AUCO Czech Economic Review 3 (2009) 270–291
Acta Universitatis Carolinae Oeconomica

Simulating Party Competition and Vote Decision under Mixed Member Electoral Systems

Susumu Shikano*

Received 22 August 2007; Accepted 27 July 2008

Abstract This paper proposes an interactive mechanism between both tiers of mixed-member electoral systems to explain high-level linkage in West Germany, that is, the dominance of district races by the candidates of the two largest parties at the national level. The distinctive feature of the model lies in interactive effects in terms of expectation formation. That is, voters under mixed systems are assumed to utilize national-level PR results to form expectations which, in turn, are used to vote strategically in the plurality tier. To sort out the independent effect of these kinds of interactions, this paper develops a computational model and examines its simulation results.

Keywords Agent-based modelling, mixed-member electoral systems, strategic voting
JEL classification D72

1. Linkage in the West German party system

There are some simple facts which are obvious to those within Germany, but are largely ignored by people outside it. Such a fact is that all of West Germany's single member districts (SMD) elected using plurality, which makes up one half of the members of the German federal parliament (Bundestag), are won by either the CDU/CSU (Christian Democrats) or the SPD (Social Democrats). No district has been won by another party since 1961. This becomes more remarkable if one observes not only gained seats, but also competitions for the seats. The solid line in Figure 1 shows the percentage of West German districts where candidates of both large parties ranked first and second. Accordingly, at the first general election 1949 only cca 60 % of districts were contested between the candidates of CDU/CSU and SPD. Thereafter, this percentage increased rapidly. Since 1965, all districts in West Germany have been dominated by the candidates of both large parties without any exceptions. To compare this with England, while in the 1950's almost all English districts were competed between the candidates of the Labour and the Conservative, the share of such districts has never reached 100%. It even decreased in the course of time and sank under 50% in the 1980's. That is, in more than half of the districts, the top two candidates were either {Conservative, one small party} or {Labour, one small party} or {one small party, another small party}.

* University of Konstanz, Faculty of Law, Economics and Politics, Department of Political Science and Management, Postbox D92, Universitätsstrasse 10, 78457 Konstanz, Germany. Phone: +49(0)7531883679, E-mail: susumu.shikano@uni-konstanz.de.

Note: West Berlin is not accounted in West Germany.

Figure 1. Percentage of districts which are competed between the two main parties at national level in West Germany and England

The results above demonstrate that the district races in Germany are more uniformly dominated by the two largest national parties than the ones in England. This should appear counter-intuitive in the eyes of most political scientists. Students in the discipline learn in their introductory course on comparative politics that there are two archetypes of democracy: majoritarian and consensual democracy. According to the celebrated work of Lijphart (1999), the majoritarian democracy is characterized by two competing large parties which alternately control governmental power. In the consensus democracy, by contrast, more parties are visible and coalition government is common. Lijphart takes the Westminster-Model of England as the epitome of the former type of democracy and (West) Germany for the latter consensual type. The findings above are, however, not in line with the classification since election results in West Germany show a more typical pattern of bipartism than England.

The goal of this paper is to give an answer to the anomaly discussed above. My supposition is that the answer can be found in the co-existence of plurality- and PR-rule under the West German electoral system, which more recently has been referred to as a “mixed-member electoral system” or simply “mixed system” (Shugart 2001). Here, I speculate that the visibility of the PR tier facilitates the uniform dominance of the two large national parties and suppresses local district races deviating from the national one. The visibility of the two large national parties also boosts the advantage of their candidates so that they enjoy a significantly large margin over other candidates. The trend toward this dualism in district races, in turn, contributes to an increasing visibility of the large two parties in the PR tier. This further amplifies the dualism of district races though the mechanism above. This spiral of dualism leads to a stable dominance of two large parties in plurality as well as in the PR tier which we have observed in West German election results.

This uniformity of district races has been studied as linkage by multiple scholars in connection with Duverger's Law (for example Cox 1997). Duverger's Law forecasts bipartism under simple plurality rule based on the wasted-vote logic. However, this wasted-vote logic can work only at the district level. Assuming N single-member districts, two candidates can gain significant amounts of votes in each of N districts, while these $2N$ candidates do not necessarily belong to one of the two main parties at the national level. As a result, up to $2N$ parties can gain significant amounts of votes. To establish bipartism at the national level, in contrast, the $2N$ candidates should be linked with each other into two large national-level parties. This process is called linkage and its perfect form has been observed in the uniformity in West Germany district races. Whereas the linkage process under plurality system has been studied by multiple scholars, there are few studies about this process under mixed systems. This paper aims to off set this research deficit.[1]

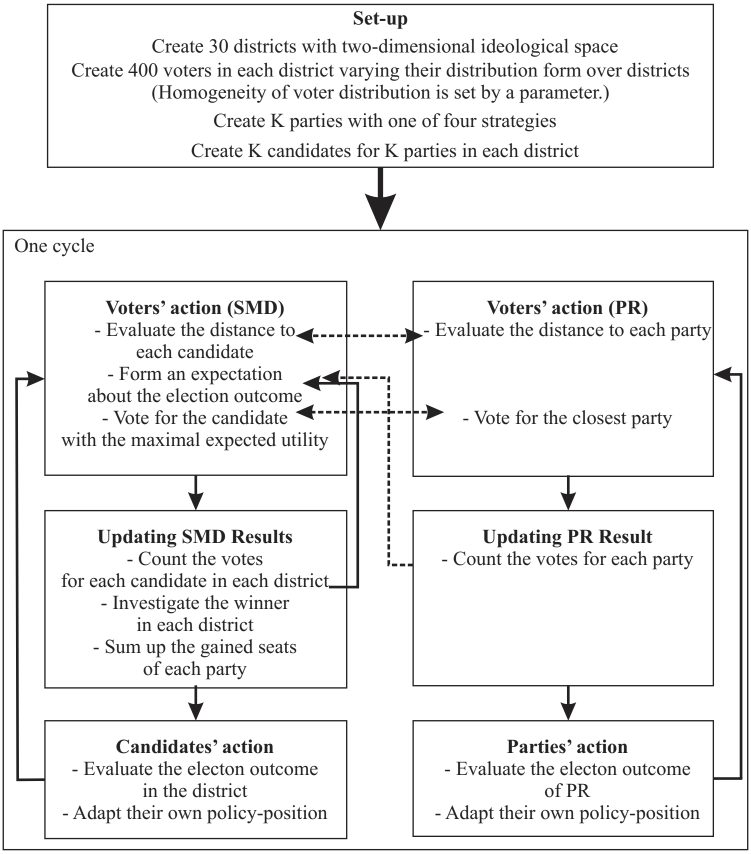
The remainder of this paper proceeds as follows: the next section introduces a model with interactive mechanism of mixed system. The third section gives the results of simulated data. The last fourth section closes this paper with some discussions on this paper's results.

2. Set-ups of computational model

This section introduces the model to be simulated below. The distinctive feature of the model is that it incorporates diverse kinds of interactive effects between two ballots under mixed systems. While one such interactive effect was investigated as contamination by Ferrara, Herron and Nishikawa (2005) extensively, this paper's model accounts for further kinds of interactive effects. Among them, interactive effects in expectation formation are most important for this paper's conjecture: national-level bipartism under mixed systems. More concretely, the use of national-level results in the PR tier in expectation formation in the plurality tier operates in favor of candidates belonging to the largest national parties and to the disadvantage of candidates of regional parties (for a less formal introduction of the model and its assumptions see Shikano 2007, Chapter 4).

More specifically, the simulation model developed here is an extension of Laver (2005). Laver's model can be summarized as follows: first, a two-dimensional ideological space is set up in which the ideal point of voters are normally distributed. Second, parties are also assigned to a certain ideological position in the same ideological space. Third, voters evaluate their distance to each party and cast their ballots for the nearest party. Fourth, cast votes are counted and the results of each party are announced. Fifth, being confronted with the election result, each party adapts their own ideological position. There are four strategies according to which each party adapt their positions: Hunter, Aggregator, Predator and Sticker. All of the strategies are adaptive, that is, they

[1] This paper's measure of linkage is different from those in conventional studies. For more details see Shikano (2007, Chapter 2).



Note: See also the overview of Laver (2005, Figure 2) for the differences from his simulation model.

Figure 2. Overview of the simulation model

need only limited information of party competition. Some of them are results-oriented or vote-maximizing and others are rather policy-oriented.[2] Repeating the third, fourth and fifth step, Laver observes the consequences of various combinations of the party strategies.

I chose Laver's model for the following reasons: first, the setting of simulation is relatively simple. For example, the ideological space is two-dimensional. It is simple; however, it also corresponds to the ideological space which is assumed for various political systems by many political scientists and the other kinds of experts. Second, there is a variety of party strategies whereby some parties are more oriented at the election result and the other more at policy. This is one of the features which are not feasible in conventional analytical models, but in computational models. Third, all

[2] Each of the strategies is described more in detail later in this section.

strategies are not optimizing, but adaptive. That is, parties have to adapt their position using only limited information concerning the election results and the position of other parties. In other computational models parties often possesses much more information, in particular prospective information, for example possible results in a hypothetical move in future. If one considers actor's bounded capacity of information processing, however, it is rather unrealistic to assume that political actors can take such kinds of strategy. Fourth, Laver's model does not aim to find equilibrium, but to describe complex and dynamic processes of party competition. This is also the aim of the current paper as discussed above.

Figure 2 gives the overview of the model. I extended Laver's model in following four points: First, Laver models party competitions under the pure proportional representation (PR)-system. The present paper, in contrast, is interested in the party competition and voting behavior under mixed systems. Correspondingly, I also model the party competition under plurality in single-member districts. Second, aside from just the party, the model here also incorporates the behavior of individual district candidates who can take a deviating policy-position from that of their own parties. The third extension concerns the action of voters. Voters in Laver's model always vote sincerely, i.e. for the closest party in the ideological space. This is, however, less plausible if one models voter decision under the plurality system. Under this system, as we have seen above, voters can strategically cast their vote for the second best candidate to maximize their expected utility (for example Cox 1997). Therefore, not only the proximity to parties, but also the expectation of the election's outcome in plurality races is also incorporated here. Fourth, the model of the present paper suggests the existence of some interactive effects in casting plurality- and PR-ballots. To observe which kind of consequences the interactive effects have, the model is also extended in this regard (dotted arrows in Figure).[3]

2.1 Ideological space and voter distribution

Voters in the model are assumed to be oriented by the policy outcome which is represented as position in a two-dimensional ideological space $\mathbb{R} \times \mathbb{R}$. Voter i's ideal point is represented as:

$$z_i = (z_{i1}, z_{i2}) \in \mathbb{R} \times \mathbb{R} \tag{1}$$

While the number of dimensions (two) is chosen arbitrarily, the number of dimensions is less relevant here. This is because the number of dimensions is equivalent to heterogeneity of voter distributions in different districts. As discussed below, the heterogeneity is parameterized in the computational model and simulates different kinds of dimensionality.

Each dimension has a length of 100 in the sense that a single dimension contains 100 possible ideological positions for voters, parties and candidates. Since the space is two-dimensional, there are 10,000 ($= 100 \times 100$) possible positions. In this sense, the

[3] The model introduced below was programmed using Repast. Repast (Recursive Porus Agent Simulation Toolkit) is a broadly used, free and open-source Java-based toolkit for agent-based modeling and further simulation techniques. The simulation program introduced here is available from the author upon request.

ideological space is not continuous, but discrete. Each discrete position in the space can be occupied by multiple actors.

In the same ideological space, political parties and their candidates compete for the vote and seat with following positions:

$$\omega_j = (\omega_{j1}, \omega_{j2}) \in \mathbb{R} \times \mathbb{R} \tag{2}$$

$$\omega'_j = (\omega_{j'1}, \omega_{j'2}) \in \mathbb{R} \times \mathbb{R} \tag{3}$$

One important difference from that of Laver (2005) is that this paper's model also accounts for the plurality race in each SMD. Therefore, while Laver's model has only one district at national level the model of this paper sets up 30 single member districts with 400 voters each. All districts share the same two-dimensional ideological space.

Voters are distributed in a limited space as defined above. To simulate different voter distributions in the limited ideological space, the Beta distribution is appropriate. This distribution can take a diverse flexible form on a limited scale.

The shape of a beta distribution is determined by two parameters α and β. Correspondingly, the ideal points of each voter $i \in I_k$ in district k on each dimension are drawn as follows:

$$z_{i1} \sim \mathscr{B}(\alpha_k, \beta_k)$$

$$z_{i2} \sim \mathscr{B}(\alpha_k, \beta_k)$$

I_k is the set of voters who cast their votes in district k.

If α_k and β_k are homogeneous over K districts, each district's voter distribution should be similar to each other since they are drawn from a similar distribution. In contrast, heterogeneity of α_k and β_k leads to heterogeneous voter distributions. To parameterize this, α_k and β_k are also drawn from another beta distribution $\mathscr{B}(10^h, 10^h)$:

$$\alpha_k \sim \mathscr{B}(10^h, 10^h) \times 5 + 2 \ (\in [2,7])$$

$$\beta_k \sim \mathscr{B}(10^h, 10^h) \times 5 + 2 \ (\in [2,7])$$

The parameter $h \in [-2, 2]$ controls the homogeneity/heterogeneity of drawn α_k and β_k. To see this, one has to know further characteristics of the beta distribution. First, if both parameters are the same, the shape of the corresponding beta distribution is symmetric. Therefore, $\mathscr{B}(10^h, 10^h)$ results in symmetrical distributions. Second, the larger the parameter values the more density around the middle of distribution. If both parameter values equal one, the beta distribution is the uniform distribution. This is realized by $h = 0$ since $10^0 = 1$. If both parameter values are less than one, the beta distribution is bimodal at both extreme values. This is the case if $h < 0$.

To summarize, the higher the assigned value of h is, the more similar the form of voter distribution in each district. For example, the distributions generated with a high homogeneity ($h = 1$) show all symmetrical and normal distributions which are similar to each other. In contrast, some of the distributions generated with a low homogeneity ($h = -1$) show skewed distributions in various directions and the other show symmetrical distributions. Among symmetrical distributions, however, some have a larger variance than the other.

2.2 Voters' actions

2.2.1 Utility and vote choice

Voters cast their ballot according to their proximity to the party/candidate. More formally, voter i's utility of party j is specified as the negative of the squared distance between the ideal point of i and that of j:

$$U_i(j) = -(\omega_j - z_i)^2 \tag{4}$$

The negative sign is used since utility declines with increasing distance. Voters decide deterministically according to the utility defined above. That it, they vote for a party whose utility is the highest among all parties. If there is a tie at the highest utility the voter decides randomly among the parties with the highest utility.

If a non-zero value is set as weight of the actual election result in the computing expectation, voters form expectations on the outcome of district race and cast their ballot based on their expected utility. In this case, they can strategically vote for a candidate whose ideological position is not next to their own positions. The details will be given in the next section.

In the standard literature about strategic voting with single vote, expected utility is defined as follows (for example Palfrey 1992):

$$EU_i(j) = \sum_{j \neq j'} p_{jj'}(U_i(j) - U_i(j')), \tag{5}$$

where $p_{jj'}$ is the expectation that the race is so close that a vote for a candidate j influences the election results. Accordingly, $p_{jj'} = 1$ when candidate j ties with another candidate j' in the same district or is one vote behind her. Otherwise, a vote for candidate j has no impact on the election result, so that $p_{jj'} = 0$.

In the spirit of adaptive actors, the simulation relaxes the assumptions based on the concept of the strict rationality. Accordingly, voters do not compare all pairs of candidates in the same district. Instead, they compare solely the incumbent denoted as 0 and the other candidates. The "adaptive" expected utility of a non-incumbent candidate j looks like the following:

$$AEU_i(j) = p_j(U_i(j) - U_i(0)) \tag{6}$$

The expected utility of incumbent is thus set to zero.

2.2.2 Expectation formation

It is tricky to endogenize the formation of expectation as subjective probability. We cannot take the approach of Savage (1954) who suggests to infer the subjective probability of a person using her material behavior. This "revealed expectation" approach is not feasible in this context since we do not have to deal with real data and we are going to "generate" behavior based on utility and expectation defined as a priori. Alternatively, we generate expectation from past experience of voters. In this regards, it is well known that the margin between the winner and the first loser at the last election predicts well the percentage of strategic voting, i.e., the smaller the margin, the more strategic voting occurs (for example Black 1978, 1980; Cain 1978; Cox 1997). Accordingly, we can assume that the last vote difference between the incumbent and candidate j' is used by voters in forming their subjective probability at election t.

$$p_{j't} = f(d_{j't-1}) \tag{7}$$

$$d_{j't-1} = e_{0't-1} - e_{j't-1}, \tag{8}$$

where $e_{0't-1}$ and $e_{j't-1}$ are the percentage of received votes of the incumbent and candidate j' respectively.

There still remains a problem of the form of f, i.e. how can we relate a past experience in the scale of vote margin with a subjective probability. Conventional analytical studies would assume the multi-nomial distribution. This paper, however, does not utilize this kind of relationship based on the multi-nomial distribution. As stated repeatedly above, voters in this paper are assumed to be limited in their capacity of information processing and only adaptively rational. The use of multi-nomial distribution is not consistent with this assumption.

Alternatively, Black (1978) proposes a simple linear relationship:

$$p_{j't} = 1 - d_{j't-1} \tag{9}$$

This function form, however, seems to be too simple to model the relationship between the vote margin and the subjective probability. Imagine following two cases: the vote margin of candidate A increases from 2% to 10% while candidate B's vote difference to the incumbent shifts from 72% to 80%. Intuitively, we would say that the subjective probability that A ends in a tie at the first place drops more drastically than that of B. According to (9), however, voters' expectation for both candidates drops by the same amount (8 percentage points). The function form of (9) does not correspond to our intuition saying that the marginal expectation also depends on the vote margin. For this reason, I use the following decreasing function:

$$p_{j't} = \frac{1}{\exp(a_t d_{j't-1})} \tag{10}$$

This function possesses some advantages. If $d_{j't-1} = 0$ then $p_{j't} = 1$. And $p_{j't}$ is monotonically decreasing along $d_{j't-1}$. In addition, this function has only one parameter which makes it easier to estimate in the simulation process and also to interpret results. Furthermore, $p_{j't}$ stays in the interval of $[0,1]$.

The next step is to determine the parameter a. To do this, I transform (10) as follows:

$$a_t = -\frac{\ln(p_{j't})}{d_{j't-1}} \tag{11}$$

It is obviously tautological to estimate a_t from a vote margin of $t-1$ and expectation at t since our goal is to estimate $p_{j't}$ using a_t. If we, however, see the situation as a repeated game and voters develop subjective expectation adaptively we can assume that the parameter a can be approximated through its past parameter value:

$$a_t \sim a_{t-1} = -\frac{\ln(p_{j't-1})}{d_{j't-2}} \tag{12}$$

We still need the real probability that the race is close at $t-1$, $p_{j't-1}$. While the objective probability that a voter is pivotal is de facto zero for a large electorate, we also relax this assumption. Instead of the one vote behind, we assume that the voter would recognize the impact of his vote as 95% if the vote share of candidate j' is within the 90% confidence interval of incumbent's (v_0). Otherwise, the level of impact is perceived as 1%.[4] Thus:

$$\widetilde{p}_{j't-1} = \begin{cases} 0.95 & \text{if} \quad d_{j't-1} \leq 1.64 \cdot \sqrt{\frac{v_{0t-1}(1-v_{0t-1})}{n}} \\ 0.01 & \text{if} \quad d_{j't-1} > 1.64 \cdot \sqrt{\frac{v_{0t-1}(1-v_{0t-1})}{n}} \end{cases} \tag{13}$$

By putting this $\widetilde{p}_{j't-1} \in \{0.01, 0.95\}$ and vote margin of individual candidates at $t-2$ into (12), $a_{j't-1}$ can be computed for individual non-incumbent candidates. For general parameter for all candidates, I take the average of individual parameters:

$$a_{t-1} = \overline{a}_{j't-1} \tag{14}$$

By putting this value as an approximate of a_t into (10), the adaptively approximated value of $p_{j't}$ can be obtained.

2.2.3 Interactive effects between ballots

The important feature of the current paper's model is interactive effects between two decision processes, that is, casting vote in the plurality and the PR tier. As discussed above, the model assumes three kinds of interactive effects:

(i) Preference formation (controlled by the parameter λ_1)

(ii) Decision making (controlled by the parameter λ_2)

(iii) Expectation formation (controlled by the parameter λ_3)

Interactive effects in preference formation between both tiers are incorporated via a perceived position of each party/candidate. If information from the PR tier is utilized by voters in building preference, one can assume that the perceived ideological position of a candidate is determined also by the position of the candidate's party, and vice versa. The degree and direction of interactive effects is parameterized via a weighting parameter, λ_1, which can take a value between -1 and 1. A positive value of λ_1 means

[4] Note that the values 95% and 1% for p and 90% for the confidence interval have nothing to do with each other. The value chosen here was found in some trial-and-error processes. If one takes a larger confidence interval, the situation in which a voter is pivot takes place more frequently. Correspondingly, estimated value for a tends to be smaller. The choice of likelihood value (here 95% and 1%) also influence estimation of a.

that the PR tier influences the plurality tier; a negative value means the reversed direction of influence. The absolute value of λ_1 corresponds to the weight of information from the influencing tier.

Accordingly, voter i's utility of party j based on perceived ideological distance is calculated as follows:

$$U_i(j) = \begin{cases} -(\omega_j - z_i)^2 & \text{if} \quad \lambda_1 \geq 0 \\ -|\lambda_1|(\omega_{j'} - z_i)^2 - (1 - |\lambda_1|)(\omega_j - z_i)^2 & \text{if} \quad \lambda_1 < 0 \end{cases} \tag{15}$$

Analogously, voter i's utility of candidate j' is calculated as follows:

$$U_i(j') = \begin{cases} U_i(j') = -(1 - \lambda_1)(\omega_{j'} - z_i)^2 - \lambda_1(\omega_j - z_i)^2 & \text{if} \quad \lambda_1 > 0 \\ U_i(j') = -(\omega_{j'} - z_i)^2 & \text{if} \quad \lambda_1 \leq 0 \end{cases} \tag{16}$$

Note that if $\lambda_1 = 0$ this kind of interactive effects does not operate between both tiers.

Interactive effects in vote decision are modeled in a similar way to those in preference buildings. Interactive effects are parameterized via $\lambda_2 \in [-1;1]$. Like λ_1, a positive value of λ_2 means an influence of vote decision in the PR tier upon that in the plurality tier; a negative value means the direction of influence is reversed. While the absolute value of λ_2 also corresponds to the degree of influence (which differs from the perceived ideological distance), vote decision cannot be modeled as a weighted mean. Instead, interactive effects in vote decision are modeled in a probabilistic way. Assuming that voter i would decide for party G and district candidate S' according to (expected) utility defined above. Denote the candidate of party G in the district of voter i G' and candidate S''s party S. Then:

$$\left.\begin{array}{lcl} \text{Prob}(i \text{ votes for } G) & = & 1 \\ \text{Prob}(i \text{ votes for } G') & = & \lambda_2 \\ \text{Prob}(i \text{ votes for } S') & = & 1 - \lambda_2 \end{array}\right\} \quad \text{if} \quad \lambda_2 \geq 0 \tag{17}$$

$$\left.\begin{array}{lcl} \text{Prob}(i \text{ votes for } G) & = & 1 - |\lambda_2| \\ \text{Prob}(i \text{ votes for } S) & = & |\lambda_2| \\ \text{Prob}(i \text{ votes for } S') & = & 1 \end{array}\right\} \quad \text{if} \quad \lambda_2 < 0 \tag{18}$$

Note that $\lambda_2 = 0$ also means no interactive effects between both tiers.[5]

Unlike the other two kinds of interactive effects, those in expectation formation have only one direction. That is, information from the PR tier is used in the plurality tier while the reversed flow of information is not modeled. This is because in the vote decision process in the PR tier expectation plays no role as defined above. Correspondingly, the parameter λ_3 for this kind of interactive effects takes only a value between 0 and 1. Like λ_1, the value corresponds to the weight of influencing information from the PR tier. As defined above, expectation is formed based on the vote margin at the last election, $d_{j't-1}$ (see p. 277). Therefore, if $\lambda_3 > 0$, not only candidate j''s vote margin but also that of her party j is considered using following equation:

$$d_{j't-1} = (1 - \lambda_3)(e_{0't-1} - e_{j't-1}) + \lambda_3(e_{0t-1} - e_{jt-1}) \tag{19}$$

[5] Concretely in the program, a random number is generated from a uniform distribution between 0 and 1 for each voter and compared with $|\lambda_2|$.

2.3 Parties'/ candidates' actions

As discussed above, the model also incorporates the (strategic) behavior of political elites. Following Laver (2005), four types of parties are generated: Hunter, Predator, Aggregator and Sticker.

Hunter is oriented toward the support for itself. If the support increased after a move, Hunter makes the same move again. Otherwise, Hunter randomly selects a move in an opposite direction of the last move. This behavior can be formally expressed:

$$\Delta\omega_{jt} = \Delta\omega_{jt-1} \qquad \text{if} \quad v_{jt-1} > v_{jt-2} \tag{20}$$

$$\left.\begin{array}{l} \Delta\omega_{j1t} \sim \mathrm{Unif}\big(0, m\cdot(-\mathrm{sign}\Delta\omega_{j1t-1})\big) \\ \Delta\omega_{j2t} \sim \mathrm{Unif}\big(0, m\cdot(-\mathrm{sign}\Delta\omega_{j2t-1})\big) \end{array}\right\} \quad \text{if} \quad v_{jt-1} \leq v_{jt-2}, \tag{21}$$

where v_{jt} is the votes which party j gains at election t, and m is the maximal distance of a move on a dimension in a single cycle. This is set as 2 throughout the following simulation runs.

Predator observes only the position of the most successful party and moves toward it. If the party itself is the largest party, it does not change its position. Formally:

$$\Delta\omega_{jt} = \begin{cases} m\cdot\mathrm{sign}(\omega_{0t-1} - \omega_{jt-1}) & \text{if} \quad \max_{j'\neq j} v_{j't-1} > v_{jt-1} \\ 0 & \text{if} \quad \max_{j'\neq j} v_{j't-1} \leq v_{jt-1}, \end{cases} \tag{22}$$

where ω_{0t} is the position of the largest parties at election t.

Aggregator cares about the distribution of the ideal position of his supporters and moves to their average position. Formally:

$$\omega_{jt} = \bar{z}_i \text{ for } i : v_{it-1} = j, \tag{23}$$

where v_{it} is voter i's decision at election t.

Sticker never changes its position in the ideological space. Accordingly:

$$\Delta\omega_{jt} = 0 \tag{24}$$

The choice of these four party types of Laver is justified by the following reasons: first, every type of parties behaves based on limited information which makes the simulation more realistic. Second, Hunter and Predator prevent reaching an equilibrium which has never existed at real elections. Furthermore, the first two parties can be classified as competition oriented or vote maximizing parties while the other two parties are policy-driven. This enables one to test the hypothesis if these two kinds of motivations affect the linkage. The number of each type of parties is randomly selected from $\{0,1,2,3\}$. This can generate a party system with minimum zero and maximum 12 parties. Since a party competition with less than two parties is not relevant and simulations with too many parties are time-consuming, the total number of parties is limited between 3 and 8.

Differently from Laver (2005), not only political parties but also district candidates compete against each other in 30 districts. In this regard, candidates for each party are generated in each district and take the strategy of their affiliated party. Therefore, the combination of strategies in each district is identical to that of the national party competition. Each party and its candidates are assigned to a randomly selected starting point in the ideological space. Once a simulation starts, however, candidate positions can deviate from their affiliated party. Each district candidate — except for those with the Sticker-strategy — takes position in reaction to its district specific circumstance. Its positioning can, however, also be influenced by the PR tier via different kinds of interactive effects in vote decision process. If election results also depend on the positioning of political parties due to interactive effects, candidates have to react correspondingly. Each candidate moves in each cycle while each party moves in every tenth cycles. This is based on the assumption that district candidates can change their position easier than political parties as collective actors.

2.4 Collecting data

The model of this paper contains a large set of parameters. Since it would take a long time to investigate every point in such a high-dimensional parameter space, 1,000 simulation runs are conducted with randomly drawn parameters. The possible value for each parameter is summarized in the following:

(i) Homogeneity $h \in \{-2,-1,0,1,2\}$

(ii) No. of strategies $\in \{0,1,2,3\}$, whereby the total no. of parties $\in \{3,4,5,6,7,8\}$

(iii) Actualizing expectation $\in \{0,0.25,0.5,0.75,1\}$

(iv) Interactive effects in preference building $\lambda_1 \in \{0,0.25,0.5,0.75,1\}$

(v) Interactive effects in vote choice $\lambda_2 \in \{-1,-0.75,-0.5,\dots,1\}$

(vi) Interactive effects in expectation formation $\lambda_3 \in \{-1,-0.75,-0.5,\dots,1\}$

As mentioned above, the starting positions of parties and candidates are randomly set in the ideological space. Also voter positions are randomly determined according to the homogeneity above.

A single simulation run longs for 500 cycles. That is, voters cast their ballot 500 times. Of 500 cycles in each simulation, the first 300 cycles are discarded as so-called burn-in since cycles in earlier phases are strongly conditioned by the starting values.[6] Furthermore, voters as well as candidates/parties need some cycles to learn adaptively to follow their strategies. After this burn-in, that is, after the results are more or less independent of the initial condition, each 5 of 200 cycles are collected. This is because, if one collects data from all cycles, each data point is not independent of each other. That is, the data set suffers a time-series problem which would make the statistical inference more complicated. Consequently, 40 data points are collected for a single simulation run. This is repeated 1,000 times which generate 40,000 data points.

The data collected in this way are suited to the conventional statistical techniques since the frequentist view postulates that the analyzed data are only one realization of infinitely repeatable data collection processes. This is, however, hardly realistic for

[6] Pilot simulation runs show that the results obtained between 300 and 500 cycles have no clear deviation from those in further cycles.

the data collected in conventional methods. The simulation technique is one of a few methods compatible with the frequentist assumption above. Furthermore, there is no risk of multi-collinearity in multivariate analysis since the random draws of parameters were conducted independently of each other and there is no correlation between parameters.

3. Evaluation of simulation results

In this section, I draw implications from the computational model which was set up above. The model is not analytically solved, but it generates data via multiple simulation runs from which implications are drawn. After the relationships are described in a bivariate manner, the second subsection analyzes the factor of high-level linkage by using multivariate statistical models.

3.1 Bivariate observation of results

3.1.1 Linkage under simple plurality rule

Before we observe the simulation results under mixed-member electoral systems, those under simple plurality rule are described as a benchmark. For this purpose, 500 simulation runs were conducted with the following parameters fixed to 0:

(i) Interactive effects concerning perceived distance (λ_1)
(ii) Interactive effects in vote decision (λ_2)
(iii) Interactive effects in expectation formation (λ_3)

By setting these parameters to 0, we can obtain simulation results of party competitions and vote decisions which independently take place in plurality and PR.

Figure 3 presents the frequency of dominated districts under different parameter conditions. An interesting result is that districts' homogeneity concerning voter distributions has no clear-cut effect on the level of linkage. In the line of arguments of classical political sociology, the number of cleavage in the society should positively influence the number of parties. Accordingly, the more cleavages exist in a society, the more heterogeneous voters' preference profiles are across districts. This should lead to success of some regional parties in certain districts, which contributes to a multi-party system at the national level. Consistent with this expectation, increasing homogeneity of voter distributions tends to show a higher level of linkage (more dominated districts); however, its magnitude of impact is somewhat moderate.

In contrast, the number of parties demonstrates more clear-cut effect upon the level of linkage. The more the parties are in the race, the lower the level of linkage. One might argue that it is self-evident that in a party competition with fewer parties, it is more likely that a same set of two parties compete in more districts. However, the degree of the impact is still clear if one compares with simulated frequencies of dominant districts where election results are randomly generated.[7] This is primarily due to the party/candidate strategies.

[7] This result is available from the author upon request.

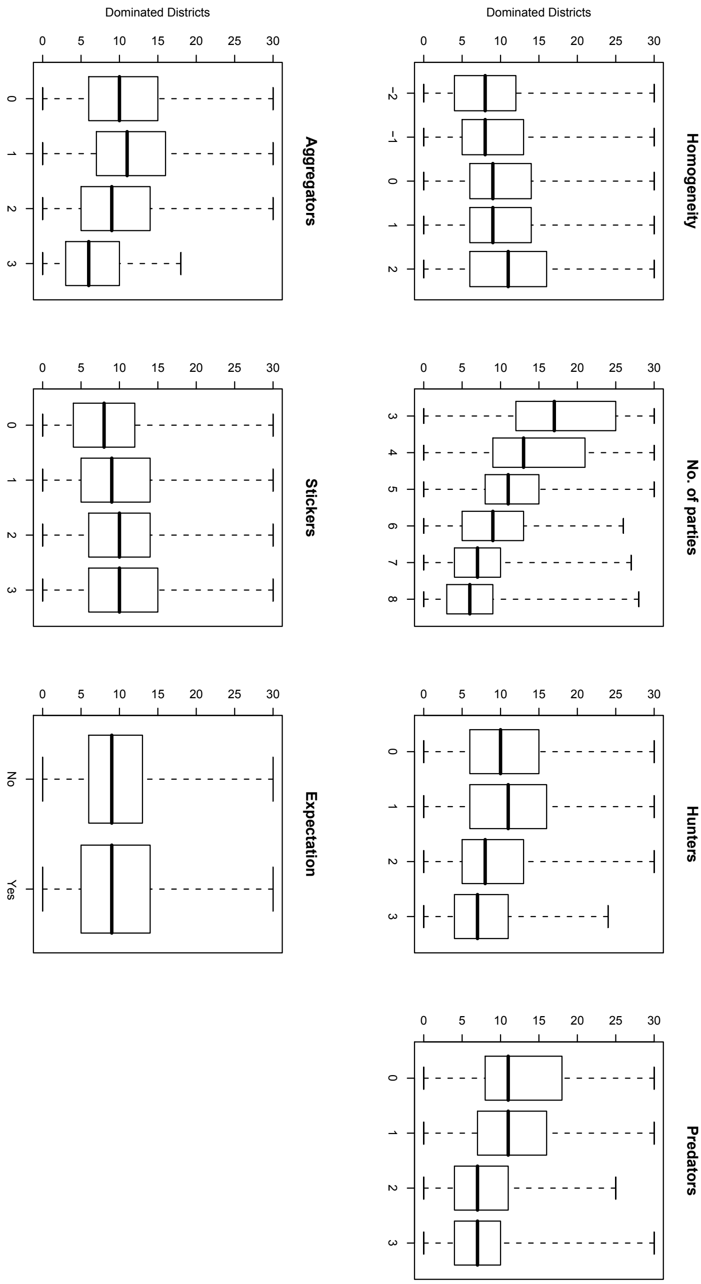


Figure 3. Frequency of dominated districts in plurality

As introduced in the set-up of the computational model, all candidates of a same party share the same type of strategy. Each strategy is different in its successfulness in the party competition. Therefore, if candidates of a part of parties take a successful strategy they can be successful across the districts. This leads to a higher level of linkage. This can be supported partly by the result that the presence of one Hunter, Aggregator or Sticker leads to higher-level linkage than the absence of the same strategy. An increasing number of Hunters, Predators and Aggregators lead to a decreasing level of linkage. This is conceivable since votes gained by two or more candidates with a same strategy should be similar so that it can differ among districts which candidates rank first and second. As a result, the level of linkage should be lowered.

Interestingly, this does not hold for Stickers. That is, the increasing number of Stickers raises the level of linkage. There are two reasons for this. First, Sticker is the least successful strategy. Therefore, the increase in their number has a less direct impact on the level of linkage since they have less chance to be one of the two largest parties at the national level. However, the increasing number of Stickers also means that the decreasing number of other strategies, and in particular, the decreasing probability that two candidates with the same successful strategy compete with each other. Due to this, the level of linkage can rise. Second — and this is less likely — if Stickers are in the two main parties at the national level, the level of linkage can also be high. Since the position of candidates with Sticker-strategy is fully identical across the districts, the candidates from the same party are likely to be consistently successful.

Whereas the level of linkage can be attributed to the number of competing parties, the use of expected utility seems to have no clear impact upon linkage (the lower-right panel of Figure 3 shows). Therefore, we can conclude that the strategic voting based on the expected utility model can have no consequence for the linkage process under a simple plurality system. This can be different under mixed systems. To see this, we observe in the next section the data generated under mixed systems with various interactive effects between both ballots.

3.1.2 Linkage under mixed systems

Figure 4 presents distributions of dominated districts' frequency at different degrees of interactive effects.

As expected above, the higher degree of interactive effects in expectation formation operates in favor of dominance by the two largest national-level parties. That is, if voters utilize more national-level PR results to form expectation, candidates from the two same parties rank first and second across districts.

Each of the further two kinds of interactive effects also has an impact upon the frequency of dominated districts. The common tendency concerning both kinds of interactive effects is that the influence of the plurality tier upon the PR tier leads to a higher frequency of dominated districts, and vice versa. In terms of interactive effects in vote decision (λ_2), this may seem to be in line with Duverger's conjecture, or the spill-over effect. He expected to see in mixed systems that the PR vote should be influenced by plurality vote, which should result in a national-level bipartism. Note, however, that Duverger did not account for the linkage process. In contrast, the results

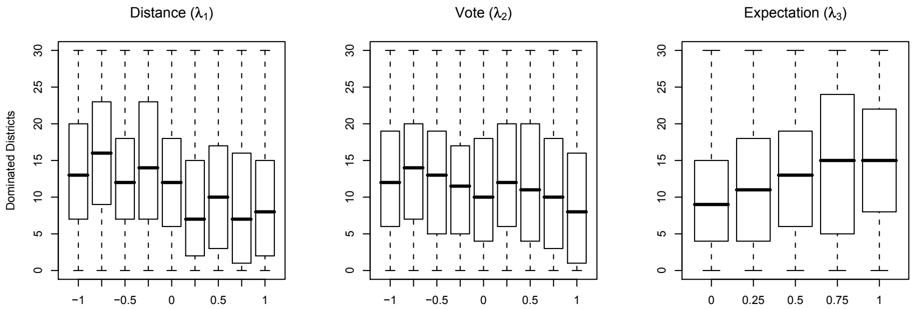


Figure 4. Frequency of dominated districts at different degrees of interactive effects under mixed systems

here demonstrate that interactive effects are causing this ignored element of Duverger's Law. Interactive effects concerning perceived ideological position λ_1 also have the same kind of impact upon the level of linkage even if it is less clear than that of λ_2.

Now, we are turning to the effect of further factors summarized in Figure 5. As discussed earlier, high-level linkage is expected to occur within homogeneous voter distributions across districts. The upper-left panel supports this expectation. The more homogeneous voter distributions are, the more likely the districts are dominated by the largest national-level parties. If one compares with the result under simple plurality (Figure 3), the frequency of dominated districts at each level of homogeneity is higher under a mixed system than under plurality. This is due to the existence of interactive effects, in particular in terms of expectation formation.

The number of parties also has an expected effect. That is, the more parties that are in competition, the less frequent dominated districts are. The degree is higher than that under simple plurality (Figure 3). Similar to the homogeneity, interactive effects under mixed systems operate in favor of a higher level of linkage.

Observation of the influence of a number of individual strategies upon the level of linkage provides similar results which have been observed under simple plurality (Figure 3). However, the variance of the linkage level is greater here for each number of strategies. This means that its effect is less clear under mixed systems than under simple plurality. The added variances are attributed to the introduction of interactive effects under mixed systems. As observed above, interactive effects also have non-trivial impacts upon the linkage level, which made the effect of further factors — here the number of individual strategies — less clear.

A reversed result can be observed in the effect of the use of expected utility in vote decision. According to the down-right panel of Figure 5, the use of expected utility has a certain impact in favor of high-level linkage. It has, in contrast, no impact under plurality (Figure 3). The elements are integrated only in the simulation under mixed systems, i.e., interactive effects play an intervening role and make the difference between simulations with and without the use of expected utility.

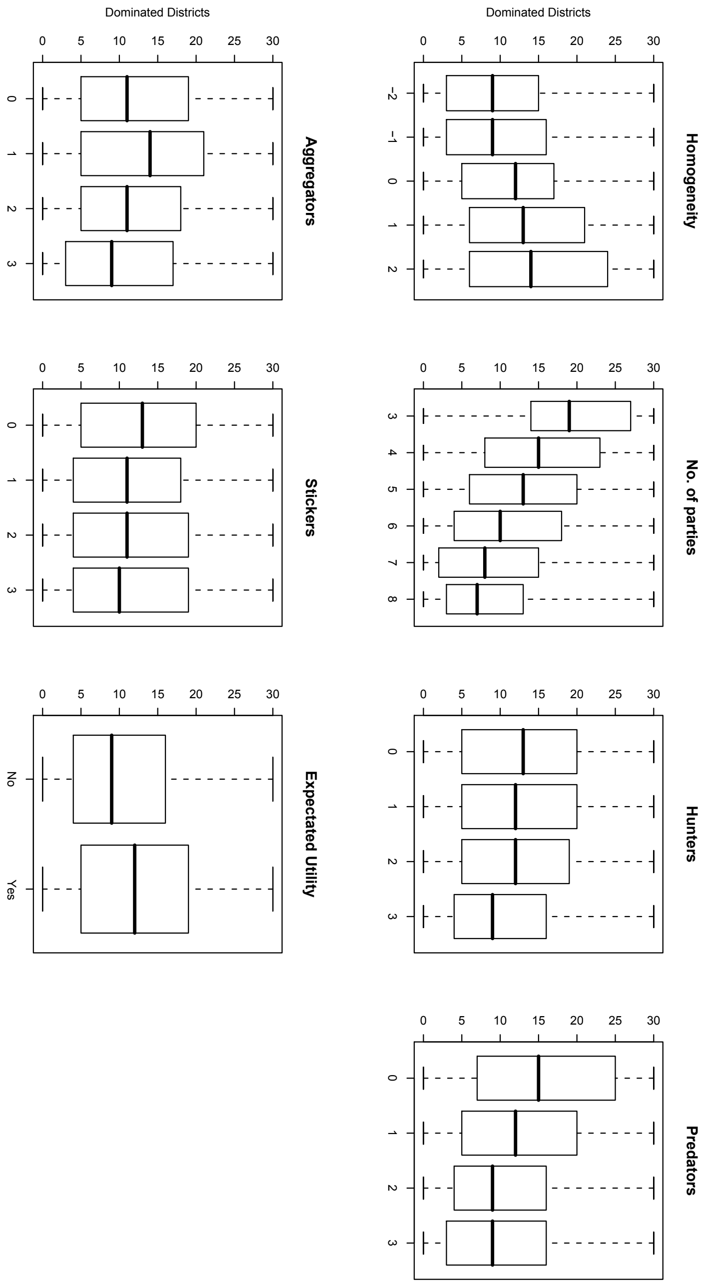


Figure 5. Frequency of dominated districts at different degrees of further factors under mixed systems

3.2 Multivariate analysis of the linkage process

In the last section, we predominantly observed bivariate relationships between frequency of dominated districts and various factors. As I mentioned in some places above, however, different levels of linkage are the products of certain interactions between diverse factors. I move, therefore, to multivariate analysis to observe the net impact of each factor.

In the following analysis, statistical estimations are conducted via a maximum likelihood approach since simulated data can be assumed to stem from repeatable observations of an identical data generating process. Furthermore, the simulated data are suited for the multivariate analysis since parameters are generated randomly and, thus, independent of each other. Consequently, there is no risk of multicolinearity among independent variables. Note that the amount of observations analyzed here is quite extensive. As noted above, 1,000 simulation runs with randomly generated parameters for each were repeated. For each run, 40 cycles were collected so that we have in total 40,000 cycles. Furthermore, the unit of observation is not a single cycle, but each 30 districts in each cycle. Therefore, we have a total of 1,200,000 observations in data sets. Since this extensive number of observations makes estimated standard errors (SE in tables) quite small, I have to note that the stated SE in following results are less conclusive. Therefore, I will concentrate on the interpretation of estimated coefficients and, needless to say, mention their significance tests.

To observe which factor determines the level of linkage, the following model is estimated.

$$\begin{aligned} d_{rck} &= \mathrm{Bern}(\pi_r) \\ \mathrm{logit}(\pi_r) &= \beta X_r, \end{aligned} \tag{25}$$

where d_{rck} is the dominance of district k at c-th cycle in r-th simulation run. It is coded one if the district is competed among candidates of two largest parties at the national level, and otherwise zero. X_r is the vector of parameters of r-th simulation run.

As discussed above, I expect that the increasing use of national-level PR tier results in expectation formation (higher λ_3 parameter) boosts the level of linkage. Also, the absence of the PR tier's influence in terms of preference formation ($\lambda_1 \leq 0$) is expected to contribute indirectly to higher-level linkage since each candidate can better adapt to each district specific voter distribution. In contrast, I have no expectation concerning interactive effects in vote decision λ_1. In addition to interactive effects between both tiers, I also expect that heterogeneity of voter distributions across districts and the number of each candidate/party strategies affect the level of linkage. High-level heterogeneity of voter distributions is expected to lead to a higher level of linkage. If preference profiles of each district are more similar to each other, it is more likely that candidates from the same set of two parties compete in every district. For each candidate strategy, one can expect that the increasing number would reduce the level of linkage.

Table 1. Estimation results of the logit model (dependent variable: dominance of districts)

	Model 1		Model 2	
	Coefficient	S.E.	Coefficient	S.E.
Intercept	1.0227	0.0083	1.0486	0.0083
Interactive effects in terms of				
Perceived Position (λ_1)	−0.4817	0.0030	−0.1908	0.0046
Vote Decision (λ_2)	−0.2517	0.0031	−0.1128	0.0046
Expectation (λ_3)	0.8976	0.0057	0.8726	0.0058
$\lambda_1 \times \lambda_3$			−0.6961	0.0083
$\lambda_2 \times \lambda_3$			−0.3638	0.0084
Actualization of expectation	−0.2361	0.0061	−0.2185	0.0061
Homogeneity	0.2078	0.0014	0.2128	0.0014
No. Hunters	−0.2500	0.0020	−0.2469	0.0020
No. Predators	−0.4380	0.0020	−0.4460	0.0020
No. Aggregators	−0.2971	0.0021	−0.3017	0.0021
No. Stickers	−0.1815	0.0019	−0.1810	0.0020

Note: $N = 1,200,000$.

Estimation results of Model 1 in Table 1 show, as expected, that interactive effects between two ballots facilitate linkage of district races. Increasing interactive effects in the expected direction between two ballots causes higher-level linkage. While I expected in λ_3 the most direct effect, the other two kind of interactive effects also seem to affect the level of linkage to a high extent. These effects are, however, highly intermediated by λ_3. If one adds to the model interaction effects of $\lambda_1 \times \lambda_3$ and $\lambda_1 \times \lambda_2$ (Model 2), the coefficients of the main effect of λ_1 and λ_2 were estimated as being much smaller. That is, these parameters boost the linkage level to a much lesser extent without λ_3. Furthermore, the interaction effects are even larger than the main effects of λ_1 and λ_2 estimated in Model 1. By contrast, the main estimated effect of λ_3 remains on the same level of Model 1. These results support that interactive effects in expectation formation are the most direct and important factor for linkage among various kinds of interactive effects between two ballots.

Turning to the other factors, homogeneity of districts regarding distribution of voter's ideal points catapults the linkage process. The estimated coefficient is smaller than those for interactive effects between two ballots. However, the parameter for the homogeneity has a larger range between -2 and 2 than other parameters. Therefore, the effect of homogeneity upon the linkage level should be interpreted correspondingly. Figure 6 clarifies the effect of homogeneity and interactive effects in expectation formation. The two lines show forecasted linkage levels as dependent on the level of λ_3 according to Model 1 in Table 1. For the above line, the homogeneity parameter h was set to 2 and for the below line $h = -2$. The difference at the two lines' level clearly demonstrates that homogeneity is a quite important factor for linkage of district races. While h was bounded between -2 and 2 in the simulation runs, one can extend the range of this parameter. Therefore, one can even expect a higher linkage level for a higher h value and vice versa. This may appear to support the notion of political

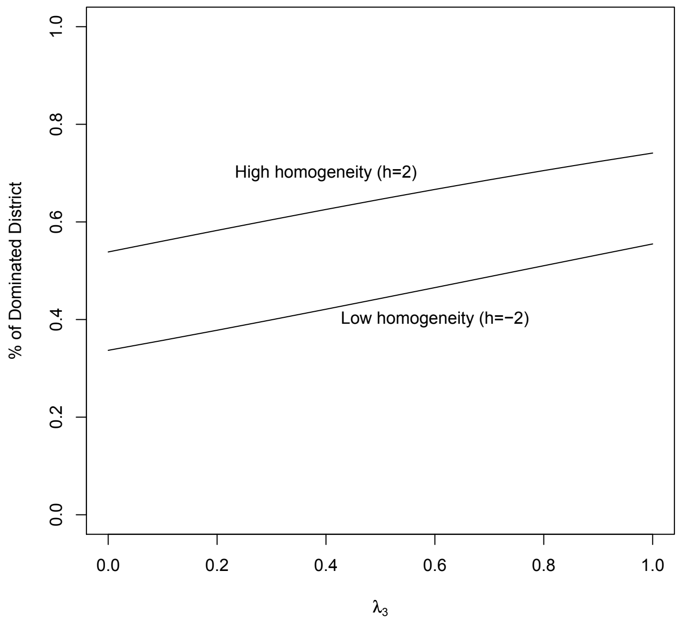


Figure 6. Forecasted percentages of dominated districts based on the simulation results

sociologists, i.e. the large impact that social contexts of political competition have on the party system. However, a more precise notion based on the results here would be to use the term "a large impact" instead of "the large impact". For interactive effects between two ballots in forming expectation and vote decision also boost the percentage of dominated districts. For a certain high level of homogeneity ($h = 2$), the levels of linkage yielded with the full use of the PR tier information in expectation formation ($\lambda_3 = 1$) and without it make up a 17%-points difference. This seems to explain why West Germany, New Zealand and Japan demonstrate increases of the linkage level in such short time periods. In contrast, it is less conceivable that for such short time spans those voter distributions across districts became so homogeneous that high-level linkage could be established.

Concerning the number of candidate strategies, an increase in number generally lowers the percentage of dominated districts. The extent is, however, different among candidate/party strategies. The increasing number of Predators shows the strongest tendency toward the absence of linkage and the number of Stickers has the least effect on the linkage level. These are consistent with the results in the bivariate analysis above (Figure 5). Note also here that the effect of the number of strategies can be larger, as was the case with the homogeneity parameter, since that parameter has a larger range of values. This, however, does not mean that the number of strategies is the dominant factor for linkage but one of them, including interactive effects and homogeneity.

4. Discussion

Beginning with an empirical puzzle concerning the West German party system, this paper proposed an interactive mechanism model of mixed-member systems which operates in favor of high-level linkage. The results of the computational model demonstrated that mixed systems have independent effects of further competing factors for high-level linkage, like social homogeneity and party competition.

One might ask whether mixed systems in the other countries also operate in a similar way to that of West Germany. At present, there are three political systems in the world which have experienced mixed systems at least for 10 years: New Zealand, Japan and East Germany. Among them, New Zealand and Japan show a rapid increase of the linkage level after the introduction of their mixed systems. Social homogeneity is less conceivable as an explaining factor for such a rapid change of the linkage level. Party competition in both countries has been relatively stable as well. Therefore, the interactive mechanism suggested in this paper seems to be most appropriate one to explain the short-termed linkage trend.

In contrast to New Zealand and Japan, East Germany shows no increasing level of linkage. Also here, we should not attribute it to social homogeneity or party competition due to its short termed change, in particular between the 2002 and 2005 election. One possible hypothesis based on this paper's model would be that East Germany has some special conditions which disrupt interactive effects in expectation formation under mixed systems. One such possible condition could be the deviating developments of East German state-level parliaments where the ex-communist PDS occasionally becomes the second-largest party. At the time of the 2005 federal election, this was the case in four of six East German state parliaments.

References

Black, J. H. (1978). The Multicandidate Calculus of Voting: Application to Canadian Federal Elections. *American Journal of Political Science*, 22(3), 609–638.

Black, J. H. (1980). The Probability-Choice Perspective in Voter Decision Making Models. *Public Choice*, 35(5), 565–574.

Cain, B. (1978). Strategic Voting in Britain. *American Journal of Political Science*, 22(3), 639–655.

Cox, G. W. (1997). *Making Votes Count: Strategic Coordination in the Worlds Electoral Systems*. Cambridge, Cambridge University Press.

Ferrara, F., Herron, E. S. and Nishikawa, M. (2005). *Mixed Electoral Systems: Contamination and Its Consequences*. New York, Palgrave.

Laver, M. (2005). Policy and the Dynamics of Political Competition. *American Political Science Review*, 99(2), 263–81.

Lijphart, A. (1999). *Patterns of Democracy: Government Forms and Performance in Thirty-Six Countries*. New Haven, Yale University Press.

Palfrey, T. R. (1992). A Mathematical Proof of Duvergers Law. In Ordeshook, P. C. (ed.), M*odels of Strategic Choice in Politics*. Ann Arbor, University of Michigan Press, 69–91.

Savage, L. J. (1954). *The Foundations of Statistics*. New York, Wiley.

Shikano, S. (2007). Interactive Mechanism of Mixed-member Electoral Systems: A Theory-driven Comparative Analysis via Computational Modeling and Bayesian Statistics. University of Mannheim, Habilitation Thesis.

Shugart, M. S. (2001). “Extreme Electoral Systems and the Appeal of the Mixed-Member Alternative. In Soberg, M. S. and Wattenberg, M. P. (eds.), *Mixed-member electoral systems: the best of both worlds?* Oxford, Oxford University Press, 25–51.