On the Extent of Strategic Voting*

Jörg L. Spenkuch
University of Chicago

November 2012

Abstract

Social scientists have long speculated about individuals' tendencies to manipulate elections by misrepresenting their preferences. The fact that preference orderings are generally unobserved, however, has made it very difficult to document strategic behavior empirically. Exploiting the incentive structure of Germany's voting system to solve the fundamental identification problem, this paper estimates the extent of strategic voting in large, real-world elections. Evidence from reduced form as well as structural methods indicates that almost one third of voters abandons their most preferred candidate if she is not in contention for victory. As predicted by theory, tactical behavior has a non-trivial impact on individual races. Yet, as one aggregates across districts, these distortions partially offset each other, resulting in considerably more modest effects on the overall distribution of seats.

*I am especially grateful to Gary Becker, Roland Fryer, Steven Levitt, and Roger Myerson for many helpful conversations and continued encouragement. Moreover, I have benefitted from comments by Scott Ashworth, Eric Budish, Ethan Bueno de Mesquita, Dana Chandler, Tony Cookson, Will Dobbie, Alex Frankel, Richard Holden, Kenneth Mirkin, Matthew Notowidigdo, Elisa Olivieri, Jesse Shapiro, Jan Stoop, Philipp Tillmann, David Toniatti, and Richard Van Weelden, as well as seminar participants at the University of Chicago and the Calvó-Armengol Prize Workshop (Barcelona GSE). I am also indebted to Gabriele Schömel at the office of the Bundeswahlleiter for assistance in acquiring the data used in this paper, and to Che-Lin Su and Jeremy Fox for supplying me with their computer programs. Steven Castongia provided excellent research assistance. Financial support from the Beryl W. Sprinkel Fund at the University of Chicago is gratefully acknowledged. All views expressed in this paper as well as any remaining errors are solely my responsibility. Correspondence can be addressed to the author at Department of Economics, University of Chicago, 1126 E 59th Street, Chicago, IL 60637, or by e-mail: jspenkuch@uchicago.edu.

1. Introduction

Democracy is rooted in the conviction that collective choices should be based on the preferences of all members of society. To provide citizens with an opportunity to participate in the decision making process, almost all democratic states rely on elections. Yet, as a mechanism for eliciting and aggregating preferences, elections are known to be flawed (Arrow 1951). Among other shortcomings, practically every reasonable voting system fails to be strategy-proof (Gibbard 1973; Satterthwaite 1975). As a result, strategic voters may be able to manipulate the outcome of an election by misrepresenting their true preferences.

Although social scientists have long speculated about individuals' proclivities to cast tactical ballots (e.g., Black 1948; Downs 1957; Duverger 1954; Farquharson 1969; Sen 1970), and despite a plethora of anecdotal evidence, it has proven extremely difficult to document strategic behavior empirically. The key problem is that voters' tastes are generally unobserved. Without imposing further assumptions it is, therefore, impossible to know whether ballots accurately reflect the underlying preference orderings and to what extent tactical voting distorts electoral outcomes. In fact, Degan and Merlo (2009) prove that *any* cross-section of votes can be rationalized by some utility function, without resorting to strategic behavior.

While there are many ways in which agents could be strategic, this paper focuses on testing one of the central predictions of rational choice voting theory. According to the theory, strategic voters choose someone other than their most preferred candidate whenever the latter is believed to have no chance of winning (see, e.g., Myerson and Weber 1993). Conversely, agents who always vote for their favorite contestant—regardless of whether she is in contention for victory—are said to be sincere.[1]

In order to resolve the fundamental identification problem the present paper exploits the structure of parliamentary elections in Germany. Under the German system individuals have two votes. Both are submitted simultaneously, but are associated with very different incentives. The *list vote* is cast for a party and counted on the national level. Up to a first-order approximation, list votes determine the distribution of seats in the Bundestag. Since mandates are awarded on a proportional basis (conditional on clearing a 5%-threshold), it is in practically every voter's best interest to reveal his true "preferences" by choosing the party he wishes to gain the marginal seat.[2]

[1] Naturally, issues of tactical voting do not arise in elections with only two candidates. In such cases even strategic voters simply select their favorite contestant.

[2] Note well, individuals' preferences over which party wins the marginal seat in parliament need not coincide with their deep ideological convictions. Nevertheless, it is useful to think of preferences in this narrowly defined way, as it conditions on expectations about post-election coalition formation, the influence of campaign activities, etc.

By contrast, the *candidate vote* is counted in a first-past-the-post system on the district level. Whoever wins the plurality of votes in a given district is automatically elected to parliament. Votes cast for any other contestant are "lost." Although the candidate vote is primarily used to determine the identity of local representatives, by securing a disproportionate share of districts parties may actually increase their seat totals (see Section 4 for details). Hence, when it comes to choosing among different candidates, voters have a clear incentive to behave tactically. As in all elections under plurality rule, they should *never* vote for their favorite party's nominee if she is known to be "out of the race."

Figure 1 demonstrates how observing the same individuals under both electoral regimes facilitates identification. For the sake of illustration, suppose that candidates are perfect representatives of their parties and that it is known a priori who has a realistic chance of winning. First, entertain the possibility that all voters cast sincere ballots. As shown in the panels on the left, if this were indeed the case, there would be a one-to-one correspondence between list and candidate votes—irrespective of whether the particular candidate is a contender.

Next, assume that all voters behave tactically. In such a world only candidates who are believed to be in contention for victory should receive any votes. Thus, for non-contenders the curve representing the relationship between list and candidate votes ought to be perfectly flat.[3]

Lastly, consider the case in which there are both sincere and strategic types. As before, sincere agents vote for their preferred candidate, but no tactical type chooses a non-contender. Consequently, the line relating non-contenders' share of the candidate vote to their parties' list votes must have a slope between zero and one. It is this slope, denoted by λ, that identifies the fraction of voters who are sincere.

Of course, not all candidates are perfect representatives of their parties—despite the fact that party platforms are much more salient in Germany than in the U.S. (see, e.g., Korte 2010a). For instance, some candidates are more charismatic, better qualified, or have a higher media profile than others. Supporters of rival parties might, therefore, be "drawn in" by a candidate's personal appeal. Conditional on being out of the race, however, theory predicts that even high valence contestants will be deserted by tactical voters. Hence, after carefully controlling for candidate quality, the partial correlation between non-contenders' list and candidate votes continues to measure the fraction of agents who stick with their preferred contestant.[4]

[3] Strictly speaking, the shape of the curve for contenders is indeterminate. While it has to lie weakly above the origin and must end at $(100\%, 100\%)$, the slope at intermediate points will generally depend on the correlation of preferences. If, however, tastes are approximately uncorrelated within precincts, then one would expect an almost one-to-one relationship between list and candidate votes.

[4] Of course, there exist more elaborate models in which voters are reluctant to abandon high valence candidates. Empirically, however, there is no evidence that this is the case. If anything, there is a small

To account for heterogeneity across candidates this paper uses previously unavailable precinct level data from the 2005 and 2009 federal elections. In Germany, precincts are the smallest administrative units at which votes are counted, and each precinct is fully contained within one electoral district. Since races take place at the district level, these data allow for λ in Figure 1 to be estimated from within-candidate variation only, thereby *conditioning* on the characteristics of candidates and their competitors.

Figure 2 provides an empirical example from the 2009 election in District 207. Each panel plots a candidate's own vote share against her party's share of the list vote in the same precinct. In contrast to the contestants in the top row, those in the bottom two rows were widely believed to have no chance of winning. As one would expect if list votes were a good proxy for individuals' preferences, after allowing for candidate specific intercepts there is an almost one-to-one relationship between list and candidate votes for contenders. Non-contenders, however, are deserted by a significant number of supporters—although by no means all.

An important remaining obstacle is that it is generally unknown which candidates are considered to be "in the race." Fortunately, the raw data are highly suggestive of focal equilibria in which voters coordinate on the most popular parties' nominees. Given this type of play it is possible to estimate who is a serious contender by drawing from the literature on structural breaks in time series data (e.g., Andrews 1993; Bai 1997; Hansen 2000). Reassuringly, the main findings are qualitatively and quantitatively robust to more than twenty alternative ways of classifying contestants.

The results in this paper show that about 30% of voters abandon their most preferred candidate if she is believed to have no chance of winning. Two key pieces of evidence indicate that defection is, in fact, driven by strategic considerations: First, not only do individuals who abandon their favorite contestant substitute toward the nominee of a potential coalition partner, but desertion rates themselves are much higher among voters faced with at least one palatable alternative than among those who can only choose among two evils. Second, estimates of strategic voting are substantially larger for elections perceived as "critical" than for ordinary ones. Alternative explanations, such as agents receiving utility from supporting the eventual winner of the election, are rejected by the data.

Although these findings may be interesting in and of themselves, they are ultimately insufficient to assess the impact of strategic voting on electoral outcomes. In order to address the question of whether manipulability of electoral systems has important real-world consequences, this paper also estimates a structural model of preferences and voting decisions.

negative correlation between the measured quality of a contestant and the fraction of voters who abandon her.

Again, list votes provide a crucial source of identifying variation.

Reduced form and structural methods yield very similar results with respect to the share of strategic voters, but the latter are also useful in constructing counterfactual election results. Beyond illustrating the importance of the voting scheme (i.e. plurality rule as opposed to proportional representation), the counterfactual experiments show that tactical behavior has a non-trivial impact on individual races. Compared to the current equilibrium, about one in ten districts would change hands if all voters were to cast sincere ballots. However, as one aggregates across districts, these distortions partially offset each other, resulting in only modest effects on the overall distribution of seats.

The results in this paper speak directly to large theoretical literatures on tactical voting (e.g., Austen-Smith and Banks 1988; Carroll 2011; Cox 1994; Feddersen and Pesendorfer 1996; Myatt 2007, 2012; Myerson 2002; Myerson and Weber 1993) and strategyproofness in social choice (see Barberà 2011 for a recent review of the latter). On a purely descriptive level, the empirical evidence indicates that the fundamental assumption in rational choice studies of voting holds for some, but not all, individuals. The evidence also suggests that is important to distinguish the impact of strategic voting on individual races from its effect on aggregate seat totals.

Moreover, the findings presented below may aid in the formulation of political economy models. Often, these models require an *assumption* regarding the voting behavior of agents; and the conclusion may depend critically on whether voters are taken to be tactical or sincere (compare, for instance, Besley and Coate 1997 with Osborne and Slivinski 1996).

The remainder of the paper proceeds as follows. The next section provides a brief review of the existing empirical literature. Section 3 introduces a formal model of sincere and strategic voting. Section 4 outlines the crucial features of Germany's electoral system, and Section 5 describes the data. The main results appear in Section 6. Section 7 estimates a structural model of voting decisions and presents several counterfactual experiments to analyze the impact of strategic behavior. The last section concludes.[5]

2. Previous Literature

There exists a large empirical literature on strategic voting and turnout. While laboratory experiments provide generally convincing evidence of tactical behavior (e.g., Duffy and Tavits 2008; Eckel and Holt 1989; Forsythe et al. 1993, 1996), real-world results tend to be mixed.

[5]There are four appendices. Appendix A contains a formal proof omitted from the body of the paper. Appendix B explains the algorithm for calculating each party's number of seats in the Bundestag. Appendix C provides precise definitions of all variables used throughout the analysis, and Appendix D elaborates on the numerical methods in the papers' structural part.

Coate et al. (2008), for instance, reject the pivotal voter model based on the finding that it is unable to replicate winning margins in Texas liquor referenda. Reed (1990) and Cox (1994), however, argue that the distribution of votes in Japan's multimember districts conforms roughly to the predictions of rational choice theory. More recently, Fujiwara (2011) uses a sharp regression discontinuity in Brazilian mayoral elections to show that third-place candidates are more likely to be deserted in races under simple plurality rule than in runoff elections. The most comprehensive study to date is Cox (1997). His findings are suggestive of strategic behavior in a number of electoral systems, but indicate a lack thereof in others.

Even less is known about the *extent* of tactical voting. Two notable exceptions are Spenkuch (2012) as well as Kawai and Watanabe (2012). Spenkuch (2012) exploits a highly unusual by-election in Germany, which allowed a party to gain one seat by receiving *fewer* votes, to put a 12% lower bound on the share of tactical agents. Kawai and Watanabe (2012) estimate a fully structural model of voting decisions in Japan's general election, concluding that between 63% and 85% of voters are strategic. Their counterfactual experiments suggest that tactical voting had a significant impact on the distribution of seats.[6]

Recall, the fundamental difficulty in inferring strategic behavior from naturally occurring variation is that voters' preferences are not observed. Thus, any conclusions must either be based on indirect tests (as in Coate et al. 2008; Cox 1997; Fujiwara 2011; Spenkuch 2012), or preferences need to be estimated in order to compare them to actual vote counts (as in Kawai and Watanabe 2012).

A separate strand of the literature tries to circumvent this problem by using survey data on voting decisions and political orientations (see, e.g., Abramson et al. 1992; Blais et al. 2001; Niemi et al. 1993; or, for Germany, Gschwend 2007; Pappi and Thurner 2002). Estimates in this tradition typically range from 3% to 17%. Wright (1990, 1992), however, points to important survey biases and raises serious doubts about conclusions based on self-reported votes. Alvarez and Nagler (2000) even show that, depending on the survey design, estimates of strategic voting differ by as much as a factor of seven.[7]

Most closely related to the present paper is work on ticket-splitting in mixed-electoral systems (for the case of Germany see Barnes et al. 1962; Fisher 1973; Gschwend 2001; Jesse

[6] Given that the estimates in Kawai and Watanabe (2012) pertain to Japan's general elections, whereas this paper studies strategic voting in Germany, the findings are not directly comparable. At least for the case of Germany, the raw data itself rule out strategic voting in excess of 50% (cf. Table 1 and the accompanying discussion).

[7] Notwithstanding survey biases, Kawai and Watanabe (2012) as well as Alvarez et al. (2006) point out that some of this literature estimates *misaligned voting* (i.e. the share of voters who actually voted for someone other than their most preferred candidate) instead of *strategic voting* (i.e. the fraction of voters who would do so if their preferred candidate was believed to be not in contention for victory). Not surprisingly, both papers argue that the latter number is significantly larger than the former. For comparison, this paper estimates the extent of misaligned voting to be about 5.8%.

1988). This literature realizes that voters may cast split ballots for strategic reasons, but it ignores the possibility that contenders differ in their personal appeal, thereby confounding tactical behavior with candidate specific preferences (see the discussion in Klingeman and Wessels 2001). Moreover, even if one accepts the mere presence of ticket splitting as evidence of tactical behavior, naïvely estimating the number of split tickets (often from survey data) may over- or understate its true extent.

3. A Model of Sincere and Strategic Voting

In order to fix ideas and frame the empirical work to follow, this section introduces a simple model of voting under plurality rule. Instead of considering the complete decision problem associated with list and candidate votes in Germany, it is more useful to focus on the race for a direct mandate in one electoral district, i.e. on a single subgame. The model is a straightforward extension of Myerson and Weber (1993) with the addition of sincere voters, stochastic turnout, and endogenous pivot probabilities.[8]

3.1. *Basic Building Blocks*

Let the set of candidates be denoted by $K = \{1, 2, \ldots, k\}$. Members of the electorate (simultaneously) cast single nontransferable votes, and the contestant with the highest vote total is declared the winner of the election. Ties are broken by the flip of a fair coin.

Voters are either sincere or tactical, $\theta \in \{s, t\}$. Sincere voters always choose their most preferred candidate, whereas tactical agents act based on personal preferences as well as their beliefs about the actions of other players in the game. The share of agents who are sincere is given by $\lambda \in [0, 1]$.

Each voter has strict preferences over candidates summarized by a vector $u = (u_1, \ldots, u_k)$ in some finite set $U \subset \mathbb{R}^k$, where u_i is the expected utility from candidate i winning the district. $f(u)$ denotes the fraction of individuals with a particular preference profile. That is, f is a probability distribution over U. A voter's type is defined as the tuple $(u, \theta) \in \mathcal{I} \equiv U \times \{s, t\}$. For simplicity, u and θ are assumed to be independent random variables.

Agents know their own type, but are uncertain about the number of other players in the game. This captures the idea that real world elections are characterized by substantial uncertainty about turnout, and that voters are typically not aware of everybody else's identity. Following Myerson (1998, 2000), assume that the total number of voters is a random variable drawn from a Poisson distribution with mean $n < \infty$.[9] n, f, as well as λ are common knowledge.

[8] See Kawai and Watanabe (2012) for a similar model with a fixed set of voters and a simple ordering restriction on beliefs about pivot probabilities.

[9] The Poisson assumption is made for convenience. None of the empirical results depend on it.

As mentioned above, strategic agents maximize expected utility taking the behavior of others into account. More specifically, tactical voters choose candidate k only if doing so maximizes

$$\overline{u}\left(k,\widetilde{\pi}|u,t\right)=\frac{1}{2}\sum_{k'\in K\backslash\{k\}}\widetilde{\pi}(k,k')\left[u_k-u_{k'}\right], \tag{1}$$

where $\widetilde{\pi}=\left(\widetilde{\pi}\left(k,k'\right)\right)_{k,k'\in K}$ denotes players' common beliefs about the probability of casting a pivotal vote.[10]

By contrast, sincere players always select their most preferred contestant. They maximize the utility function:

$$\overline{u}\left(k,\widetilde{\pi}|u,s\right)=u_k.$$

3.2. *Equilibrium*

Let $\sigma\left(k|u,\theta\right)$ denote voters' strategies. That is, $\sigma:\mathcal{I}\rightarrow\Delta\left(K\right)$ specifies the probability that a type (u,θ) voter casts a ballot for candidate k. In equilibrium it must be the case that, for all $(u,\theta)\in\mathcal{I}$,

$$\sigma\left(k|u,\theta\right)>0\quad\text{only if}\quad k\in\arg\max_{k'\in K}\overline{u}\left(k,\widetilde{\pi}|u,\theta\right).$$

Given σ, realized vote totals, $\upsilon=\left(\upsilon\left(k\right)\right)_{k\in K}$, are random variables with means, $\mu=\left(\mu\left(k\right)\right)_{k\in K}$, equal to

$$\mu\left(k\right)=n\sum_{u\in U}\left[\lambda\sigma\left(k|u,s\right)+\left(1-\lambda\right)\sigma\left(k|u,t\right)\right]f\left(u\right). \tag{2}$$

From the Poisson assumption it follows that the elements of υ are independently distributed (see Myerson 1998 for a proof), which allows the probability of casting the pivotal vote to be expressed in a transparent way. More specifically, not knowing the exact number of players in the game, the ex ante probability of candidate k being tied for first or one vote behind k' is given by

$$\pi\left(k,k'\right)=\sum_{\tau=1}^{\infty}\left[\psi\left(\upsilon=\tau|\mu\left(k\right)\right)\left(\sum_{\tau'=\tau}^{\tau+1}\psi\left(\upsilon=\tau'|\mu\left(k'\right)\right)\right)\left(\prod_{k''\in K\backslash\{k,k'\}}\sum_{\tau''=0}^{\tau-1}\psi\left(\upsilon=\tau''|\mu\left(k''\right)\right)\right)\right],$$

[10]To see that (1) results from expected utility maximization note that an individual's vote affects his payoff only if it changes the outcome of the election, i.e. if two candidates are either tied for first or one vote apart. If candidate k and k' are tied, then voting for the former results in an expected utility gain of $u_k-\frac{1}{2}\left(u_k+u_{k'}\right)$. If k is one vote behind k', then choosing k changes payoffs by $\frac{1}{2}\left(u_k+u_{k'}\right)-u_{k'}$, which is the same as the previous expression. Summing over all candidate pairs and weighting by $\widetilde{\pi}$ gives (1).

where $\psi(\upsilon = \tau | \mu(k))$ denotes the probability of a Poisson random variable υ with parameter $\mu(k)$ being equal to τ.[11]

Definition: Given the Poisson game $\Gamma(K, \mathcal{I}, n, f, \lambda)$, a *voting equilibrium* consists of a *strategy function* σ satisfying, for all $(u, \theta) \in \mathcal{I}$,

(i) $\sigma(k|u,\theta) \geq 0 \quad \forall k \in K$,

(ii) $\sum_{k \in K} \sigma(k|u,\theta) = 1$, and

(iii) $\sigma(k|u,\theta) > 0$ only if $k \in \arg\max_{k' \in K} \overline{u}(k', \widetilde{\pi}|u,\theta)$;

as well as a set of *beliefs* such that

(iv) $\widetilde{\pi}(k,k') = \pi(k,k') \quad \forall k,k' \in K$.

Proposition 1: *The set of voting equilibria is always non-empty.*

Proof: See Appendix A.

To get a sense of what equilibrium play looks like note that strategic voters' utility function is homogenous in $\widetilde{\pi}$. Hence, tactical voting decisions are determined by the relative—not absolute—size of perceived pivot probabilities. From the magnitude theorem in Myerson (2000) it follows that some pivot probabilities are going to be much larger than others; although for large electorates all elements of π will be very close to zero. That is, as $n \to \infty$ most pivot probabilities become infinitesimal relative to, at most, a few remaining ones. Intuitively, this is because homogeneity of the utility function implies that $\widetilde{\pi}(k,k')$ can be rewritten as the probability of k and k' running neck-and-neck ahead of all other contestants, *conditional* on the election being tied in the first place. Such a tie, however, is substantially more likely to involve the two front-runners than an underdog. Hence, almost all of the probability mass must be concentrated in one or two candidate pairs. Put differently, in large elections only a subset of candidates will be "in the race," and strategic voters behave as if choosing only among those who are believed to be serious contenders.

Since tactical agents become more inclined to select a particular candidate as they form favorable beliefs about her being in contention for victory—say, because her standing in pre-election polls improves, or due to campaign activities that manipulate voters' perception of candidate viability—the model above exhibits the potential for *bandwagon effects* and *self-fulfilling prophecies* (Simon 1954). This property can result in multiple equilibria, and any candidate that is not a Condorcet loser may be the sole likely winner under plurality rule (cf. Myerson and Weber 1993). Thus, without further refinement the model makes no

[11] As is typical in the literature on strategic voting, the probability of three-way ties is assumed to be negligible.

prediction on the set of candidates who will be "in the race," i.e. on which of the many possible equilibria is being played.

In reality, however, one may suspect that tactical voters play focal equilibria, especially in cases in which the share of sincere agents is large. When many voters simply select their most preferred candidate, then tastes alone will lead to substantial differences in vote shares, thereby "pre-determining" who is a serious contender. A number of potential equilibria may even be ruled out completely. Consequently, strategic agents would only want to choose among the most popular contestants. As argued in Section 6.1 the data do, indeed, suggest focal equilibria of this type.

Lastly, note how equation (2) gives rise to the estimation strategy outlined in the introduction. Since tactical voters would never support a candidate who is "out of the race" it must be the case that $\sigma(k|u,t) = 0$ for all non-contenders. $\sigma(k|u,s)$, however, equals either 0 or 1, depending on whether type (u,s) agents prefer k over every other contestant. Given these strategies, the expression for non-contenders' expected vote shares simplifies to

$$\mu(k)/n = \lambda \sum_{\widetilde{u} \in \{u \in U | u_k > u_{k'} \forall k'\}} f(\widetilde{u}).$$

Thus, if one accepts the assumption that list votes are a not only a measure of voters' preferences over parties, but also a (potentially noisy) proxy for the fraction of individuals who favor the respective nominees, i.e. for $\sum_{\widetilde{u}} f(\widetilde{u})$, then λ can be identified from the candidate-list vote gradient among non-contenders. Of course, for this assumption to be reasonable it is important to properly account for systematic differences in candidates' idiosyncratic appeal.

3.3. *A Note on Turnout*

Before proceeding to the empirical application, it is important to discuss one of the model's limitations. Instead of considering the decision to go to the polls explicitly, turnout is assumed to be stochastic. Although it is natural to think of n as the (expected) number of agents for whom the cost of voting are smaller than the benefits, this assumption does entail a loss of generality. In particular, in the formulation above it is not the case that turnout varies with the probability of being pivotal.

One may rationalize this modeling choice by appealing to the "paradox of voting" (Downs 1957; Riker and Ordeshook 1968). To generate high turnout in light of small pivot probabilities agents have to derive a psychic benefit from going to the polls.[12] If the net utility gain is sufficiently large, then tactical considerations matter very little for turnout, but may nevertheless determine for whom strategic agents vote. That is, the utility from fulfilling

[12] For an alternative approach relying on aggregate uncertainty see Myatt (2012).

their 'civic duty' might bring voters to the polls, but conditional on voting there is little that keeps them from behaving tactically.

Naturally, there exist alternative models in which strategic considerations do not determine voting decisions, or only if they are sufficiently important. For instance, agents might derive utility from voting for the "right candidate," and this utility may easily outweigh the expected gain from influencing the outcome of the election. Thus, whether individuals vote strategically is ultimately an empirical question.

4. Germany's Electoral System

In order to shed light on the extent of strategic voting the present paper exploits the structure of parliamentary elections in Germany. Elections to the Bundestag are held according to a mixed member system with approximately proportional representation. Except for minor modifications, the same system has been in place since 1953 (see Bawn 1993 for an account of its inception).[13]

As mentioned in the introduction, each voter casts two different votes. The first vote, or *candidate vote* (Erststimme), is used to elect a constituency representative in each of 299 single-member districts. The second vote, or *list vote* (Zweitstimme) is cast for a party and counted on the national level.

Figure 3 depicts a sample ballot. The column on the right enumerates all electable parties in a particular state together with the first few individuals on the respective party's list. The column on the left contains the names of all candidates running for district representative. Contestants always appear directly next to the party that nominated them, and both the candidate's last name as well as the respective party acronym are printed in bold font. The close connection between candidates and parties reflects the fact that most contestants simply campaign on their own party's platform.[14]

District representatives are determined in a first-past-the-post system. That is, whichever contestant achieves the plurality of candidate votes in a given district is automatically awarded a seat in the legislature. Winners are said to hold *direct mandates*, and votes cast for any other candidate are discarded.

Figure 4 shows the party affiliation of all district winners in the 2005 and 2009 elections. Although the CDU/CSU secured the majority of direct mandates in both years, there remains ample geographic variation, as well as some variation over time. Of the five major parties only the FDP did not win any districts—despite gaining a non-trivial fraction of votes.

To still achieve approximately proportional representation of all parties clearing a 5%-

[13]In describing the German electoral system this section borrows from Korte (2010b) and Spenkuch (2012).
[14]Since the introduction of the two-ballot system in 1953 no independent candidate has ever won a district.

threshold, the German electoral system also awards *list mandates*. First, all list votes are aggregated up to the national level and a total of 598 preliminary seats are distributed to parties on a proportional basis. Each party's allotment is then broken down to the state level, and compared to its number of direct mandates in the same state. Whichever number is greater determines how many seats the party will actually receive in that state.

More formally, let $d_{p,s}$ denote the number of districts that party p won in state s, and let $l_{p,s}$ be how many mandates it would have received in the same state under proportional representation. Then, the final number of seats that p retains in s equals

$$n_{p,s} = \max\{d_{p,s}, l_{p,s}\},$$

and its total in the Bundestag is given by $n_p = \sum_s n_{p,s}$ (see Appendix B for a more detailed description of the algorithm used to allocate seats).

If $d_{p,s} < l_{p,s}$, then in addition to the district winners the first $l_{p,s} - d_{p,s}$ candidates on p's list are elected as well. Otherwise, only holders of direct mandates receive a seat. Parties are said to win *overhang mandates* (Überhangmandate) whenever $d_{p,s} > l_{p,s}$. In such cases the total number of seats in the Bundestag increases beyond 598.

Note, however, the number of mandates awarded under proportional representation, i.e. $\sum_p \sum_s l_{p,s}$, exceeds the number of districts, $\sum_p \sum_s d_{p,s}$, by a factor of two. Thus, situations in which $d_{p,s} > l_{p,s}$ are a fairly rare. Instead, the list vote determines the overall distribution of seats—at least up to a first-order approximation.[15]

Critical for the purposes of this paper are the incentives associated with each vote. First, consider the list vote. Due to the fact that seat totals are approximately proportional to parties' vote shares (and given that it is nearly impossible to predict rounding), most agents can be expected to choose their "preferred" party, defined as the one they would like to win the marginal seat in parliament.

In fact, it is easy to see that voting for one's preferred party is a strictly dominant strategy whenever $\mathbb{E}[d_{p,s}] < \mathbb{E}[l_{p,s}]$. Only if one anticipates that $d_{p,s} > l_{p,s}$ may it not be optimal to do so. However, most voters cannot reasonably assign a high probability to this case. Historically, it occurred for less than 10% of state-party combinations; and FDP, Greens, as well as The Left have never won any overhang mandates. The vast majority of agents has, therefore, an ex ante incentive to cast "truthful" list votes.[16]

When it comes to the candidate vote, however, agents should behave tactically—as in

[15] Theoretically, overhang mandates may lead to departures from proportionality of up to one third of all seats. In the current equilibrium deviations are much smaller. For instance, relative to its share of the list vote, the CDU/CSU received an additional 7 mandates in 2005, whereas the SPD secured 9 extra seats.

[16] The results in this paper are qualitatively and quantitatively robust to focusing on state-party combinantions for which $d_{p,s} \leq l_{p,s}$.

all elections under plurality rule. If their preferred candidate is believed to be a contender, then voting for her is a dominant strategy. After all, there is a small (but strictly positive) probability that doing so will lead to an extra seat in the Bundestag, and even in cases in which one's preferred party is unlikely to win an additional mandate, supporting the party's nominee can prevent a rival faction from increasing its seat total. If, however, a candidate is known to be "out of the race," then one can always do better by voting for somebody else. Substituting toward a contestant who actually has a chance of winning could benefit a coalition partner, or it might spoil a particularly disliked party's victory. Even if individuals took the aggregate distribution of seats as fixed, by casting tactical ballots they may elect "better" local representatives. Although expected payoffs may not be very large, as long as strategic agents are not completely indifferent to who carries their district, they should *never* waste their vote on a candidate who is "out of the race."[17]

Table 1 shows the relative frequency of different list and candidate vote combinations in the 2009 federal election.[18] First and foremost, the evidence indicates that some, but not all voters desert non-contenders. Although candidates of FDP, Greens, and other minor parties are rarely in contention for victory, they are abandoned by only about half of their followers. At the same time, the numbers show that about 46% of all FDP supporters voted for a candidate of the CDU—its coalition partner—whereas 33% of Green Party adherents chose an SPD nominee. It therefore appears that voters who do desert non-contenders substitute toward close political allies.

By itself, however, Table 1 is insufficient to determine the extent of strategic voting. For instance, some FDP supporters might have chosen CDU candidates not because of tactical considerations, but because they are better qualified or more charismatic. Also, not all CDU and SPD adherents voted for their own party's nominee. In fact, almost one third of those who deviate end up picking a political rival. While it is possible that these voters chose among the lesser of two evils in districts in which the CDU or the SPD candidate happened to be "out of the race," it is equally plausible that their voting decisions were based on candidate idiosyncrasies.

Simply computing the share of "ticket splitters" among supporters of FDP, Greens, and other minor parties would put the share of strategic voters at slightly above 50%. Compared to the 30% estimate based on the identification strategy outlined above, this naïve approach overstates the extent of strategic voting substantially.

[17] At first, the German system may seem very complicated. Yet, the results of Spenkuch (2012) show that voters are keenly aware of the incentives at work.

[18] Table 1 is based on a 3.9% random sample of actual votes. German electoral law requires the Federal Returning Officer to publish descriptive statistics on vote combinations, as well as voting behavior by age and gender (see Bundeswahlleiter 2010). Unfortunately, the micro-data are not publicly accessible.

5. Data Sources and Descriptive Statistics

In order to estimate the share of strategic voters, it is useful to focus on the 2005 and 2009 federal elections. In these years all of the five major parties were widely expected to clear the 5%-threshold.[19] Since voters could be virtually certain that their preferred party would be represented in parliament, list votes should reflect party preferences more accurately in 2005 and 2009 than in other years.

In contrast to prior work on ticket splitting in Germany, the analysis in this paper relies on official election results by polling precinct (Wahlbezirk). These data have been obtained from the Federal Returning Officer and were until recently not publicly available. In Germany, precincts are the smallest administrative units at which votes are counted. Each precinct is fully contained within an electoral district and associated with one polling station where a returning officer oversees the election. By law, no precinct can contain more than 2,500 eligible voters. As of 2009 there were 299 electoral districts and almost 89,000 precincts.

Differentiating between East and West Germany as well as election year, Table 2A displays summary statistics for all precinct level variables used throughout the analysis. Compared to the U.S., turnout is fairly high. Averaging across 2005 and 2009, almost 75% of the electorate went to the polls. Together with a mean size of 821 eligible voters, this means that precincts handle about 615 votes on average.

As is well known, CDU, SPD, FDP, and the Green Party fare substantially better in West Germany than in the East. The opposite is true for The Left—the successor of the East German communist party. Moreover, CDU and SPD receive more candidate than list votes. Since nominees of these two parties are serious contenders in most districts, this "surplus" of candidate votes may be indicative of tactical behavior (cf. Cox 1997).

At the same time, CDU and SPD might simply field "better" contestants. In fact, the descriptive statistics in Table 2B demonstrate that candidates differ along important dimensions.[20] For instance, only 19% of CDU candidates are female, compared to 35% of Social Democrats and 34% of Green Party nominees. Moreover, relative to their FDP, Left, or Green Party counterparts, contestants of CDU and SPD are about four times more likely to be a current member of parliament, and more than forty times as likely to be an incumbent. Therefore, any argument linking differences in the distribution of list and candidate votes to strategic behavior must be based on an econometric strategy that carefully controls for candidates' idiosyncratic appeal.

[19] For instance, more than 90% of adults sampled in the 2009 pre-election survey of the German Longitudinal Election Study (GLES) expected the FDP and Green Party to receive more than five percent of the list vote.

[20] The information in Table 2B has been compiled from official publications by the Federal Returning Officer (Bundeswahlleiter 2005c, 2009b).

6. Estimating the Extent of Strategic Voting

6.1. *Econometric Approach*

Following the argument in the introduction, the fraction of sincere voters can be identified from the partial correlation between non-contenders' share of the candidate vote and their parties' share of the list vote, after accounting for candidate quality. Hence, with ideal data it would be possible to estimate the parameter of interest, λ, from

$$(3) \qquad v^C_{k,r,d,t} = \chi_{k,t} + \lambda v^L_{k,r,d,t} \times 1\left[k \notin C_t\left(d\right)\right] + \gamma v^L_{k,r,d,t} \times 1\left[k \in C_t\left(d\right)\right] + \epsilon_{k,r,d,t},$$

where $v^C_{k,r,d,t}$ denotes contestant k's share of the candidate vote in precinct r during election year t, and $v^L_{k,r,d,t}$ is her party's share of the list vote in the same precinct. $C_t\left(d\right)$ stands for the set of candidates who are believed to be contenders in district d, and $\chi_{k,t}$ marks a year specific candidate fixed effect. N.B.: As races take place at the level of the electoral district, precinct level data allow for identification using only within-candidate variation, thereby conditioning on candidate quality, beliefs about tie probabilities, etc.

There does remain one obstacle. Since $C_t\left(d\right)$ is unobserved, it is not clear which candidates were considered to be "in the race." If agents had perfect foresight then only contestants who ended up finishing first or second would have been a priori serious contenders. In reality, however, voters are likely to have noisy priors and may believe more than just two candidates to be in contention for the direct mandate (Cox 1994 shows that this can be an equilibrium).[21]

Fortunately, the raw data suggest an intuitively appealing heuristic for classifying contenders. To see this consider Table 3, which contains a cross-tabulation of candidates' own rank (based on the candidate vote) against the ranking of their party among voters in the same district (based on the list vote). Out of the 598 contestants whose party placed first, only 41 did not win a direct mandate and a mere 2 finished third or lower. In contrast, none of the candidates who ran for a party ranked fourth or worse came in first, and only 3 finished second. Overall, the correlation between list and candidate vote based rank is .93. The evidence is, therefore, consistent with voters playing focal equilibria in which they coordinate on the most popular parties' nominees.

This impression is further strengthened by the regression results in Table 4. The estimated specifications are analogous to (3), but allow for the slope to vary by preference, i.e. list vote, based rank. As should be the case if list votes were not only a measure of individuals' preferences over parties but also a good proxy for voters' assessment of the respective nominees,

[21] Unfortunately, pre-election surveys in Germany are too small to derive reliable estimates of voters' expectations. For instance, in only 50 electoral districts did the German Longitudinal Election Study (GLES)—the best available data source—survey more than 15 adults prior to the 2009 elections.

the R^2 of these specifications is extraordinarily high. Moreover, there is a near one-to-one correspondence between list and candidate votes for the first two contestants. For for those ranked third and below, however, the candidate-list vote gradient is significantly smaller, suggesting that voters desert (only) candidates of "unpopular" parties.[22]

Based on the results in Tables 3 and 4, it seems reasonable to assume that contestants backed by one of a district's two favored parties are serious contenders, whereas candidates of parties ranked fourth and below should be considered "out of the race." The only ambiguity arises with respect to those in third place. In practice, almost 10% of third ranked contestants finish first or second. Hence, one would want to classify some (but not all) of them as serious competitors, especially in cases in which only a few percentage points separate their own party from the one in second place.

Drawing from the literature on structural breaks in time series data (e.g., Bai 1997; Hansen 2000) it is possible to estimate a cutoff value, κ, separating candidates into contenders and non-contenders.[23] More specifically, a candidate is said to be a contender if and only if her party trails a district's second most popular one by less than κ percentage points.

With this simple heuristic in hand, the estimating equation becomes

$$(4) \qquad v^C_{k,r,d,t} = \chi_{k,t} + \lambda v^L_{k,r,d,t} \times 1\left[\overline{v}^{L,2^{nd}}_{d,t} - \overline{v}^L_{k,d,t} > \kappa\right] + \gamma v^L_{k,r,d,t} \times 1\left[\overline{v}^{L,2^{nd}}_{d,t} - \overline{v}^L_{k,d,t} \leq \kappa\right] + \epsilon_{k,r,d,t}.$$

Here, $\overline{v}^L_{k,d,t}$ denotes the list vote share of candidate k's party in district d, and $\overline{v}^{L,2^{nd}}_{d,t}$ is that of the second most popular party in the same district. All other symbols are as defined above.

If (4) is correctly specified, then searching for the value of κ that maximizes the R^2 yields a super-consistent estimate of the true break point (Hansen 2000). Moreover, under the null hypothesis that such a point exists, estimates of the model's other parameters are normally distributed and standard errors need not be adjusted for sampling variability in the location of the break (see, e.g., Bai 1997).[24]

Although intuitively appealing, there is no guarantee that this method classifies all candidates correctly. For this reason Section 6.3 performs a series of robustness checks, demonstrating that the main findings are qualitatively and quantitatively robust to employing more

[22] One can reject the null hypothesis of equal slopes at the 1%-level.

[23] Card et al. (2008) use the same approach in combination with regression discontinuity methods to test for "tipping points" in the dynamics of neighborhoods' racial composition.

[24] Complications arise under the null of no break. In this case, λ will have a non-standard distribution, and conventional hypothesis tests will have incorrect size. One way to correct for this problem would be to follow Card et al. (2008) and split the data into a sample used to estimate the location of the change point, and another one for further analyses. Given the evidence in Tables 3 and 4, the possibility of no break, i.e. that all contenders suffer from equal desertion rates, seems unlikely in the present context. Moreover, the robustness checks in Table 8 indicate that the results are qualitatively and quantitatively robust to misspecification of the change point.

than twenty alternative definitions.

6.2. *Reduced Form Results*

Focusing on candidates of the five major parties, Table 5 shows results from estimating equation (4) by ordinary least squares. To allow for arbitrary forms of autocorrelation in the residuals as well as for correlation within and across districts, standard errors are clustered by state. Going from the left to the right the set of fixed effects grows steadily. The most inclusive specification contains candidate-year fixed effects, thus controlling nonparametrically for the appeal of individual candidates, as well as that of their competitors. The structural break in this specification is estimated to occur at 6.4%, implying that voters treat "close" third ranked contestants as if they are in the race.[25]

More to the point, the first row of Table 5 presents estimates of the share of voters who stick with their preferred candidate despite her having no chance of winning. Estimates of λ range from .657 to .724, and are fairly precise. Taken at face value, these numbers indicate that about two thirds of all voters are sincere, whereas the remaining third abandons their favorite contestant whenever she is "out of the race."

The crucial question, however, is whether voters do so for strategic reasons. Strictly speaking, any model would be consistent with the evidence in Table 5 as long as it predicts the candidate-list vote gradient for non-contenders to be smaller than one. For instance, some individuals may simply vote for whichever candidate advertises the most, and advertising expenditures might be highly correlated with who remains in contention for victory. It would, therefore, appear as if voters abandon weak candidates, despite the fact that no agent behaves tactically.

In order rule out mechanical explanations of this kind, Table 6 compares the extent of strategic voting, i.e. $1 - \lambda$, across a number of different settings. The first set of results demonstrates that individuals' tendency to desert weak contestants depends on *who* remains in contention for victory. Voters are more than three times as likely to abandon their favorite candidate when that of an allied party is still "in the race" than when faced with the choice among two evils.[26]

Moreover, distinguishing between races that were "close" and those that were not, strategic voting appears to have been somewhat more prevalent in the former—although the difference is not statistically significant; and disaggregating the data by election year shows that desertion of non-contenders was more common in 2005 than in 2009. The 2005 election followed

[25] Bootstrapping suggests that κ is statistically different from zero.

[26] The following parties are defined as close substitutes: CDU and FDP, SPD and Green Party, The Left and SPD. The difference in Table 6 is even starker when supporters of The Left are assumed to have no substitute available.

a failed motion of confidence that triggered the dissolution of the Bundestag and was widely perceived to be a "critical election" (Korte 2009).[27] It is, therefore, hardly surprising that a greater number of voters behaved tactically when the stakes were particularly high.[28]

Taken together these findings suggest that it might be more appropriate to consider strategic behavior a conscious decision as opposed to an agent's type. That is, all agents may be capable of voting tactically, but only for a subset of them do the subjective benefits outweigh the (psychic) costs.

More importantly, the evidence in Table 6 is at odds with many alternative theories for why voters abandon candidates who are "out of the race." Any model in which voters defect candidates for non-strategic reasons would not only have to predict a correlation between desertion rates and a contestant's chance of winning, but it would also have to explain why defection is more common when the stakes are higher, and why it depends on who remains in contention for victory. The patterns in Table 6 as well as the fact that voters who do cast split-tickets substitute toward candidates of a potential coalition partner (cf. Table 1) suggest that the observed behavior is, in fact, driven by strategic considerations.[29]

In order to strengthen this claim, Table 7 provides an explicit test of the pivotal voter model against the alternative that individuals receive utility from voting for the eventual winner of the election. The entries correspond to the candidate-list vote gradient among second ranked candidates, by distance to the first ranked one. If agents did, indeed, have a taste for supporting the likely winner, then one would expect runner-ups to be abandoned as well, especially those that trail far behind. By contrast, the model in Section 3 predicts that voters do *not* abandon the runner-up, even if her chances of winning are very small. This is because if a race were to be tied—however unlikely that may be—the tie would almost certainly involve the second ranked candidate, in which case voting for her would change the outcome of the election. Although counterintuitive, the evidence in Table 7 does support the

[27] Campaigning to stay in office, Chancellor Schröder and his SPD-Green coalition promised to undo some of their unpopular labor market and welfare reforms, while raising taxes on the rich. In stark contrast, led by Angela Merkel, the conservative-libertarian bloc sought to further increase the pace and scope of deregulation, slashing income taxes and public spending in the process.

[28] Official statistics analogous to those in Table 1 show a substantially higher fraction of split tickets in 2005 (cf. Bundeswahlleiter 2006, 2010). Although abstention was approximately 7 percentage points lower than in 2009, the difference in turnout alone is too small to account for the finding in Table 6.

[29] Ancillary results (available from the author upon request) demonstrate that these substitution patters are, in fact, driven by expextations regarding post-election coalition formation. The empirical approach exploits the small, libertarian FDP's position as "kingmaker" in much of the post-World War II period. From the 1960s until the advent of the Green Party in the 1980s the FDP was the *only* other major party represented in parliament and participated in every government from 1969 to 1998. Crucially, it switched allegiances halfway through this era. Although parties' positions in policy space remained almost unchanged, after the end of the SPD-FDP coalition in 1982, a substantial number of FDP adherents switched from voting for SPD candidates to supporting those of its new coalition partner, the CDU.

pivotal voter model.

6.3. *Sensitivity and Robustness Checks*

For the results above to correctly identify the shares of each type of voter it must be the case that the regressors are uncorrelated with the error term. One obvious source of bias may be systematic misclassification of contenders. To ameliorate this concern, Table 8 presents estimates of the share of strategic agents employing a number of alternative definitions (listed in the column on the left). For each definition, the table shows two estimates: one based on candidate fixed effects, and another one using candidate-year fixed effects. For comparison, the top row displays the main results from Table 5.

Although individual point estimates do, of course, vary, the majority of them are fairly close to their baseline values. In all cases overlap the 95%-confidence intervals substantially with those in Table 5. On the whole, the evidence suggests that misclassification of contestants is not a serious problem.

However, there are at least two other threats to identification. First, the argument laid out in the introduction assumes that the share of agents who desert non-contenders does not depend on precincts' ideological composition (i.e. the position on the x-axis in Figure 1). This assumption is not innocuous. For instance, it rules out peer effects in voting, certain forms of sorting across precincts, or that the perception of candidate quality varies systematically with voters' political views. Fortunately, the assumption is testable.

To see this, note that all of the threats mentioned above imply a non-linear relationship between list and candidate votes for non-contenders. That is, if social interactions or differences in candidate perception were important, then one would expect desertion rates (and hence the slope of the line in Figure 1) to vary systematically with the position on the x-axis.

Yet, estimating λ separately for each candidate and regressing it on her own party's vote share in the same district yields no indication that voters are more or less likely to abandon a contestant when they live around others with similar political views. In fact, for non-contenders the correlation between λ_k and $\overline{v}^L_{k,d,t}$ is almost exactly zero. Another, perhaps more powerful test would be to estimate the relationship between list and candidate votes nonparametrically. As shown in Figure 5, there is little reason to believe that the linearity assumption is violated.

The other concern relates to the behavior of political parties. For instance, whether parties nominate a particularly "good" or "bad" candidate might depend on the anticipated likelihood of winning the district. Since the empirical strategy relies on within-candidate variation, this sort of behavior poses a problem (only) if candidate quality *interacts* with the share of voters who behave sincerely—say, because voters might be reluctant to abandon very charismatic contestants.[30] Although plausible, the data do not suggest that "good" candidates (as measured by $\chi_{k,t}$) are less likely to be deserted when they are "out of the race." If anything, estimates of the covariance between $\chi_{k,t}$ and a non-contender's λ_k are slightly negative.

Table 9 performs a series of additional robustness checks. Again, the results do not depend on the weighting scheme or on whether one also includes candidates of "micro parties." However, two additional pieces of evidence ought to be pointed out explicitly.

As discussed in Section 4, a small numbers of voters might anticipate overhang mandates, and thus cast tactical list votes. Although one would surely want to classify these individuals as being strategic, it useful to probe the robustness of the results with respect to this potential confound. The fourth row in Table 9 restricts attention to states in which overhang mandates had never occurred before 2005. Reassuringly, the resulting point estimates are very similar to their baseline values.

Since list votes are hardly a *perfect* proxy for preferences over candidates, one may also worry about measurement error. The resulting attenuation bias would lead to estimates of λ that are smaller than one and, therefore, to incorrect inference about the prevalence of tactical behavior. Note, however, measurement error should not only affect the estimated candidate-list vote gradient among non-contenders, but also that for contenders. Thus, looking at the *difference* in slopes for both sets of contestants, i.e. $\gamma - \lambda$, provides a way to address this concern. As the fifth row in Table 9 demonstrates, attenuation bias cannot the explain the results. In fact, the estimated extent of strategic voting remains almost unchanged.

6.4. *Correlates of Strategic Voting*

Taken together the results above show that a non-trivial fraction of agents behaves strategically. Yet, simple averages may conceal considerable heterogeneity across individuals. Since strategic voters are more likely to be pivotal and thus exert a disproportionate influence on the positioning of political candidates, it is also important to understand *who* votes tactically.

In order to infer whether λ varies with the characteristics of voters, the present paper relies on official statistics for the universe of German cities and villages, published by the Federal Statistical Office and the statistical offices of the Länder (Statistische Ämter des Bundes und der Länder 2007, 2011).[31] After aggregating election results to the village level and focusing on the set of municipalities that are fully contained within an electoral district,

[30] If this were the case, then the estimates above would simply describe voters' behavior in the current equilibrium, *conditional* on the actions of parties and candidates.

[31] Unfortunately, comparable data for polling precincts do not exist. Polling precincts are too small to produce reliable estimates from existing data sets.

it is straightforward to estimate specifications that allow for λ to increase or decrease in some village characteristic.

Table 10 displays the results. The first column demonstrates that aggregation to the municipality level does not affect the estimated share of sincere voters. The remaining four columns examine how λ changes with population density, income tax revenue per capita, as well as the gender and age composition of the electorate. For ease of interpretation covariates have been demeaned, so that the estimates in the second row refer to the share of sincere voters at the sample average.[32]

Interestingly, urban voters are not more strategic than rural ones; nor is there a significant gender gap—although lack of variation in the composition of villages makes the latter conclusion somewhat more speculative. There are, however, important differences with respect to socio-economic status (as proxied by income tax revenue per capita) as well as age.

Since the income tax variable captures only revenues that accrue to the respective municipalities and given that the German tax system is highly non-linear, it is easiest to judge the magnitude of the coefficient by an example. Consider two villages: one's per capita income tax revenue is a standard deviation below the mean, while that of the other village is one standard deviation above the sample mean.[33] The share of sincere voters is estimated to be almost 6 percentage points higher in the latter.

Disparities by age are even larger. Taken at face value the coefficient in column (5) suggests that the fraction of strategic agents is 24.4 percentage points lower among voters below the age of 30.[34] Of course, this estimate is based on limited variation, and therefore not very precise. But, together with the results in column (4), it suggests that sophistication and learning affect tactical behavior.

6.5. *Strategic Voting and Democratic Experience*

In order to further investigate the hypothesis of learning, this subsection uses the German Reunification as a natural experiment.

Although the German Democratic Republic (GDR) held regular, formal elections to the *Volkskammer* (People's Chamber), they were effectively meaningless. East Germans could only choose from candidates on a single list controlled by the Socialist Unity Party (SED), and it was customary to cast one's ballot in public, simply accepting all nominated candidates. Not surprisingly, official approval rates often exceeded 99%.[35] Free, democratic elec-

[32] Estimating models that are fully interacted yields qualitatively very similar results.

[33] On the average municipalities receive about 13% of all income tax revenues. Thus, the per capita sample mean is 260 EUR and the standard deviation equals 110 EUR.

[34] The legal voting age on the federal level is 18.

[35] Accusations of election fraud by the SED regime are widespread. For actual evidence see Herz and Rothe

tions were only held on March 18, 1990—after months of peaceful political protest. The newly elected government then negotiated the end of the GDR.

In stark contrast, citizens of the Federal Republic of Germany (FRG) have had the opportunity to participate in free elections since 1949; and from 1953 on under a two-ballot system almost identical to the current one. Thus, they had more than 40 years of democratic experience by the time the GDR joined the West.

The first parliamentary elections in unified Germany were held on December 2, 1990, and were subject to (essentially) the same rules that had previously been used in the West and that continue to be in place today.[36] If learning and familiarity with the electoral system do indeed matter, then one would expect large initial differences in the share of strategic voters, which should disappear over time.

This prediction corresponds exactly to the findings in Figure 6. For each election since 1990, the figure plots the estimated difference in the share of tactical voters between East and West Germany. Negative values indicate fewer strategic agents among residents of the former GDR.[37]

Just two months after reunification, East Germans were almost 18 percentage points less likely to vote strategically than their Western counterparts. By 2005, however, the gap had practically vanished. Although none of the point estimates is very precise, one can nevertheless reject the null hypothesis of a constant difference at the 1%-confidence level. Moreover, both the initial gap as well as the speed of convergence are in line with the "age effect" in Table 9. Thus, as in many laboratory experiments, sophistication and learning appear to be important determinants of tactical behavior.[38]

7. Structural Analysis and Counterfactual Seat Distributions

Although the reduced form results provide strong evidence of strategic voting, they are subject to a number of limitations. For instance, the assumption that candidate quality enters (3) linearly might be overly restrictive. Taken literally, linearity could lead to predicted vote

(2004).

[36]The most important exception was that the 5%-threshold applied separately to East and West Germany. Thus, in 1990 a party had to gain more than 5% of the list vote in only one of the two regions to enter the Bundestag.

[37]The specification on which the estimates are based is analogous to equation (4), but allows for different slopes and cutoff values in East and West Germany. A qualitatively similar picture emerges if one were to restrict the cutoff to be the same in both regions.

[38]While the evidence is fully consistent with this interpretation, the findings in Figure 6 should be taken with a grain of salt. During this time period, East Germans gained not only familiarity with the electoral system, but other economic factors changed as well. If these changes had an independent effect on the propensity to cast tactical ballots, then the estimates in Figure 6 need not capture the true impact of learning. The convergence in income per capita, however, is almost an order of magnitude too small to explain the results.

shares that are greater than one or even negative. Moreover, by themselves the ordinary least squares estimates do not shed any light on which candidates strategic agents actually vote for (after deserting non-contenders). But, knowledge of substitution patterns is essential to construct appropriate counterfactuals by which to judge whether misrepresentation of preferences is of first order importance for electoral outcomes. To properly account for the drawbacks of the reduced form analysis and to assess the impact of strategic behavior, this section estimates a structural model of voting decisions in the 2009 federal election. Again, list votes provide a crucial source of identifying variation.[39]

7.1. *Adding Structure*

The theoretical model in Section 3 shows that voters will generally group contestants into two categories: candidates who are "in the race" and those who are not. It is, therefore, natural to model agents' decisions as a discrete choice problem in which sincere and strategic voters face different equilibrium choice sets. The former choose among all contestant in a particular district, whereas the latter consider only candidates who are in contention for victory. When it comes to the list vote, however, all voters pick from the set of major parties.

In order to represent agents' preference profiles in a tractable yet flexible fashion, assume that individual i receives utility

$$u^L_{i,p} = \xi_{p,m} + \zeta_{i,p} + \eta_{i,p} \tag{5}$$

from voting for party p's list. Here, $\xi_{p,m}$ denotes the average utility that agents living in municipality m derive from voting for p, and $\zeta_{i,p}$ are individual specific deviations from the mean. $\eta_{i,p}$ is an i.i.d. type-I extreme value (T1EV) taste shock.

Moreover, define the underlying utility from casting one's candidate vote for the nominee of party p to equal

$$u^C_{i,k} = \xi_{p,m} + \zeta_{i,p} + \chi_k + \varepsilon_{i,k}, \tag{6}$$

where k indexes candidates, and χ_k is voters' assessment of k relative to that of her party. $\varepsilon_{i,k}$ denotes another i.i.d. T1EV shock.[40]

It is critical to note that $\xi_{p,m}$ and $\zeta_{i,p}$ appear in both (5) and (6), implying that official party positions influence not only voters' perceptions of the respective organizations, but

[39] Results are qualitatively and quantitatively similar when looking at the 2005 election instead.

[40] The mean utility from abstaining is normalized to zero. Since the available data do not allow turnout to be calculated for individual precincts, the analysis in this section is conducted at the municipality level instead.

also that with respect to their candidates. This assumption captures the fact that German politicians campaign heavily on their own party's platform and it introduces the correlation between list and candidate votes that has been the identifying source of variation above.

To allow individuals' preferences to systematically deviate from the average in their municipality, $\zeta_{i,p}$ is assumed to follow a multivariate normal distribution with an unrestricted covariance matrix. That is, $(\zeta_{i,p})_{p \in P} \sim N(0, \Sigma)$.[41] Hence, supporters of the conservative CDU may, for example, also have a taste for the FDP, while holding more negative views of the communist Left.

While $\xi_{p,m}$ and $\zeta_{i,p}$ model commonalities in voters' assessments of parties and the respective contestants, $\eta_{i,p}$ and $\varepsilon_{i,k}$ admit differences in tastes that go beyond the common perception of candidate quality, i.e. χ_k. The T1EV assumption is convenient because it results in a smooth closed form representation of individual choice probabilities (see, e.g., McFadden 1974 or Train 2009 for derivations).

Given the structure of preferences, party p's expected share of the list vote in municipality m equals

$$(7) \qquad \widehat{v}^{L}_{p,m} = \int \frac{\exp\left(\xi_{p,m} + \zeta_{i,p}\right)}{1 + \sum_{p' \in P} \exp\left(\xi_{p',m} + \zeta_{i,p'}\right)} d\Phi\left(\zeta\right),$$

where P denotes the set of electable parties. Note that $\widehat{v}^{L}_{p,m}$ does not depend on the share of strategic voters. After all, even tactical agents have an ex ante incentive to cast truthful list votes.

This is not true when it comes to the candidate vote. The candidate vote is a mixture of sincere and strategic ballots:

$$\widehat{v}^{C}_{k,m} = \lambda \widehat{v}^{C,S}_{k,m} + (1 - \lambda)\, \widehat{v}^{C,T}_{k,m}.$$

Here, $\widehat{v}^{C,S}_{k,m}$ denotes candidate k's share among sincere voters, and $\widehat{v}^{C,T}_{k,m}$ that among tactical ones. As before, λ is the fraction of agents who are sincere.

Since sincere voters consider every candidate, $\widehat{v}^{C,S}_{k,m}$ is given by

$$(8) \qquad \widehat{v}^{C,S}_{k,m} = \int \frac{\exp\left(\xi_{p,m} + \zeta_{i,p} + \chi_k\right)}{1 + \sum_{k' \in K(d)} \exp\left(\xi_{p',m} + \zeta_{i,p'} + \chi_{k'}\right)} d\Phi\left(\zeta\right),$$

where $K(d)$ marks the set of all contestants in district d. Tactical agents, however, behave as if they are choosing only among the set of serious contenders, $C(d)$. That is, irrespective

[41] As the variance of $\zeta_{i,p}$ is determined by the distribution of the logit error, it is not necessary to impose a normalization.

of the underlying utility in (6), strategic voters disregard all candidates that are not "in the race." Consequently, k's share among strategic individuals equals

$$(9) \qquad \widehat{v}_{k,m}^{C,T}=\begin{cases}\int \frac{\exp(\xi_{p,m}+\zeta_{i,p}+\chi_k)}{1+\sum_{k'\in C(d)}\exp\left(\xi_{p',m}+\zeta_{i,p'}+\chi_{k'}\right)}d\Phi\left(\zeta\right) & \text{if } k\in C(d)\\ 0 & \text{otherwise}\end{cases}.$$

A seemingly natural way to estimate $(\xi,\chi,\Sigma,\lambda)$ would be to find the parameter combination that produces the best fit between predicted vote shares and the data. This, however, entails that preferences would be partially identified from candidate votes, which may confound strategic desertion with simple distaste. In order to avoid this problem, electorates' average tastes should be inferred solely from list votes.

Accordingly, with data on $C(d)$ and actual vote shares in hand, estimates of $(\xi,\chi,\Sigma,\lambda)$ could be obtained by minimizing the objective function:

$$(10) \qquad SSR\left(\xi,\chi,\Sigma,\lambda|v^C,v^L\right)=\sum_{d\in D}\sum_{m\in M(d)}\sum_{k\in K(d)}\left(\widehat{v}_{k,m}^{C}-v_{k,m}^{C}\right)^2$$

subject to the set of constraints

$$(11) \qquad \widehat{v}_{p,m}^{L}=v_{p,m}^{L} \quad \forall p,m,d.$$

Yet, as $C(d)$ is not observed, it needs to be estimated as well. Based on the evidence of focal equilibria in Section 6.1, a candidate is assumed to be a contender if, and only if, her party trails the district's second most popular one by less than κ percentage points. Thus, estimating $C(d)$ adds the following set of equilibrium constraints

$$(12) \qquad \overline{v}_{d}^{L,2^{nd}}-\overline{v}_{k,d}^{L}\leq\kappa \quad \forall k\in C\left(d\right),\forall d$$

$$(13) \qquad \overline{v}_{d}^{L,2^{nd}}-\overline{v}_{k,d}^{L}>\kappa \quad \forall k\notin C\left(d\right),\forall d$$

as well as the additional parameter κ.[42]

Given the granularity of the data, the optimization problem defined by equations (10)–(13) is extremely large. Finding the solution involves optimizing over more than 63,000 parameters, solving about 61,500 non-linear constraints, and approximating roughly 120,000 different five dimensional integrals. To keep the computational burden manageable without compromising the quality of the solution, this paper uses state-of-the-art numerical methods, such as integration on sparse-grids (Heiss and Winschel 2008) and mathematical program-

[42]Experimentation with a subset of the contender classifcations in Table 8 yielded very similar results.

ming with equality constraints (Dube et al. 2012; Su and Judd 2012). For a description of these methods see Appendix D.

Before proceeding to the results it is useful to provide some intuition on how the parameters are identified. Identification of $\xi_{p,m}$ is straightforward. From Berry (1994) it follows that, for every Σ, there exists a unique vector ξ which solves (11). Economically, this means that the list vote pins down the average taste in different markets, as desired.

Akin to the analysis in the previous section, identification of λ is based on the intuition in Figure 1. That is, the share of sincere voters can be inferred from the ratio of non-contenders' observed vote shares (depicted on the y-axis) to those they would receive if all agents acted solely based on their preferences (proxied by the position on the x-axis). As the list vote determines mean preferences, structural and reduced form methods rely ultimately on very similar variation.

Candidate quality, i.e. χ_k, can be gleaned by comparing contestant's actual performance in different municipalities with predictions thereof based on party preferences and λ. χ_k will be positive for candidates whose vote shares systematically exceed their predicted values, and negative for those who underperform.

Lastly, Σ is identified from the empirical covariance between non-contenders' list and contenders' candidate votes. Take, for instance, a district in which the FDP candidate is out of the race, while the nominee of the CDU is a contender. If the latter receives, on average, more votes in villages that have a greater taste for the FDP, then the respective parameter in the covariance matrix will be positive. Analogous arguments apply to the remaining elements of Σ.

7.2. *Results and Counterfactual Experiments*

With 26.3% (and a standard error of 7.8%) the estimated share of strategic voters, i.e. $1-\lambda$, is strikingly close to the reduced form results in Tables 5 and 6. Unfortunately, few of the model's other parameters are easily interpretable by themselves. Thus, instead of listing parameter estimates, the following discussion presents results in a way that relates straightforwardly to common intuition.[43]

In order to judge the model's fit consider Figure 7. The upper two panels contrast the true marginal distributions of candidate and list votes (dark bars) with those predicted by the model (light bars). Given that $(\xi, \chi, \Sigma, \lambda, \kappa)$ have been chosen to mimic these data, there are practically no discernible differences.

The lower panel depicts the frequency of valid list and candidate vote *combinations*. It is important to note that information on the joint distribution of votes come from an indepen-

[43] A list of all estimates is available from the author upon request.

dent source (Bundeswahlleiter 2009a, 2010) and were *not* used to fit the model. Thus, the lower panel of Figure 7 provides a strong quasi-out-of-sample test of whether the estimation results are reasonable. Although there do exist differences, on the whole the predicted distribution matches the qualitative features of its real world counterpart fairly well, lending credibility to the structural estimates.

Table 11 compares actual and simulated outcomes of district level races. As can be seen from the entries on the diagonal, the model does an excellent job at ranking candidates. In particular, it predicts almost 95% of winners correctly.

While Figure 7 and Table 11 are useful in evaluating the goodness of fit, a more interesting question might be for whom supporters of different parties would vote if their preferred candidate was out of the race. In order to shed light on the ordering of preferences, Table 12 shows the frequency with which voters would substitute toward the candidate of any other party, assuming that all but their preferred contestant were still in the race. Thus, the entries correspond to the probability of some other party's candidate being "the next best choice." The model predicts FDP adherents to substitute toward candidates of the CDU, whereas most supporters of the Green Party and The Left would choose SPD contenders instead. Given parties' ideological positions, these patterns conform exactly to what one would expect.

Based on the structural estimates, Figure 8 presents several counterfactual election results by which to judge the impact of strategic voting.[44] The top left panel shows the actual distribution of seats in the Bundestag, whereas the panel on the right displays the distribution that would prevail if mandates were awarded based solely on a single vote counted under proportionality rule with a 5%-threshold, i.e. the list vote. Evidently, the current Bundestag mirrors a parliament formed under proportional representation fairly closely: all five major parties are represented, with more than 60% of seats accruing to the CDU and the SPD. In the current equilibrium, distortions introduced through strategic candidate votes are very small.

The remaining two panels assume a single vote counted under plurality rule on the district level (akin to the candidate vote, or elections to the House of Representatives in the U.S.) The counterfactual on the bottom left shows the model's predictions for such a first-past-the-post scheme with 26.3% of voters behaving strategically. In the panel on the bottom right all voters are sincere.

In line with common intuition, relative to proportional representation a "winner-take-all" system would result in dramatic losses for small parties. However, as comparing the panels on the right shows, these losses are due to the way different electoral rules map vote shares

[44] For details on the computation of these counterfactuals see Appendix D.

into mandates and not to tactical behavior.

The impact of strategic voting can be gleaned from comparing the two counterfactuals on the bottom. Given its estimated extent, tactical voting has only a modest effect on the overall allocation of seats. Not a single party's share of seats would change by more than 5 percentage points, often substantially less. Although these results depend on the distribution of preferences, at first glance it would appear that lack of strategyproofness has only a modest effect on electoral outcomes.

Yet, looking only at seat totals misses an important point. As the evidence in Table 13 indicates, compared to the current equilibrium, about one in ten districts would change hands if all voters were to cast truthful ballots. Thus, even moderate degrees of tactical voting may have a non-negligible impact on a given race—as predicted by theory. But as the number of heterogenous districts becomes large, distortions tend to average out. In fact, the effect of tactical behavior on aggregate seat totals will always be (weakly) lower than that on individual races.

8. Concluding Remarks

Whether individuals vote strategically is one of the most important questions at the intersection of economics and political science. Although tactical voting has interested social scientists for decades, it has proven very difficult to estimate its extent. The fundamental problem is that voters' preferences are generally unobserved, which makes it nearly impossible to compare them to actual votes without imposing strong assumptions.

The present paper resolves this problem by exploiting the structure of parliamentary elections in Germany. The German system allows for preferences to be held fixed, while contrasting the behavior of the same individuals under different electoral regimes. Reduced form as well as structural estimates indicate that approximately 30% of voters behave as predicted by rational choice theory. That is, almost one third of voters abandons their most preferred candidate if she is believed to have no chance of winning. Interestingly, strategic voting is more common among individuals of high socio-economic status and those more familiar with the electoral rules, suggesting that sophistication and learning play an important role.

Although strategic voting has a non-trivial impact on individual races, its effect on the overall distribution of seats is considerably more modest. No party's seat total would change by more than a few percentage points if all voters behaved sincerely. Depending on the application it may, thus, be important to distinguish between the "micro" and "macro" effects of tactical behavior.

With more than a dozen countries using electoral systems similar to the German one future research may follow the approach in this paper and construct comparable estimates for a number of different settings. The findings from such a cross-country analysis could further enhance our understanding of tactical voting and shed light on the vulnerability of elections around the world.

Appendix A: Proofs

Proof of Proposition 1: To demonstrate existence of voting equilibria this proof adapts the fixed point argument in Myerson (1998, 2000).

First, let ex ante beliefs about pivot probabilities be accurate and express expected utility as a function of voters' strategies:

$$\widehat{u}_\imath(\delta,\sigma)=\sum_{k\in K}\left[\delta(k|\imath)\sum_{\tau\in Z(K)}\left(\prod_{k'\in K}\psi\left(\tau(k')\,|\mu(k'|\sigma)\right)\right)p(k,\tau|\imath)\right],$$

with payoffs given by $p(k,\tau|u,s)=u_k$ and

(14)

$$p(k,\tau|u,t)=\begin{cases}\frac{1}{2}(u_k-u_{k'}) & \text{if } \exists k'\in K \text{ s.t. } \tau(k')=\tau(k)\wedge\tau(k)\geq\tau(k'')\,\forall k''\in K\backslash\{k'\}\\ \frac{1}{2}(u_k-u_{k'}) & \text{if } \exists k'\in K \text{ s.t. } \tau(k')=\tau(k)+1\wedge\tau(k)\geq\tau(k'')\,\forall k''\in K\backslash\{k'\}\\ 0 & \text{otherwise}\end{cases}.$$

Here $\widehat{u}_\imath$ denotes the expected utility of a type $\imath\in\mathcal{I}$ voter who pursues strategy $\delta\in\Delta(K)$ when everybody else plays $\sigma\in\Delta(K)^{\mathcal{I}}$, $\mu(k|\sigma)\equiv n\sum_{u\in U}\left[\lambda\sigma(k|u,s)+(1-\lambda)\sigma(k|u,t)\right]f(u)$, and $Z(K)$ is the set of all possible action profiles, i.e. the number of other players who choose actions $(\tau(k))_{k\in K}$.[45] Next, define the best response correspondence $B:\Delta(K)^{\mathcal{I}}\rightrightarrows\Delta(K)^{\mathcal{I}}$ such that, for all σ, $B=(B_\imath(\sigma))_{\imath\in\mathcal{I}}$, where $B_\imath(\sigma)\equiv\{\delta|\widehat{u}_\imath(\delta,\sigma)\geq\widehat{u}_\imath(\delta',\sigma)\,\forall\delta'\in\Delta(K)\}$. With these definitions in hand it suffices to show that there exists a set of mixed strategies for which $\sigma\in B(\sigma)$. The remainder of the proof verifies that the conditions for Kakutani's Fixed Point Theorem are satisfied.

By definition of $\Delta(K)$ as k-dimensional simplex, $\Delta(K)^{\mathcal{I}}$ is a non-empty, compact, and convex subset of a finite dimensional Euclidean space.

As $\widehat{u}=(\widehat{u}_\imath)_{\imath\in\mathcal{I}}$ is continuous, it follows from the definition of B, compactness of $\Delta(K)^{\mathcal{I}}$, and the Weierstrass Extreme Value Theorem that $B(\sigma)$ is non-empty.

[45] Note that (14) could have alternatively been expressed as

$$p(k,\tau|u,t)=\begin{cases}u_k & \text{if } \tau(k)\geq\tau(k')+1\;\forall k'\in K\backslash\{k\}\\ \frac{1}{2}(u_k-u_{k'}) & \text{if } \exists k'\in K \text{ s.t. } \tau(k')=\tau(k)\wedge\tau(k)\geq\tau(k'')\,\forall k''\in K\backslash\{k'\}\\ \frac{1}{2}(u_k-u_{k'}) & \text{if } \exists k'\in K \text{ s.t. } \tau(k')=\tau(k)+1\wedge\tau(k)\geq\tau(k'')\,\forall k''\in K\backslash\{k'\}\\ u_{k'} & \text{if } \exists k'\in K \text{ s.t. } \tau(k')>\tau(k)+1\wedge\tau(k')\geq\tau(k'')\,\forall k''\in K\end{cases}.$$

In (14) strategic voters care only about influencing the outcome of the election, whereas in the formulation above they derive utility from whichever candidate wins the elections. Since payoffs differ only by a constant, the two specifications are equivalent.

Moreover, $B(\sigma)$ is convex-valued. To see this note that if $\delta', \delta'' \in B_\imath(\sigma)$, then, for all $\alpha \in [0,1]$, $\alpha\widehat{u}_\imath(\delta',\sigma) + (1-\alpha)\widehat{u}_\imath(\delta'',\sigma) \geq \widehat{u}_\imath(\delta''',\sigma)\,\forall\delta''' \in \Delta(K)$. Also, $\alpha\widehat{u}_\imath(\delta',\sigma) + (1-\alpha)\widehat{u}_\imath(\delta'',\sigma) = \widehat{u}_\imath(\alpha\delta' + (1-\alpha)\delta'',\sigma)$ because $\widehat{u}_\imath$ is linear in δ. Together with the previous statement this implies $\alpha\delta' + (1-\alpha)\delta'' \in B_\imath(\sigma)$, as required.

By way of contradiction suppose that $B(\sigma)$ does not have a closed graph. Then there exists a sequence $(\sigma_n,\delta_n) \to (\sigma,\delta)$ such that $\delta_n \in B(\sigma_n)$ but $\delta \notin B(\sigma)$. Hence, $\exists\delta' \in \Delta(K)$ and some $\epsilon > 0$ for which $u_\imath(\delta',\sigma) > u_\imath(\delta,\sigma) + 3\epsilon$. From continuity of $\widehat{u}_\imath$ and since $\sigma_n \to \sigma$, it follows that for n large enough, $u_\imath(\delta',\sigma_n) > u_\imath(\delta',\sigma) - \epsilon$. Combining the two preceding inequalities gives $u_\imath(\delta',\sigma_n) > u_\imath(\delta,\sigma) + 2\epsilon$, and continuity of $\widehat{u}_\imath$ implies that $u_\imath(\delta,\sigma) + 2\epsilon > u_\imath(\delta_n,\sigma_n) + \epsilon$. But $u_\imath(\delta',\sigma_n) > u_\imath(\delta_n,\sigma_n) + \epsilon$ contradicts $\delta_n \in B(\sigma_n)$. Hence, $B(\sigma)$ must have a closed graph.

From Kakutani's Fixed Point Theorem it then follows that there exists at least one $\sigma \in \Delta(K)^{\mathcal{I}}$ for which $\sigma \in B(\sigma)$, as desired. *Q.E.D.*

Appendix B: Calculating a Party's Number of Seats

Following Spenkuch (2012), this appendix explains the algorithm that is currently used to calculate a party's number of seats in the Bundestag. Let $d_{p,s}$ denote the number of direct mandates accruing to party p in state s. $v_{p,s}$ is the number of list votes that p received in s, with the equivalent number on the national level given by $\overline{v}_p = \sum_s v_{p,s}$. With this notation in hand, party p's seat total is calculated in three steps:

Step 1: Proportional Allocation of List Mandates to Parties. Absent overhang mandates, there are 598 seats in the Bundestag. These are allocated by proportionality rule to the set of parties clearing the 5%-threshold or winning at least three direct mandates. That is, the number of list mandates of party p equals

$$\bar{l}_p \cong \begin{cases} 598\frac{\overline{v}_p}{\sum_{p'\in\widetilde{P}}\overline{v}_{p'}} & \text{if } p \in \widetilde{P} \\ 0 & \text{otherwise} \end{cases},$$

where $\widetilde{P} = \left\{p|\frac{\overline{v}_p}{\sum_{p'}\overline{v}_{p'}} \geq .05 \vee \sum_s d_{p,s} \geq 3\right\}$ and $\cong$ represents equality after rounding according to the Sainte-Laguë method, which ensures that $\sum_p \bar{l}_p = 598$.[46]

Step 2: Proportional Allocation of Mandates to State Lists. German electoral law requires parties to compete with different lists in each state. Therefore, list mandates need to be allocated to the respective state lists. In practice, the number of mandates awarded to a party's state list is proportional to the list's contribution to the party's vote total. More precisely, for all s and all p,

$$l_{p,s} \cong \begin{cases} \bar{l}_p\frac{v_{p,s}}{\overline{v}_p} & \text{if } p \in \widetilde{P} \\ 0 & \text{otherwise} \end{cases},$$

[46] In 2005 the method of Hare-Niemeyer was used instead.

where $\cong$ is defined as above.

Step 3: Determination of the Actual Number of Seats. However, the *actual* number of seats that party p receives in state s is given by

$$n_{p,s} = \max\left\{d_{p,s}, l_{p,s}\right\}.$$

If $d_{p,s} < l_{p,s}$ then, in addition to the district winners, the first $l_{p,s} - d_{p,s}$ candidates on p's list in s are elected to the Bundestag as well. Otherwise, only holders of direct mandates receive a seat.

Note that only if $d_{p,s} \leq l_{p,s}$ for all s, will party p's seat total, $\overline{n}_p = \sum_s n_{p,s}$, be equal to the number of seats it would be assigned under proportional representation, i.e. $\bar{l}_p$.

Appendix C: Variable Definitions

This appendix provides a description of all data used in the paper, as well as precise definitions together with the sources of all variables.

C.1. *Election Results*

Data containing the official results of the 1980, 1983, 1987, 1990, 1994, and 1998 federal elections by municipality (Gemeinde) as well as the 2002, 2005, and 2009 elections by polling precinct (Wahlbezirk) have been purchased from the Federal Returning Officer. These data include information on the number of list and candidate votes for each party and each candidate, the number of eligible voters, as well as the number of valid and invalid votes. In 2009 there were approximately 89,000 precincts. Whenever necessary precinct level numbers are aggregated using the municipality identifiers contained in the raw data. Municipalities spanning multiple districts are discarded. Throughout the analysis the following variables are used:

Number of Eligible Voters is defined as the number of residents of each precinct that were allowed to vote in the particular year. In general this encompasses all German citizens over the age of 18, who have not been declared mentally unfit, or whose voting rights have not been suspended due to criminal behavior.

Turnout is defined as the number of actual voters over the number of eligible voters. This number cannot be calculated for absentee precincts, as absentee voters are included in the number of eligible voters in their district of residence. Hence, in-person turnout in each district needs to be adjusted for absentee voters. In practice, this is done by multiplying the number of issued absentee ballots by .95 (which corresponds to the empirical frequency with which they are cast) and adding them to the ballots that are cast in person.

Share of List Vote is defined as the portion of all valid list votes (in %) that are cast for a particular party. "Micro parties", i.e. those not clearing the 5%-threshold, are grouped together.

Share of Candidate Vote is defined as the portion of all valid candidate votes (in %) that are cast for the candidate of a particular party. Votes for candidates of "micro parties" are pooled.

Absentee Precinct is an indicator variable equal to one if a given precinct handles only absentee ballots.

C.2. *Candidate Characteristics*

Prior to every election to the Bundestag the Federal Returning Officer publishes information on certain characteristics of all official list and direct candidates. This paper focuses only on the latter. The data have been compiled from Bundeswahlleiter (2005c, 2009b). Throughout the analysis the following variables are used:

Age at the time of the election is defined as election year minus year of birth.

Female is an indicator variable equal to one if a candidate is female, and zero otherwise.

Doctorate is an indicator variable equal to one if a candidate holds a doctoral degree and/or a professorship, and zero otherwise. As doctoral degrees are part of Germans' official names, this variable has been created using a text search for "Dr." and "Prof.".

Currently Member of Parliament is an indicator variable equal to one if the candidate holds a list or direct mandate, and zero otherwise.

Holds Direct Mandate is an indicator variable equal to one if the candidate holds a direct mandate, and zero otherwise.

Also on Part List is an indicator variable equal to one if the candidate does not only run in the district race, but is also on her party's state list (and could thus enter the Bundestag either way).

Position on Party List denotes the candidate's rank on her party's state list (conditional on having been placed on the list).

C.3. *Municipality Characteristics*

Information on municipalities' demographic and socio-economic characteristics is taken from *Statistik lokal 2007* and *Statistik lokal 2011* (Statistische Ämter des Bundes und der Länder 2007, 2011). *Statistik lokal* is an annual publication of the German Federal Statistical Office and the statistical offices of the Länder containing data on various characteristics of approximately 12,000 municipalities and administrative units in Germany as of about 2 years before to the publication date. These data have been linked with the election results described above using the municipality identifier (Allgemeiner Gemeindeschlüssel) contained in both data sets. Below follows a brief description of all municipality level variables used throughout the analysis.

Population Density is defined as a municipality's total average population (in thousands) per square kilometer during the respective calendar year.

Fraction of Voters Female is defined as the share of women among a municipality's population over the age of 18.

Income Tax Revenue Per Capita is defined as the total income tax receipts (in 1,000 EUR) accruing to the respective municipality divided by its population during the same calendar year.

Fraction of Voters under Age 30 is defined as the fraction of individuals aged 18–30 among those over the age of 18.

Appendix D: Numerical Methods

This appendix describes the numerical methods used to solve the optimization problem defined by equations (10)–(13) as well as the construction of counterfactual election results in Section 7.2.

D.1. *Mathematical Programming with Equality Constraints*

Typically, to recover mean utilities in models of discrete choice (i.e. $\xi_{p,m}$) researchers turn to inverting the system of non-linear markets share equations via the nested fixed point (NFP) algorithm in Berry (1994) and Berry et al. (1995). Recently, however, Su and Judd (2012) and Dube et al. (2012) have shown how to recast extremum estimators in general, and the one in the Berry et al. (1995) in particular, as a mathematical programming problem with equality constraints (MPEC).

Key to the MPEC approach is the insight that the inner loop can be eliminated entirely by recasting the estimator as an optimization problem subject to a set of non-linear constraints, i.e. (11), which require predicted market shares to equal observed ones.

Since objective function and market share equations are usually smooth, one can rely on state-of-art optimization software, such as KNITRO or SNOPT, to find candidate solutions. Moreover, dispensing with the inner loops avoids numerical problems associated with loose inner loop error tolerances (see Dube et al. 2012 for a discussion of the NFP algorithm's numerical properties), and it may significantly increase computational speed because the system of market share equations does not have to be solved exactly at each iteration. (The constraints have to be satisfied only at the solution.) Importantly, Su and Judd (2012) prove that MPEC and NFP solve the same problem, yielding the same estimates with the same statistical properties.

The implementation of MPEC in this paper is based on the MATLAB code of Dube et al. (2012), using both of the KNITRO solver's interior-point and active set algorithms (Byrd et al. 1999, 2004, 2006). To improve numerical accuracy as well as computational performance, KNITRO is provided with hand-coded first-order analytical derivatives of the objective function and the constraints, second order derivatives, as well as the sparsity patterns of the constraint Jacobian and the Hessian. Since the Hessian contains almost 4×10^9 elements of which only about 1.8×10^6 are non-zero, supplying the solver with the sparsity pattern is critical in order to economize on memory usage and time. To increase the likelihood of finding the global optimum five different starting points are used. Relative optimality and feasibility error tolerances, i.e. the maximum violation of the first order conditions and the constraints, have each been set to 10^{-6}. Reported standard errors are based on the block-bootstrap with 100 iterations.

In order to provide the solver with a completely smooth optimization problem, the constraints in (12)–(13) have been rewritten as an indicator function for each candidate, $c_k(\kappa)$, and are numerically approximated by the hyperbolic tangent. That is, $c_k(\kappa) = \frac{1}{2} + \frac{1}{2}\tanh\left(\rho\left(\kappa + \overline{v}^L_{k,d} - \overline{v}^{L,2^{nd}}_d\right)\right)$ for $\rho = 5{,}000$. Thus, equation (9) becomes $\widehat{v}^{C,T}_{k,m} = \int \frac{c_k \exp(\xi_{p,m}+\zeta_{i,p}+\chi_k)}{1+\sum_{k'\in K} c_k \exp\left(\xi_{p',m}+\zeta_{i,p'}+\chi_{k'}\right)} d\Phi(\zeta)$.

D.2. *Sparse Grid Integration*

Instead of solving the approximately 120,000 five dimensional integrals in equations (7), (8), and (9) using simulation methods, the present paper relies on sparse grid integration (SGI), introduced into economics by Heiss and Winschel (2008). SGI provides a way to approximate integrals numerically avoiding the curse of dimensionality associated with conventional quadrature rules (see Judd 1998). Monte Carlo evidence by Skrainka and Judd (2011) indicates that SGI imposes a significantly lower computational burden than simulation methods achieving the same level of accuracy.

SGI is closely related to conventional Gaussian quadrature rules, but by exploiting symmetry properties it relies only on a small subset of nodes and (appropriately rescaled) weights. This paper uses a Konrad-Patterson rule with Gaussian kernel for choosing nodes, as explained in Heiss and Winschel (2008). This particular rule has only 151 nodes; yet it exactly integrates (over five dimensions) all complete polynomials of total order less than 7. Experimentation with more accurate rules yielded essentially the same point estimates, but required significantly more CPU time.

D.3. *Construction of Counterfactuals*

The counterfactual election results in Section 7.2 of this paper have been constructed by simulation. More specifically, for each municipality in the data 100 times its actual number of voters have been simulated by randomly drawing ζ, η, and ε from the respective (estimated) distributions. A fraction λ of simulated voters (rounded to the nearest integer) are designated to behave tactically. Next, each voter's candidate and party specific utilities are calculated and his (partial) preference orderings for the list and candidate vote are determined. Naturally, sincere voters consider all candidates, whereas tactical voters choose only among those contestants who are estimated to be contenders. Election results are then constructed by aggregating votes to the appropriate level, and applying the specified electoral rule.

References

Abramson, P. R., J. H. Aldrich, P. Paolino, and D. W. Rohde (1992). "Sophisticated Voting in the 1988 Presidential Primaries," *American Political Science Review*, 86, 55–69.

Alvarez, R. M., and J. Nagler (2000). "A New Approach for Modelling Strategic Voting in Multiparty Elections." *British Journal of Political Science*, 30, 57–75.

——— , F. J. Boehmke, and J. Nagler (2006). "Strategic Voting in British Elections." *Electoral Studies*, 25, 1–19.

Andrews, D. W. K. (1993). "Tests for Parameter Instability and Structural Change with Unknown Change Point." *Econometrica*, 61, 821–856.

Arrow, K. J. (1951). *Social Choice and Individual Values.* New York: Wiley.

Austen-Smith, D., and J. S. Banks (1988). "Elections, Coalitions, and Legislative Outcomes." *American Political Science Review*, 82, 405–422.

Bai, J. (1997). "Estimation of a Change Point in Multiple Regression Models." *Review of Economics and Statistics*, 79, 551–563.

Barberà, S. (2011). "Strategyproof Social Choice," (pp. 731–831) in K. J. Arrow, A. K. Sen, and K. Suzumaru (eds.), *Handbook of Social Choice and Welfare, Vol. 2.* Amsterdam: Elsevier.

Barnes, S. H., F. Grace, J. K. Pollock, and P. W. Sperlich (1962). "The German Party System and the 1961 Federal Election." *American Political Science Review*, 56, 899–914.

Bawn, K. (1993). "The Logic of Institutional Preferences: German Electoral Law as a Social Choice Outcome." *American Journal of Political Science*, 34, 965–989.

Berry, S. T. (1994) "Estimating Discrete-Choice Models of Product Differentiation." *RAND Journal of Economics*, 25, 242–262.

——— , J. A. Levinsohn, and A. Pakes (1995). "Automobile Prices in Market Equilibrium." *Econometrica*, 63, 841–890.

Besley, T., and S. Coate (1997). "An Economic Model of Representative Democracy." *Quarterly Journal of Economics*, 112, 85–114.

Black, D. (1948). "On the Rationale of Group Decision-making." *Journal of Political Economy*, 56, 23–34.

Blais, A., R. Nadeau, E. Gidengil, and N. Nevitte (2001). "Measuring Strategic Voting in Multiparty Elections." *Electoral Studies*, 20, 343–352.

Bundeswahlleiter (2005a). *Wahl zum 16. Deutschen Bundestag am 18. September 2005. Heft 3: Endgültige Ergebnisse nach Wahlkreisen.* Wiesbaden: Statistisches Bundesamt.

——— (2005b). *Wahlkreiskarte für die Wahl zum 16. Deutschen Bundestag.* Wiesbaden: Statistisches Bundesamt.

——— (2005c). *Wahl zum 16. Deutschen Bundestag am 18. September 2005. Sonderheft: Die Wahlbewerber für die Wahl zum 16. Deutschen Bundestag 2005.* Wiesbaden: Statistisches Bundesamt.

——— (2006). *Wahl zum 16. Deutschen Bundestag am 18. September 2005. Heft 4: Wahlbeteiligung und Stimmabgabe der Männer und Frauen nach Altersgruppen.* Wiesbaden: Statistisches Bundesamt.

——— (2008). *Wahlkreiskarte für die Wahl zum 17. Deutschen Bundestag.* Wiesbaden: Statistisches Bundesamt.

——— (2009a). *Wahl zum 17. Deutschen Bundestag am 27. September 2009. Heft 3: Endgültige*

Ergebnisse nach Wahlkreisen. Wiesbaden: Statistisches Bundesamt.

——— (2009b). *Wahl zum 17. Deutschen Bundestag am 27. September 2009. Sonderheft: Die Wahlbewerber für die Wahl zum 17. Deutschen Bundestag 2009*. Wiesbaden: Statistisches Bundesamt.

——— (2010). *Wahl zum 17. Deutschen Bundestag am 27. September 2009. Heft 4: Wahlbeteiligung und Stimmabgabe der Männer und Frauen nach Altersgruppen*. Wiesbaden: Statistisches Bundesamt.

Byrd, R. H., M. E. Hribar, and J. Nocedal (1999). "An Interior Point Algorithm for Large Scale Nonlinear Programming." *SIAM Journal on Optimization*, 9, 877–900.

——— , N. I. M. Gould, J. Nocedal, and R. A. Waltz (2004). "An Algorithm for Nonlinear Optimization using Linear Programming and Equality Constrained Subproblems." *Mathematical Programming, Series B*, 100, 27–48.

——— , J. Nocedal, and R. A. Waltz (2006). "KNITRO: An Integrated Package for Nonlinear Optimization," (pp. 35–59) in G. di Pillo and M. Roma (eds.), *Large-Scale Nonlinear Optimization*, New York: Springer.

Card, D., A. Mas, and J. Rothstein (2008). "Tipping and the Dynamics of Segregation." *Quarterly Journal of Economics*, 123, 177–218.

Carroll, G. (2011). "A Quantitative Approach to Incentives: Application to Voting Rules." Unpublished Manuscript. Massachusetts Institute of Technology.

Coate, S., M. Conlin, and A. Moro (2008). "The Performance of Pivotal-Voter Models in Small-Scale Elections: Evidence from Texas Liquor Referenda." *Journal of Public Economics*, 92, 582–596.

Cox, G. W. (1994). "Strategic Voting Equilibria Under the Single Nontransferable Vote." *American Political Science Review*, 88, 608–621.

——— (1997). *Making Votes Count*. Cambridge, UK: Cambridge University Press.

Degan, A., and A. Merlo (2009). "Do Voters Vote Ideologically?" *Journal of Economic Theory*, 144, 1868–1894.

Downs, A. (1957). *An Economic Theory of Democracy*. New York: Harper & Row.

Duverger, M. (1954). *Political Parties: Their Organization and Activity in the Modern State*. New York: Wiley.

Dube, J.-P. H., J. T. Fox, and C.-L. Su (2012). "Improving the Numerical Performance of BLP Static and Dynamic Discrete Choice Random Coefficients Demand Estimation." *Econometrica*, 80, 2231–2267.

Duffy, J., and M. Tavits (2008). "Beliefs and Voting Decisions: A Test of the Pivotal Voter Model." *American Journal of Political Science*, 52, 603–618.

Eckel, C., and C. A. Holt (1989). "Strategic Voting in Agenda-Controlled Experiments." *American Economic Review*, 79, 763–773.

Farquharson, R. (1969). *Theory of Voting*. Oxford: Blackwell.

Feddersen, T. J., and W. Pesendorfer (1996). "The Swing Voter's Curse." *American Economic Review*, 86, 408–424.

Fisher, S. L. (1973). "The Wasted Vote Thesis." *Comparative Politics*, 5, 293–299.

Forsythe, R., R. B. Myerson, T. A. Rietz, and R. J. Weber (1993). "An Experiment on Coordination in Multi-Candidate Elections: The Importance of Polls and Election Histories." *Social Choice and Welfare*, 10, 223–247.

——— , T. A. Rietz, R. B. Myerson, and R. J. Weber (1996). "An Experimental Study of Voting Rules and Polls in Three-Candidate Elections." *International Journal of Game Theory*, 25, 355–383.

Fujiwara, T. (2011). "A Regression Discontinuity Test of Strategic Voting and Duverger's Law." *Quarterly Journal of Political Science*, 6, 197–233.

Gibbard, A. (1973). "Manipulation of Voting Schemes: A General Result." *Econometrica*, 41, 587–601.

Gschwend, T. (2001). *Strategic Voting in Mixed Electoral Systems*. Doctoral Dissertation. State University of New York at Stony Brook.

——— (2007). "Ticket-splitting and strategic voting under mixed electoral rules: Evidence from Germany." *European Journal of Political Research*, 46, 1–23.

Hansen, B. E. (2000). "Sample Splitting and Threshold Estimation." *Econometrica*, 68, 575–603.

Herz, A., and A. Rothe (2004). *Wahl und Wahlbetrug im Mai 1989*. Erfurt: Landesbeauftragter Thüringen für die Unterlagen der Staatssicherheitsdienstes der ehemaligen DDR.

Heiss, F., and V. Winschel (2008). "Likelihood Approximation by Numerical Integration on Sparse Grids." *Journal of Econometrics*, 144, 62–80.

Jesse, E. (1988). "Split-Voting in the Federal Republic of Germany: An Analysis of the Federal Elections from 1953 to 1987." *Electoral Studies*, 7, 104–124.

Judd, K. L. (1998). *Numerical Methods in Economics*. Cambridge, MA: MIT Press.

Kawai, K., and Y. Watanabe (2012). "Inferring Strategic Voting." *American Economic Review*, forthcoming.

Klingeman, H.-D., and B. Wessels (2001). "The Political Consequences of Germany's Mixed-Member System: Personalization at the Grass Roots?" (pp. 279–296) in M. S. Shugart and M. P. Wattenberg (eds.), *Mixed-Member Electoral Systems: The Best of Both Worlds?* Oxford: Oxford University Press.

Korte, K.-R. (2009). "Die Bundestagswahlen 2005 als Critical Elections." *Der Bürger im Staat*, 59, 68–73.

——— (ed.) (2010a). *Die Bundestagswahl 2009: Analysen der Wahl-, Parteien-, Kommunikations und Regierungsforschung*. Wiesbaden: VS Verlag für Sozialwissenschaften.

——— (2010b). *Wahlen in Deutschland*. Bonn: Bundeszentrale für Politische Bildung.

McFadden, D. L. (1974). "Conditional Logit Analysis of Qualitative Choice Behavior," (pp. 105–142) in P. Zarembka (ed.), *Frontiers in Econometrics*. New York: Academic Press.

Myatt, D. P. (2007). “On the Theory of Strategic Voting.” *Review of Economic Studies*, 74, 255–281.

——— (2012). “A Rational Choice Theory of Voter Turnout.” Unpublished Manuscript. London Business School.

Myerson, R. B. (1998). “Population Uncertainty and Poisson Games.” *International Journal of Game Theory*, 27, 375–392.

——— (2000). “Large Poisson Games.” *Journal of Economic Theory*, 94, 7–45.

——— (2002). “Comparison of Scoring Rules in Poisson Voting Games.” *Journal of Economic Theory*, 103, 219–251.

———, and Robert J. Weber (1993). “A Theory of Voting Equilibria.” *American Political Science Review*, 87, 102–114.

Niemi, R. G., G. Whitten, and M. N. Franklin (1993). “Constituency Characteristics, Individual Characteristics and Tactical Voting in the 1987 British General Election.” *British Journal of Political Science*, 22, 229–240.

Osborne, M. J., and A. Slivinski (1996). “A Model of Political Competition with Citizen Candidates.” *Quarterly Journal of Economics*, 111, 65–96.

Pappi, F. U., and P. W. Thurner (2002). “Electoral Behaviour in a Two-Vote System: Incentives for Ticket Splitting in German Bundestag Elections.” *European Journal of Political Research*, 41, 207–232.

Reed, S. R. (1990). “Structure and Behaviour: Extending Duverger’s Law to the Japanese Case.” *British Journal of Political Science*, 20, 335–356.

Riker, W. H., and P. C. Ordeshook (1968). “A Theory of the Calculus of Voting.” *American Political Science Review*, 62, 25–42.

Satterthwaite, M. A. (1975). “Strategy-proofness and Arrow’s Conditions: Existence and Correspondence Theorems for Voting Procedures and Social Welfare Functions.” *Journal of Economic Theory*, 10, 187–217.

Sen, A. K. (1970). *Collective Choice and Social Welfare*. San Francisco: Holden-Day.

Simon, H. A. (1954). “Bandwagon and Underdog Effects and the Possibility of Election Predictions.” *Public Opinion Quarterly*, 18, 245–253.

Spenkuch, J. L. (2012). “Please Don’t Vote for Me: Strategic Voting in a Natural Experiment with Perverse Incentives.” Unpublished Manuscript. University of Chicago.

Statistische Ämter des Bundes und der Länder (2007). *Statistik lokal 2007: Daten für die Kreise, kreisfreien Städte und Gemeinden Deutschlands*. Wiesbaden: Statistisches Bundesamt.

——— (2011). *Statistik lokal 2011: Daten für die Kreise, kreisfreien Städte und Gemeinden Deutschlands*. Wiesbaden: Statistisches Bundesamt.

Su, C.-L., and K. L. Judd (2012). “Constrained Optimization Approaches to Estimation of Structural Models.” *Econometrica*, 80, 2213–2230.

Skrainka, B. S., and K. L. Judd (2011). “High Performance Quadrature Rules: How Numerical Integration Affects a Popular Model of Product Differentiation." Unpublished Manuscript. University College London.

TRAIN, K. E. (2009). *Discrete Choice Methods with Simulation* (2nd ed.). New York: Cambridge University Press.

WRIGHT, G. C. (1990). "Misreports of Vote Choice in the 1988 NES Senate Election Study." *Legislative Studies Quarterly*, 15, 543–63.

——— (1992). "Reported Versus Actual Vote: There Is a Difference and It Matters." *Legislative Studies Quarterly*, 17, 131–42.

YATCHEW, A. (1998). "Nonparametric Regression Techniques in Economics." *Journal of Economic Literature*, 36, 669–721.

Figure 1: Theoretical Predictions under Sincere and Strategic Voting

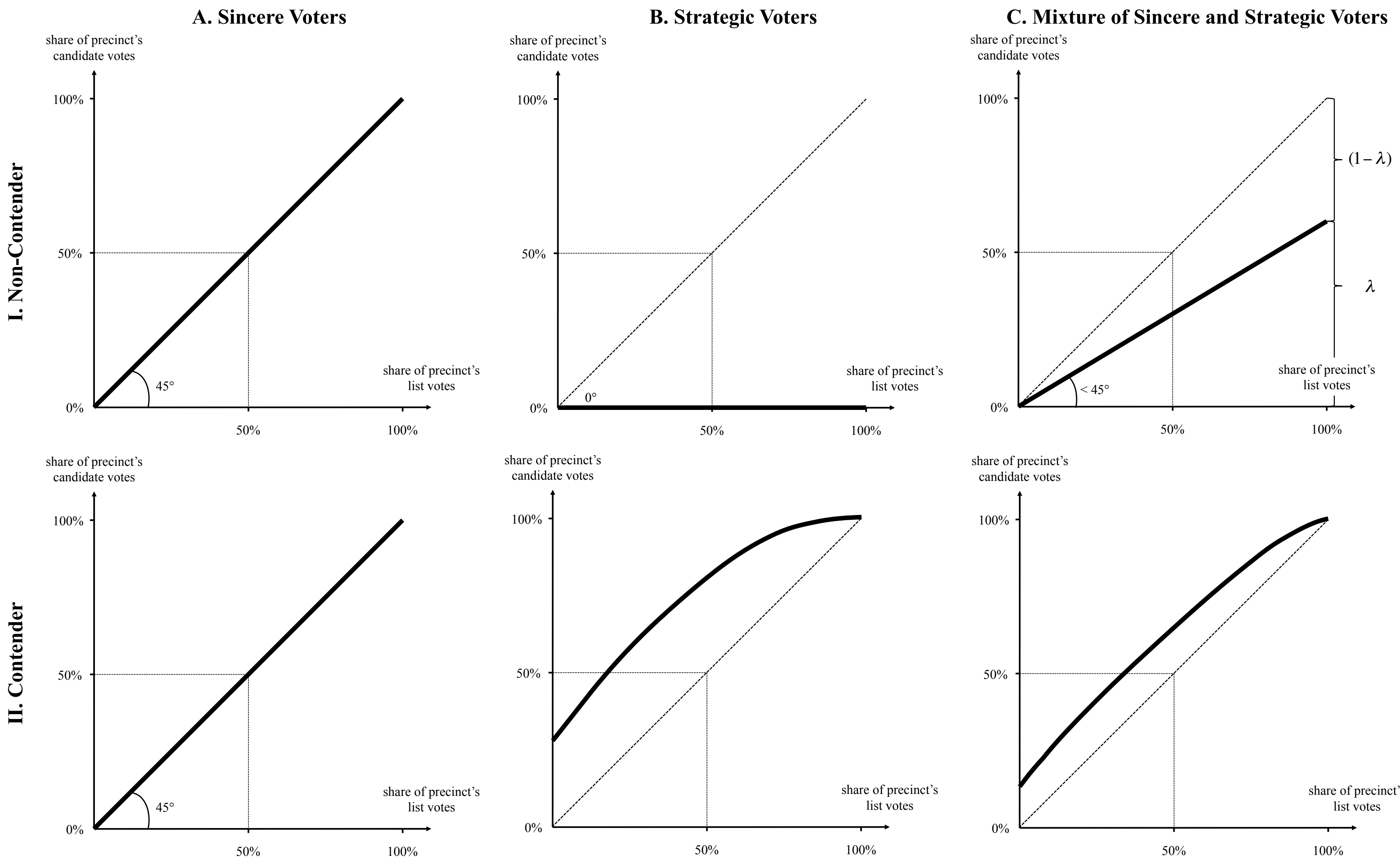
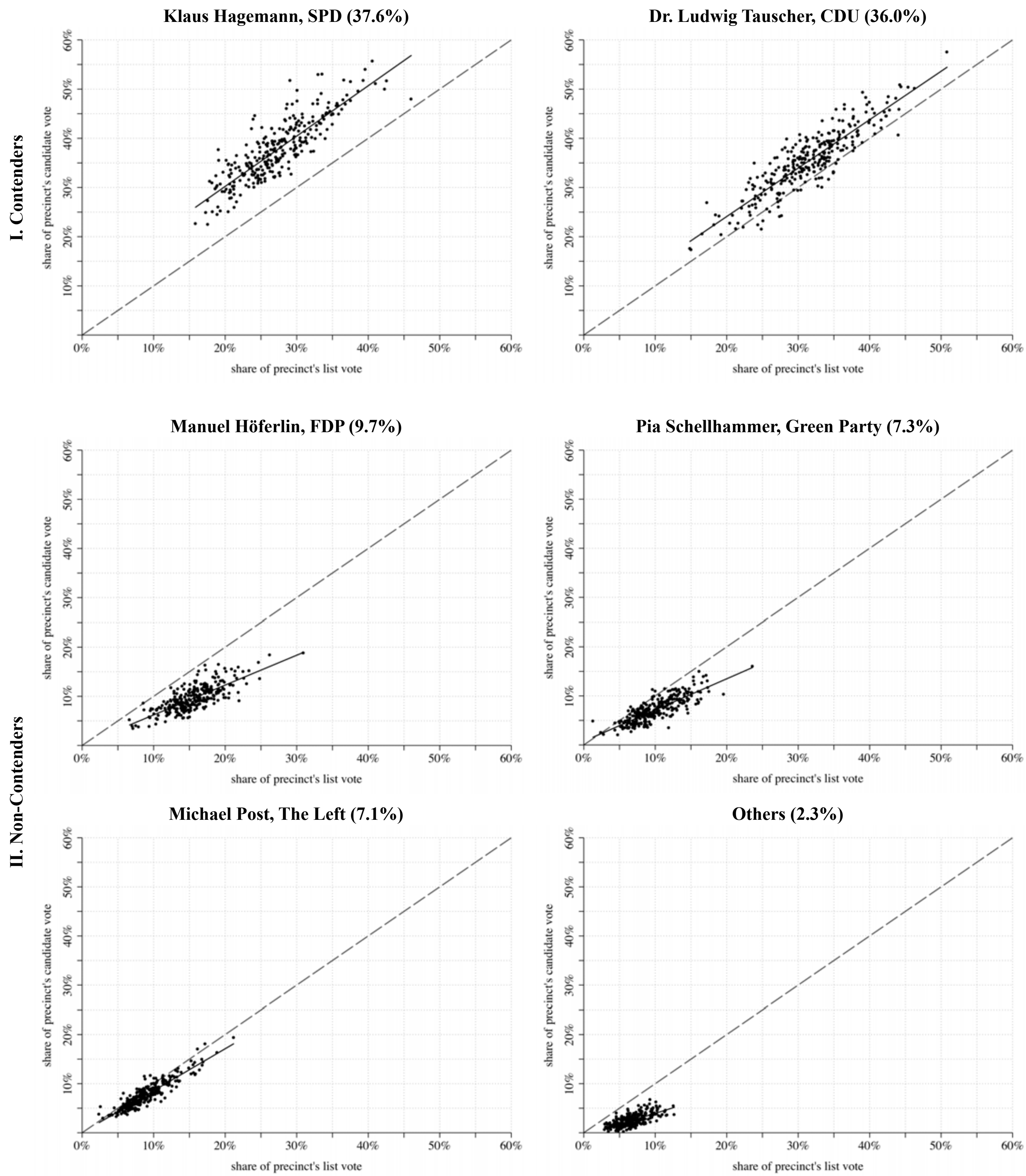
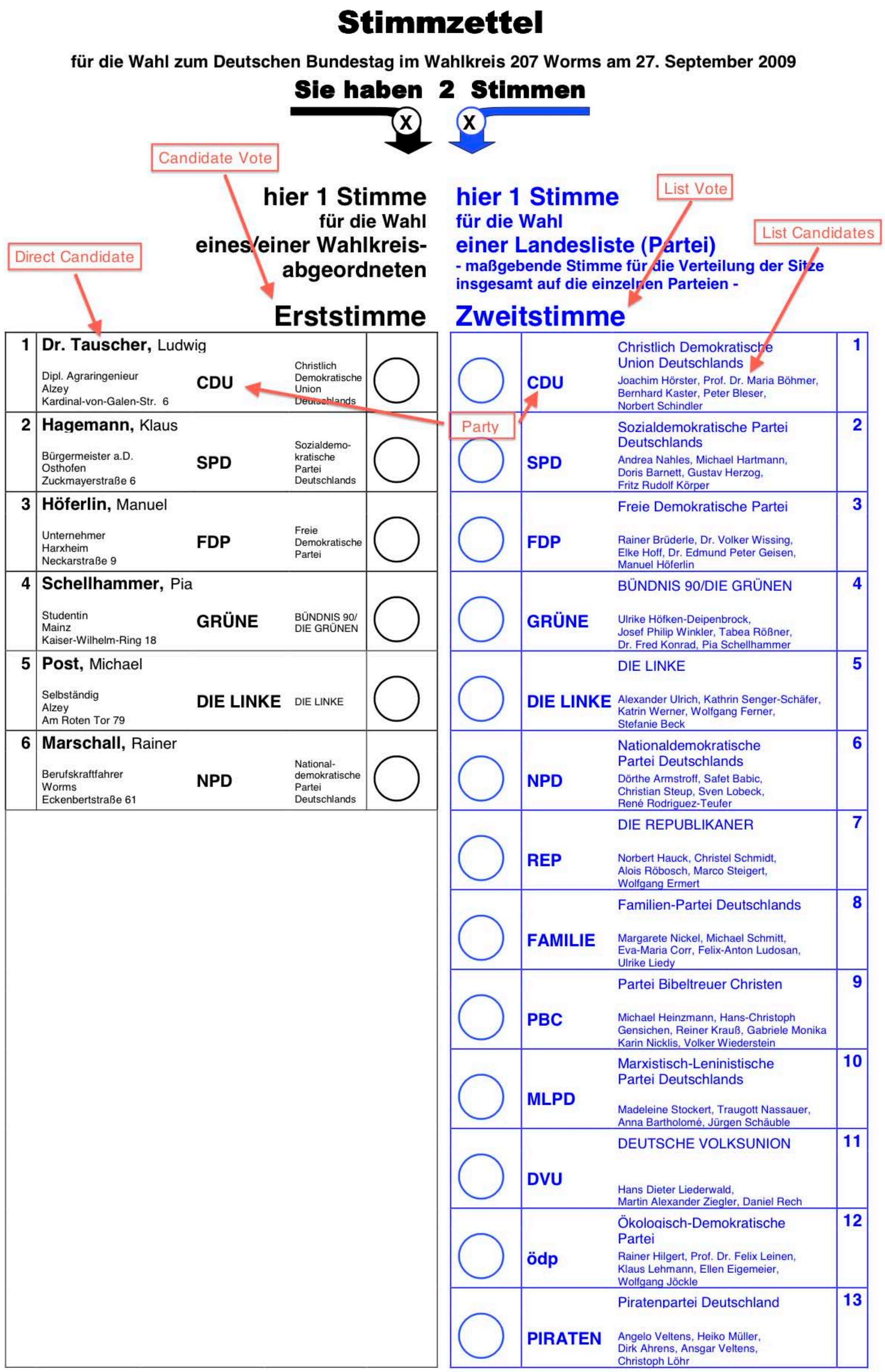


Figure 2: Empirical Example, 2009 Federal Election in Electoral District 207



Notes: Figure depicts the relationship between list votes and candidate votes for contenders (upper two panels) as well as non-contenders (lower four panels) in District 207 during the 2009 federal election. Each dot corresponds to one electoral precinct.

Figure 3: Ballot for the 2009 Federal Election (Electoral District 207)



Stimmzettel

für die Wahl zum Deutschen Bundestag im Wahlkreis 207 Worms am 27. September 2009

Sie haben 2 Stimmen

hier 1 Stimme für die Wahl eines/einer Wahlkreis-abgeordneten	hier 1 Stimme für die Wahl einer Landesliste (Partei) - maßgebende Stimme für die Verteilung der Sitze insgesamt auf die einzelnen Parteien -
Erststimme	**Zweitstimme**

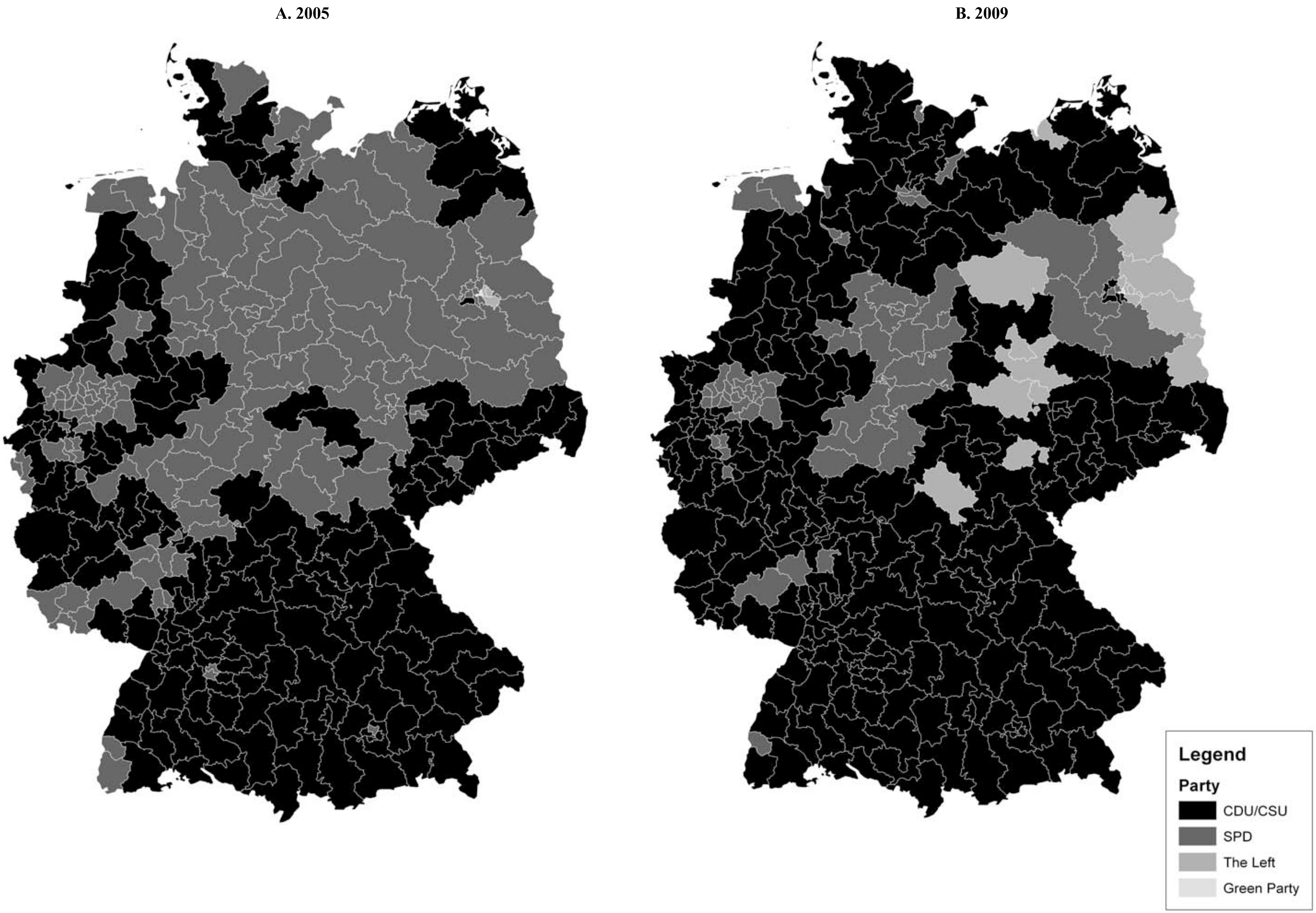
Erststimme

1	**Dr. Tauscher,** Ludwig Dipl. Agraringenieur Alzey Kardinal-von-Galen-Str. 6	**CDU**	Christlich Demokratische Union Deutschlands
2	**Hagemann,** Klaus Bürgermeister a.D. Osthofen Zuckmayerstraße 6	**SPD**	Sozialdemo-kratische Partei Deutschlands
3	**Höferlin,** Manuel Unternehmer Harxheim Neckarstraße 9	**FDP**	Freie Demokratische Partei
4	**Schellhammer,** Pia Studentin Mainz Kaiser-Wilhelm-Ring 18	**GRÜNE**	BÜNDNIS 90/ DIE GRÜNEN
5	**Post,** Michael Selbständig Alzey Am Roten Tor 79	**DIE LINKE**	DIE LINKE
6	**Marschall,** Rainer Berufskraftfahrer Worms Eckenbertstraße 61	**NPD**	National-demokratische Partei Deutschlands

Zweitstimme

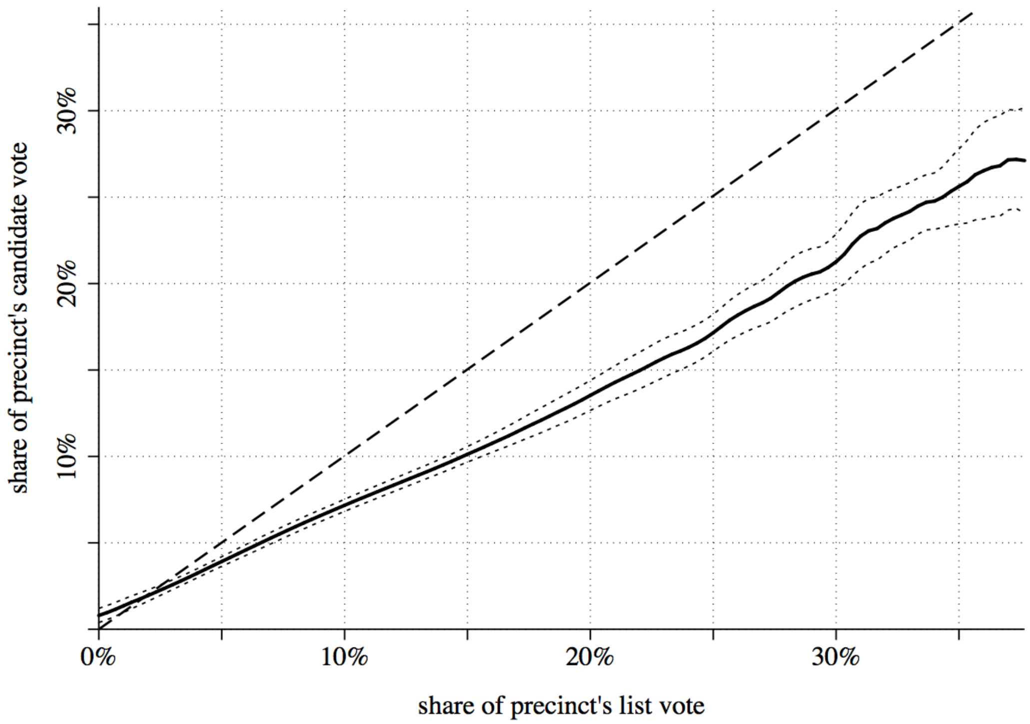
CDU	Christlich Demokratische Union Deutschlands Joachim Hörster, Prof. Dr. Maria Böhmer, Bernhard Kaster, Peter Bleser, Norbert Schindler	1
SPD	Sozialdemokratische Partei Deutschlands Andrea Nahles, Michael Hartmann, Doris Barnett, Gustav Herzog, Fritz Rudolf Körper	2
FDP	Freie Demokratische Partei Rainer Brüderle, Dr. Volker Wissing, Elke Hoff, Dr. Edmund Peter Geisen, Manuel Höferlin	3
GRÜNE	BÜNDNIS 90/DIE GRÜNEN Ulrike Höfken-Deipenbrock, Josef Philip Winkler, Tabea Rößner, Dr. Fred Konrad, Pia Schellhammer	4
DIE LINKE	DIE LINKE Alexander Ulrich, Kathrin Senger-Schäfer, Katrin Werner, Wolfgang Ferner, Stefanie Beck	5
NPD	Nationaldemokratische Partei Deutschlands Dörthe Armstroff, Safet Babic, Christian Steup, Sven Lobeck, René Rodriguez-Teufer	6
REP	DIE REPUBLIKANER Norbert Hauck, Christel Schmidt, Alois Röbosch, Marco Steigert, Wolfgang Ermert	7
FAMILIE	Familien-Partei Deutschlands Margarete Nickel, Michael Schmitt, Eva-Maria Corr, Felix-Anton Ludosan, Ulrike Liedy	8
PBC	Partei Bibeltreuer Christen Michael Heinzmann, Hans-Christoph Gensichen, Reiner Krauß, Gabriele Monika Karin Nicklis, Volker Wiederstein	9
MLPD	Marxistisch-Leninistische Partei Deutschlands Madeleine Stockert, Traugott Nassauer, Anna Bartholomé, Jürgen Schäuble	10
DVU	DEUTSCHE VOLKSUNION Hans Dieter Liederwald, Martin Alexander Ziegler, Daniel Rech	11
ödp	Ökologisch-Demokratische Partei Rainer Hilgert, Prof. Dr. Felix Leinen, Klaus Lehmann, Ellen Eigemeier, Wolfgang Jöckle	12
PIRATEN	Piratenpartei Deutschland Angelo Veltens, Heiko Müller, Dirk Ahrens, Ansgar Veltens, Christoph Löhr	13

Figure 4: Distribution of Direct Mandates in the 2005 and 2009 Federal Elections



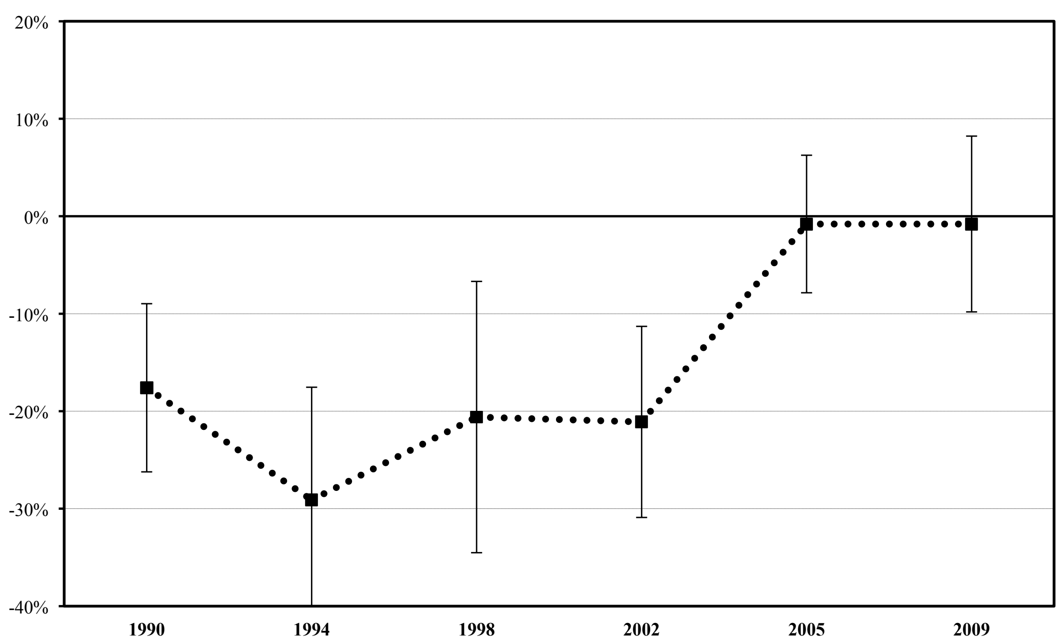
Notes: Figure depicts majority of the candidate vote by electoral district and candidates' party affiliation in the 2005 (left) and 2009 (right) federal elections. In the 2005 (2009) election, candidates running for the CDU/CSU won the plurality of votes in 150 (218) out of 299 electoral districts. SPD candidates gained 145 (64) direct mandates. Candidates of the The Left won 3 (16) districts, and the Green Party achieved 1 (1) direct mandate. No FDP contestant won a district race.
Sources: Based on Bundeswahlleiter (2005a, 2005b, 2008, 2009a).

Figure 5: Relationship between List and Candidate Votes for Non-Contenders



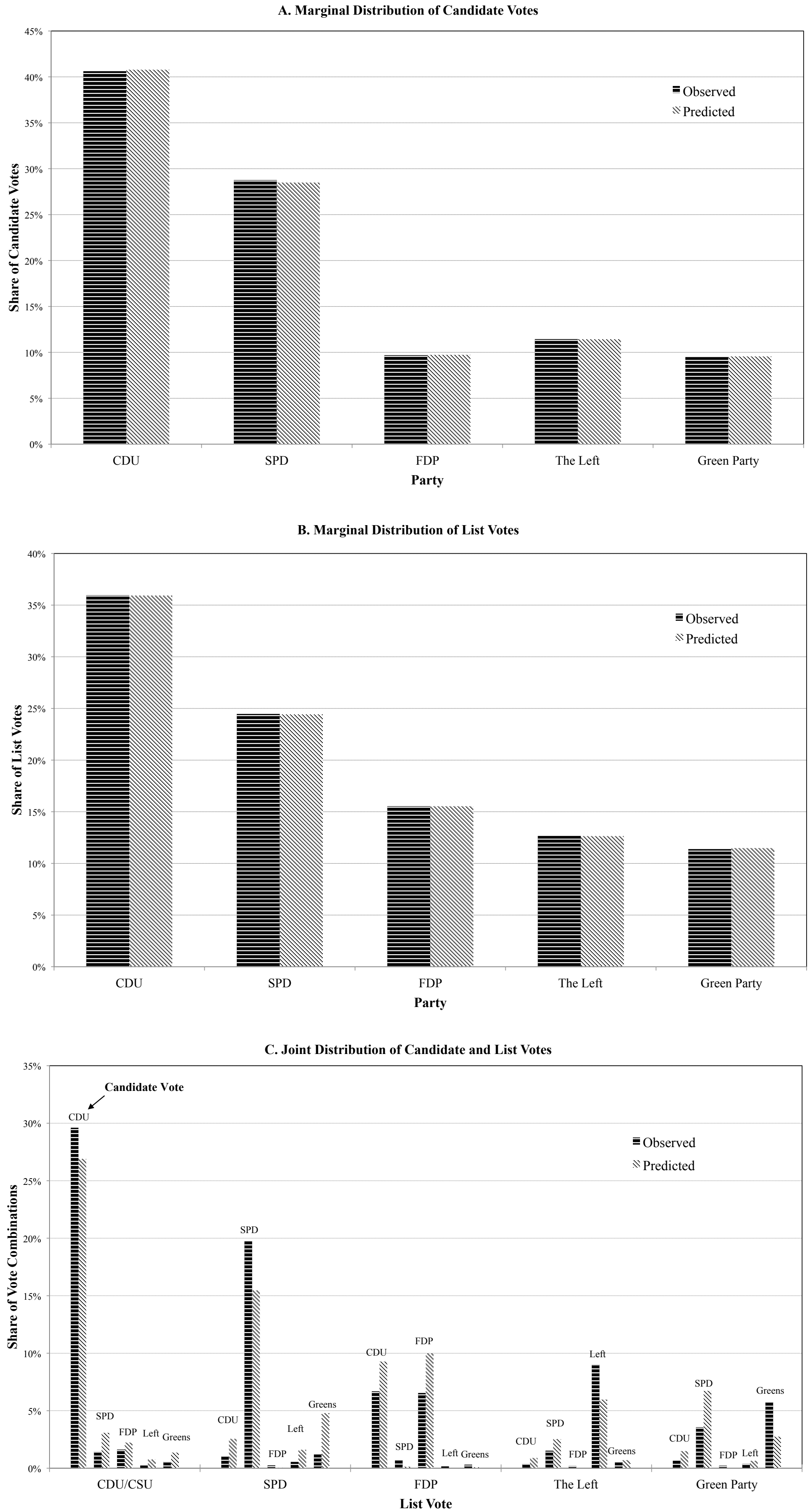
Notes: Figure shows a semi-parametric estimate of the relationship between list and candidate votes for non-contenders and the associated 95%-confidence interval. Estimates are obtained by the differencing procedure in Yatchew (1998) and local-mean smoothing using an Epanechnikov kernel with a bandwidth of .55%.

Figure 6: Difference in the Share of Strategic Voters between East and West Germany, 1990–2009



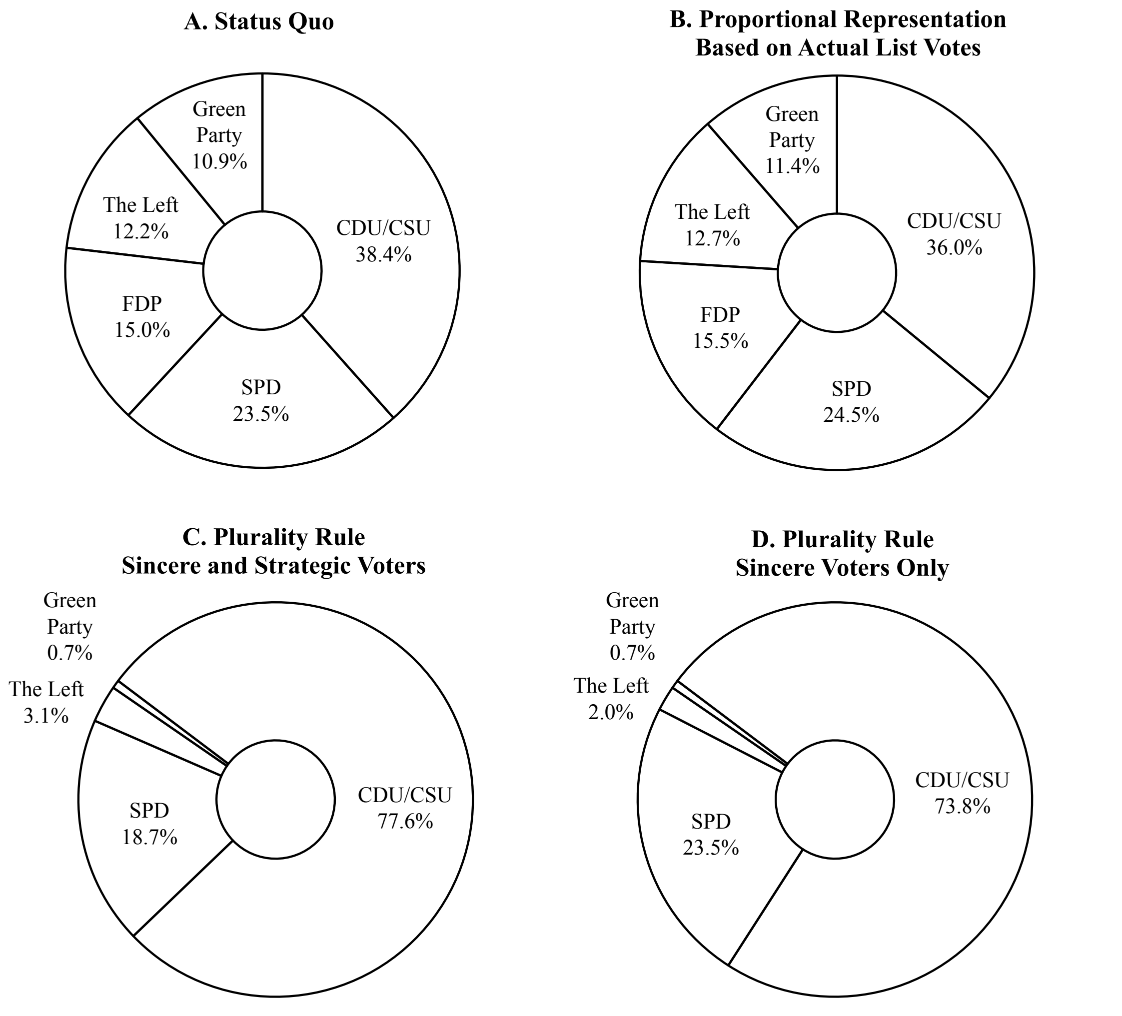
Notes: Figure shows the percentage point difference in the share of strategic voters between East and West Germany for each federal election from 1990 to 2009 as well as the associated 95%-confidence intervals. Negative values indicate fewer strategic voters among residents of the former GDR. The null hypothesis of a constant difference across all years can be rejected at the 1%-confidence level, and that of an equal difference in 1990 and 2009 is rejected at the 5%-level.

Figure 7: Observed vs Predicted Distribution of Votes, 2009 Federal Elections



Notes: Figure depicts actual and predicted vote shares in the 2009 federal election. Panel A shows the marginal distribution of candidate votes, and panel B that of list votes. Panel C depicts the frequency of valid list and candidate vote combinations, i.e. their joint distribution. Dark columns are based on official statistics by the Federal Returning Officer (Bundeswahlleiter 2009a, 2010). Light columns corresponds to the predictions of the structural model in Section 7.

Figure 8: Counterfactual Seat Distributions in the 17th Bundestag



Notes: Figure depicts counterfactual seat distributions in the Bundestag following the 2009 federal election. See the main text for a description of the assumptions underlying each panel.

Table 1: Distribution of List and Candidate Votes in the 2009 Federal Election

Party	Share of List Vote	Share of Candidate Vote	Number of Direct Mandates	Candidate Vote as Fraction of Party's List Vote						
				CDU/CSU	SPD	FDP	The Left	Green Party	Others	Invalid
CDU/CSU	33.3%	38.7%	218	**.876**	.042	.048	.007	.017	.005	.006
SPD	22.7%	27.5%	64	.045	**.858**	.011	.024	.052	.004	.006
FDP	14.4%	9.3%	0	.458	.048	**.448**	.011	.021	.009	.005
The Left	11.7%	10.9%	16	.031	.128	.017	**.757**	.048	.014	.005
Green Party	10.6%	9.0%	1	.061	.333	.021	.036	**.536**	.008	.004
Others	5.9%	2.9%	0	.133	.130	.090	.114	.125	**.378**	.030
Invalid	1.4%	1.7%	--	.117	.079	.021	.025	.013	.013	**.732**

Notes: Entries denote each party's number of direct mandates, share of the list and candidate vote, as well as the frequency of different list and candidate vote combinations (calculated as fraction of a party's list vote) in the 2009 federal election. Due to rounding entries may not add up to unity.
Source: Author's calculations based on Bundeswahlleiter (2009a, 2010).

Table 2A: Summary Statistics for Electoral Precincts

Variable	Full Sample	West Germany 2005	West Germany 2009	East Germany 2005	East Germany 2009
Number of Eligible Voters	820.7	821.5	834.2	782.9	802.5
	(406.1)	(385.4)	(387.6)	(460.3)	(487.7)
Turnout	.747	.789	.727	.751	.658
	(.087)	(.071)	(.083)	(.069)	(.084)
Share of Candidate Vote (in %):					
CDU/CSU	41.07	44.81	41.94	29.65	32.94
	(13.02)	(13.33)	(11.49)	(9.93)	(10.32)
SPD	32.23	38.73	28.80	31.35	20.03
	(12.61)	(12.66)	(10.71)	(8.28)	(7.54)
FDP	7.04	4.59	9.66	5.16	8.17
	(3.93)	(2.25)	(3.74)	(2.85)	(3.73)
The Left	9.66	3.95	7.35	24.86	28.61
	(9.67)	(3.06)	(4.01)	(7.37)	(8.46)
Green Party	6.87	5.46	9.32	3.76	5.53
	(5.29)	(3.78)	(5.23)	(5.80)	(6.08)
Others	3.08	2.38	2.88	5.22	4.72
	(2.88)	(2.75)	(2.59)	(3.10)	(2.70)
Share of List Vote (in %):					
CDU/CSU	35.47	38.67	35.59	26.21	30.65
	(11.60)	(12.48)	(10.25)	(8.59)	(8.92)
SPD	27.98	34.22	23.62	29.96	17.95
	(10.91)	(10.84)	(8.63)	(7.16)	(5.94)
FDP	12.01	10.10	15.18	8.00	10.57
	(4.83)	(3.56)	(4.61)	(3.25)	(3.76)
The Left	10.43	4.83	8.40	25.05	28.27
	(9.23)	(3.09)	(4.26)	(6.26)	(7.56)
Green Party	8.83	8.38	10.93	4.78	5.90
	(5.38)	(4.87)	(5.25)	(4.13)	(4.93)
Others	5.23	3.71	6.23	6.01	6.66
	(2.85)	(2.03)	(2.76)	(3.03)	(3.03)
Absentee Precinct	.148	.155	.166	.090	.098
	(.355)	(.362)	(.372)	(.286)	(.297)
	177,425	71,614	72,056	17,110	16,645

Notes: Entries are means and standard deviations for all precinct level variables used in the analysis, differentiating between East and West Germany as well as election year. See the Data Appendix for a precise definition of each variable.

Table 2B: Characteristics of Direct Candidates

Variable	Full Sample	Party Affiliation: CDU/CSU	SPD	FDP	The Left	Green Party	Others
Age	47.16	49.14	48.87	44.96	49.29	44.01	46.94
	(11.97)	(9.72)	(9.83)	(11.29)	(10.48)	(10.97)	(14.45)
Female	.226	.191	.353	.169	.259	.344	.139
	(.418)	(.393)	(.478)	(.375)	(.438)	(.475)	(.346)
Doctorate	.109	.204	.134	.161	.090	.105	.041
	(.312)	(.403)	(.341)	(.367)	(.287)	(.306)	(.199)
Currently Member of Parliament	.231	.652	.602	.161	.083	.148	.002
	(.422)	(.477)	(.490)	(.367)	(.277)	(.356)	(.039)
Holds Direct Mandate	.111	.376	.403	.000	.009	.003	.001
	(.315)	(.485)	(.491)	(.000)	(.092)	(.058)	(.028)
Also on Party List	.626	.759	.950	.888	.434	.546	.414
	(.484)	(.428)	(.218)	(.316)	(.496)	(.498)	(.493)
Position on Party List | Also List Candidate	12.89	13.26	17.36	17.47	9.40	8.89	7.32
	(12.86)	(12.79)	(15.46)	(15.32)	(8.19)	(6.94)	(7.02)
	4,257	598	598	598	587	593	1,283

Notes: Entries are means and standard deviations of characteristics of direct candidates running in the 2005 or 2009 federal elections, by party affiliation. See the Data Appendix for the precise definition and source of each variable.

Table 3: Ranking of Candidates in the 2005 & 2009 Federal Elections

Rank Based on Candidate Vote	Rank Based on List Vote 1	2	3	4	5	6
1	**557**	38	3	0	0	0
2	39	**502**	54	3	0	0
3	2	44	**369**	139	39	5
4	0	14	131	**306**	138	9
5	0	0	39	139	**332**	87
6	0	0	2	11	88	**474**
Total	598	598	598	598	597	575
First or Runner-Up	99.7%	90.3%	9.5%	0.5%	0.0%	0.0%

Notes: Entries denote the number of candidates for each combination of own rank based on received candidate votes (left column) and the within district ranking of the associated party based on the list vote in the same year (top row). For instance, out of the 598 candidates whose party received the most list votes within a particular district, 557 won the direct mandate for that district, 39 candidates finished in second place, and 2 ended up third. The rank order correlation within districts is .93.

Table 4: Preference Based Rank and the Candidate-List-Vote Gradient

	Share of Candidate Vote				
Independent Variable	(1)	(2)	(3)	(4)	(5)
Share of List Vote		1.023	1.004	.983	.998
× Ranked First		(.024)	(.019)	(.008)	(.011)
Share of List Vote		1.141	1.062	1.011	1.037
× Ranked Second		(.019)	(.012)	(.017)	(.015)
Share of List Vote		1.066	.852	.780	.730
× Ranked Third		(.075)	(.053)	(.045)	(.052)
Share of List Vote		.707	.711	.691	.653
× Ranked Fourth		(.025)	(.026)	(.030)	(.031)
Share of List Vote		.809	.846	.795	.782
× Ranked Fifth		(.018)	(.014)	(.014)	(.014)
Share of List Vote		.817	.823	.787	.765
× Ranked Sixth		(.044)	(.042)	(.045)	(.047)
Preference Based Rank:					
First	44.518	4.999	6.174	1.197	
	(1.949)	(1.238)	(.944)	(.304)	
Second	29.490	-.494	3.093	.926	
	(1.840)	(.857)	(.482)	(.286)	
Third	11.205	-4.331	5.053	5.107	
	(1.385)	(.718)	(1.054)	(.786)	
Fourth	7.002	.259	6.944	6.147	
	(.423)	(.241)	(.676)	(.815)	
Fifth	5.502	.078	5.507	5.123	
	(.117)	(.093)	(.574)	(.536)	
Sixth	3.885	.200	5.900	5.400	
	(.239)	(.078)	(.686)	(.695)	
Party Fixed Effects	No	No	Yes	No	No
Candidate Fixed Effects	No	No	No	Yes	No
Candidate × Year Fixed Effects	No	No	No	No	Yes
R-Squared	.777	.953	.965	.980	.982
Number of Observations	882,061	882,061	882,061	882,061	882,061

Notes: Entries are coefficients and standard errors from regressing the variables shown on the left on a candidate's vote share. Heteroskedasticity robust standard errors are clustered by state and reported in parentheses. See the Data Appendix for the precise definition and source of each variable.

Table 5: Estimating the Extent of Strategic Voting

	Share of Candidate Vote				
Independent Variable	(1)	(2)	(3)	(4)	(5)
Share of List Vote		.688	.724	.698	.657
× Non-Contender		(.050)	(.016)	(.020)	(.019)
Share of List Vote		1.165	1.063	.982	1.004
× Contender		(.018)	(.014)	(.010)	(.012)
Non-Contender	6.366	.303	6.673	.698	
	(.221)	(.328)	(.731)	(.020)	
Contender	33.678	-1.353	3.402	.982	
	(1.275)	(.553)	(.595)	(.010)	
Party Fixed Effects	No	No	Yes	No	No
Candidate Fixed Effects	No	No	No	Yes	No
Candidate × Year Fixed Effects	No	No	No	No	Yes
R-Squared	.620	.950	.965	.980	.982
Number of Observations	882,061	882,061	882,061	882,061	882,061

Notes: Entries are coefficients and standard errors from estimating equation (4) by ordinary least squares. The structural break is estimated to occur at κ=.064. Heteroskedasticity robust standard errors are clustered by state and reported in parentheses. See the Data Appendix for the precise definition and source of each variable.

Table 6: Comparative Statics

Restriction	Share of Strategic Voters: Candidate Fixed Effects	Share of Strategic Voters: Candidate × Year Fixed Effects
Baseline	.302	.343
	(.020)	(.019)
By Availability of Close Substitute:		
Allied Party's Candidate in the Race	.328	.362
	(.015)	(.015)
Only Rival Parties' Candidates in the Race	.093	.093
	(.019)	(.019)
By Closeness of Race Based on Preferences:		
less than 5% between First and Second	.351	.350
	(.027)	(.026)
more than 5% between First and Second	.317	.339
	(.021)	(.022)
By Ex Post Closeness of Race:		
less than 5% between First and Second	.350	.360
	(.026)	(.026)
more than 5% between First and Second	.307	.336
	(.021)	(.022)
By Year:		
2005	.452	--
	(.027)	
2009	.251	--
	(.020)	

Notes: Entries are coefficients and standard errors on 1-λ using different subsamples of the data. The respective restriction is indicated on the left of each row. See the Data Appendix for the precise definition and source of each variable.

Table 7: Candidate-List-Vote Gradient among Second Ranked Candidates

Distance to First Ranked Candidate	Slope: Candidate Fixed Effects	Slope: Candidate × Year Fixed Effects
Based on Preferences:		
< 2% behind First Ranked Candidate	1.026	1.025
	(.013)	(.013)
2% to 5% behind First Ranked Candidate	1.019	1.019
	(.012)	(.012)
5% to 10% behind First Ranked Candidate	1.057	1.058
	(.015)	(.015)
10% to 15% behind First Ranked Candidate	1.041	1.044
	(.021)	(.021)
> 15% behind First Ranked Candidate	1.018	1.033
	(.025)	(.023)
Based on Ex Post Outcome of Races:		
< 2% behind First Ranked Candidate	1.022	1.025
	(.020)	(.020)
2% to 5% behind First Ranked Candidate	1.043	1.042
	(.014)	(.014)
5% to 10% behind First Ranked Candidate	1.047	1.049
	(.014)	(.014)
10% to 15% behind First Ranked Candidate	1.058	1.061
	(.022)	(.024)
> 15% behind First Ranked Candidate	1.007	1.022
	(.018)	(.016)

Notes: Entries denote the candidate-list-vote gradient for second ranked candidates, by distance to the first ranked one. The respective cutoffs are shown in the column on the left. Heteroskedasticity robust standard errors are clustered by state and reported in parentheses.

Table 8: Sensitivity Analysis Using Alternative Classifications of Contenders

	Share of Strategic Voters	
Classification of Contenders	Candidate Fixed Effects	Candidate × Year Fixed Effects
Baseline (Preference Based, Original Cutoff)	.302	.343
	(.020)	(.019)
Ex Post Outcome of Races (Original Cutoff)	.299	.345
	(.016)	(.018)
Preference Based Using Different Cutoffs:		
< 1% behind Second Ranked Candidate	.255	.295
	(.026)	(.029)
< 2% behind Second Ranked Candidate	.268	.307
	(.024)	(.027)
< 5% behind Second Ranked Candidate	.293	.331
	(.021)	(.021)
< 8% behind Second Ranked Candidate	.306	.352
	(.019)	(.017)
Electoral District × Year Specific Cutoffs	.238	.282
	(.026)	(.030)
Ex Post Outcome of Races Using Different Cutoffs:		
< 1% behind Second Ranked Candidate	.269	.310
	(.023)	(.027)
< 2% behind Second Ranked Candidate	.274	.317
	(.022)	(.026)
< 5% behind Second Ranked Candidate	.290	.337
	(.018)	(.020)
< 8% behind Second Ranked Candidate	.308	.350
	(.015)	(.018)
Electoral District × Year Specific Cutoffs	.277	.350
	(.021)	(.018)
Ranked First or Second Based on Preferences	.243	.284
	(.027)	(.030)
Ranked First, Second, or Third Based on Preferences	.269	.299
	(.023)	(.023)
Ranked First or Second Based on Ex Post Outcome	.264	.305
	(.024)	(.029)
Ranked First, Second, or Third Based on Ex Post Outcome	.334	.371
	(.013)	(.013)
Finished First or Second in Last Federal Election	.250	.287
	(.027)	(.032)
Finished First, Second, or Third in Last Federal Election	.289	.319
	(.014)	(.015)
Finish in Last Federal Election (Original Cutoff)	.274	.316
	(.022)	(.026)
Finish in Last Federal Election Using Different Cutoffs:		
< 1% behind Second Ranked Candidate	.252	.291
	(.026)	(.031)
< 2% behind Second Ranked Candidate	.256	.296
	(.026)	(.030)
< 5% behind Second Ranked Candidate	.271	.313
	(.023)	(.027)
< 8% behind Second Ranked Candidate	.279	.322
	(.021)	(.024)
Estimated Electoral District × Year Specific Cutoffs	.252	.293
	(.023)	(.026)

Notes: Entries are coefficients and standard errors on 1-λ using alternative classifications of "contender." The respective definition is shown in the column on the left. Heteroskedasticity robust standard errors are clustered by state and reported in parentheses.

Table 9: Additional Sensitivity and Robustness Checks

Restriction	Share of Strategic Voters: Candidate Fixed Effects	Share of Strategic Voters: Candidate × Year Fixed Effects
Baseline	.302	.343
	(.020)	(.019)
Weighted by Number of Party Supporters	.295	.334
	(.032)	(.028)
Including "Other" Party Candidates	.319	.351
	(.021)	(.019)
In States without Overhang Mandates	.346	.378
	(.014)	(.013)
Difference Estimator	.285	.347
	(.025)	(.027)

Notes: Entries are coefficients and standard errors on the share of strategic voters, i.e. 1-λ, using different subsamples of the data and weighting schemes. The respective restriction is indicated on the left of each row. See the Data Appendix for the precise definition and source of each variable.

Table 10: Correlates of Strategic Voting

	Share of Candidate Vote				
Independent Variable	(1)	(2)	(3)	(4)	(5)
Share of List Vote	1.008	1.008	1.008	1.008	1.008
× Contender	(.009)	(.009)	(.009)	(.009)	(.009)
Share of List Vote	.666	.644	.664	.605	.679
× Non-Contender	(.031)	(.029)	(.030)	(.033)	(.022)
Share of List Vote		-.005			
× Non-Contender × Population Density		(.007)			
Share of List Vote			.019		
× Non-Contender × Fraction of Voters Female			(.023)		
Share of List Vote				-.271	
× Non-Contender × Income Tax Revenue per Capita				(.078)	
Share of List Vote					.244
× Non-Contender × Fraction of Voters under Age 30					(.074)
Candidate × Year Fixed Effects	Yes	Yes	Yes	Yes	Yes
R-Squared	.984	.984	.984	.984	.984
Number of Observations	120,700	120,700	120,700	120,700	120,700

Notes: Entries are coefficients and standard errors from estimating specifications analogous to equation (4) by ordinary least squares, with a structural break at κ=.064. Heteroskedasticity robust standard errors are clustered by state and reported in parentheses. For ease of interpretation, covariates interacted with Share of List Vote × Non-Contender have been demeaned. In addition to the variables shown in the table, indicator variables for missing values on each covariate are also included in the regressions. See the Data Appendix for the precise definition and source of each variable.

Table 11: Actual vs Predicted Ranking of Candidates, 2009 Federal Elections

Actual Rank	Predicted Rank (as Fraction of Actual Rank)				
	1	2	3	4	5
1	**94.6%**	5.1%	0.3%	0.0%	0.0%
2	5.1%	**91.8%**	3.1%	0.0%	0.0%
3	0.3%	3.1%	**88.1%**	7.5%	1.0%
4	0.0%	0.0%	8.2%	**86.1%**	5.8%
5	0.0%	0.0%	0.3%	6.5%	**93.2%**

Notes: Entries denote the frequency with which the predictions of the structural model in Section 7 coincide with observed outcomes, considering only candidates of the 5 major parties.

Table 12: Voters' Partial Preference Orderings

First Choice Candidate	Second Choice Candidate (as Fraction of First Choice)				
	CDU/CSU	SPD	FDP	The Left	Green Party
CDU/CSU	--	26.8%	50.0%	8.1%	15.1%
SPD	16.1%	--	0.3%	15.5%	68.0%
FDP	98.9%	0.6%	--	0.1%	0.4%
The Left	18.0%	60.4%	0.1%	--	21.5%
Green Party	14.5%	77.1%	0.4%	8.1%	--

Notes: Entries denote the simulated relative frequency of voters' second choice candidate, conditional on their first choice. See Appendix D for details on the simulation.

Table 13: Joint Distribution of District Winners under Sincere and Strategic Voting

District Winner with Sincere Voters	District Winner with Sincere and Strategic Voters				
	CDU/CSU	SPD	FDP	The Left	Green Party
CDU/CSU	**70.8%**	1.7%	0.0%	1.0%	0.3%
SPD	6.5%	**16.7%**	0.0%	0.3%	0.0%
FDP	0.0%	0.0%	**0.0%**	0.0%	0.0%
The Left	0.0%	0.3%	0.0%	**1.7%**	0.0%
Green Party	0.3%	0.0%	0.0%	0.0%	**0.3%**

Notes: Entries compare the simulated distribution of district winners in a first-past-the-post system with only sincere voters (left column) to the one that would obtain with a mixture of types (top row). Summing across columns gives the percentage of districts that would acrue to a particular party if all voters behaved sincerely, whereas summing across rows gives a party's share of districts if 26.3% of voters behaved strategically. Consequently, adding the entries on the diagonal shows that about 90% of districts would accrue to the same party. See Appendix D for details on the simulation.