon}$. The known $D(J-1)\times 1$ vector $\boldsymbol{\delta}^{(\mathrm{hyp})}$ consists of constants (*J*-1 of them stacked *D* times) used to model statewide partisan swing.

Replacing $\mathbf{X}$ with $\mathbf{X}^{(\mathrm{hyp})}$, allows us to specify different counterfactual conditions, such as no incumbents running, under which to simulate hypothetical elections. This is useful for evaluating the causal effects of such changes, but for the seats-votes calculations in this paper, I will always use $\mathbf{X}^{(\mathrm{hyp})} = \mathbf{X}$. (See Monroe 1998 for an example where $\mathbf{X}^{(\mathrm{hyp})} \neq \mathbf{X}$.) The constants, $\boldsymbol{\delta}^{(\mathrm{hyp})}$, are added to allow for statewide partisan swings.[23]

The parameters to be estimated are $\boldsymbol{\Sigma}$, $\boldsymbol{\Lambda}$, and $\boldsymbol{\beta}$. These are not of direct interest themselves, but are used to determine the conditional distribution of $\mathbf{y}^{(\mathrm{hyp})}$. This, in turn is used to estimate the seats-votes curve and quantities of interest, like responsiveness and partisan bias.

I do not wish to digress too far here on interpretations of this model. I do so at greater length in the companion paper (Monroe 1998) and the discussion by Gelman and King (1994: pp. 521-6) applies more or less in its entirety. Some digression on the error components, and particularly $\boldsymbol{\Lambda}$, is worthwhile however, especially as it links the variable swing method with the uniform swing method already discussed.

Consider the extreme case in which $\boldsymbol{\Lambda}$ has all elements equal to 0. That means that $\boldsymbol{\Sigma}_{\gamma}$ is also a matrix of zeros. All error is then random error, $\boldsymbol{\varepsilon}$. Any two districts with identical values of $\mathbf{X}$ then have exactly the same distribution of hypothetical election results. Even if we observe different election results in the two districts, we ignore those differences in calculating hypothetical results; the explanatory variables in $\mathbf{X}$ are perfect.

At the other extreme, $\boldsymbol{\Lambda} = \mathbf{I}$, so that now $\boldsymbol{\Sigma}_{\varepsilon}$ is a matrix of zeros. Now every hypothetical election result, $\mathbf{y}^{(\mathrm{hyp})}$, must be exactly the same as the observed result, $\mathbf{y}$, except for the national

---

[23] We could instead specify $E(\mathbf{y}^{(\mathrm{hyp})})$ and calculate swings as $\boldsymbol{\delta}^{(\mathrm{hyp})} = E(\mathbf{y}^{(\mathrm{hyp})}) - (1/D)\sum_{i=1}^{D}(\mathbf{X}^{(\mathrm{hyp})}\boldsymbol{\beta})_i$, but for the purposes in this paper there is no reason to do so.

swing, $\boldsymbol{\delta}^{(\mathrm{hyp})}$; we completely ignore the covariates. This is uniform swing. That is, we can think of the uniform swing method outlined in the previous section as a special case of the model described here, in which we do not estimate $\boldsymbol{\Lambda}$, but rather set it equal to $\mathbf{I}$.

Preliminary Estimation

Estimation of $\boldsymbol{\beta}$ and $\boldsymbol{\Sigma}$ is a straightforward application of the seemingly unrelated regression model summarized in equation 18. We can use feasible generalized least squares (as summarized, for instance, in Greene 1997, 674-88) to obtain $\hat{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{\Sigma}}_{\beta}$. This is done with either the current election results (if observed) or the most recent election available, if not. The elements of $\boldsymbol{\Sigma}$ are estimated (consistently) by the sample covariances.[24]

Estimation of $\boldsymbol{\Lambda}$ is much more difficult, presenting one of the major technical difficulties in generalizing to the multiparty system. Gelman and King discuss two different methods for the two party context, where $\boldsymbol{\Lambda}$ is a constant, λ. First, they note that λ is the coefficient on v in the expected value of the $v^{(\mathrm{hyp})}$ distribution. From this observation, a natural approach is to use two sequential elections and estimate the regression

$$v_{t+1} = \lambda v_t + X_t \beta + u\,, \tag{22}$$

taking v_{t+1} as an observation of $v^{(\mathrm{hyp})}$, and truncating λ to zero or one if necessary. A second approach is based on the observation that the theoretical correlation between $v^{(\mathrm{hyp})}$ and v is λ. The estimator based on this observation is the correlation between the residuals of any two sequential elections, truncated to zero if necessary. Gelman and King report that these provide similar estimates in their empirical examples.

Changing v to $\mathbf{y}$, these methods generalize quite easily to the multiparty context, but do not have similarly nice properties. This problem is closely related to that of estimating a dynamic error components model for panel data. Dynamics make decomposition of the error term problematic (Amemiya 1971; Mátyás 1996). The two methods here produce very different results in many cases. In particular, because

---

[24] This is a minor deviation from the Gelman and King model, in that they incorporate a degrees-of-freedom correction. Degrees of freedom corrections are problematic in this context, and the difference is typically negligible (Judge, et al. 1985, p. 469; Greene 1997: 676-7).

of high temporal correlation, the regression method tends to yield estimates of $\hat{\mathbf{\Lambda}} \approx \mathbf{I}$, regardless of the specification of $\mathbf{X}$. The correlation method, $\hat{\lambda}_{jk} = s_{[jt][k(t+1)]}$, appears to yield more sensible estimates.[25]

Note further, however, that $\hat{\mathbf{\Lambda}}$ cannot be an arbitrary $(J-1)\times(J-1)$ matrix. We must find $\hat{\mathbf{\Lambda}}$ such that $\mathbf{\Sigma}_\gamma = \hat{\mathbf{\Lambda}}\mathbf{\Sigma}$ and $\mathbf{\Sigma}_\gamma = (\mathbf{I} - \hat{\mathbf{\Lambda}})\mathbf{\Sigma} = \mathbf{\Sigma} - \mathbf{\Sigma}_\varepsilon$ are symmetric and positive definite. We can do so in two ways. Define $\hat{\mathbf{\Lambda}}_0$ as the estimate of $\mathbf{\Lambda}$ constructed from the residual correlations above. The first method is to find the $\hat{\mathbf{\Lambda}}$ "closest" to $\hat{\mathbf{\Lambda}}_0$ (by, say, a least-squares loss function on each element) that satisfies these conditions. The second is to assume that $\mathbf{\Lambda} = \lambda\mathbf{I}$, constraining the off-diagonal elements of $\hat{\mathbf{\Lambda}}$ to zero and the diagonal elements to, say, the average diagonal element of $\hat{\mathbf{\Lambda}}_0$, $\hat{\lambda}_{jj}$. The latter is a little bit easier, and the resulting $\hat{\mathbf{\Lambda}}$ is easier to interpret. This does impose the unlikely constraint that our covariates are equally useful in explaining the performance of each party. In practice, however, the two often yield similar results.

As do Gelman and King, I pool estimates of $\mathbf{\Sigma}$ and $\mathbf{\Lambda}$ across election years where appropriate. Such pooling is appropriate if we have data from multiple elections run under similar circumstances (e.g., the same electoral system and districts are used, and the same explanatory variables are available). After estimation and pooling, I assume $\mathbf{\Sigma} = \hat{\mathbf{\Sigma}}$ and $\mathbf{\Lambda} = \hat{\mathbf{\Lambda}}$. That is, I ignore any uncertainty in the estimates $\mathbf{\Sigma}$ and $\mathbf{\Lambda}$ in the analysis that follows. The estimates of $\boldsymbol{\beta}$ are not pooled.

The Distribution of Hypothetical Votes

I omit here entirely the derivation of the distribution of hypothetical votes and present only the final result. The derivation follows almost exactly that of Gelman and King and is given in its entirety in Monroe (1998). There are two distributions of interest. The first is the *predictive distribution*. In this case, we have not observed $\mathbf{y}$ $(\mathbf{v})$ and wish to know $P(\mathbf{y}^{(\text{hyp})})$. This distribution is

$$P\left(\mathbf{y}^{(\text{hyp})}\right) = N\left(\mathbf{y}^{(\text{hyp})} \mid \mathbf{X}^{(\text{hyp})}\hat{\boldsymbol{\beta}} + \boldsymbol{\delta}^{(\text{hyp})}, \mathbf{X}^{(\text{hyp})}\hat{\mathbf{\Sigma}}_\beta \mathbf{X}^{(\text{hyp})\prime} + \mathbf{\Sigma} \otimes \mathbf{I}\right). \tag{23}$$

Expected values are merely the usual predicted values plus partisan swings. Note that variance emerges both from the inherent error, $\mathbf{\Sigma}$, and from error in the estimation of $\boldsymbol{\beta}$, and that $\boldsymbol{\gamma}$ and $\mathbf{\Lambda}$ do not enter the distribution.

---

[25] On the other hand, when a system has strong intertemporal correlation, like the British system does, the estimates of $\mathbf{\Lambda}$ tend to zero. Some theoretical support for this estimation method would, of course, be helpful.

When evaluating an existing electoral system (under actual or counterfactual conditions), we will have observed **y**. We therefore need the distribution of $P(\mathbf{y}^{(\text{hyp})} \mid \mathbf{y})$, which is

$$P(\mathbf{y}^{(\text{hyp})} \mid \mathbf{y}) = N\left(\mathbf{y}^{(\text{hyp})} \mid \mathbf{\Lambda y} + (\mathbf{X} - \mathbf{\Lambda X}^{(\text{hyp})})\hat{\boldsymbol{\beta}} + \boldsymbol{\delta}^{(\text{hyp})}, (\mathbf{X}^{(\text{hyp})} - \mathbf{LX})\hat{\mathbf{\Sigma}}_{\beta}(\mathbf{X}^{(\text{hyp})} - \mathbf{LX})' + (\mathbf{I} - \mathbf{\Lambda\Lambda})\mathbf{\Sigma} \otimes \mathbf{I}\right) \quad (24)$$

This simplifies somewhat when $\mathbf{X}^{(\text{hyp})} = \mathbf{X}$, as is always the case in this paper. Now,

$$P(\mathbf{y}^{(\text{hyp})} \mid \mathbf{y}) = N\left(\mathbf{y}^{(\text{hyp})} \mid \mathbf{\Lambda y} + (\mathbf{I} - \mathbf{\Lambda})\mathbf{X}\hat{\boldsymbol{\beta}} + \boldsymbol{\delta}^{(\text{hyp})}, (\mathbf{I} - \mathbf{\Lambda})\mathbf{X}\hat{\mathbf{\Sigma}}_{\beta}\mathbf{X}'(\mathbf{I} - \mathbf{\Lambda})' + (\mathbf{I} - \mathbf{\Lambda\Lambda})\mathbf{\Sigma} \otimes \mathbf{I}\right) \quad (25)$$

In this case, expectation is just a matrix-weighted average of actual outcome and predicted outcome (plus swing), with $\mathbf{\Lambda}$ as the weighting matrix. At one extreme, if $\mathbf{\Lambda} = \mathbf{I}$, then the uniform swing assumption is correct. Now, one can use current votes and swing alone and ignore the predictions, $\mathbf{X}\hat{\boldsymbol{\beta}}$, and any variance. If, on the other hand, $\mathbf{\Lambda} = \mathbf{0}$, one uses only $\mathbf{X}\hat{\boldsymbol{\beta}}$, concluding that all excess variation is nonsystematic and discarding the information contained in the observed votes (the latter is, in effect, the approach in King 1989).

Calculating Seats-Votes Curves, Bias and Responsiveness

The basic approach, which would apply to the calculation of any quantity of interest, is Monte Carlo simulation. This involves three steps:

1. Generate *M* hypothetical election results, $\mathbf{y}^{(\text{hyp})}$, by drawing from the appropriate multivariate normal distribution (Equation 23 or 24).
2. For each hypothetical election result, *m*, calculate the quantities of interest (e.g., number of winning incumbents, vote in district *i*, bias, etc.), $\mathbf{Q}^{(\text{hyp})m}$.
3. Calculate the point estimate, $\hat{\mathbf{Q}}$, as the mean,

$$\hat{\mathbf{Q}} = \overline{\mathbf{Q}} = \frac{1}{M}\sum_{m=1}^{M}\mathbf{Q}^{(\text{hyp})m},$$

and for any one quantity, *Q*, the standard error, $\hat{\sigma}_Q$, as the square root of the variance,

$$\hat{\sigma}_Q = \sqrt{\text{Var}(Q)} = \sqrt{\frac{1}{M-1}\sum_{m=1}^{M}\left(Q^{(\text{hyp})m} - \overline{Q}\right)^2}.$$ [26]

[26] There are other ways to define the standard errors from the simulated outcomes (see, e.g., Mooney 1996).

Note that some quantities of interest are multidimensional. Expected seats and partisan bias, for example, are vectors with J-1 quantities. The full posterior distribution of Q can be illustrated in a J-1-dimensional scatter plot or histogram (perfectly feasible for $J \leq 3$; problematic for $J \geq 4$).

Step 1 cannot be accomplished directly, as the covariance matrices are singular. Instead, generate a hypothetical election in steps:

1. Draw one random vector, $\tilde{\boldsymbol{\beta}}$, from the (posterior) distribution for β:

$$P(\boldsymbol{\beta}) = N\left(\boldsymbol{\beta} \mid \hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\Sigma}}_{\beta}\right)$$

2. If $\mathbf{y}$ has been observed, insert $\tilde{\boldsymbol{\beta}}$ into the distribution for $\boldsymbol{\gamma}_i \mid \boldsymbol{\beta}, \mathbf{y}_i$,

$$P(\boldsymbol{\gamma}_i \mid \boldsymbol{\beta}, \mathbf{y}_i) = N\left(\boldsymbol{\gamma}_i \mid \boldsymbol{\Lambda}\left(\mathbf{y}_i - \mathbf{X}_i^{(\mathrm{hyp})}\right)\boldsymbol{\beta}, \boldsymbol{\Lambda}(\mathbf{I} - \boldsymbol{\Lambda})\boldsymbol{\Sigma}\right),$$

and draw one $(J-1)\times 1$ vector, $\tilde{\boldsymbol{\gamma}}_i$, for each district i. If $\mathbf{y}$ has not been observed, $\tilde{\boldsymbol{\gamma}}_i = \mathbf{0}$ (and $\boldsymbol{\Lambda} = \mathbf{0}$).

3. Insert $\tilde{\boldsymbol{\beta}}$ and $\tilde{\boldsymbol{\gamma}}_i$ into the distribution for $\mathbf{y}_i^{(\mathrm{hyp})} \mid \boldsymbol{\beta}, \boldsymbol{\gamma}_i, \mathbf{y}_i$,

$$P\left(\mathbf{y}_i^{(\mathrm{hyp})} \mid \boldsymbol{\beta}, \boldsymbol{\gamma}_i, \mathbf{y}_i\right) = N\left(\mathbf{y}_i^{(\mathrm{hyp})} \mid \mathbf{X}_i^{(\mathrm{hyp})}\boldsymbol{\beta} + \boldsymbol{\gamma}_i + \boldsymbol{\delta}_i^{(\mathrm{hyp})}, (\mathbf{I} - \boldsymbol{\Lambda})\boldsymbol{\Sigma}\right),$$

set partisan swings, $\boldsymbol{\delta}_i^{(\mathrm{hyp})}$, as you choose, and draw one $(J-1)\times 1$ vector, $\mathbf{y}_i^{(\mathrm{hyp})}$, for each district i.

The final issue is how to generate seats-votes curves and derive estimates of responsiveness and bias, the quantities of interest here. This is done through repeated application of the uniform swing procedure outlined in the previous section. We calculate M hypothetical election results for $\boldsymbol{\delta}^{(\mathrm{hyp})} = -\bar{\mathbf{y}}$ (that is, at the origin) and then, with each one, apply uniform swing over the desired domain to determine M hypothetical seats-votes curves. Each of these produces a single point estimate of responsiveness and the bias vector, through the formulae discussed under uniform swing. Final point estimates and standard errors for bias and responsiveness follow as summaries of the simulated distributions. It is difficult to display all of the information that the full seats-votes curve simulation provide, but we can describe the relationship between expected votes and expected seats as before (for $J \leq 4$). If we would like to see the confidence bounds on the seats-votes curve, we can plot the histogram of seat simulations for a given vote, rather than the single-point tip of the arrow as before. Such plots are cluttered, however, unless we look at a very few points on the seats-votes curve.

Calculating Seats-Votes Curves, Bias and Responsiveness

Let us return to the running example of British General elections in English constituencies. I have calculated seats votes curves for 1992, using election results from 1983-1992 for the calculation of $\mathbf{\Sigma}$ and $\mathbf{\Lambda}$. Explanatory variables included in the model are indicators of party control after the previous election, indicators of candidate incumbency for each party, and lagged values of the vote logratios. Figure 7 shows the curve calculated using the most general estimator $\mathbf{\Lambda} = \hat{\mathbf{\Lambda}}$.[27] For 1992, the model produces a responsiveness estimate of $\hat{\rho} = 2.52$ $(\hat{\rho}^* = 0.65)$ with a standard error of $\hat{\sigma}_\rho = 0.09$, indicating strong majoritarianism, although not to the extent found using uniform swing. The estimated bias vector is $\hat{\mathbf{b}} = (+.229, +.003)$ with standard errors $\hat{\boldsymbol{\sigma}}_b = (0.033, 0.0002)$, indicating a pro-Labour bias, although again not to the extent found with uniform swing. Note that the curve again looks generally like a power law curve with some party specific distortions. An example of this is that the Liberal Democratic logratio of seats seem to be more responsive than are Labour seats. This reflects the common wisdom that the Liberal Democrat's tendency to have their votes spread evenly throughout the country hurts them badly while they are in third, but would benefit them tremendously if they were to move into second or first.

One red flag in these results is the small standard errors, which are probably too small to be reasonable. Indeed, the confidence regions around any particular vote-seat mapping are not much different from point predictions. In my concluding remarks, I discuss some generalizations and extensions of what has been done here that might give us more confidence in our measures of confidence.

Although our intuition should suggest that the variable swing approach provides the most reliable estimates, we really have no independent confirmation that this is so. To offer some evidence that this is indeed so, I have calculated the seat prediction, given the vote outcome, for the 1997 elections. In each case, I have used only information that was available before the election. For the multiyear and index methods, the results from 1950-1992 were used. For uniform and variable swing methods, the result of the 1992 election was used. It should be noted that this is an abuse of the uniform and variable swing models, as the U.K. was redistricted between 1992 and 1997, but the models have been run on the old districts.

Figure 8 illustrates the point predictions for seats given the 1997 votes, for the power law relationship implied by each set of bias and responsiveness estimates. The 1950-1992 results are also shown for reference. (We could, for variable and uniform swing methods use the *exact* relationship on the

---

[27] Lagged values are such a strong predictor of relative performance in British electoral districts that $\mathbf{\Lambda} \approx \mathbf{0}$ when lags are included in the model, as they are here.

calculated curve, rather than the power law approximation, but this would be doubly abusive given the districting changes.) The swing and index methods all get the direction approximately right. Uniform swing overestimates the Labour advantage and Liberal Democratic disadvantage; the index is almost spot on. All three of the multiyear methods underestimate the size of the Labour win and overestimate the size of the Liberal Democratic defeat.

We can see here fairly dramatically the difference in the approach of the three types of methods. The multiyear methods have nothing to use but the national results imposed on a weak functional form, and they do all look like they are about the same length and end in approximately the same location. The swing methods, on the other hand, have a wealth of information about the structure of the electoral system, and this can be used to extrapolate sensibly beyond the domain of observed election results. The index has the same information as the multiyear methods plus one assumption: an unbiased power law relationship. If we do not care to calculate bias, and there is not much of it, the index will be reasonably accurate.[28] Particularly if we have no choice — where we have no district-level data, for example — the index is the best choice. There seems little reason to accept the outcome from any of the multiyear methods as useful.

Further Examples

I offer here a quick set of further examples to demonstrate the range (and limitations) of these various methods.

Mauritius: A Highly Majoritarian Electoral System?

Mauritius is an interesting case for comparative electoral scholars. In the context of complex political culture, the electoral system is the unusual "block voting" system, a multimember plurality system in which voters have as many votes as there are seats. Specifically, Mauritius has 20 districts of three seats each and one district of two seats (Rodrigues, which often has its own unique parties competing); there are also six seats distributed via a "best loser" system to candidates of designated ethnic groups.

The block voting system is usually thought of as extremely majoritarian. Parties run three candidates each and most voters vote in a party block, so all three seats can change hands on a very small

---

[28] There are a number of reasons to think that the 1997 electoral system was fairly unbiased, especially just within England (malapportionment introduces significant biases when Scottish, Welsh, and Northern Ireland seats are included). First, our estimates from 1992 indicate only a small amount of bias in the first place. Second, asserted relative losses of voter rolls for the Labour party, due to poll tax avoidance, would move the bias in the electoral system away from Labour. Third, the most recent redistricting was arguably the first one in which the Boundary Commissions were closely watched and held accountable on a national basis.

swing in average district vote. Mauritius has the further distinction of being (to my knowledge) the only democracy in which one party, the Mauritian Militant Movement (MMM) out of government after one election (1976), proceeded to win *every single seat* the following election (1982), only to fall into opposition the next (1987). For the unusual sweep election in 1982, the index can only give an estimate of $\hat{\rho} = \infty \ (\hat{\rho}^* = 1)$ — this is as majoritarian as they come.

There are reasons, however, to think that this is a misleading picture. First, in the 1982 election, the MMM won with 69% of the vote and a 44% margin. That will translate into a high proportion of seats in any system that is even slightly majoritarian. Second, if voters vote *completely* in a party block, then the system is no different from a 20-district plurality system. The relatively small number of districts might increase the majoritarianism of the system relative to a system like Canada or the UK with hundreds of districts, but it should not be as qualitatively different as if it had only a single district. Deviations from perfect party block voting then decrease majoritarianism as they allow for the possibility of split delegations.

The variable swing model has been estimated for the 1982 legislative election using data for 1976 and 1982 to calculate the variance matrices. Mauritius had, for these two elections, a three party system with the MMM, the Independence Party (IP), and the Mauritian Social Democratic Party (PMSD). Data included were party control after the previous election (number of seats held), and candidate incumbency (number of incumbents for each party running again).[29] Rodgrigues and best loser seats were ignored, leaving a total of 60 seats. Block voting requires that we make an assumption about the distribution of a party's swing among its candidates. In the presence of insufficient data on which to base any candidate-specific stochastic model (and the belief that it would make no difference whatsoever), the assumption used was one of uniform swing within the party. This means that if a candidate received more votes than other candidates of the same party in the observed elections she will continue to do so in every hypothetical election, and that any observed tendency to vote in party blocks (or not) persists in the hypothetical elections.

The calculated curve is shown in Figure 9. It is clearly majoritarian, but does not appear to be dramatically more so than the UK. Indeed, our responsiveness estimates for the two are nearly identical: $\hat{\rho} = 2.49 \ (\hat{\rho}^* = 0.64)$ with $\hat{\sigma}_{\rho} = 0.35$. The estimated bias vector is $\hat{\mathbf{b}} = (\hat{b}_{PMSD}, \hat{b}_{IP}) = (+.153, +.447)$ with standard errors $\hat{\boldsymbol{\sigma}}_b = (0.07, 0.13)$, indicating a bias *for* the two parties who won *no* seats. This is really not so incredible when we consider the national vote shares, $\mathbf{v} = (v_{MMM}, v_{PMSD}, v_{IP}) = (69\%, 6\%, 25\%)$. It just means that at equal vote shares, the MMM would have been disadvantaged; or equivalently that the PMSD or more so the IP could have won all of the seats on less than a 44% advantage. The key substantive conclusion here is that the Mauritius' block voting system is not the democratic outrage many have made it

[29] Data were constructed from those available in Mathur (1991).

out to be. The key methodological insight is that we would not have been able to extract this kind of information, especially with confidence, from any other method.

A Proportional Electoral System: Costa Rica

[I intend for the final paper to include a Costa Rican example.]

Comparative Geographic Representation

I have mentioned several times, including in the title, that there are types of representation other than that of parties relative to their votes — representation of geographic regions, ethnic groups, and so on — that are of practical and scientific interest, and which seem at a conceptual level to be similar problems. We can think of groups [parties, regions] as having some measure of size that indicates their claim on representation [votes, population]. We can think of such group representation as being majoritarian, proportional, or consensual; we can think of it as being biased or unbiased. Can the methods described here be used for such a purpose?

For the most part, no. Swing methods cannot be applied because we do not have the equivalent of an electoral system that determines group representation as groups change in size. We also do not have the equivalent of an election, not only to provide an official count of group size, but to let us know an observation has been made. This rules out multiyear methods, even if we found them acceptable otherwise.

Consider the case of geographic representation and, specifically, the apportionment of Congressional seats to the U.S. states. If we treat each election as an observation, we will have seats remaining unchanged for five observations, while population shifts around. The result would be estimates of proportional responsiveness with 49 distinct bias terms, all correlated with size (e.g., the system is specifically biased against the states that happen to be the largest.) We could, on the other hand, simply take the decennial apportionment as the observation. Periodic lawsuits aside, we know the formula by which seats are apportioned and could even develop a uniform swing method of a sort. But then we lose the ability to compare. Many countries apportion by a much more vague process and less often (both are true in the U.K.); others have no process at all. Iceland, for example, is clearly malapportioned in favor of small rural districts, because it has never been reapportioned. But we have no "observations" to put through our models.

This is where index methods are of great use. If we are willing to forego estimates of bias, and ignore the index when obvious bias forces nonsense answers, indices can provide us with reasonable answers where no other is available. In Monroe 1997b, I argue that the way a particular set of groups will be represented (majoritarian, etc.) is predictable according to certain types of collective action capabilities held by groups and factions thereof. For some types of groups the pattern is always the same. Party representation is predicted to be always majoritarian. Geographic representation is predicted to be always consensual. For other types of groups, the pattern depends on the distribution of group-specific characteristics. Ethnic group representation might then be consensual or majoritarian. Indices can be used to provide a rough guide to the accuracy of these hypotheses. For example, I report there the responsiveness index for geographic representation in 82 countries, finding 75 of them (93%) to fit the prediction, $\hat{\rho} < 1 \left(\hat{\rho}^* < 0\right)$. The handful that do not fit exactly come very close (see Figure 10).

Given the errors that will be introduced when bias is present, this seems a reasonably good correspondence. It would be worthwhile, of course, to examine a large number of party representation cases to see how highly index and variable swing estimates are correlated, as well as to identify some correspondence between the presence of bias and mistakes in the index estimates. Then we would be even more confident about using the index in these other types of situations where we have no other choice.

Conclusion

Rather than summarize the discussion to this point, I would like to conclude by highlighting several directions in which the methods described here could be further improved. Some of these are addressed in the companion paper, but several await further research. I am concerned here, in particular, with improvements that can be made in the variable swing model.

There are a number of changes in the specification of the model described here that might be fruitful. Katz and King (1997), for example, have argued that the appropriate error distribution for district-level electoral data in additive logistic form is the multivariate-T, rather than the multivariate normal described here. They, in fact, offer convincing evidence with the British general election data that this is so. A more fundamental change, which would be a departure from Gelman and King's two-party Judgeit as well, might be to take more seriously the underlying panel structure of the district-level electoral data. Since the true data-generating process is probably something like a multiple equation error components model with dynamics, and the data are rarely ideal for estimating such a model, this is not a particularly straightforward task.

The solution to the primary compositional data problem, logistic transformation, introduces problems that have been avoided in the examples here, but cannot be in most applications. The most important of these is the presence of zeros in the data: party system changes between elections, noncontesting parties, and parties that compete only regionally. In the two-party world, uncontested seats alter apparent relationships and affect our inferences if we do not correct for them; in the multiparty world, they make analysis impossible, because the logratios are undefined. The problem of noncontesting parties must be dealt with by imputing "effective vote", the vote shares or logratios that would have been won if all parties had contested. A variety of methods are available along a complexity-accuracy tradeoff from simply imputing a reasonable value to the method used by Katz and King or the MCMC method described by them (Katz and King 1997).[30] Regional parties are a different matter, however. We should not impute the effective vote of the Scottish Nationalist Party in an English seat. We need rather to recognize explicitly the regional differences in the party system when they exist. I address this issue in Monroe 1998.

These and other barriers aside, the discussion here moves us forward on a number of fronts in the comparative study of elections and representation. The variable swing method has been extended from the two-party world to the vast majority of multiparty electoral systems; extension to virtually all electoral systems is primarily a matter of finding appropriate models of votes and seats at the district level. Multiyear methods have been found wanting, but index and uniform swing methods may have some new life in them for particular applications, and should not be encouraged solely on the grounds of their simplicity or deterministic implications. While the multiparty seats-votes graphs introduced here are unlikely to become as ubiquitous as ecological inference plots, they should prove a useful tool for scholars attempting to analyze and discuss comparative electoral systems. As resources allow, it is my intention to produce and distribute a software package incorporating the various methods and graphics described herein.

[30] Which becomes computationally intensive with as few as four parties competing.

References

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

Amemiya, T. 1971. "The Estimation of Variances in Variance Components Model." *International Economic Review* 12: 1-13.

Ansolabehere, Stephen and Gary King. 1990. "Measuring the Consequences of Delegate Selection Rules in Presidential Nominations." *Journal of Politics* 52(2): 609-21.

Brookes, Ralph H. 1959. "Legislative Representation and Party Vote in New Zealand: Reflections on the March Analysis." *Public Opinion Quarterly* 23: 288-91.

Butler, David E. 1947. "Appendix III: The Relation of Seats to Votes." In R. B. McCallum and Alison ?? (eds.) *The British General Election of 1945*London: Oxford University Press.

Butler, David E. 1951. "Appendix." In H. G. Nicholas (ed.) *The British General Election of 1950* London: MacMillan.

Butler, David E. 1963. *The Electoral System in Britain since 1918* London.

Casstevens, Thomas and William D. Morris. 1972. "The Cube Law and the Decomposed System." *Canadian Journal of Political Science* 5: 521-31.

Christensen, Raymond V. and Paul E. Johnson. 1995. "Toward a Context-Rich Analysis of Electoral Systems: The Japanese Example." *American Journal of Political Science* 39(3): 575-98.

Dixon, Robert G., Jr. 1968. *Democratic Representation: Reapportionment in Law and Politics* New York: Oxford University Press.

Edgeworth, F. Y. 1898. "Miscellaneous Applications of the Calculus of Probabilities." *Journal of the Royal Statistical Society* 61: 534-42.

Eldersveld, Samuel J. 1951. "Polling Results and Prediction Techniques in the British General Election of 1950." In James K. Pollock (ed.) *British Election Studies* Ann Arbor: George Wahr.

Garand, James C. and T. Wayne Parent. 1991. "Representation, Swing, and Bias in U. S. Presidential Elections, 1872-1988." *American Journal of Political Science* 35(4): 1011-31.

Gelman, Andrew and Gary King. 1994. "A Unified Method of Evaluating Electoral Systems and Redistricting Plans." *American Journal of Political Science* 38(2): 514-54.

Greene, William H. 1997. *Econometric Analysis* (3rd ed.) Upper Saddle River, NJ: Prentice Hall.

Grofman, Bernard, Thomas Brunell, and Janet Campagna. 1997. "Distinguishing Between the Effects of Swing Ratio and Bias on Outcomes in the U. S. Electoral College, 1900-1992." *Electoral Studies*

Gudgin, G. and P. Taylor. 1979. *Seats, Votes and the Spatial Organisation of Elections* London: Pion.

Guinier, Lani. 1994. *The Tyranny of the Majority: Fundamental Fairness and Representative Democracy* New York: Free Press.

Jackman, Simon. 1994. "Measuring Electoral Bias: Australia, 1949-93." *British Journal of Political Science* 24: 319-57.

Judge, G., C. Hill, W. Griffiths, and T. Lee. 1985. *The Theory and Practice of Econometrics* New York: Wiley.

Katz, Jonathan N. and Gary King. 1997. "A Statistical Model for Multiparty Electoral Data." Paper presented to Midwest Political Science Association. Chicago, April.

Katz, Richard. 1997. *Democracy and Elections* Princeton: Princeton University Press.

Kendall, M. G. and A. Stuart. 1950. "The Law of Cubic Proportions in Election Results." *British Journal of Sociology* 1: 183-96.

Kendall, M. G. and A. Stuart. 1952. "La Loi du Cube dans les Elections Britanniques." *Revue Française de Science Politique* 2: 270-6.

King, Gary and Andrew Gelman. 1991. "Systemic Consequences of Incumbency Advantage in U. S. House Elections." *American Journal of Political Science* 35(1): 110-38.

King, Gary and Robert X Browning. 1987. "Democratic Representation and Partisan Bias in Congressional Elections." *The American Political Science Review* 81(4): 1251-73.

King, Gary. 1989. "Representation through Legislative Redistricting: A Stochastic Model." *American Journal of Political Science* 33(4): 787-824.

King, Gary. 1990. "Electoral Responsiveness and Partisan Bias in Multiparty Democracies." *Legislative Studies Quarterly* 15(2): 159-81.

Laakso, Markku. 1979. "Should a Two-and-a-Half Law Replace the Cube Law in British Elections?" *British Journal of Political Science* 9: 355-62.

Lijphart, Arend. 1984. *Democracies: Patterns of Majoritarian and Consensus Government in Twenty-One Countries* New Haven: Yale University Press.

Linehan, William J. and Philip A. Schrodt. 1978. " New Test of the Cube Law." *Political Methodology* 4: 353-67.

March James G. 1957. "Party Legislative Representation as a Function of Election Results." *Public Opinion Quarterly* 21(4): 521-42.

Mathur, Hansraj. 1991. *Parliament in Mauritius* Mauritius: Editions de l'Ocean Indien.

Mátyás, László. 1996. "Error Components Models." In László Mátyás and Patrick Sevestre (eds.), *The Econometrics of Panel Data: A Handbook with Applications* (2d ed.). Dordrecht: Kluwer.

McHenry, Dean. 1955. "The Australian General Election of 1954." *Australian Quarterly* 27: 1: 14-23.

McLean, Iain. 197?. *Political Studies*

Monroe, Burt L. 1997a. "Compositional Data Analysis: Methodology for Distributive Political Data." Paper presented to Midwest Political Science Association. Chicago, April.

Monroe, Burt L. 1997b. "Mirror Representation in the Funhouse: Systematic Distortions in the Legislative Representation of Groups." Paper presented to the International Political Science Association. Seoul, August.

Monroe, Burt L. 1998. "A Method for Evaluating Multiparty Electoral Systems and Redistricting Plans." Typescript, Indiana University.

Mooney, Christopher J. 1996. "Bootstrap Statistical Inference: Example and Evaluations for Political Science." *American Journal of Political Science* 40(2): 570-602.

Niemi, Richard and Patrick Fett. 1986. "The Swing Ratio: An Explanation and an Assessment." *Legislative Studies Quarterly* 11: 75-90.

Niemi, Richard and Simon Jackman. 1991. "Bias and Responsiveness in State Legislative Redistricting." *Legislative Studies Quarterly* 16: 183-202.

Niemi, Richard G. and John Deegan, Jr. 1978. "A Theory of Political Districting." *American Political Science Review* 72(4): 1304-23.

Qualter, Terence H. 1968. "Seats and Votes: An Application of the Cube Law to the Canadian Electoral System." *Canadian Journal of Political Science* 1(3): 336-44.

Rydon, Joan. 1957. "The Relation of Votes to Seats in Elections for the Australian House of Representatives, 1949-1954." *Political Science* 9(2): 49-61.

Schrodt, Philip A. 1981. "A Statistical Study of the Cube Law in Five Electoral Systems." *Political Methodology* 7: 31-53.

Smith, James Parker (Rt. Hon.). 1909. *Minutes of Evidence before the Royal Commission on Systems of Elections*, 19 May. (U.K. House of Commons.) [Smith attributes the cube law to Maj. P. A. MacMahon.]

Taagepera, Rein and Marku Laakso. 1980. "Proportionality Profiles of west European Electoral Systems." *European Journal of Political Research* 8: 423-46.

Taagepera, Rein. 1973. "Seats and Votes: A Generalization of the Cube Law of Elections." *Social Science Research* 2: 257-75.

Taagepera, Rein. 1986. "Reformulating the Cube Law for Proportional Representation Elections." *The American Political Science Review* 80(2): 489-504.

Taagepera, Reing and Matthew Soberg Shugart. 1989. *Seats and Votes: The Effects and Determinants of Electoral Systems* New Haven: Yale University Press.

The Economist 1950 (7 January). "Electoral Facts." pp. 5-7.

Theil, Henri. 1969. "The Desired Political Entropy." *American Political Science Review* 63(2): 521-5.

Theil, Henri. 1970. "The Cube Law Revisited." *Journal of American Statistical Association* 65: 1213-19.

Tufte, Edward R. 1973. "The Relationship between Seats and Votes in Two-Party Systems." *American Political Science Review* 67(2): 542-54.

Table 1 - Multiyear Methods

Three-Party British General Elections, 1950-1992

Multiparty Ordinary Least Squares (MPOLS)

Parameters	**Estimate (s.e.)**
ρ	1.31 (0.05)
b_2 (Bias for Labour)	0.01 (0.14)
b_3 (Bias for Liberals)	-1.44 (0.15)

$N = 13$

$R^2 = 0.97$

Multinomial Bilogit (MNBL)

Parameters	**Estimate (s.e)**
ρ	1.02 (0.15)
b_2 (Bias for Labour)	-0.05 (0.04)
b_3 (Bias for Liberals)	-2.12 (0.14)

$N = 13$

$\ln L = -475.5$

Constrained Seemingly-Unrelated Regression (CSURE)

Parameters	**Estimate**
ρ	0.57 (0.24)
b_2 (Bias for Labour)	-0.09 (0.15)
b_3 (Bias for Liberals)	-2.60 (0.41)

$N = 13$

$\ln L = -0.177$

(The CSURE model also estimates three ancillary parameters not reported here.)

Figure 1 -- Responsiveness Index
Three-Party British General Elections, 1950-1992

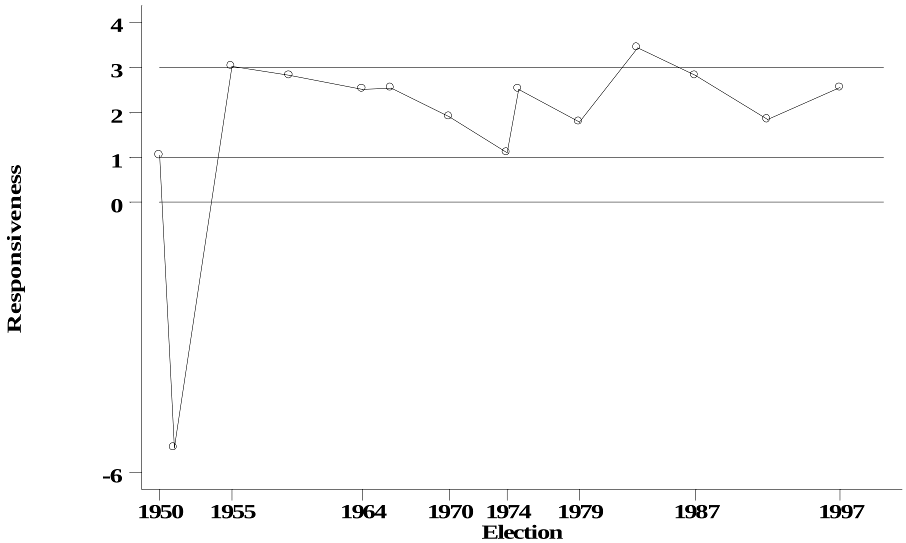

Figure 10 -- Geographic Representation, 82 countries