Understanding People's Choice When They Have Two Votes

Ludovic Rheault, University of Toronto
André Blais, University of Montreal
John Aldrich, Duke University
Thomas Gschwend, University of Mannheim

Abstract

This paper introduces a model of vote choice in mixed electoral systems where electors cast two votes. The model is tested using survey data collected at the time of the 2013 federal election in Germany. We start by studying a general problem of vote choice where parties are running in multiple elections concurrently. We propose a new approach to account for contamination, a phenomenon that we define as a causal influence across the vote decisions. Since causality entails a time-ordering, we argue that contamination arises only when voters choose sequentially. By making use of information on the timing of vote decisions, we are then able to estimate the magnitude of those contamination effects in the German case. We show that such effects are present only among voters with lower levels of education. Moreover, the magnitude of the contamination effect from the list vote toward the candidate vote appears greater than the other way around. We also test a number of predictions about the determinants of the two vote choices in mixed systems. Our findings suggest that local chances of winning and local candidate ratings affect mostly the candidate vote, whereas party, leader and coalition ratings influence more strongly the list vote.

Understanding People's Choice When They Have Two Votes

Understanding why people vote the way they do is the central question in the field of electoral studies. There is a vast literature on the factors that lead citizens to vote for a particular party or candidate in a given election. The question that we address here is how to make sense of the choices that people make when they have two votes. There is surprisingly little research on this question in spite of the fact that the number of elections where people make two choices is growing. According to Carter and Farrell, 35 democracies, about one out of four, use a mixed system for their legislative elections, and the great majority of them have two votes, almost always a party list vote and a vote for a candidate in a local constituency (Carter and Farrell, 2010, 27; Massicotte and Blais, 1999). Instances of parties competing for multiple offices concurrently also extend beyond mixed electoral systems. Many countries feature elections in which parties are running at the same time for senate and lower house seats, or federal and subnational parliaments. Although the model that we develop in this study is applied to mixed electoral systems, its main features are exportable to a wider range of elections with multiple votes.

The presence of two votes in mixed systems has spurred interest about the magnitude and determinants of ticket-splitting, that is, people voting for a candidate associated with a party different from the one that they support on the list vote (Burden, 2009; Gschwend, 2007; Gschwend, Johnston, and Pattie, 2003; Helmke, 2009; Karp et al., 2002; Moser and Scheiner, 2005, 2009; Pappi and Thurner, 2002). A number of studies have also investigated the presence or absence of contamination effects (or spillover effects), that is, whether vote choice in one election affects the choice in the other election (Cox and Schoppa, 2002; Ferrara, Herron, and Nishikawa, 2005; Hainmueller and Kern, 2008; Herron and Nishikawa, 2001; Karp, 2009). Yet, as far as we can tell, there has been no systematic, comparative investigation of the factors that affect each vote when people have two decisions to make, one with respect to the choice of a candidate in the local constituency and one with respect to the choice of a party list. We intend to fill that gap here by formulating a number of predictions and testing them using data collected during the 2013 German federal election. Germany uses a mixture of proportional representation (PR) and first-past-the-post (FPTP) rules since 1953. This makes the country a natural choice to test our model, since we can reasonably expect the German public to have well established strategies for casting the two votes, unlike voters in newly adopted mixed systems.

A key contribution of this study is to introduce a new framework to estimate contamination effects at the micro-level. So far, most of the debate surrounding the existence of contamination effects between electoral systems has relied upon macro evidence (see e.g. Cox and Schoppa, 2002; Herron and Nishikawa, 2001; Ferrara, 2004; Moser and Scheiner, 2004; Maeda, 2008; Hainmueller and Kern, 2008; Crisp, Potter, and Lee, 2012). Building upon Duvergerian principles, many of those studies are looking to assess contamination by testing whether mixed electoral systems constrain the number of parties the same way plurality systems do, as opposed to pure PR systems. However, since mechanisms leading to contamination ultimately depend on the decisions of voters, the macroscopic approach leaves a large part of the phenomenon unobserved. Existing attempts to account for contamination at the micro level are reviewed in the next section, but overall, most of them use indirect estimates. Spillovers across vote decisions have also been studied in contexts where presidential elections are held concurrently with other national or subnational elections (Ferejohn and Calvert, 1984; Campbell and Sumners, 1990; Ames, 1994; Samuels, 2000), then again, most of the time using aggregate data or indirect measures of coattail effects. Our approach provides a direct measure of causal effects across vote decisions, making it possible to overcome identification issues while remaining general enough to be implemented in most situations involving simultaneous elections.

A Model of Multiple Vote Choice

We begin by introducing a general model in which voters must choose multiple times between a set of parties competing in concurrent elections. Our model accounts for the possibility of "contamination" between the choices of individual voters. We define contamination as a causal effect between the vote decisions, which requires establishing a time-ordering between the choices made by voters. We justify this modeling choice and define the relevant concepts more formally in this section, before making predictions regarding the determinants of voting in mixed systems more specifically.

In general terms, we are considering situations in which a voter forms preferences over J parties indexed by $j = \{1, ..., J\}$ running in T concurrent elections, indexed as $t = \{1, ..., T\}$. For simplicity of notation, we do not use an index to designate voters in what follows. We assume that preferences are complete and transitive and we represent them using latent utilities denoted u_{tj}. The decision process across the T elections can be expressed as a discrete optimization problem of the form:

$$\underset{y_{tj}\ \forall\ t,j}{\operatorname{argmax}} \sum_t \sum_j u_{tj} y_{tj}, \tag{1}$$

subject to

$$y_{tj} \in \{0,1\} \qquad \text{for } t = \{1,...,T\} \text{ and } j = \{1,...,J\}, \tag{2}$$

$$\sum_j y_{tj} = 1 \qquad \text{for } t = \{1,...,T\}, \tag{3}$$

where y_{jt} equals one if the voter chooses party j in context t, and zero otherwise. This problem means that a voter chooses her preferred alternative in each election t, and the constraints impose that she cast exactly one vote in each election.[1] Notice the particularity of the constraints $\sum_j y_{tj} = 1$ for all t. Unlike the classical consumer problem with multiple goods in economics, where consumers are subjected to a budget constraint across goods, there is no constraint *across* elections in the multiple election problem. Put another way, voters cannot substitute a vote in election s for a vote in election t the way a consumer could substitute goods to choose an optimal bundle. This particular institutional context justifies our reliance on separate and additive utility functions for each election.

On the other hand, we want to account for the possibility that a decision in the election s has an influence on the decision in election t. We define a *contamination effect* as the causal effect of a voter's decision to support party j in election s on the preference of that voter for the same party j in election t (where $t \neq s$). We adopt a strict interpretation of causality that entails the existence of a time ordering between vote decisions: a condition for the existence of a contamination effect from election s to t is that the voter's decision in election s has been made before the decision in election t. Consequently, our definition rules out the possibility of simultaneous contamination for individual voters. This avoids postulating the existence of simultaneous causation, a problematic concept at the philosophical level (Granger, 1969; Mellor, 2002, Ch. 17), while being consistent with the nature of the vote problem laid out in (1)–(3). Our model involves discrete choices that are made only once in each election, which precludes the possibility of a feedback from t to s once the decision in election s has been made. For each pair $\{s, t\}$, either a voter makes one decision before the other, or else both decisions are made at the same time, in which case we rule out the possibility of contamination. As we illustrate later on in this study, it is possible to establish the sequence of the vote decisions for individual voters by collecting information about the timing of those choices in surveys.

Let us denote the event of a voter choosing in election s before election t as $s \to t$.

[1] An extension of this problem could account for electoral rules in which multiple votes can be cast within elections, for instance approval voting or preferential voting. We do not consider this possibility here.

Furthermore, we denote the contamination effect of a vote decision y_{sj} on the preference for party j in election t as θ_{st}. This allows us to write the utilities of a voter as:

$$u_{tj} = \boldsymbol{\beta}_{tj}\boldsymbol{x}'_{tj} + \sum_{s \neq t} \theta_{st}\mathbf{1}(s \to t)y_{sj} + \varepsilon_{tj}, \tag{4}$$

where $\boldsymbol{x}_{tj}$ is a vector of covariates (some of which may be voter-specific, some party-specific or election-specific), $\boldsymbol{\beta}_{tj}$ a vector of parameters, $\mathbf{1}(.)$ is the indicator function equaling one if the expression between brackets is true and zero otherwise, and ε_{tj} represents a random disturbance term. For the purpose of this paper, we assume that disturbances are multivariate normally distributed within each election t, with mean vector zero and covariance matrix $\boldsymbol{\Sigma}_t$. Our notation $\theta_{st}\mathbf{1}(s \to t)y_{sj}$ reflects the definition introduced above: a contamination effect θ_{st} is the impact of y_{sj} on u_{tj}, given that the decision y_{sj} has been made *before* the decision in election t.[2] Since the events $(s \to t)$ and $(t \to s)$ cannot be true at the same time, our model is "recursive": the choice y_{sj} cannot depend in turn on u_{tj} if $(s \to t)$ is true. By construction, the terms $\mathbf{1}(s \to t)y_{sj}$ represent the part of the variation in y_{sj} that is strictly exogenous to u_{tj}. This means that the structural form of this model can be estimated with standard statistical methods.

The natural expectation is that the contamination effects θ_{st} are positive. In other words, choosing j in election s increases the likelihood of choosing j in election t. We believe that this captures the intuition behind the contamination argument as commonly invoked in the previously cited literature on electoral systems. If contamination occurs, the outcome of election t becomes more similar to that of election s than it would otherwise have been. Thus, it must be the case that for some voters, the decision to vote for j in context s increases the chance of voting for j in context t.

We note that alternative theoretical predictions are possible. It could be that a decision to vote for j in election s reduces the likelihood of voting for the same party in election t, for instance if we assume that voters cast split-ticket votes to balance out the composition of government (see e.g. Lewis-Beck and Nadeau, 2004). Such a hypothesis would have important implications for the study of electoral systems, since it could lead to an aggregate outcome in election t that becomes more dissimilar to that of election s, compared the outcome of t that would have been observed had that election taken place alone. An additional benefit of our approach is that we can test the relevance of such an alternative expectation by examining the sign of the estimated θ_{st} parameters. In particular, we expect sophisticated voters to split their vote more frequently, which has

[2] Notice that the contamination effect parameters θ_{st} are not specific to party. In multi-party elections, such models will likely be estimated with multinomial models in which party-specific variables have constrained coefficients across the alternatives j. As a result, there would be only one parameter θ_{st} for each pair of elections $\{s, t\}$.

consequences for the sign of those estimates and, as a result, the nature of contamination effects. We account for this possibility in our empirical section.

With regard to mixed electoral systems more specifically, our approach expands on existing attempts to measure contamination effects, which have been mostly indirect so far. For instance, Ferrara, Herron, and Nishikawa examined whether local candidate ratings have an independent impact on the list vote choice in the 1999 New Zealand elections, controlling for a host of other factors (Ferrara, Herron, and Nishikawa, 2005, Ch. 5). They found that they do and inferred the existence of a contamination effect of the candidate vote on the list vote (Ferrara, Herron, and Nishikawa, 2005, 73, Table 5.1). This effect is estimated to be rather weak. For his part, Karp (2009) used data on the 2002 New Zealand election and determined whether the presence of an incumbent increases the propensity to cast a list vote for the incumbent's party in the case of the two main parties (Labour and National) and whether the presence of a local candidate enhances the likelihood of supporting the party's candidate in the PR vote. The author reported a small incumbency effect for Labour vote but none for National and no candidate presence effect for the minor parties. His conclusion is that "the overall impact of candidate effects appears to be quite small" (Karp, 2009, 49).

As far as we can tell, no previous empirical study has attempted to measure contamination effects *from* the list vote *toward* the candidate vote per se. To the extent that the list vote is more important than the candidate vote—since it is the former that determines how many seats a party obtains—voters should focus their attention on party lists first and foremost. We thus expect a stronger contamination from the list vote toward the candidate vote than the other way around. We also expect contamination effects to occur mostly among the less sophisticated voters. Sophisticated voters should not only be more likely to cast split-ticket votes, a pattern that would offset contamination effects, they should also rely more easily on distinct evaluative criteria for each of the two votes.

Determinants of Vote Choice in Mixed Electoral Systems

In addition to measuring contamination effects, we make predictions about the factors affecting vote choice in a mixed electoral system with a list vote and a local candidate vote, such as the German system. This problem is a special case of the general model introduced in equations (1) to (4) above with $T = 2$. We rewrite the latent utilities of a voter as:

$$u_{Lj} = \boldsymbol{\alpha}_j \boldsymbol{x}_j' + \theta_{CL} \mathbf{1}(C \to L) y_{Cj} + \varepsilon_{Lj}, \tag{5}$$

$$u_{Cj} = \boldsymbol{\beta}_j \boldsymbol{x}_j' + \theta_{LC} \mathbf{1}(L \to C) y_{Lj} + \varepsilon_{Cj}, \tag{6}$$

where the notation is similar as before except that we use the subscript L to indicate the *List* vote and the subscript C to denote the *Candidate* vote. We denote the relevant vectors of parameters $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ to emphasize the distinction between the two elections, the former referring to the list vote and the latter to the candidate vote.

Our model comprises five basic proximate determinants of vote choice: party ratings, leader ratings, local candidate ratings, local chances of winning, and coalition ratings. Put simply, we propose that a voter is more likely to vote for a party when she thinks of herself as close to a party, when she likes the party, its leader and its local candidate, when she believes that the party's candidate has some chance of winning locally, and when she likes the coalition that the party is associated with.

We understand that the vote decision is affected by other considerations, perhaps the most obvious being issue positions and ideology. But we assume that these are more distant factors whose effect is basically indirect. We assume that voters' ideology and attitudes affect how much they like or dislike the various parties and leaders, and that these likes and dislikes in turn determine the final vote choice. This research strategy was followed by Page and Jones (1979) and Rahn et al. (1990), among others. We focus here on the more proximate factors.

The question that we address is whether these factors have a similar influence on the two votes. Put another way, we want to know how $\boldsymbol{\alpha}$ differs from $\boldsymbol{\beta}$. The first prediction concerns the party's perceived chances in the local constituency and local candidate ratings. These two considerations should affect the candidate vote choice but they should have no impact on the list vote. Logically, whether one likes the local candidate should have little influence on the decision to support a party list, all else equal. To be sure, Klingemann and Wessels (2001) noted previously that the German mixed system "[...] was specifically designed to strengthen personal ties between representatives and constituents (p. 279)." However, citing empirical evidence, the authors also argue that such a personalization of politics through local candidates probably has little effect on the list vote, and is limited to the candidate vote (see Klingemann and Wessels, 2001, 279-280). In the same manner, since the number of seats a party gets from the list vote is practically unaffected by the number of votes it gets in a constituency, there is no reason to defect from a party list simply because that party is unlikely to win in the local constituency. We thus predict these two factors to affect only the candidate vote.

The second prediction has to do with party and leader ratings. We expect these two factors to have a stronger effect on the list than on the candidate vote. The list vote entails expressing support for a given party, and we should thus observe that how much one likes a party has a strong effect on the propensity to vote for that party list. The marginal effect should be weaker in the case of the candidate vote since people are explicitly asked to express support for a person. We do expect a positive association, however, between party ratings and the candidate vote since people may well prefer to be represented in their local constituency by a person associated with a party that they like and trust even if they do not particularly like that person.

The same should apply to leader ratings. Poguntke and Webb (2005) argued that leaders have become increasingly important in contemporary political parties. They are the public face of the party during election campaigns, they exercise considerable control over the extra parliamentary party and its resources, and they lead the elected members in the legislature. Voting for a party is thus also implicitly voting for its leader. Again we expect a weaker effect of leader ratings on the candidate vote because the latter entails supporting a particular person in the constituency. We nevertheless anticipate a positive association because, everything else being equal, one should prefer the local candidate to be under the direction of a "good" party leader.

The third prediction is about coalition ratings. The hypothesis is that coalition preferences affect only the list vote. There is empirical evidence that in countries where coalition governments are the norm, people's vote choice depends not only on how they feel about the parties but also on how they feel about the coalitions that could be formed after the election (Abramson et al., 2008; Bargsted and Kedar, 2009; Blais et al., 2006; Gschwend and Hooghe, 2008; Meffert and Gschwend, 2010). Such considerations, however, should enter the calculus decision only for the list vote since the candidate vote has no consequence on the number of seats won by the various parties (and thus on the likelihood of different coalitions).

To sum up, we expect that vote choice in mixed systems is characterized by

$$\text{Local Chances of Winning: } \beta^1 > \alpha^1 = 0; \tag{7}$$

$$\text{Local Candidate Ratings: } \beta^2 > \alpha^2 = 0; \tag{8}$$

$$\text{Party Ratings: } \alpha^3 > \beta^3 > 0; \tag{9}$$

$$\text{Leader Ratings: } \alpha^4 > \beta^4 > 0. \tag{10}$$

Again, the parameters β are marginal effects in the latent utilities for the candidate vote, whereas the parameters α are the corresponding effects in the list vote. Since all those variables are party-specific, we denote them with a single coefficient. For coalition ratings, we expect that

$$\text{Coalition Ratings: } |\alpha_j^5| > |\beta_j^5| = 0, \ \forall\ j. \tag{11}$$

Notice that coalition ratings are voter-specific variables, which implies one parameter per party. We use absolute values since the signs can vary depending on which coalition is considered, and which party is used as a reference for comparisons. Finally, we expect that contamination effects are more important from the list vote toward the candidate vote, rather than the other way around:

$$\text{Contamination: } \theta_{LC} > \theta_{CL} > 0. \tag{12}$$

As explained before, we also expect contamination effects to be largest among the less sophisticated voters.

We test all these predictions with a special data set that includes all the necessary information about the five proximate determinants of vote choice identified above and the timing of vote decisions.

The 2013 German Federal Election

Our empirical tests rely upon survey data on the 2013 German federal election. As mentioned in the introduction to this paper, Germany uses a mixed compensatory system where voters cast two votes on the same ballot, one for a local candidate in their constituency and one for a party list.[3] The party list election takes place under a PR system whereas the local candidate election uses a FPTP system. Our survey includes samples from two major Landers, Bavaria and Lower Saxony, encompassing both variants of the German party system: in Bavaria, the Christian Social Union (CSU) runs in lieu of its sister party, the Christian Democratic Union (CDU), although both formations are united at the federal level under a single leader.

The 2013 election gave the center-right Christian Democratic Union/Christian Social Union (hereafter CDU for short) a plurality of seats, for the third consecutive time. In terms of popular vote, the CDU received above 40% of the ballots in both election types, distancing the center-left Social Democratic Party (SPD) by some 10 percentage points.[4] Angela Merkel's party also secured a comfortable advance from the second-runner in terms of the seat distribution, obtaining slightly less than 50% of the seats compared to 30.5% for the SPD. On the other hand, the Christian Democrats' traditional coalition partner, the Free Democratic Party (FDP), did not garner enough list votes to meet the legal threshold of 5%, and as a result, was not allocated any seat. The two other minor parties having received seats (the Greens and the Left, with about 10% of the deputation

[3] Notice that electoral agencies in Germany refer to the candidate and list votes as the "first" and "second" votes, respectively. We do not use those labels here.

[4] Official results are taken from the web site of the German Federal Returning Officer (www.bundeswahlleiter.de).

each) being ideologically at odds with Merkel's party, this complicated the formation of a coalition controlling a majority of seats in the Bundestag. The two largest parties eventually formed an unlikely coalition government.

Our data come from Internet campaign surveys conducted by Harris/Decima. The Bavarian survey was conducted between September 16 and 21, 2014, with 4,762 respondents, and a post-election wave conducted between September 23 and 28, with 4,041 from the first-wave responding to the second wave questionnaire. The Lower Saxon survey was conducted between September 12 and 19 with 1001 respondents. A post-election wave was held between September 23 and 30, with 789 respondents. Overall, the contact rate was 12% and the response rate 11.5%. In both cases, a stratified, quota based sampling approach was used, the quotas being established for age, gender, and education. The vote distributions in the sample closely match the observed vote distribution in the total population of voters, especially in the case of Bavaria. As a result, our multivariate analysis does not include sample weights. Our empirical models will contain a dummy variable called Bavaria, equaling one if a respondent is from Bavaria and zero if from Lower Saxony.[5]

The dependent variables are the two vote choices as reported during the post-election waves. For the purpose of our analysis, we focus on the five main parties in each election: CDU, SPD, Greens, FDP, and Left. The other parties garnered marginal vote shares, and we do not have measures of party and leader ratings for those smaller parties. The cross-tabulation in Table 1 shows the relationship between the two vote choices. As can be seen, most voters opt for a straight-line ticket, the main diagonal containing the largest proportions. Some proportion of Green and SPD voters were also keen on splitting their vote between those parties, which are usually expected to become coalition partners should the opportunity to form the government come about. Likewise, voters supporting the FDP list were somewhat likely to pick a CDU candidate in their local constituency, and vice-versa.

As mentioned above, our model includes five proximate determinants of the vote: party ratings, leader ratings, local candidate ratings, local chances, and coalition preferences. We also consider age, education, gender, and party identification as control variables. Table 2 reports descriptive statistics for those variables, which have all been rescaled between 0 and 1.

Party and leader ratings are based on questions asking respondents how much they like or dislike the various parties and leaders on a 0 to 10 scale (later rescaled from 0

[5]We tested whether our main results are affected by the combination of two Landers. The results presented below were replicated after including interaction terms with the Bavaria dummy variable. Those interaction effects were for the most part insignificant, suggesting that our results are not driven by one of the two regions.

Table 1: Cross-Tabulation of Observed List and Candidate Votes

		List Vote				
		CDU	SPD	Greens	FDP	Left
Candidate Vote	CDU	87%	3%	1%	8%	1%
	SPD	4	78	14	1	4
	Greens	6	16	69	3	7
	FDP	19	3	0	78	0
	Left	3	10	0	0	87
Total (N = 2,662)		1,300	754	266	185	157

The table reports the percentage distribution of the list vote across the candidate vote choice in the survey sample. Percentages may not add up to 100 due to rounding. The bottom row reports the frequencies by list vote choice.

to 1). Table 2 shows that the CDU was the most liked party while the Left was the most disliked. As for party leaders, Angela Merkel enjoyed an advantage with an average rating of 0.687 (about 7 out of 10), followed by Peer Steinbrück from the SPD. Overall, the ordering of average ratings in the sample is consistent with the support each party received at the time of the 2013 federal election.

The measurement of local candidate ratings differs from the previous two variables. We make use of responses to a survey question asking whether there is a candidate that the respondent particularly likes in the constituency, and if yes, from which party. Only around 30% of those who voted mentioned a candidate, almost all of them referring to a CDU or SPD candidate. The Local Chances variable is the score given by respondents to the perceived chances of each party winning in their constituency, on a 0 to 10 scale (again rescaled to run from 0 to 1). Typically, CDU and SPD candidates were perceived to have the best chances of winning in the constituency.

Coalition preferences were tapped through questions asking people how much they like or dislike (on a 0 to 10 scale) different government coalitions that could be formed after the election. For the purpose of this study, we focus on the two most plausible coalitions, those involving CDU with the FDP on the one hand and SPD with the Greens on the other hand. The actual turn of events leading to a Grand Coalition between the CDU and SPD was an unlikely outcome, which is why it would make little sense to explain vote choice using preferences over such an unusual coalition. To create our coalition ratings variable, we simply subtract the score given to the SPD–Greens coalition from the score given to the CDU–FDP coalition. The resulting variable is rescaled into the $[0, 1]$ range. The mean score is 0.523, suggesting that the CDU and FDP formed a slightly more popular coalition than did the SPD and Greens (0.5 indicates indifference between the coalitions).

Table 2: Descriptive Statistics

Variable	Party	Mean	Std. Err.
Local Chances	CDU	0.815	0.191
	SPD	0.547	0.252
	Greens	0.339	0.253
	FDP	0.206	0.220
	Left	0.142	0.196
Local Ratings	CDU	0.182	0.386
	SPD	0.086	0.280
	Greens	0.017	0.129
	FDP	0.006	0.077
	Left	0.006	0.077
Party Ratings	CDU	0.632	0.296
	SPD	0.596	0.254
	Greens	0.481	0.287
	FDP	0.305	0.278
	Left	0.229	0.286
Leader Ratings	CDU	0.687	0.317
	SPD	0.521	0.304
	Greens	0.391	0.289
	FDP	0.294	0.257
	Left	0.331	0.306
Coalition Ratings		0.523	0.299
Party ID	CDU	0.320	0.467
	SPD	0.178	0.382
	Greens	0.068	0.252
	FDP	0.020	0.141
	Left	0.024	0.152
Age		0.429	0.199
Education		0.610	0.381
Gender (Female = 1)		0.438	0.496

The table reports descriptive statistics for the 2,662 respondents with non-missing observations on all variables. All variables are scaled between 0 and 1.

Next, to account for contamination effects, we create time-ordered vote choice variables using information on the reported timing of the vote decisions. All post-election wave respondents were asked squarely whether they made one of the two vote decisions before the other. The wording of that survey question is "Which vote did you decide upon first?" and the response categories included "The candidate vote", "The list vote", "I decided both at the same time" and "Don't know" (translated from the German original version). Table 3 shows the distribution of responses to this question. Overall, about 36% of the respondents made their decision regarding the candidate vote before they chose which party list to support, whereas approximately 18% did the opposite. We measure the expressions $\mathbf{1}(L \rightarrow C)y_{Lj}$ from equation (6) using binary variables measuring the list vote choice conditional on having made a decision regarding party lists before the decision regarding local candidates. The expressions $\mathbf{1}(C \rightarrow L)y_{Cj}$ are constructed in a similar fashion, the other way around.

Table 3: Timing of the Vote Decisions

Timing	Frequency	Percentage
Candidate Vote First	951	36%
Both at the Same Time	1,227	46%
List Vote First	484	18%
Total	2,662	

Distribution of responses to the survey question tapping the timing of the two vote decisions.

In the last part of our empirical analysis, we reassess our hypotheses after accounting for our respondents' level of sophistication. We measure sophistication using educational attainment. Since survey quotas were established in part on this variable, its distribution closely matches that of the actual German population, making it the most reliable indicator of sophistication at our disposal. This variable contains three categories that account for the specificities of the German education system, in which different types of high schools coexist. The first category contains respondents with lower secondary or incomplete secondary schooling (about 20% of our final sample), the second category contains those with standard secondary schooling or technical degrees (37% of the sample), while the third category contains respondents with high secondary degrees or college education (42% of our sample). We create interaction terms by multiplying this educational attainment measure with our party-specific variables. For reasons exposed previously, we are especially interested in testing for the existence of contamination effects conditional on the level of education.

The choice of empirical estimators for our vote models requires some consideration. We started by considering multinomial logit models with alternative-specific variables (conditional logit models) but preliminary tests indicated that the assumption of independence of irrelevant alternatives cannot be met. Instead, we estimate the parameters of the two elections using separate multinomial probit models. Even though our models contain recursive elements (the contamination effects), the time-ordering restrictions that we impose in (4) ensure that the components $\mathbf{1}(s \rightarrow t)y_{sj}$ on the right-hand side of the latent utility functions do not introduce a correlation with the error terms ε_{tj}. This follows from the fact that $(s \rightarrow t)$ and $(t \rightarrow s)$ cannot be true at the same time, eliminating the possibility of feedback across equations. Previous research on binary probit models with endogenous dummy regressors suggest that when exogeneity is met, univariate estimators are to be preferred (Monfardini and Radice, 2007). We adopt a similar strategy with our multinomial estimators here.

For both our estimation methods, we rely upon Bayesian implementations of the multinomial probit model using the marginal data-augmentation algorithm proposed by Imai and van Dyk (2005*a*,*b*). This approach makes use of a Gibbs sampler to successively draw the latent utilities and the parameters, allowing to find credible posterior distributions of the form $p(\alpha|D)$, where we denote our data as D for short. For each model, we use noninformative priors for the parameters $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ and the θ's, namely a multivariate normal distribution with mean zero and an identity covariance matrix $\boldsymbol{I}$. The priors for the covariance matrices $\boldsymbol{\Sigma}_L$, $\boldsymbol{\Sigma}_C$ are drawn from the inverse-Wishart distribution using the default values proposed by Imai and van Dyk (2005*b*), the degrees of freedom being set to five—the number of alternatives—and the scale parameter to one. Notice that covariances matrices are of size $(J-1) \times (J-1)$ since one alternative serves as the base category. Moreover, the first diagonal elements of the covariance matrices are constrained to one to provide a normalizing scale.

Findings

We start by computing three samples of 1,000,000 Markov chain Monte Carlo (MCMC) draws, using different sets of starting values each time, for each of the two vote models. In both cases, a first MCMC sample is computed after setting the starting values of our parameters to zero. The next two samples use overdispersed starting values following the sequences $(-1, 1, -1, ...)$ and $(1, -1, 1, ...)$, respectively. This leaves us with a total of six million MCMC draws. To proceed with our empirical analysis, we make use of the last 500,000 draws from each of the three MCMC samples, once again for both the list vote and candidate vote equations.

Using three different samples allows the computation of the Gelman-Rubin potential scale reduction factors (PSRF) to assess whether the Markov chains have converged to a stationary distribution (Gelman and Rubin, 1992). Table 9, in the Appendix, reports those statistics for the parameters of both vote equations. A value close to one indicates that convergence has been achieved. As can be seen, for both equations, all values fall well below the conventional benchmark of 1.1, suggesting that we have successfully reached convergence. The conclusion holds when considering the upper limit of the 97.5% credible interval for this statistic. Table 9 also reports the p-values from the Heidelberger and Welch (1983) convergence tests. In all cases, we obtain values larger than 0.05, supporting the (null) hypothesis of stationary distributions and strengthening the conclusion that we have successfully reached convergence.

Table 4 reports the mean and the 95% credible intervals of the posterior distributions of parameters. The left section of the table reports statistics for the list vote equation, whereas the right section reports candidate vote estimates. Overall, the models perform well, correctly predicting 84.4% and 84.0% of actual vote choices, respectively for the list and candidate votes. This represents of a 70% proportional reduction in errors for the list vote model, and of 66% for the candidate vote model. Notice that the CDU serves as the reference category for voter-specific variables. Moreover, since all explanatory factors are on the same $[0,1]$ scale, the means and quantiles of the posterior distributions can be compared in size.

Table 4: Bayesian Multinomial Probit Models of List and Candidate Votes

		List Vote		Candidate Vote	
Party	Variable	Mean	Credible Interval	Mean	Credible Interval
	Local Chances (α^1, β^1)	–0.062	[–0.319,0.197]	0.452	[0.210,0.700]
	Local Ratings (α^2, β^2)	0.256	[0.084,0.432]	0.706	[0.512,0.904]
	Party Ratings (α^3, β^3)	3.103	[2.638,3.592]	2.423	[2.000,2.858]
	Leader Ratings (α^4, β^4)	1.288	[0.985,1.607]	0.813	[0.542,1.094]
	Candidate $\rightarrow$ List (θ_{CL})	0.200	[0.066,0.342]		
	List $\rightarrow$ Candidate (θ_{LC})			0.330	[0.171,0.498]
	Party ID	0.802	[0.651,0.967]	0.516	[0.368,0.668]
SPD	Coalition Ratings (α^5_2, β^5_2)	–1.245	[–1.899,–0.586]	–1.769	[–2.394,–1.170]
	Age	0.032	[–0.466,0.531]	0.195	[–0.256,0.647]
	Education	–0.233	[–0.509,0.041]	0.253	[0.004,0.503]
	Gender	0.073	[–0.128,0.276]	–0.005	[–0.185,0.176]
	Bavaria	–0.013	[–0.283,0.256]	0.054	[–0.193,0.302]
	Intercept	0.733	[0.284,1.180]	0.584	[0.168,1.005]
Greens	Coalition Ratings (α^5_3, β^5_3)	–1.076	[–1.856,–0.331]	–0.564	[–1.309,0.128]
	Age	–0.426	[–1.041,0.180]	0.108	[–0.467,0.674]
	Education	0.164	[–0.184,0.525]	0.203	[–0.113,0.529]
	Gender	0.072	[–0.175,0.323]	0.125	[–0.096,0.356]
	Bavaria	0.139	[–0.205,0.485]	0.148	[–0.160,0.463]
	Intercept	0.186	[–0.370,0.728]	–0.430	[–0.986,0.100]
FDP	Coalition Ratings (α^5_4, β^5_4)	1.391	[0.476,2.309]	0.043	[–0.987,1.032]
	Age	–0.405	[–1.060,0.251]	–1.265	[–2.021,–0.541]
	Education	0.059	[–0.340,0.473]	–0.109	[–0.542,0.343]
	Gender	–0.161	[–0.457,0.123]	–0.111	[–0.434,0.199]
	Bavaria	0.552	[0.095,1.026]	0.773	[0.257,1.328]
	Intercept	–1.229	[–2.035,–0.474]	–0.693	[–1.610,0.160]
Left	Coalition Ratings (α^5_5, β^5_5)	–0.973	[–1.873,–0.079]	–0.872	[–1.764,–0.002]
	Age	–0.206	[–0.989,0.575]	0.142	[–0.643,0.934]
	Education	–0.462	[–0.916,–0.010]	–0.220	[–0.664,0.223]
	Gender	0.031	[–0.304,0.368]	0.093	[–0.233,0.423]
	Bavaria	0.069	[–0.339,0.475]	0.320	[–0.079,0.725]
	Intercept	0.410	[–0.256,1.057]	–0.420	[–1.095,0.222]
% Correctly Predicted		84.5%		84.0%	
Observations		2,662		2,662	
Monte Carlo Draws		1,500,000		1,500,000	

Summary statistics of the posterior predictive distributions of parameters from the list and candidate vote equations, estimated with Bayesian multinomial probit models. The 95% credible intervals are reported between brackets.

Our first prediction is that local chances of winning and local candidate ratings affect the candidate vote but not the list vote, as stated in equations (7) and (8). The findings are consistent with this prediction when considering local chances of winning.

As shown in the left section of Table 4, the mean of the posterior distribution for the local chances parameter α^1 falls close to zero in the list vote equation. In contrast, the posterior distribution of the corresponding parameter β^1 in the candidate vote equation appears unambiguously positive. Our results suggest that an important reason why the major parties do better than their junior partners in the candidate election is that some supporters of small parties are willing to defect at the constituency level, because they do not want to waste their vote on a candidate who is unlikely to win. To facilitate visualization of this result, Figure 1a reports the posterior distributions for both parameters, stacked in the same graph.

However, we need concrete tests before drawing definitive conclusions involving comparisons across models. To achieve this, we use two different approaches. First, we evaluate the prediction that $\beta^1 > \alpha^1$ by testing whether the posterior distribution $p(\beta^1|D)$ is larger than the most credible value of α^1, namely the median of its posterior distribution, which we denote by $\tilde{\alpha}^1$. Because both models include the same variables and are estimated using identical priors and specifications, the coefficients are on the same scale, allowing comparisons based on point estimates. We posit the hypothesis $H_1 : \beta^1 > \alpha^1$ against the null $H_0 : \beta^1 \leq \alpha^1$, and rely upon non-informative priors $Pr(H_0) = 0.5$ and $Pr(H_1) = 0.5$. We then evaluate the posterior probability $Pr(H_1|D)$ numerically using the MCMC draws, which corresponds to

$$Pr(\beta^1 > \tilde{\alpha}^1|D) = \int_{\tilde{\alpha}^1}^{\infty} p(\beta^1|D)d\beta^1 \approx 0.99999. \tag{13}$$

The probability is close to 1, which means compelling evidence that $\beta^1 > \alpha^1$ is true. We may also compute the Bayes factor B_{10} as

$$B_{10} = \frac{Pr(H_1|D)}{Pr(H_0|D)} \bigg/ \frac{Pr(H_1)}{Pr(H_0)}, \tag{14}$$

representing the odds of observing the data given that our hypothesis is true, relative to the null hypothesis. Following usual conventions, this value can be assessed using the scale proposed by Kass and Raftery (1995) for twice the log Bayes factors.[6] In this case, we find that $2\ln(B_{10}) = 24.05$, confirming that we have very strong evidence in favor of H_1.

Second, since our posterior distributions are approximately normal, the difference $\beta^1 - \alpha^1$ is also normally distributed, and we can invoke the exchangeability assumption to

[6] The evidence in favor of H_1 is considered very strong if $2\ln(B_{10})$ is greater than 10, strong if between 6 and 10, positive if between 2 and 6, and barely worth mentioning if between 0 and 2 (Kass and Raftery, 1995, 777). The null is supported if the value is negative.

compute the posterior distribution of this difference numerically, using the MCMC draws of both parameters. This method accounts for the variance of both parameters. However, we need to make the strong assumption that the correlation between the coefficients is zero, which was not necessary using the first approach. Despite the caveat with the second approach, replicating the tests with two different methods is a way to assess the robustness of our findings. Using the posterior distribution of differences between parameters, the probability $Pr(H_1|D)$ becomes:

$$Pr(\beta^1 - \alpha^1 > 0|D) = \int_0^\infty p(\beta^1 - \alpha^1|D)d(\beta^1 - \alpha^1) \approx 0.998, \tag{15}$$

and the log Bayes factors can be computed as before, yielding $2\ln(B_{10}) = 12.07$. This confirms that we have strong evidence in favor of H_1. The summary of our tests, computed in a similar fashion for all of our predictions, is reported in Table 5.

Table 5: Hypothesis Testing

			Method 1		Method 2	
Hypothesis	H_1	H_0	$Pr(H_1\|D)$	$2\ln(B_{10})$	$Pr(H_1\|D)$	$2\ln(B_{10})$
Local Chances	$\beta^1 > \alpha^1$	$\beta^1 \leq \alpha^1$	0.99999	24.05	0.99762	12.07
Local Ratings	$\beta^2 > \alpha^2$	$\beta^2 \leq \alpha^2$	0.99930	14.52	0.99880	13.45
Party Ratings	$\alpha^3 > \beta^3$	$\alpha^3 \leq \beta^3$	0.99820	12.64	0.98215	8.02
Leader Ratings	$\alpha^4 > \beta^4$	$\alpha^4 \leq \beta^4$	0.99925	14.39	0.98809	8.84
Contamination	$\theta_{LC} > \theta_{CL}$	$\theta_{LC} \leq \theta_{CL}$	0.94526	5.70	0.88222	4.03
Coalition: SPD	$\|\alpha_2^5\| > \|\beta_2^5\|$	$\|\alpha_2^5\| \leq \|\beta_2^5\|$	0.05926	–5.53	0.12475	–3.90
Coalition: Greens	$\|\alpha_3^5\| > \|\beta_3^5\|$	$\|\alpha_3^5\| \leq \|\beta_3^5\|$	0.91485	4.75	0.83344	3.22
Coalition: FDP	$\|\alpha_4^5\| > \|\beta_4^5\|$	$\|\alpha_4^5\| \leq \|\beta_4^5\|$	0.99755	12.02	0.97593	7.40
Coalition: Left	$\|\alpha_5^5\| > \|\beta_5^5\|$	$\|\alpha_5^5\| \leq \|\beta_5^5\|$	0.59143	0.74	0.56360	0.51

Bayesian hypothesis tests based on the models reported in Table 4. Integrals are computed numerically using the MCMC draws. Method 1 compares the posterior distribution of the parameter of a vote model against the most credible value of the parameter in the other vote model (the median). Method 2 uses the difference between MCMC draws across vote models, restricting the correlation of coefficients to zero. In all cases, we use non-informative prior probabilities $Pr(H_0) = 0.5$ and $Pr(H_1) = 0.5$.

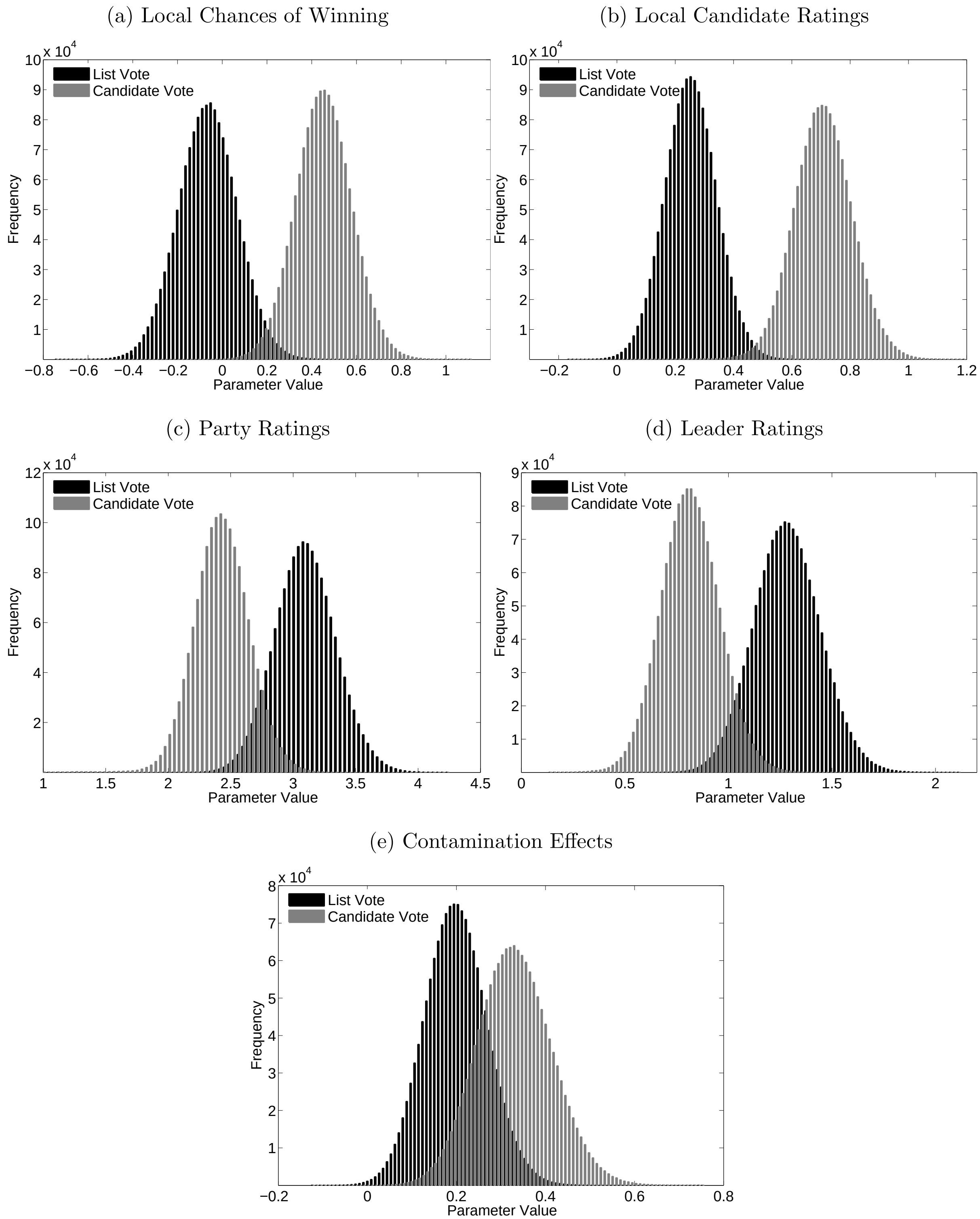
We reach similar conclusions when considering local candidate ratings. The means of the two posterior densities are once again far apart, which is apparent when plotting the distributions in Figure 1b. Accordingly, the probability that they differ in size is again close to 1, computed with either of the two methods introduced above, while the log Bayes factors suggest very strong evidence in favor of our prediction (second row of Table 5). However, returning to the left section of Table 4, we note that the credible interval does not contain the value of zero for local candidate ratings in the list vote equation. In other words, although we are confident in claiming that local candidate ratings are most important for the candidate vote, those evaluations do matter a little for the list vote as well, which is not what we expected. This suggests that the process of personalization of politics in the German system, as discussed in the previous section, may be relevant to some extant, even though this effect appears weak.

The second set of predictions, according to which party and leader ratings should have a stronger impact on the list vote than on the constituency vote, is also supported. In both cases, the means of the posterior distributions are clearly lower in the candidate vote equation than in the list vote's. Starting with party ratings, the mean of $p(\alpha^3|D)$ is greater than the 97.5 percentile of the corresponding parameter's density in the candidate vote equation, as shown in the third row of Table 4. The log Bayes factor is above 10 using the first approach to hypothesis testing, and between 5 and 10 using the second approach, which suggests strong evidence in favor of our hypothesis in the most conservative case. Again, we report the densities in Figure 1c. We note that party ratings do influence the candidate vote, though to a lesser extent, which is consistent with our assumption that some people simply vote for their preferred party's candidate in the local constituency. The conclusions are very similar for leader ratings. The probability $Pr(\alpha^4 > \beta^4|D)$ is close to one, substantiating our prediction that leader evaluations have a stronger effect on the list vote.

On the other hand, the third prediction, that coalition preferences disproportionately affect the list vote, is only partially supported. Table 4 shows that quantiles of the posterior densities in the list vote equation are larger (in absolute terms) than those of the local vote equation, except for the SPD. Everything else being equal, the more one prefers the CDU–FDP coalition over the SPD–Green coalition, the less likely one is to vote for the SPD, Green and Left party lists, compared to the CDU. On the other hand, the FDP appears to benefit from those coalition preferences, respondents more supportive of the CDU–FDP alliance being even more likely to choose the FDP's list than the CDU's, everything else being equal. However, our results indicate that coalition preferences also affect the SPD candidate vote, which runs contrary to our initial expectation. Using Bayes factors (rows 6 to 9 of Table 5), we find supporting evidence of our hypothesis only when considering the Greens and the FDP, and strong evidence only for the FDP.

Finally, our results provide a concrete assessment of contamination effects across votes. As explained earlier, the predetermined vote choice variables enter as exogenous regressors in each equation, and they represent estimates of the causal effect of a previously reached vote decision on the other vote. We do find evidence of contamination effects in both directions and report the posterior densities in Figure 1e. As can be seen by comparing the fifth and sixth rows of Table 4, those effects appear larger from the list vote toward the candidate vote than the other way around. Interestingly, even though more respondents declared making up their mind about the local vote first (see Table 3), this pattern does

Figure 1: Posterior Densities of List and Candidate Vote Parameters



not induce a larger contamination effect going in that direction. In fact, the opposite effect prevails: the smaller number of voters who chose in the list vote first were more likely to support the same party in their local constituency. Although consistent with our expectation, the evidence is not strong, as can be seen in the fifth row of Table 5. The posterior probability $P(\theta_{LC} > \theta_{CL}|D)$ approximates to 0.95 using our first approach, and only 0.88 using the second approach, meaning that we are only 88% percent confident that θ_{LC} is larger than θ_{CL} in the more conservative case.

To illustrate the size of contamination effects, we compute changes in the predicted probability of choosing a local candidate (or party list), without and with an anterior decision to support the list (or local candidate) of that party. Table 6 reports those values. In each case, the posterior predicted probabilities are computed after setting all other variables at their sample means. As can be seen in the upper portion of Table 6, a voter having made a prior decision to support the CDU list is more likely to vote for a CDU local candidate. The marginal effect is reported in the upper-left cell of Table 6, and corresponds to approximately +12 percentage points. Obviously, this implies that the same voter is less likely to support candidates from the other parties. The size of the contamination effect is similar for a voter having decided to support the SPD list, about +12 percentage points. On the other hand, the spillovers are weaker for smaller parties. The bottom part of the table reports the change in predicted probabilities associated with contamination effects in the opposite direction. As anticipated from the previous discussion, those effects are weaker in magnitude.

The results differ when considering the level of sophistication of voters. To show this, we reassess our last prediction after including interaction terms between the level of education (labeled x_E) and each of the party-specific indicators. The rest of the specification is the same as before. Table 7 reports the results. The lower panel of the table is produced by adding the MCMC draws for the coefficients of each variable and those of their respective interactions with education. Notice that, for simplicity, we only report the summary of posterior distributions for variables with interactions. As can be seen by comparing the two panels of Table 7, contamination effects are essentially driven by voters with lower levels of education. Figure 2 compares the posterior densities in both vote models, for lowly and highly educated voters, which helps to summarize our findings. When considering respondents with a lower level of education, our estimates now clearly support our initial expectation that contamination effects are more important from the list vote toward the candidate vote than the other way around. The log Bayes factor is 8.87, suggesting strong evidence in favor of our initial hypothesis, as can be observed in Table 8. The same is not true for voters with a high level of education (see Figure 2b, showing that posterior densities are very close to each other). We also note that, apart

Table 6: Estimated Size of Contamination Effects

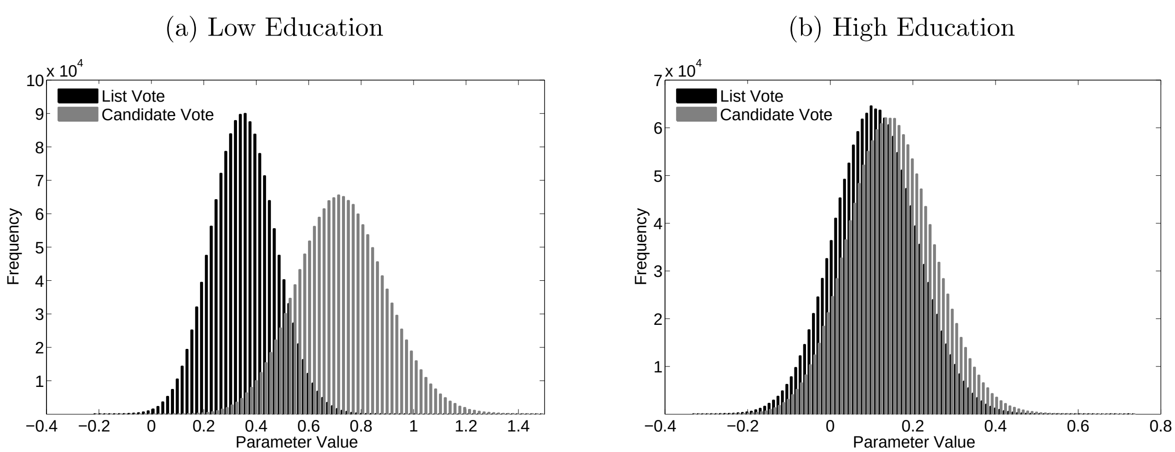
	List → Candidate				
	Change in Candidate Vote Probability				
List Vote	CDU	SPD	Greens	FDP	Left
CDU	0.117				
SPD		0.121			
Greens			0.046		
FDP				0.011	
Left					0.010

	Candidate → List				
	Change in List Vote Probability				
Candidate Vote	CDU	SPD	Greens	FDP	Left
CDU	0.075				
SPD		0.074			
Greens			0.028		
FDP				0.019	
Left					0.009

Marginal effects representing the change in the out-of-sample posterior predictive probability of choosing a party given that a voter previously made a decision to vote for the party indicated in the row header in the other election type. Probabilities are computed after setting all other explanatory variables of the models at their mean values.

from contamination effects, most of our other predictions are substantiated even more clearly when focusing on voters with a high level of education (see Table 8). Those results reinforce the idea that sophisticated voters are more likely to consider separate criteria for each the two vote decisions. In contrast, since the contamination argument implies a very simple decision-making heuristic, the fact that educated voters do not exhibit this behavior makes sense, substantively speaking.

Figure 2: Posterior Densities of Contamination Effects, by Education Level



Overall, our results on contamination suggest that such effects are rather modest in scope. They affect mostly the decision process of less sophisticated voters, and contami-

Table 7: Bayesian Multinomial Probit Models with Education Interactions

Variable	List Vote Mean	List Vote [Credible Interval]	Candidate Vote Mean	Candidate Vote [Credible Interval]
		Low Education (Education = 0)		
Local Chances	0.317	[–0.151,0.791]	0.513	[0.065,0.981]
Local Ratings	0.398	[0.089,0.717]	0.525	[0.216,0.846]
Party Ratings	2.540	[1.936,3.177]	2.052	[1.497,2.627]
Leader Ratings	1.597	[1.097,2.118]	1.089	[0.628,1.563]
Candidate → List	0.354	[0.118,0.599]		
List → Candidate			0.724	[0.400,1.066]
Party ID	0.466	[0.229,0.712]	0.412	[0.178,0.654]
		High Education (Education = 1)		
Local Chances	–0.245	[–0.584,0.092]	0.444	[0.141,0.755]
Local Ratings	0.170	[–0.080,0.426]	0.805	[0.531,1.092]
Party Ratings	3.675	[3.020,4.354]	2.800	[2.185,3.414]
Leader Ratings	1.062	[0.658,1.484]	0.630	[0.272,1.004]
Candidate → List	0.112	[–0.082,0.312]		
List → Candidate			0.146	[–0.053,0.356]
Party ID	0.981	[0.777,1.203]	0.567	[0.379,0.763]
% Correctly Predicted	84.5%		85.0%	
Observations	2,662		2,662	
Monte Carlo Draws	1,500,000		1,500,000	

Summary statistics of the posterior predictive distributions of parameters from the list and candidate vote equations, estimated with Bayesian multinomial probit models after including interaction variables between Education and each of the party-specific covariates. Only estimates for party-specific covariates are reported for simplicity, but the models include the same controls as in Table 4. The 95% credible intervals are reported between brackets. The Gelman-Rubin multivariate PSRF statistic approximates to 1.01 for both models.

nation from the list vote toward the candidate vote turns out to be more sizable. This last finding lends credence to the view that the list vote, which determines how many seats the party gets in the legislature, is the most important of the two. Our results are also consistent with the literature at the party level, which shows that contamination effects are clearly from the PR component to the FPTP component. That is, in compensatory systems where the PR dimension dominates (as in the German case), small parties tend to nominate more candidates in the local constituencies than they would "normally" do in a FPTP election, because they have already decided to present them in the PR election (Ferrara, Herron, and Nishikawa, 2005, 63). Our own conclusion is that contamination effects, when they arise, are more likely to run in the same direction for voters too, from the more "important" component (PR) to the less important (FPTP).

Table 8: Hypothesis Testing, by Level of Education

Hypothesis	H_1	H_0	Method 1 $Pr(H_1\|D)$	Method 1 $2\ln(B_{10})$	Method 2 $Pr(H_1\|D)$	Method 2 $2\ln(B_{10})$
Low Education (Education = 0)						
Local Chances	$\beta^1 > \alpha^1$	$\beta^1 \leq \alpha^1$	0.79948	2.77	0.72032	1.89
Local Ratings	$\beta^2 > \alpha^2$	$\beta^2 \leq \alpha^2$	0.78726	2.62	0.71193	1.81
Party Ratings	$\alpha^3 > \beta^3$	$\alpha^3 \leq \beta^3$	0.94291	5.61	0.87346	3.86
Leader Ratings	$\alpha^4 > \beta^4$	$\alpha^4 \leq \beta^4$	0.97722	7.52	0.92486	5.02
Contamination	$\theta_{LC} > \theta_{CL}$	$\theta_{LC} \leq \theta_{CL}$	0.98826	8.87	0.96248	6.49
High Education (Education = 1)						
Local Chances	$\beta^1 > \alpha^1$	$\beta^1 \leq \alpha^1$	0.99999	23.31	0.99863	13.18
Local Ratings	$\beta^2 > \alpha^2$	$\beta^2 \leq \alpha^2$	0.99997	21.01	0.99923	14.35
Party Ratings	$\alpha^3 > \beta^3$	$\alpha^3 \leq \beta^3$	0.99565	10.86	0.97237	7.12
Leader Ratings	$\alpha^4 > \beta^4$	$\alpha^4 \leq \beta^4$	0.98260	8.07	0.93729	5.41
Contamination	$\theta_{LC} > \theta_{CL}$	$\theta_{LC} \leq \theta_{CL}$	0.62846	1.05	0.59293	0.75

Bayesian hypothesis tests based on the models reported in Table 7. Integrals are computed numerically using the MCMC draws. Method 1 compares the posterior distribution of the parameter of a vote model against the most credible value of the parameter in the other vote model (the median). Method 2 uses the difference between MCMC draws across vote models, restricting the correlation of coefficients to zero. In all cases, we use non-informative prior probabilities $Pr(H_0) = 0.5$ and $Pr(H_1) = 0.5$.

Conclusion

In many elections voters are faced with two decisions when they are making up their mind how to vote: which party list to support and which candidate to support in the constituency. As far as we can tell, no previous study has examined the determinants of the two votes using an integrated framework.

We formulated hypotheses regarding five proximate determinants of voting and their relative influence on the two vote decisions in mixed systems. As predicted, local chances of winning and local candidate ratings affect mostly the candidate vote, while party and leader ratings affect more strongly the list vote. We have also found coalition preferences to affect mostly (although not uniquely) the list vote. We believe that the methodology proposed here is the most logical way to proceed. There are two votes, and scholars must provide an explanation for each. Because the factors driving voter behavior are expected to be similar for both votes, we argue in favor of a general model comprising the same variables in each vote equation. The goal is to determine whether some of the factors have a greater influence on the candidate vote than on the list vote.

We have found some differences, and these differences make sense. The most important is that perceptions of local chances affect the candidate vote but not the list vote. Moreover, our analysis helps to understand why ticket splitting is not more widespread. The fact is, for both votes, the most crucial proximate factor is how one feels about the parties. This is obvious with respect to the list vote but less so for the candidate vote. The two votes are first and foremost an expression of party preferences, and this is why most people support the same party with their two votes.

Moreover, our methodology allows us to make a contribution to the micro-foundations of contamination effects. We developed a general framework to estimate potential contamination effects between votes, and relied upon a direct measure of those effects by making use of information on the timing of vote decisions. Our findings suggest that contamination effects are limited to voters with lower levels of education. Previous studies have examined whether the list vote is influenced by the local candidate vote and they have found weak contamination effects. We observe a weak contamination effect in that direction using our survey data at the time of the 2013 German federal election. But spillovers can go in the opposite direction, from the list vote to the candidate vote, and we have found stronger evidence of this type of effect. Overall, we believe that those findings make sense. We should expect people to pay greater attention to the most "important" vote, and the influence should flow mainly from the most to the least important decision.

References

Abramson, Paul R., John H. Aldrich, André Blais, Daniel Lee, and Renan Levine. 2008. "Coalition Considerations and the Vote." In *The Elections in Israel—2006*, ed. Asher Arian, and Shamir Michal. New Brunswick: Transaction Books.

Ames, Barry. 1994. "The Reverse Coattails Effect: Local Party Organization in the 1989 Brazilian Presidential Election." *American Political Science Review* 88(1): 95–111.

Bargsted, Matias A., and Orit Kedar. 2009. "Coalition-Targeted Duvergerian Voting: How Expectations Affect Voter Choice under Proportional Representation." *American Journal of Political Science* 53(2): 307–323.

Blais, André, John H. Aldrich, Indridi H. Indridason, and Renan Levine. 2006. "Do Voters Vote For Government Coalitions?: Testing Downs' Pessimistic Conclusion." *Party Politics* 12(6): 691–705.

Burden, Barry C. 2009. "Candidate-Driven Ticket Splitting in the 2000 Japanese Elections." *Electoral Studies* 28(1): 33–40.

Campbell, James E., and Joe A. Sumners. 1990. "Presidential Coattails in Senate Elections." *American Political Science Review* 84(2): 513–524.

Carter, Elisabeth, and David Farrell. 2010. "Electoral Systems and Election Management." In *Contemporary Democracies 3: Elections and Voting in the 21st Century*, ed. Lawrence LeDuc, Richard G. Niemi, and Norris Pippa. London: Sage.

Cox, Karen E., and Leonard J. Schoppa. 2002. "Interaction Effects in Mixed-Member Electoral Systems: Theory and Evidence From Germany, Japan, and Italy." *Comparative Political Studies* 35(9): 1027–1053.

Crisp, Brian F., Joshua D. Potter, and John J. W. Lee. 2012. "Entry and Coordination in Mixed-Member Systems: A Controlled Comparison Testing the Contamination Hypothesis." *The Journal of Politics* 74(4): 571–583.

Ferejohn, John A., and Randall L. Calvert. 1984. "Presidential Coattails in Historical Perspective." *American Journal of Political Science* 28(1): 127–146.

Ferrara, Federico. 2004. "Electoral Coordination and the Strategic Desertion of Strong Parties in Compensatory Mixed Systems with Negative Vote Transfers." *Electoral Studies* 23(3): 391–413.

Ferrara, Federico, Erik S. Herron, and Misa Nishikawa. 2005. *Mixed Electoral Systems: Contamination and Its Consequences*. New York: Palgrave MacMillan.

Gelman, Andrew, and Donald B. Rubin. 1992. "Inference from Iterative Simulation using Multiple Sequences." *Statistical Science* 7(4): 457–511.

Granger, Clive W. J. 1969. "Investigating Causal Relations by Econometric Models and Cross-spectral Methods." *Econometrica* 37(3): 424–438.

Gschwend, Thomas. 2007. "Ticket-Splitting and Strategic Voting under Mixed Electoral Rules: Evidence from Germany." *European Journal of Political Research* 46(1): 1–23.

Gschwend, Thomas, and Marc Hooghe. 2008. "Should I Stay or Should I Go? An Experimental Study on Voter Responses to Pre-Electoral Coalitions." *European Journal of Political Research* 47(5): 556–577.

Gschwend, Thomas, Ron Johnston, and Charles Pattie. 2003. "Split-Ticket Patterns in Mixed-Member Proportional Election Systems: Estimates and Analyses of Their Spatial Variation at the German Federal Election, 1998." *British Journal of Political Science* 33(1): 109–127.

Hainmueller, Jens, and Holger Lutz Kern. 2008. "Incumbency as a Source of Spillover Effects in Mixed Electoral Systems: Evidence from a Regression-Discontinuity Design." *Electoral Studies* 27(2): 213–227.

Heidelberger, Philip, and Peter D. Welch. 1983. "Simulation Run Length Control in the Presence of an Initial Transient." *Operations Research* 31(6): 1109–1144.

Helmke, Gretchen. 2009. "Ticket Splitting as Electoral Insurance: The Mexico 2000 Elections." *Electoral Studies* 28(1): 70–78.

Herron, Erik S., and Misa Nishikawa. 2001. "Contamination Effects and the Number of Parties in Mixed-Superposition Electoral Systems." *Electoral Studies* 20(1): 63–86.

Imai, Kosuke, and David A. van Dyk. 2005*a*. "A Bayesian Analysis of the Multinomial Probit Model Using Marginal Data Augmentation." *Journal of Econometrics* 124(2): 311–334.

Imai, Kosuke, and David A. van Dyk. 2005*b*. "MNP: R Package for Fitting the Multinomial Probit Model." *Journal of Statistical Software* 14(3): 1–32.

Kass, Robert E., and Adrian E. Raftery. 1995. "Bayes Factors." *Journal of the American Statistical Association* 90(130): 773–795.

Karp, Jeffrey A. 2009. "Candidate Effects and Spill-Over in Mixed Systems: Evidence from New Zealand." *Electoral Studies* 28(1): 41–50.

Karp, Jeffrey A., Jack Vowles, Susan A. Banducci, and Todd Donovan. 2002. "Strategic Voting, Party Activity, and Candidate Effects: Testing Explanations for Split Voting in New Zealand's New Mixed System." *Electoral Studies* 21(1): 1–22.

Klingemann, Hans-Dieter, and Bernhard Wessels. 2001. "The Political Consequences of Germany's Mixed-Member System: Personalization at the Grass Roots?" In *Mixed-Member Electoral Systems: The Best of Both Worlds?*, ed. Matthew Soberg Shugart, and Martin P. Wattenberg. Oxford: Oxford University Press pp. 279–296.

Lewis-Beck, Michael S., and Richard Nadeau. 2004. "Split-Ticket Voting: The Effects of Cognitive Madisonianism." *The Journal of Politics* 66(1): 97–112.

Maeda, Ko. 2008. "Re-Examining the Contamination Effect of Japans Mixed Electoral System Using the Treatment-Effects Model." *Electoral Studies* 27(4): 723–731.

Massicotte, Louis, and André Blais. 1999. "Mixed Electoral Systems: A Conceptual and Empirical Survey." *Electoral Studies* 18(3): 341–366.

Meffert, Michael F., and Thomas Gschwend. 2010. "Strategic Coalition Voting: Evidence from Austria." *Electoral Studies* 29(3): 339–349.

Mellor, David Hugh. 2002. *The Facts of Causation.* New York: Routledge.

Monfardini, Chiara, and Rosalba Radice. 2007. "Testing Exogeneity in the Bivariate Probit Model: A Monte Carlo Study." *Oxford Bulletin of Economics and Statistics* 70(2): 271282.

Moser, Robert G., and Ethan Scheiner. 2004. "Mixed Electoral Systems and Electoral System Effects: Controlled Comparison and Cross-National Analysis." *Electoral Studies* 23(4): 575–599.

Moser, Robert G., and Ethan Scheiner. 2005. "Strategic Ticket Splitting and the Personal Vote in Mixed-Member Electoral Systems." *Legislative Studies Quarterly* 30(2): 259–276.

Moser, Robert G., and Ethan Scheiner. 2009. "Strategic Voting in Established and New Democracies: Ticket Splitting in Mixed-Member Electoral Systems." *Electoral Studies* 28(1): 51–61.

Page, Benjamin I., and Calvin C. Jones. 1979. "Reciprocal Effects of Policy Preferences, Party Loyalties and the Vote." *The American Political Science Review* 73(4): 1071–1089.

Pappi, Frank U., and Paul 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.

Poguntke, Thomas, and Paul Webb. 2005. *The Presidentialization of Politics.* Oxford: Oxford University Press.

Rahn, Wendy M., John H. Aldrich, Eugene Borgida, and John Sullivan. 1990. "A Social Cognitive Model of Candidate Appraisal." In *Information and Democratic Processes*, ed. John Ferejohn, and James Kuklinski. Champaign: University of Illinois Press pp. 187–206.

Samuels, David J. 2000. "The Gubernatorial Coattails Effect: Federalism and Congressional Elections in Brazil." *The Journal of Politics* 62(1): 240–253.

Appendix

Table 9: Convergence Diagnostic Statistics

		List Vote			Candidate Vote		
Party	Parameter	GR	GR (97.5%)	HW-p	GR	GR (97.5%)	HW-p
	Local Chances	1.0002	1.0005	0.33	1.0004	1.0011	0.06
	Local Ratings	1.0000	1.0001	0.32	1.0018	1.0019	0.52
	Party Ratings	1.0003	1.0009	0.16	1.0061	1.0072	0.62
	Leader Ratings	1.0001	1.0005	0.10	1.0005	1.0005	0.52
	Contamination Effects	1.0000	1.0001	0.45	1.0001	1.0003	0.61
	Party ID	1.0001	1.0002	0.68	1.0006	1.0007	0.32
SPD	Coalition Ratings	1.0002	1.0004	0.81	1.0021	1.0021	0.27
	Age	1.0000	1.0001	0.09	1.0001	1.0003	0.14
	Education	1.0001	1.0004	0.22	1.0004	1.0013	0.16
	Gender	1.0001	1.0002	0.74	1.0001	1.0004	0.21
	Bavaria	1.0000	1.0000	0.97	1.0002	1.0006	0.24
	Intercept	1.0001	1.0002	0.94	1.0004	1.0005	0.26
Greens	Coalition Ratings	1.0000	1.0001	0.23	1.0007	1.0009	0.11
	Age	1.0000	1.0000	0.19	1.0002	1.0005	0.49
	Education	1.0000	1.0001	0.47	1.0005	1.0019	0.33
	Gender	1.0001	1.0004	0.05	1.0002	1.0008	0.08
	Bavaria	1.0000	1.0001	0.92	1.0003	1.0011	0.29
	Intercept	1.0000	1.0001	0.84	1.0002	1.0004	0.18
FDP	Coalition Ratings	1.0004	1.0014	0.05	1.0019	1.0069	0.32
	Age	1.0000	1.0000	0.63	1.0002	1.0008	0.58
	Education	1.0001	1.0002	0.44	1.0004	1.0012	0.22
	Gender	1.0000	1.0001	0.97	1.0004	1.0011	0.16
	Bavaria	1.0002	1.0005	0.15	1.0006	1.0021	0.20
	Intercept	1.0004	1.0015	0.21	1.0013	1.0044	0.39
Left	Coalition Ratings	1.0004	1.0014	0.06	1.0015	1.0049	0.36
	Age	1.0002	1.0008	0.80	1.0003	1.0007	0.17
	Education	1.0002	1.0007	0.59	1.0003	1.0010	0.65
	Gender	1.0002	1.0006	0.38	1.0002	1.0008	0.73
	Bavaria	1.0001	1.0002	0.48	1.0003	1.0009	0.69
	Intercept	1.0002	1.0005	0.09	1.0001	1.0004	0.11
Cov. Matrix	SPD:Greens	1.0005	1.0013	0.22	1.0004	1.0006	0.07
	SPD:FDP	1.0005	1.0015	0.28	1.0041	1.0117	0.86
	SPD:Left	1.0003	1.0009	0.16	1.0009	1.0017	0.74
	Greens:Greens	1.0002	1.0008	0.19	1.0009	1.0011	0.55
	Greens:FDP	1.0003	1.0009	0.40	1.0006	1.0019	0.63
	Greens:Left	1.0006	1.0016	0.07	1.0002	1.0004	0.95
	FDP:FDP	1.0002	1.0005	0.89	1.0033	1.0040	0.67
	FDP:Left	1.0008	1.0022	0.31	1.0004	1.0007	0.29
	Left:Left	1.0004	1.0009	0.06	1.0039	1.0059	0.88
Multivariate PSRF		1.0022			1.0063		

The first two columns for each equation report the Gelman-Rubin (GR) potential scale reduction factor (PSRF) statistics along with the 97.5% upper bound of their credible interval. The multivariate PSRF statistics are reported in the last row of the table. The third column reports the p-value of the Heidelberger-Welch (HW) diagnostic test of stationary distributions. Values larger than 0.05 indicate that the test is passed, supporting the conclusion of a stationary distribution. The tests are performed on a combined sample of 1,500,000 MCMC draws for each vote model.