13

District Magnitude and the Comparative Study of Strategic Voting

Thomas Gschwend

Introduction[1]

Do electoral systems matter? The political consequences of electoral laws fall in two distinct categories. They encompass direct as well as indirect effects. The particular rules which determine how votes are generated into legislative seats have a direct impact on the number and the type of parties in a given polity. This has profound and well-known consequences for the type of government and the nature of representation in general. It is well known that the same distribution of votes can be translated in totally different distributions of seats in parliament using different electoral rules. If the outcome of an election is not just a foregone conclusion then the differences in the way votes are translated into seats may be a crucial determinant deciding who will govern and who has to stay put.

What is the impact of electoral rules, though, on the way people make decisions in the voting booth? Do voters actually care about electoral rules? Do such rules in some sense shape their electoral choice because they anticipate the outcome of an election and include these expectations in their decision calculus? If voters are systematically drawn away from their most preferred party, just because they realize that supporting a marginal party might be equivalent to wasting their vote given a particular electoral institution, then we speak of an indirect effect. Duverger's (1954) "psychological effects" are the prime example for these types of effects. In order to avoid wasting their votes, voters cast a *strategic* vote for a viable party (or candidate)[2] although they most prefer another one.

Duverger suggested that this logic should not apply to PR systems, since even marginal parties can expect to gain seats in such a system.

Contrary to Duverger's propositions, Leys (1959) and Sartori (1968) expect significant amounts of strategic voting even in PR systems—the more the smaller the district magnitude, that is, the less seats are awarded per electoral district. The Leys–Sartori conjecture posits that the various electoral institutions can be arrayed along a single dimension defined by the district magnitude and predicts that the smaller the district magnitude the more strategic voting we should expect at the primary district level, that is, at the level of the smallest geographic unit in which seats are allocated. The consequences of the frequency of strategic voting given varying district magnitudes, to my knowledge, have never been tested comparatively. Does the frequency of strategic voting at the electoral district depend on the number of seats that are awarded? In order to answer this question this chapter considers first the individual level and then aggregate voting decisions with regard to electoral district in order to be able to estimate the impact of district magnitude on the frequency of strategic voting.

The contributions of this chapter are threefold. First, based on a theory of how voters form expectations about the election outcome in their electoral district I propose a measure to operationalize strategic voting across more than 30 election studies using the Comparative Study of Electoral Systems (CSES) data Module 1. Second, I will test hypotheses about the relationship between the frequency of strategic voting and institutional incentives that are channeled cross-nationally through district magnitude. Third, I further provide some evidence that speaks directly to the controversy in the literature surrounding the question of how district magnitude effects should be modeled. Results yield support for the claim that district magnitude and frequency of strategic voting at the district level are negatively correlated.

A Comparative Look at Strategic Voting—Some Micro-Foundations

No matter whether you believe in the Columbia, Michigan, or Rochester school of thought, traditional theories of voting behavior have in common the prediction that voters should end up casting a vote for their most preferred party (or candidate). This is called a *sincere vote*. Students of strategic voting point out that we nevertheless observe systematic

deviations from these traditional vote-choice predictions. In an attempt to model these deviations they suggest that voters do not merely take into account the utility that a voter derives from voting for her most preferred party (U_{pref}) but also the expectation about the outcome of the election, for instance whether the most preferred party is actually a viable alternative to win a seat in her primary electoral district (Blais 2002; Blais et al. 2001; Cox 1997; Fisher 2004). It is far from clear how voters actually form and weigh their expectations against their preferences. It is quite likely that different voters employ different decision rules.

The particular approach followed here is to assume that a voter's decision rule is to vote for a party that maximizes her expected utility from voting for viable parties. Thus, a *strategic voter* is someone who votes for a less-preferred party if the expected utility that this party is likely to gain a seat in their district is higher than the expected utility derived from a sincere vote, namely that their most preferred party has a viable shot at a seat in that district.

The probability that a voter expects her most preferred party to be viable to win a seat is denoted by p_{pref}. The expected utility, EU(pref), that her most preferred party is competitive as a viable alternative to gain a seat combines the traditional utility component weighted by the voter's expectation. Thus EU(pref) = $p_{pref} \cdot U_{pref}$. This also implies that with probability $1 - p_{pref}$ no gain will be realized from voting for her most preferred party. If the voter does not expect his or her most preferred party to be viable then he or she might cast a strategic vote for a less-preferred party that is expected to be viable (i.e., $p_{pref} < p_{strat}$) in order to avoid wasting his or her vote. Given that a strategic choice cannot be the voter's most preferred option, the utility from voting strategically has to be lower, that is, $U_{strat} \leq U_{pref}$.

Moreover, *not* voting for someone's most preferred party might induce cognitive dissonances (Festinger 1957), although voters, of course, are motivated to avoid that. In general perceived cognitive dissonance does not need to have behavioral consequences per se. For instance, from public opinion polls we know that people value public spending and hate paying taxes. People appear too easily to square with facts that stay in logical contrast. What is needed for a cognitive dissonance to arouse and to yield behavioral consequences? There has to be an "aversive event" (Cooper and Fazio 1984: 232) that the voter expects to happen when casting a strategic vote. Arousal of cognitive dissonance might actually prevent such a behavior at the polls if despite being important to the voter, the perceived consequences of such a vote-choice are deemed to be

rather unfavorable. Such an "aversive event" could be that *not* voting for the voter's most preferred party is perceived as a threat to the voter's self-esteem or is expected to lead to an outcome that is counter to the voter's self-interest. Voters, however, might be able to a priori reduce the costs of a strategic vote. They could justify their voting behavior by attributing the responsibility—not voting for their most preferred party—to the specific decision-making situation. Clearly, some voters are likely to perceive the decision-making situation, which is prestructured by the electoral rules, as being in some way coercive (Cooper and Fazio 1984: 236–7).

Therefore it appears safe to assume that voting for a party other than the most preferred party imposes additional costs (c) to the voter independent of the expected outcome of the election. The expected utility of a strategic vote depends, consequently, on the expected gain and the costs of a strategic vote. Thus $EU(strat) = p_{\text{strat}} \cdot U_{\text{strat}} - c$.

When can we expect a voter to deviate from their most preferred party? Following the expectation maximization decision logic, a voter casts a strategic vote if and only if $EU(strat) > EU(pref)$, that is, if:

$$p_{\text{strat}} \cdot U_{\text{strat}} - c > p_{\text{pref}} \cdot U_{\text{pref}} \tag{1}$$

or equivalently, if

$$(p_{\text{strat}} \cdot U_{\text{strat}} - c)/p_{\text{pref}} > U_{\text{pref}} \tag{2}$$

The left hand side of this inequality can be interpreted as the *risk* of casting a strategic vote. Voters, then, are predicted to cast a strategic vote if these risks outweigh the potential gains from a sincere vote. Given the utility and the costs that are expected to come with a strategic vote as opposed to a sincere vote, the crucial factor for voters in deciding whether to desert or to stick with their most preferred party is the expected probability of their most preferred party's chances for winning a seat in their electoral district relative to the expected probability that a strategic vote is not wasted. Assuming that voters consider only viable parties as potential beneficiaries of a strategic choice, that is, they expect $p_{\text{strat}} = 1$, and holding utilities as well as costs constant, the key result from Eq. (2) is as follows: the more uncertain voters are whether their most preferred party is likely to win, that is, the lower the expected probability p_{pref}, the greater the left hand side of this inequality and, consequently, the more likely strategic a vote becomes.

What factors determine these expectations? Voting behavior is no different from any other behavior in that it can be explained by institutional as well as dispositional factors. I am going to distinguish between

dispositional and institutional criteria of how voters generate expectations about the probability that their preferred party is likely to gain a seat. Dispositional criteria have on the one hand to do with intrapersonal psychological motivations, with the ability to understand various institutional factors and employ them in the decision-making process. On the other hand, voting decisions have to do with the use of appropriate decision heuristics. Party elites or the media are likely to provide voters with cues, and as "cognitive misers" (Fiske and Taylor 1991) voters can simply rely on various heuristics to simplify the decision-making process (Gschwend 2004: 22–4). Dispositional factors are necessary in order to explain the variance of how voters generate their preferences and costs, as well as estimate the expected probabilities.

Here, however, individual-level determinants are taken as a starting point and aggregate the respondents' vote choices to the electoral district level. In doing so omitted dispositional effects are implicitly averaged over in order to try to predict the causal effect of institutional criteria. The purpose is to see if incentives of a given institutional design make the use of the wasted vote strategy at the electoral district level more or less likely.

Institutional Criteria and the Duvergerian Logic

Can we predict the level of strategic voting that should occur in a given decision context? Contrary to the approach taken here, the literature on institutional effects on elections typically does not focus on voters themselves but merely on the predictive implications of their hypothetical strategic behavior on the number of parties (e.g., Amorim Neto and Cox 1997; Clark and Golder 2006; Mozaffar, Scarritt, and Galaich 2003). The first reference point in the literature is Duverger (1954) who discusses the impact of institutional factors. In particular, he focuses upon the reductive effect of electoral systems on party systems due to the mechanism whereby voters try to avoid wasting their vote and cast a strategic vote for a less-preferred party which they believe has a chance of gaining representation. Given the workings of Duverger's proposed dichotomy—plurality systems produce strategic voting while PR systems do not—the "psychological" effects anticipating the "mechanical" effects of a given institutional decision context should operate at least on two levels: party elites and voters. Parties have to decide whether to compete in a given election, form a preelectoral coalition (Golder 2006; Gschwend and Hooghe 2008) or endorse yet another party or coalition that is effectively

competing for seats. Depending on how party elites coordinate their entry into the electoral market the menu or choice-set (Ben-Akiva and Boccara 1995) may differ even within the same institutional context. For voters the expected probability p_{pref} that their most preferred party is viable therefore depends on the choices offered to them on the ballot.

Duverger would nevertheless predict that the expected probability p_{pref} is constant within an electoral system while in terms of disposition there should be variance of how voters generate their preferences, costs and how they estimate expected probabilities. Implicitly averaging those dispositional factors, Duverger's theory would predict that the expected probability p_{pref} that a given party is viable is higher in PR systems than in plurality systems.

Contrary to Duverger, Leys (1959: 13.3) suggests that the effect of institutional factors varies across districts because a vote for a nationally small party might not be automatically wasted in every electoral district. Electoral support for a given party is often not uniformly distributed across all electoral districts. There are electoral strongholds where even a nationally small party is likely to gain seats. Consequently Leys would predict that the expected probability p_{pref} that an average voter's most preferred party in dispositional terms is viable should vary across electoral districts even within the same electoral system. Sartori (1968: 278) similarly argues that "... the influence of PR merely represents an enfeeblement of the same influence that is exerted by the plurality systems." He thus expects significant amounts of strategic voting even in PR systems.

The Leys–Sartori conjecture becomes relevant for the discussion of institutional factors that influence voters' expectations of the probability that their vote is not wasted on their preferred party. It posits that various electoral institutions can be arrayed along a single dimension defined by the district magnitude (i.e., by the number of seats awarded in each electoral district). The prediction is that the higher the district magnitude, the less likely voters are to avoid wasting their vote for smaller parties and, hence, the less strategic voting is expected to occur in that district. To put it differently, the larger the district magnitude the higher the expected probability p_{pref} that an average voter (in terms of potential dispositional criteria) believes their most preferred party is viable. According to Eq. (2), the higher the expected probability p_{pref} the less likely such an average voter will be to deviate from their most preferred party in order to cast a strategic vote.

Finally, it is thought that forming expectations as to whether a particular party is viable is a difficult task for voters. Some scholars argue

that strategic voting should fade out when district magnitude is greater than 5 because it becomes too complicated to generate expectations about which party is able to gain representation (Cox 1997: 100; Cox and Shugart 1996: 311). Evidence to support this claim comes from empirical regularities based on Japanese and Colombian district level results (Cox 1997; Cox and Shugart 1996) as well as electoral returns in Spanish districts (Cox 1997: 115–7; Gunther 1989). Despite the evidence it remains somewhat unclear, however, why voters in larger districts suddenly systematically overestimated the expected probability p_{pref} of their preferred party's electoral viability in order to vote sincerely for their preferred party rather than strategically. To sum up, the literature agrees that there is a hypothetical negative relationship between district magnitude as the institutional criterion and the frequency of strategic voting in determining voters' behavior.

Besides this general trend with regard to district magnitude and strategic voting, the literature elaborates on two different functional forms of this relationship. Some scholars assume a simple linear relationship (e.g., Cox 1997; Cox and Shugart 1996) while others argue (e.g., Benoit 2001; Ferrara, Herron, and Nishikawa 2005; Monroe and Rose 2002; Taagepera and Shugart 1989) that the marginal effect of district magnitude on the frequency of strategic voting will diminish as the magnitude increases. This is consistent with the idea that the expected difference in the frequency of strategic voting between a single-member district (as for districts in the United States, UK, or Canada) and a district with magnitude of 11 (as in some districts of Slovenia, Belgium, Sweden, or Spain) is more consequential and not at all negligible than in large districts. In districts with a magnitude of, say 30 or 40, voters should expect their most preferred party to gain representation anyway. No strong reduction in the frequency of strategic voting is expected.

Data and Measurement

The Leys–Sartori conjecture has never been tested with individual-level data. Most studies in the literature on institutional effects on elections employ cross-national data in order to pin down the relationship between district magnitude and the size of the party system (e.g., Amorim Neto and Cox 1997; Clark and Golder 2006; Mozaffar, Scarritt, and Galaich 2003). Scholars who look more closely at strategic voting use district level rather than national level data (e.g., Cox 1997; Cox and Shugart 1996;

Gschwend 2007; Herron and Nishikawa 2001). Nevertheless, employing district level data is only an indirect way to assess an individual level phenomenon like strategic voting. Heroic assumptions about voters' preferences as well as the well-known problems of ecological fallacy plague the process of making inferences based on such a research design. Moreover, different strategic voting patterns might even cancel out in the aggregate and are therefore lost from any analysis geared at this level of observation. Thus, on theoretical grounds, if one is interested in investigating effects of electoral institutions on voting behavior, the individual level is the preferred level of observation to carry out analyses of strategic voting. With survey data it is possible to measure (sincere) preferences of a given respondent directly and compare it to their stated voting behavior. This is a great advantage compared to all studies that look only at aggregated election results because one does not need to make any additional assumption about voters' preferences in order to distinguish strategic from other voting behavior.

The CSES project is an ideal data set for this approach. It is a cross-national project with election studies across countries with great variance in their electoral institutions, variance which also provides comparable individual level data. Moreover, systematic information about characteristics of the primary electoral districts as well as the electoral system at large is merged to the individual data. Thus, the CSES data (Module 1) is especially suitable for study of the effects of electoral institutions on citizens' attitudes and behavior.

The comparative literature on strategic voting and electoral systems traditionally speaks to the (primary electoral) district level because this is the level where the institutional effects should operate. I will choose the same level of observation in order to assess the consequences of varying district magnitude on the frequency of strategic voting.

The dependent variable is the fraction of all voters per electoral district who cast a strategic vote. In order to construct this variable it is necessary to derive voters' preference rankings of parties which actually field lists or candidates in a particular electoral district, that is, *after* elite coordination took place that might have reduced the number of available options on the ballot.[3] This accounts for the complications that even within the same country voters do not necessarily have the same choice-set and that their vote choices might be menu dependent. Party preferences are measured by standard 10-point party like–dislike scales and ranked accordingly for each respondent.[4]

According to my conceptualization, a strategic vote following the Duvergerian logic is a vote for a less-preferred party if the expected utility that this party is electorally viable is higher than the expected utility of the preferred party gaining a seat in the district. Unfortunately, it cannot be directly assessed how individuals form their expectations about the viability of a party, no matter how they weigh their preferences against those expectations. This holds in most CSES countries where the common module was administered as part of a postelection study. Thus I have to employ some simplifying assumptions.

In a single-member district the two parties expected to be first and second are considered viable to gain this seat (Cox 1997).The larger the district magnitude the more parties will be viable. Conceptually, voters have to calculate the expected probabilities for their most preferred party to get the last seat in a multimember district in order to decide whether their vote might be wasted. Particularly in large districts with many parties this will be quite difficult. Given the complexity it might be more reasonable to assume that "viability" of a given party is perceived differently in such districts. It is assumed that voters simply form expectations, whether or not parties gain a seat in a particular electoral district. As such, parties that are expected to win a seat are perceived as viable parties in that electoral district. Employing this heuristic is easier than calculating expected probabilities for parties winning the last district seat and, moreover, it is easily available since voters can infer this from previous election results.

There is also a methodological warning associated with attempts to operationalize the concept of "viability" for parties in multimember districts using CSES data. These data cover vote shares of up to six parties at the district level. There is no information in the CSES data, however, as to whether those parties in fact actually came first, second, third, and so forth. The parties covered by CSES are not automatically the most successful parties in every electoral district. There is always the possibility that independent candidates or parties not covered by the CSES could have been more successful in a particular district than parties that are covered by the CSES. Thus from ranking district-level results of the available parties one cannot reliably asses the "viability" of a given party.

In order to get a measure for voters' expectations about a party's electoral viability the concept was defined as "coming in first or second" in single-member districts and as "gaining at least a seat" in multimember districts. Consistent with prior research (e.g., Cain 1978; Gschwend 2004; Karp et al. 2002: 8), it is assumed that on average voters' expectations are

correct, that is, they expect a party to be viable (or to gain a seat) if the party actually ends up first or second (or winning a seat) in that district.

Consequently, the dependent variable is coded as the proportion of respondents per electoral district who cast their vote for a less-preferred party if that party comes in first or second (in single-member districts), or wins a seat in their electoral district (in multimember districts) when the preferred party does not. This group of strategic voters is most likely to follow the Duvergerian logic to avoid wasting their vote.[5] The advantage of such strategic voting is that it disentangles strategic voters following a wasted-vote strategy from voting behavior that can be interpreted as a result of other strategies (Blais et al. 2001). Thus the frequencies of strategic voting are not falsely magnified as if we would take, for instance, simply every deviation from someone's most preferred party as a strategic vote. In order to construct a measure of which party gained seats in a given electoral district this information was compiled separately from country-specific data sources and merged with the CSES data. The group of nonstrategic voters is comprised of all other voters, for example, sincere voters or voters of a party that is on the ballot in a respective electoral district but not being evaluated on the corresponding party like–dislike scale.

Some Descriptive Results

The empirical section of this chapter begins by providing an overview of the independent and the dependent variables. In the following analysis, all CSES election studies are included which passed a data consistency test and provided the relevant variables. Thus, countries without any parliamentary vote-choice variable were not included in the analysis (Belarus, Chile, Lithuania, Peru 2000); nor were countries where district level information is not available (Taiwan, Korea, Russia, Ukraine, and Thailand).

There are 1,949 electoral districts in the CSES election studies where seats are distributed at least partly on the local district level. The district magnitude varies between 1 and 48. The distribution of this variable is extremely skewed. About 80 percent of the observations have a district magnitude of 1. However, the respective seat allocation rules that determine the winner in such districts vary to some extent. Besides the single-member plurality districts of Canada, the UK, and the United States, there are also Australian alternative vote districts as well as the SMD-tier

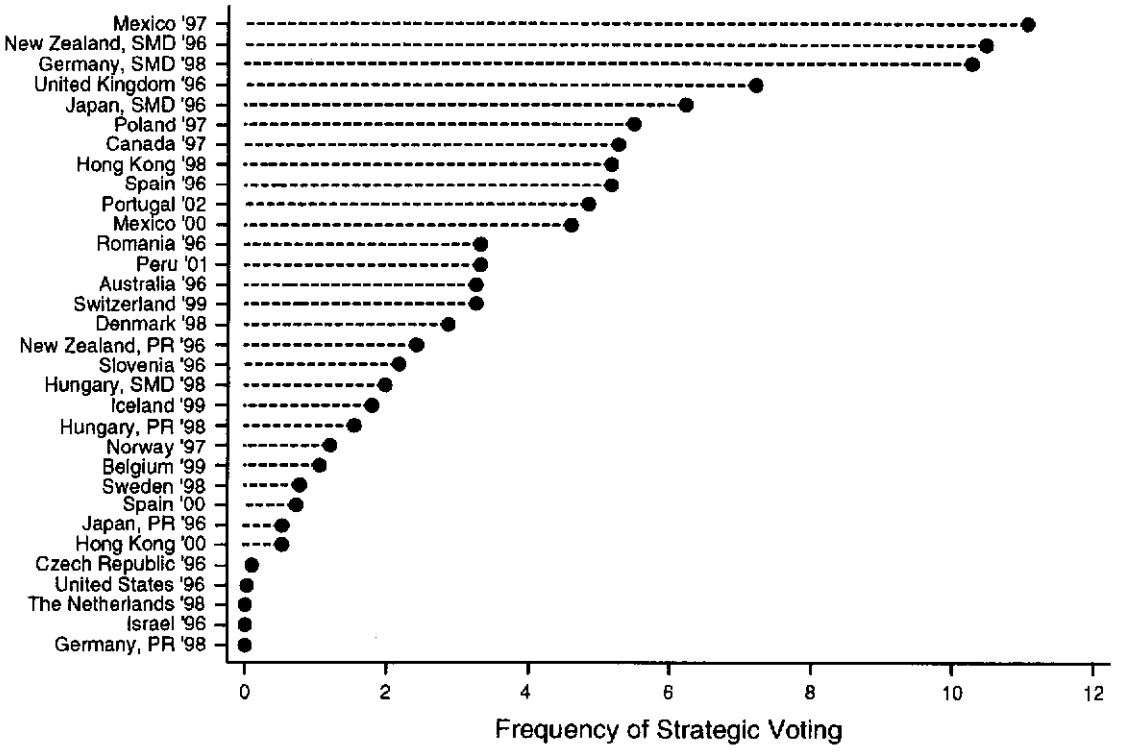

Figure 13.1. Frequency of strategic voting by election study—CSES Module 1

districts of all mixed-electoral systems (including the single ballot system of Mexico). If the Netherlands and Israel are also included—two countries where the primary electoral district is at the national level, and the available PR-tier districts of the two-ballot mixed electoral systems that are covered by CSES Module 1 (Germany, Hungary, Japan, and New Zealand)—the number of observations increases by 35 to 1,984. Those PR-tier districts have district magnitudes that range from 4 (Hungarians' PR-tier is composed of regional multimember districts) to 656 (Germany).

The dependent variable is the proportion of respondents per electoral district who cast their vote for a less-preferred party if that party comes in first or second (in single-member districts), or a party which wins a seat in their electoral district (in multimember districts) when the preferred party does not. In order to capture the distribution Figure 13.1 provides summary statistics while summing up the observed levels of strategic voting at the electoral districts within every election study.

Figure 13.1 shows that there is considerable variation in the frequency of strategic voting even on a more aggregate level. It is reassuring that based on my measurement strategy one does not find any strategic voting where votes are essentially never wasted. In neither of two PR systems

with very low thresholds—Israel and the Netherlands—is strategic voting discernable. At most 10–12 percent of the voters follow the wasted-vote logic. These high rates are observed in the SMD-tier of some mixed-electoral systems. The variation is even stronger at the electoral district level. In the following section electoral districts are chosen as the level of analysis because the hypothesized institutional effects should be present at this level.

District Magnitude and the Frequency of Strategic Voting

What is the relationship between district magnitude and the frequency of strategic voting? Theory suggests that it should be a negative relationship: the lower the district magnitude the higher a voter's expectation that their vote will be wasted because parties find it more difficult to win seats. So far, there is no agreement reached about the functional form. Moreover we should expect a sudden decline of strategic voting in electoral districts with a district magnitude greater than 5 if the "fading-out" argument is correct. The CSES provides an opportunity to examine these issues empirically.

The fraction of strategic voting per district is calculated over a different number of grouped individuals and bounded between 0 and 1. I follow the advice of the econometric literature on how to deal with this type of response data (e.g., Papke and Wooldridge 1996) and will later employ a generalized linear model (GLM) with a logit link. This particular estimation strategy makes it possible to appropriately model the binomial data generation process at the electoral district level, while at the same time accounting for the fact that the precision of those fractions depends on the number of respondents within each electoral district. The logit link finally makes sure that the model predictions are bounded between 0 and 1.

When modeling the fraction of strategic voting per district the current theory does not offer any clear guidance as to which functional form for the district magnitude should be used. Therefore, I start by fitting a slightly more flexible model, a generalized additive model (GAM) (Beck and Jackman 1998), to the data to avoid any parametric restrictions for district magnitude as the sole predictor of the expected frequency of strategic voting at the district level.

Figure 13.2 displays the fractions of strategic voting as predicted by a smooth function (estimated through a cubic smoothing spline; based

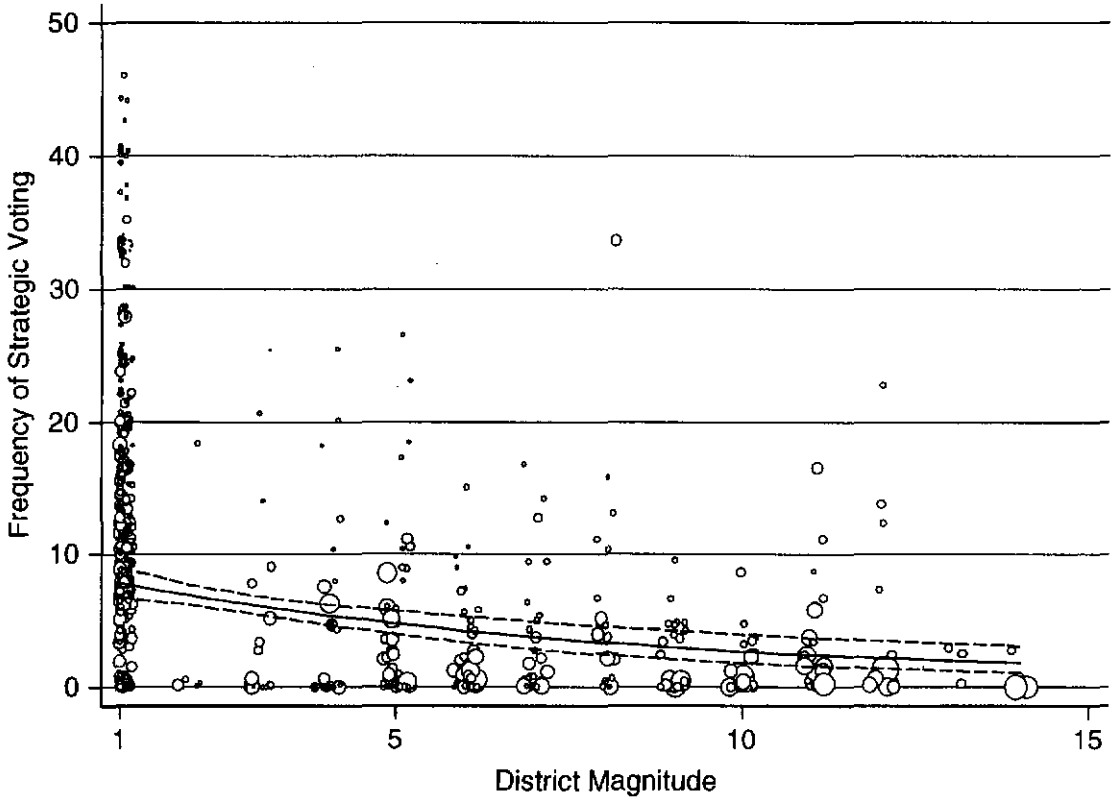

Figure 13.2. Smooth function of district magnitude and the frequency of strategic voting

on 5 df) of the district magnitude together with pointwise 95 percent confidence intervals as solid and dashed lines, respectively. While the circles represent the actual district level fractions of strategic voting, the size of the circles is proportional to the number of respondents that were interviewed in that district. To maximize readability outlying districts (about 6 percent) are excluded from Figure 13.2.

The analysis reveals that it is very difficult to obtain a precise prediction of the level of strategic voting in the districts based on the district magnitude as sole predictor. The variability of the observed fraction of strategic voting is quite high even for electoral districts of the same district magnitude. Nevertheless, electoral districts that have unexpectedly high fractions of strategic voting are mostly displayed with small circles indicating that these fractions are based on small numbers of respondents only.

In general though, the figure supports the expectation that there is a negative relationship between district magnitude and the number of strategic voters in a given electoral district. This is consistent with the theory that the larger the district magnitude the higher voters' expectations

p_{pref} that their most preferred party is viable in that district and, consequently, the less likely they are to cast a strategic vote. Moreover, the expected decrease of strategic voting seems to be rather smooth. There is no evidence, at least in these data, for the argument advanced by Cox and Shugart (1996) and Cox (1997) of a sudden decline of strategic voting in electoral districts with a district magnitude greater than 5. Instead, there is some strategic voting even in electoral districts of large district magnitude. The analysis thus far has moved beyond the dominant case study design logic that characterizes the literature on strategic voting, establishing that there is a negative relationship between district magnitude and the frequency of strategic voting even if one looks at electoral districts cross-nationally.

What can be said about the functional form of the relationship between district magnitude and the frequency of strategic voting? The literature does not offer clear guidance. Comparing the model fit of a GLM with a GAM using the same dependent and independent variables allows one to assess how reasonable the linearity constraint is for the predictors *DISTRICT MAGNITUDE* or *log(DISTRICT MAGNITUDE)* in a GLM. If all predictors in a GAM are modeled linearly (i.e., df = 1) then such a model is equivalent to a GLM. Now, if the deviance increases (significantly) when a linear predictor is used instead of a smooth function (i.e., df > 1), that is, the model fit gets worse, then the smooth functions of the predictors show significant signs of nonlinearity. Appropriate significance tests show that there are neither significant nonlinearities when one uses *DISTRICT MAGNITUDE* nor *log(DISTRICT MAGNITUDE)* for electoral districts where seats are distributed at least partly on the local district level and the district magnitude varies between 1 and 48. Consequently for such electoral districts the linearity constraint of the predictors is not really consequential substantively. Scholars can employ either functional form, *DISTRICT MAGNITUDE* or *log(DISTRICT MAGNITUDE)*. No gain can be made by going nonparametric. However, if the PR-tier districts are included, and, consequently, the district magnitude ranges between 1 and 656, *DISTRICT MAGNITUDE* shows signs of nonlinearity while *log(DISTRICT MAGNITUDE)* does not. This implies that when adding those 35 PR-tier districts to the sample—some of which have very large district magnitudes (New Zealand: 120; Israel: 120; the Netherlands: 150; Germany: 656)—scholars should rather use *log(DISTRICT MAGNITUDE)* instead of *DISTRICT MAGNITUDE* as a predictor when modeling such effects.

All told, the assessment of the controversy in the literature about the appropriate functional form when modeling district magnitude yields

Table 13.1. Generalized linear models predicting the frequency of strategic voting as a function of district magnitude

	Dependent variable: fraction of strategic voting			
	Excluding PR	All districts	Excluding PR	All districts
DISTRICT MAGNITUDE	−0.104 (0.032)**	−0.015 (0.010)		
log(DISTRICT MAGNITUDE)			−0.583** (0.071)	−0.492 (0.115)**
Constant	−2.622 (0.097)**	−3.036 (0.081)**	−2.631 (0.046)**	−2.671 (0.062)**
AIC	2.20	2.45	2.16	2.24
BIC	−11,996	−11,740	−12,074	−12,157
N	1949	1984	1949	1984

Robust standard errors in parentheses; * Significant at 5%; ** Significant at 1%.

a Solomonic sentence at least in light of the dependent variable used here. As long as the district magnitudes of the electoral districts are not greater than 50, that is, for almost all electoral districts that are covered by the CSES module, it does not make a huge difference whether *DISTRICT MAGNITUDE* or *log(DISTRICT MAGNITUDE)* is used. This is true as long as there is an appropriate link function that permits out-of-bound predictions.

Finally it is worthwhile to look at the estimation results from a GLM predicting the level of strategic voting conditional on institutional effects that get channeled through the district magnitude. To facilitate a comparison across functional forms—either *DISTRICT MAGNITUDE* or *log(DISTRICT MAGNITUDE)* as predictor, as well as samples which either exclude ($n = 1,949$) or include ($n = 1,984$) the PR-tier districts—Table 13.1 presents the estimation results across all four models.

Three out of four models yield essentially the same result. The incentives that get channeled through the district magnitude are in fact systematically related to the frequency of strategic voting at the electoral district level. No matter which functional form is used for the predictor variable, the relationship is negative: lower district magnitudes yield more strategic voting. Merely the inclusion of the large PR-tier districts of New Zealand, Israel, the Netherlands, and Germany cause problems when one attempts to model *DISTRICT MAGNITUDE* without transforming it. Moreover, the smaller samples excluding all PR-tier districts always yield better predictions given the presented fit indices even when the same functional form is used. Smaller values for the Akaike (AIC) as well as the Bayesian (BIC) information criterion, indicate better fitting models. Finally, the model fit

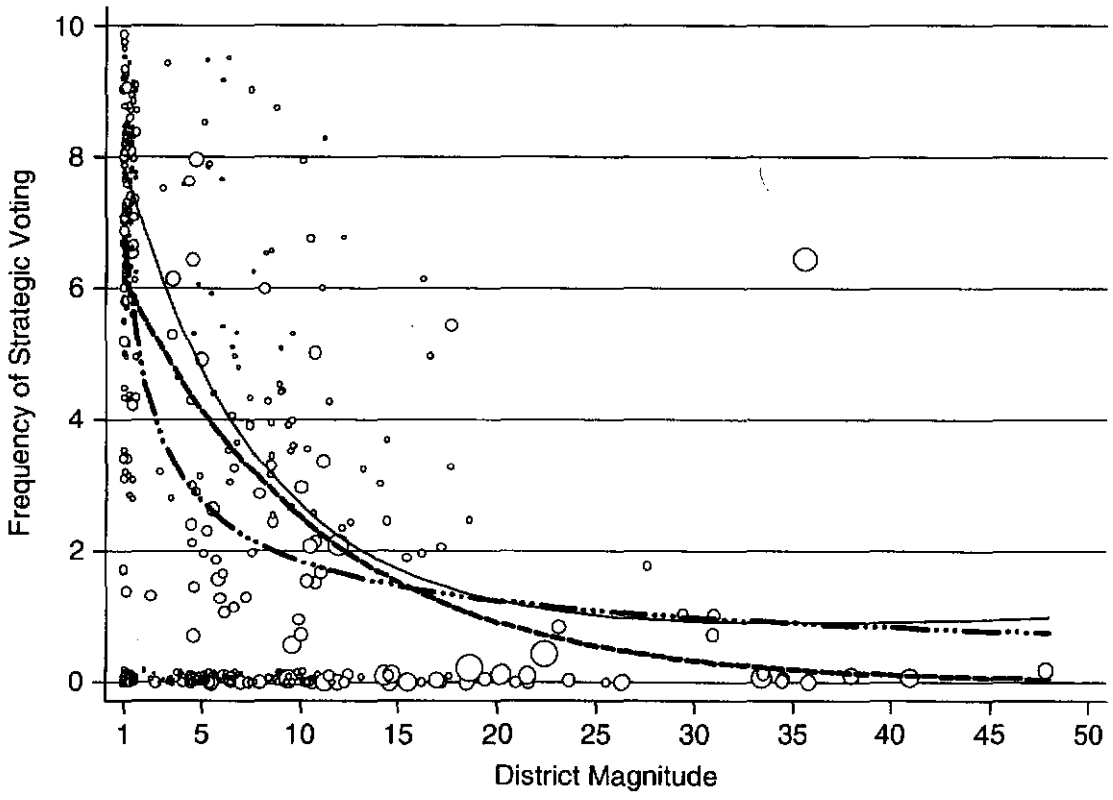

Figure 13.3. Comparison of model predictions across three estimated relationships between district magnitude and the frequency of strategic voting

is consistently slightly better when one uses *log(DISTRICT MAGNITUDE)* as the independent variable.

How large are the predicted differences in strategic voting across different models? In Figure 13.3, three functions are plotted to predict the frequency of strategic voting across a wide range of district magnitude. The range of the independent variable, which is shown on the horizontal axes, comprises more than 99 percent of the electoral districts in the CSES Module 1 data. There are two thick lines. The dashed line represents the GLM predictions based on the model where *DISTRICT MAGNITUDE* is untransformed while the dashed line separated by dots represents the respective predictions where *log(DISTRICT MAGNITUDE)* is the independent variable. The solid thin line corresponds to the GAM predictions from before.

Almost consistently the GAM predictions yield the highest district-level estimates for strategic voting across all three models. The GLM predictions with *DISTRICT MAGNITUDE* as untransformed predictor suggest the lowest levels of strategic voting across all three models when the district magnitude is larger than 15 and, consequently, the model with *log(DISTRICT MAGNITUDE)* as independent variable predicts comparatively the lowest

levels of strategic voting for districts with smaller district magnitudes. The predicted differences across those models are largest in small districts with a district magnitude between 2 and 4 and differ not more than by about 2.5 percentage points. For more than 90 percent of the districts in the sample the model predictions differ by less than 1 percentage point from one another. Thus in most situations the differences across the three models have little substantive relevance. Depending on how the relationship between district magnitude and the frequency of strategic voting is modeled one can expect on average around 6–8 percent of strategic voting in single member districts, while for electoral districts with a district magnitude of greater than 10 we should not expect to find more than about 2 percent strategic voters.

Conclusion

The workings of electoral laws have profound and well-known consequences for the party system, the type of government, and the nature of representation in general. It also has an impact on the way people make decisions in the voting booth. Some voters anticipate the outcome of an election because they form expectations about it and act accordingly. The ways these expectations play out seem to be systematically related to institutional factors that prestructure a voter's choice situation. Since voting behavior is not only determined by institutions I conceptually distinguished institutional and dispositional criteria of how voters generate expectations about the probability of a vote cast for their preferred party going to waste. For this study I focused on the institutional criteria that operate at the primary district level, possibly moderating voters' expectations and thus causing them to deviate from supporting their preferred party. These individual-level mechanisms have predictable implications for the frequency of strategic voting at the electoral district level.

The results of this study provide evidence that the level of strategic voting at the district level is related to district magnitude. Leys (1959) and Sartori (1968) suggested this long ago: The higher the district magnitude the less strategic voting we should expect. For the first time this study provides evidence for this relationship that holds across various electoral systems. Depending on how this relationship is modeled one can expect on average around 6–8 percent strategic voters in single-member districts. Nevertheless even in districts with a large district magnitude, contrary to

what Duverger would have predicted, one can still systematically observe strategic voting although at a very low level.

Although this negative relationship seems to be quite robust, there is still a great deal of variance that is not accounted for even when comparing the levels of strategic voting in electoral districts of the same district magnitude. It might be that the institutional incentives that are channeled through the district magnitude and supposedly moderate a voter's expectation formation process differ across types of electoral systems. For instance, are the incentives to cast a strategic vote in a single-member district in Australia (employing an alternative vote system) systematically different from the ones in Canada, the UK, or the United States or even from the SMD-tier districts in mixed electoral systems? Further research should seek to identify the mechanism by which other institutional effects potentially moderate the incentives that are channeled through the district magnitude.

The controversy in the literature surrounding the functional form of those district magnitude effects appears to be somewhat suspect. At least based on the analysis of the Module 1 CSES data the basis of disagreement is lost. It simply does not make a significant difference whether the district magnitude is logistically transformed or not. My sense is that the controversy should be rather around how we model the dependent variable on which the district magnitude should have an impact. A typical dependent variable in this controversy is certainly the "effective number of parties" (Laakso and Taagepera 1979). It may be more useful to theorize about the data generating mechanism behind such a concept rather than arguing about transformations of independent variables. This would also seem to be a more promising approach with regard to policy applications of the research.

Notes

1. I thank Kerstin Hönig for valuable research assistance and Martin Elff for helpful comments.
2. To simplify language I will just refer to political parties, even if voters can explicitly vote for candidates. Since I am looking at parliamentary elections, candidates are typically affiliated with a party list.
3. If the mechanism behind the Leys–Sartori conjecture were merely driven by elite coordination instead of strategic behavior of voters, marginal parties would not even contest an election. The implication for voters would be that they have no opportunity to waste their votes in the first place. Thus, the observed level

of strategic voting is driven by strategic behavior of voters in anticipation of the decision context and cannot be attributed to elite coordination.

4. In mixed systems I take the SMD vote as relevant vote choice since only in the majoritarian tier one expects an impact of the district magnitude. In order to do that I assume that party and candidate preferences coincide for voters who do not vote for the candidate of their most preferred party but for a more viable candidate.
5. If respondents simultaneously prefer two parties when one party is expected to be viable and the other party is not such a vote is counted as having been cast for the viable party as a strategic vote since not including expectations in voters' decision calculus could have resulted in a vote for a party with a lower expected utility.