Toward theory-driven comparative analysis via computational modeling and Bayesian statistics

Susumu Shikano
University of Mannheim

Tagung der Sektion „Vergleichende Politikwissenschaft"
10. Nov. 2007

Abstract

Many comparativists are often confronted with at least two problems concerning data availability: small n and absence of individual-level data. As solution, this paper introduces an approach which combines Bayesian statistics and computational modeling. It is exemplified by an analysis of the effect of mixed electoral systems boosting the linkage process of plurality races. The central hypothesis is that voters tend to use more national-level information to form expectation concerning their district races under mixed systems. The suggested approach provides results supporting the hypothesis without using any individual-level measures.

1 Introduction

Many comparativists are often confronted with at least two problems concerning data availability: small n and absence of individual-level data. The first problem is the limited number of observations. Differently from other subfield of political science like political sociology, American politics etc., the comparative political studies have less to do with survey data. Instead, aggregate-level data, or mostly, country-level data are typical basis for comparativists. Since there are a limited number of countries which deliver data for comparativists' purpose, most datasets in the comparative political studies have a small n. This is problematic in the use of statistical methods based on conventional OLS or maximum-likelihood which rely on some asymptotic properties. The other problem with which comparativists are confronted is the absence of comparable individual-level data. While such kind of data are essential for testing theoretical models based on individual actions, it is not easy to collect them in a comparative way. The standard method for systematic collection of individual-level data is the random sampling survey. Since a survey requires scholars a large amount of cost, time and energy, it is not easy to coordinate international scholars in favor of a comparative survey design. For this reason, there are only few individual-level datasets available for comparative purposes. Therefore, one has to conduct some kinds of latent variable models in order to test some individual-level hypothesis. For this purpose, however, the conventional statistical techniques are often overloaded.

As solution for the problems above, this paper suggests to use Bayesian statistics in combination with computational modeling. In comparison with conventional statistics, Bayesian statistics has the advantage in coping small-N data and latent variables. This is possible since Bayesian statistics complements the information from empirical data with a priori information. A priori information can be simply a subjective belief possessed by a researcher while it has to be formally expressed in a form of distribution. In order to achieve such a priori information this paper recommends formal modeling since it is more transparent for wider range of researchers. More specifically, this paper advocates the use of agent-based modeling, a variant of computational modeling, since this technique allows researchers to derive probabilistic implications from more flexible models based on the interaction between individual actors.

This paper introduces an approach which combines Bayesian statistics and computational modeling. It is exemplified by an analysis of effects of electoral systems. More specifically, the effect of mixed electoral systems on the linkage process of plurality races is at stake. In this example, the first problem (small n) is not relevant, but only the second one (absence of individual-level data). However, this approach can also be applied to the first problem.

2 Problem: Linkage of plurality races

Duverger's Law is one of the most famous laws in social sciences. In this law, the relationship between electoral systems and party systems is at stake (Duverger, 1954):

> The simple-majority single-ballot system favors the two-party system. Of all the hypotheses that have been defined in this book, this approaches the most nearly perhaps to a true sociological law. (217)

According to Duverger, voters have aversion against wasting their votes. Therefore, the adherents of candidates who have no chance to win the race tend to cast their ballot for a less evil candidate under those with a significant winning chance. While this so-called “wasted-voting”-logic and resulting Duvergerian equilibrium in favor of two leading candidates can be founded by the expected utility model (McKelvey and Ordeshook, 1972; Black, 1978), it has been confronted with at least two critiques which are also consistent with the expected utility model. One is the possible existence of Non-Duvergerian equilibrium and the other is the linkage problem.

While under Duvergerian equilibrium voters clearly know which two candidates are leading the race, it can also happen that voters are not sure which candidate concurs against the leading one. In the latter case, votes are not concentrated on the top two candidates. In the corresponding equilibrium situation, the first and second loser (and eventually further losers) receive the same amount of votes. This is Non-Duvergerian equilibrium (Myerson and Weber, 1993; Cox, 1997; Fey, 1997).

The second critique concerning linkage can be well described with Canada. Canada is one of the well-known exceptions to Duverger’s Law, which is preferably mentioned in various critiques against Duverger’s Law. Whereas Canada has adopted plurality in single member districts (SMD) for a quite long period, its party system is not characterized by bipartism. Duverger also was conscious of the existence of this deviating case:

> Non the less, the increased centralization of organization within the parties and the consequent tendency to see political problems from the wider, national standpoint tend of themselves to project on to the entire country the localized two-party system brought about by the ballot procedure: however, the true effect of the simple-majority system is limited to local bi-partism.

Consequently, we should differentiate district- and national-level bipartism and conclude that Duverger’s Law cannot be applied at the national level. Assuming that there are K districts and Duverger’s Law is operating at the district level. In each district two candidates of different sets of two parties can contest against each other so that we can observe a $2K$ party system at national level. Having stated that, we are still interested in the formation process of national bipartism from local bipartism. Cox uses the term linkage to address the issue why each district race is connected (or not) with each other through two common parties (1997: 181ff).

Which factor facilitates or prevents the linkage process? Duverger argued that linkage is established through the increased centralization of party organization. Chhibber and Kollman

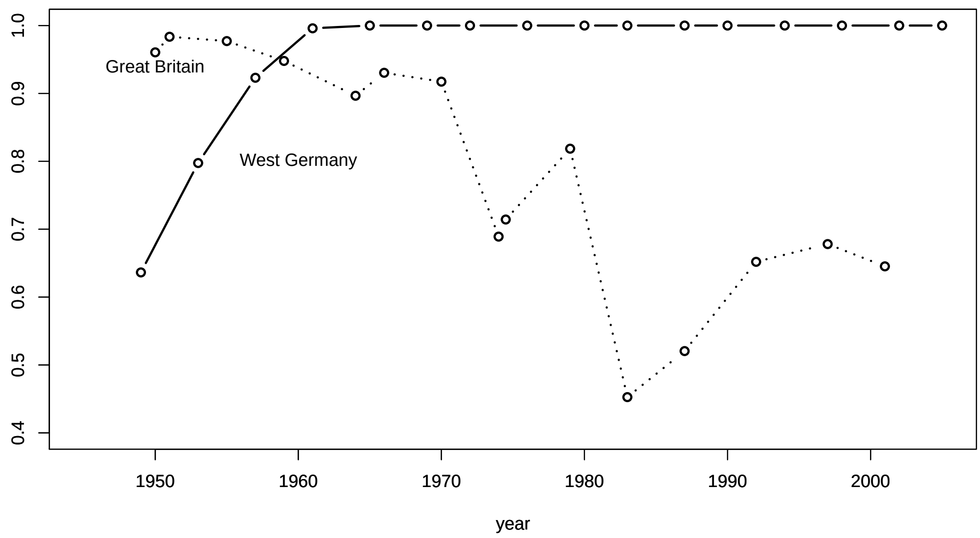
(2004) also argue that importance of the national government as exogenous factor explains the degree of linkage. The more important the national government is, the more national parties gain votes and the higher the level of linkage is, and vice versa. In their comparative study using data from the United States, Canada, Great Britain and India, they succeed to explain the temporal change of the party system at the national level by the political and economic centralization of each country.

In the traditional line of studies about linkage, Sartori (1968) calls attention to structured party systems which facilitate linkage between districts. With "structured party system", he means a nationwide organization and ideological reputation. Cox criticizes "Sartori begin[s his] argument by assuming that parties have nationalized" (1997: 185). Prior to Sartori, Leys (1959) argues that voters strategically cast their ballot in favor of the largest two national parties under single member plurality systems. Here, we have to recall that strategic voting matters in realization of bipartism whereby vote decision in this context has two components, utility and expectation. For realization of district-level bipartism rule the expectation is much more important than the utility. This should also be the case in the realization of bipartism at the national level through linkage of district races. The problem of Leys' argument is, as Cox (1997) noted, that Leys clearly assumes that voters are nationally oriented. So, we have to ask what causes the orientation at national party system. Cox himself suggests five goals of parties that might facilitate linkage: achieving a national policy, winning a presidency, winning a premiership, gaining upper tier seats, and gaining campaign finance (1997: 186ff).

Besides to the traditional explanations above, this paper suggests that mixed-member electoral system could deliver another possible cause for the voter orientation at national party system. Mixed systems combine plurality in SMD with proportional representation (PR) whereby each voter has two ballots for both tier. Under this kind of electoral systems, the visibility of PR-tier facilitates the voter orientation at national party system and suppresses local district races deviating from the national one. The visibility of the two national large parties also boosts the advantage of their candidates so that they enjoy a significantly large margin to the other candidates. As a result, a high-level linkage can be expected under mixed systems than under simple plurality.

We can easily observe some kinds of evidence for it. The solid line in Figure 1 shows the percentage of West German districts where candidates of both large parties ranked first and second. I call hereafter this kind of districts competed between two main national parties as "dominated". Accordingly, at the first general election 1949 only ca. 60% of districts were contested between the candidates of CDU/CSU and SPD. Thereafter, this percentage increased rapidly. In the 1961 general election, the SPD and CDU/CSU-candidates fought all districts except one (Crailsheim). In Crailsheim, the FDP candidate came second after the successful CDU candidate and the SPD candidate came third. At the 1965 election, however, Crailsheim saw the SPD candidate beat the FDP candidate into third place and since then all districts in West Germany have been dominated by the candidates of the two main parties without exception. To compare this with the United Kingdom, whilst in the 1950s almost all British districts saw Labour and Conservative candidates take either first or second place, this was never true of 100% of all districts. In the course of time this has actually decreased, and in the 1980s it even sank below 50%. 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}. From these results, it is obvious that West Germany's mixed system leads to a higher degree of linkage than British simple plurality.[1]

Figure 1: Percentage of districts where the two main parties at the national level came first and second in West Germany and the United Kingdom

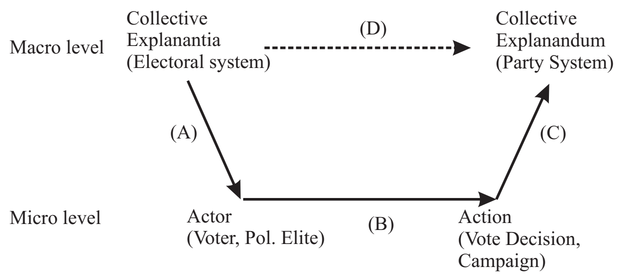


West Berlin is not included in the figure for West Germany. North Ireland is not included in the figure for the United Kingdom.

[1]The measurement of linkage used here is unconventional one. Contrary, the widely used measurement of linkage is based on the difference between the effective number of parties on national level and the average effective number of candidates in districts (Chhibber and Kollman, 1998). This measurement, however, is problematic for this paper's purpose since the measurement tells nothing about in which districts party systems are (not) linked with the national party systems. This differentiation is, as shown later, important for the implications of this paper's theoretical model.

The evidence shown above, however, confirms only the macro-level relationship between mixed systems and high-level linkage of plurality races. Figure 2 visualizes the basic logic of explanation of methodological individualism. Accordingly, Figure 1 shows only the relationship (D), but says nothing about the micro-level mechanism assumed in the explanation above (A through C in Figure 2). Whereas there are a series of survey data in West Germany and the United Kingdom, they are collected for diverse interests and contain only limited variables which can be analyzed for comparative purposes. For the micro-level mechanism discussed above, the key variable would be the use of national-level information in expectation formation concerning district races. This kind of variable is, unfortunately, hardly to find in comparative survey studies.[2]

Figure 2: Basic model of explanations in social sciences



Adapted from Esser (1999) ; Coleman (1986); McClelland (1967).

3 Solution

If a scholar has no data, should she keep still about the issue? This paper's answer is no. I agree with de Marchi (2005) when he writes: "when the data are less than perfect, logical implications can serve as excellent substitutes" (171). Accordingly, we can substitute theoretical implications for the missing link between mixed systems and high-level linkage. One such theoretical implication can be found regarding the district race where the large parties' candidates compete with each other and voters use national-level information in expectation formation. In this kind of districts, we can expect a stronger tendency toward bipartism since

[2]Recently, there are also some cross-national survey projects. One of such projects "Comparative Study of Electoral System (CSES)" may appear to be appropriate for this paper's purposes. This is, unfortunately, not the case, since measures of some important model components, e.g. perception of district candidate position, are absent. A further (and more crucial) problem of CSES is its post-election survey design. This makes it impossible to measure or impute the key variable in this paper, that is, expectation concerning election outcomes.

two factors consistently work in favor of the large parties' candidates. Consequently, they are expected to be characterized by more Duvergerian equilibrium and less Non-Duvergerian equilibrium. This implication is, however, not specific enough to serve as substitutes in analyzing empirical data.

To obtain more specific implications, we need to formalize the theoretical model (Morton 1999). For this purpose, researchers conventionally formulate models in the mathematical language and analytically solve it to achieve the implications. One drawback of this approach is that the more realistic and correspondingly complicated a model becomes, the less likely the implication can be achieved via analytical solution. To avoid this problem, this paper utilizes computational modeling as alternative to the conventional formal modeling.

3.1 Computational modeling

Computational modeling is a technique to draw implications from theoretical models using computer simulations (Troizsch, 1997; de Marchi, 2005; EITM-Report, 2002, Appendix B, Comments by Carl Simon). In this sense, this approach follows the deductive logic and is closely related to conventional deductive formal modeling which sets micro assumptions forth and analytically derives equilibrium as conclusions/predictions. Computational modeling is, however, free from restrictive assumptions and oversimplified models, which are often necessary in order to obtain any solutions or consequences. Instead to solve models analytically, computational modeling generates data via simulations and analyzed them inductively to reach implications.[3]

In terms of the problem of the absence of individual-level data, a variant of computational modeling techniques, agent-based modeling, offers a solution (see e.g. Kollman, Miller and Page, 1992; Cederman, 1997; Clough, 2007). Agent based models consist of a number of agents whose characteristics and decision-making rules can be set by theoretical models. In computational simulations, each agent repeatedly makes decisions interacting with other agents as well as its environment and learns from results of the interactions. In the context of research on electoral systems, the agents can be voters, parties or candidates like in the computational model presented later. In other topics of political science, parliament members, interest groups, ministry or a whole state can be modeled depending on the purpose of study. This kind of modeling allows researchers to observe macro-phenomena which emerge from interactions between micro agents. Emergence implies in this context aggregation of individual

[3]This has another advantage, that is, consequences of models can be also dynamic processes instead of static equilibrium which is often criticized as unrealistic.

behavior which is more or less unexpected from individual decision rules.

This kind of modeling technique not only relates two different macro-phenomena with each other via micro-level mechanisms, but also frees researchers from the restriction of available data. They need only explicitly expressed models about the data generating process (DGP) at stake which can be programmed in a computer language. Accordingly, we build a model to capture something systematic in the DGP and draw from the model some implications concerning the available data (see also Figure in de Marchi, 2005, 23). In the context of this study, the implications are testable using aggregate-level empirical data and the DGP is the entire process including micro mechanisms in Figure 2. The computational model tries to capture the unmeasured parts of DGP (A through C in Figure 2) and produce a macro-level result which can be measured in the form of election results.

3.2 Bayesian statistics

Modeling DGP via computational modeling has also a consequence for the analysis of empirical data. By seeing DGP behind empirical data, our interest should lie in how likely the model is true given certain empirical data. This view is fully consistent with a view of Bayesian statistics (Western and Jackman, 1994; Jackman, 2000; Gill, 2002). Note that this is the just reversed way of thinking to conventional empirical modeling embodied in OLS or maximum likelihood. In this view, researchers are interested in how likely the observed data are sampled from the population given a certain model. Besides to the epistemological consistency, the use of Bayesian statistics is also advantageous since it integrates the implications derived from computational modeling systematically in analyzing empirical data. This further enables statistical inference using data with small n.

The conventional statistical approach is well represented in the maximum likelihood method. This method maximizes the following likelihood function:

$$L(\theta|D) = \prod p(D|\theta) \tag{1}$$

Contrary, the logic based on the Bayesian view is represented as follows:

$$p(\theta|D) \propto p(\theta)p(D|\theta) \tag{2}$$

This means that the probability of estimates conditioned on data is a product of the likelihood function $p(D|\theta)$ and $p(\theta)$, that is, prior information about θ. Comparing with Equation 1 for the maximum-likelihood approach, Bayesian analysis also considers the prior information concerning parameters $p(\theta)$ in the way "the posterior is proportional to likelihood times prior". $p(\theta|D)$ is called posterior information since it is achieved, contrary to $p(\theta)$, after observing data.

This use of prior information is one of major critiques against the Bayesian approach. Results based on small n samples can be sensitive to the choice of prior distributions.[4] For researcher with traditional frequentist view and/or empiricist with inductive view, data should be analyzed without any prejudices. As Morton (1999, 44) points out, however, they rely on "the ability of researchers to measure the right data about the real world precisely and neutrally" and "a researcher's competence to use statistical methods to analyze data objectively and with little error". We should not forget that also this kind of empirical works is making a series of assumption in measuring and analyzing data. If one conducts e.g. regression analysis via OLS one makes an assumption that the errors are distributed according to the normal distribution. If one choose a set of independent variables, it is also an assumption of one's empirical model. If another set of independent variables is included in estimation, it would produce a different result. In this sense, results obtained using the conventional frequentist approach also sensitive to a number of assumptions, which are often made implicitly.

Contrary, estimations via Bayesian approach force researchers to choose more consciously certain prior information. The choice should be justified based on the theoretical model to be tested. Formal modeling is in this term in advantage since it can provide more clear and quantified implications than non-formal modeling. In the context of this study, the results of computational models deliver prior information in a form of distributions for the parameters of interest.

4 Applying the solution

This section applies the approach introduced in the last section to the problem sketched above. After brief introduction of the computational model,[5] I will derive implications from the model. The implications are then used to analyze empirical data to estimate some latent variables of which measures are not available.

[4] In increasing size of samples, the sensitivity comes to ignorable and the results converges to those estimated via frequentist approach. This is because sample information dominates the estimated posterior distribution.

[5] For more details see Shikano (2007)

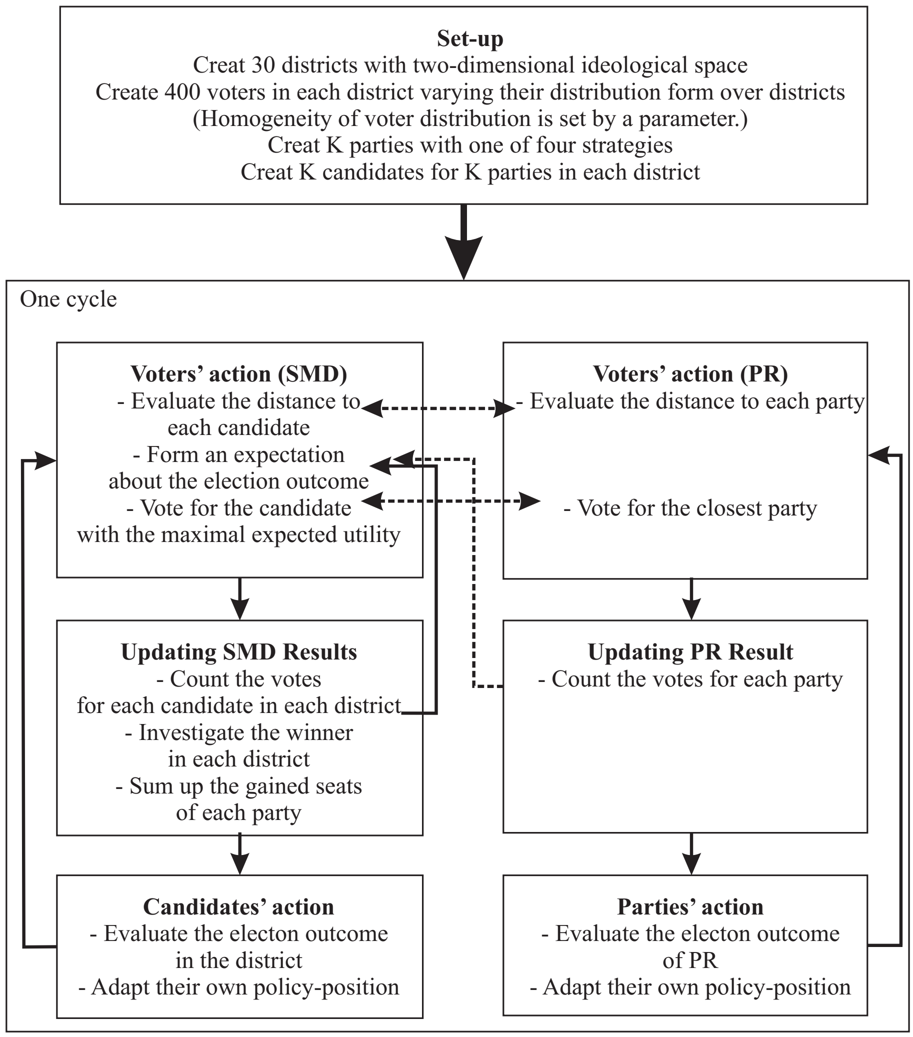
4.1 Set-ups of the computational model

The simulation model utilized 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 result of each party is 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 need only limited information of party competition. Some of them are results-oriented or vote-maximizing and the others are rather policy-oriented. Repeating the third, fourth and fifth step, Laver observes the consequences of various combinations of the party strategies.[6]

As summarized in Figure 3, this paper extends Laver's model in the following four points: First, Laver models party competitions under the pure proportional representation (PR)-system. This paper, contrary, is interested in the party competition and voting behavior under mixed systems. Correspondingly, I model also the party competition under the plurality in single-member districts. Second, besides the party the model here 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 a not most preferred candidate to maximize their expected utility. Therefore, not only the closeness to parties, but also the expectation on the election outcome in plurality races is also incorporated here. Fourth, the model of the present study suggests the existence of some interactive effects in casting plurality- and

[6] The choice of Laver's model is legitimized 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 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, especially prospective information, e.g. 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 an equilibrium, but to describe complex and dynamic processes of party competition.

Figure 3: Overview of the simulation model



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

PR-ballots. To observe which kind of consequences the interactive effects have, the model is also in this regard extended (dotted arrows in Figure 3).[7] In terms of the effect of mixed systems upon linkage, the interaction in expectation formation is relevant. The degree of the interaction λ_3 is modeled as weight between district-level and national-level information in expectation formation. If $\lambda_3 = 0$ voters form their expectation about the district race solely based on the past election results in the district. If $\lambda_3 = 1$ expectation is formed using solely the past election results at the national level. λ_3 can also take the value of 0.25, 0.50 and 0.75 and district- and national-level information is correspondingly weighted in expectation formation.

The model contains a large set of parameters. Since it would take a long time to investigate every point in such a high-dimensional parameter space, 1000 simulation runs are conducted with randomly drawn parameters.

Figure 4 gives a sample simulation run. A single simulation run lasts for 500 cycles. That is, voters cast their ballot 500 times. Four graphic windows in Figure 4 show the development of different kinds of statistics during a single simulation run. Of 500 cycles in each simulation, the first 300 cycles are discarded as so-called burn-in since cycles in earlier phase are strongly conditioned by the starting values.[8] 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 collect 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 1000 times which generate 40000 data points.[9]

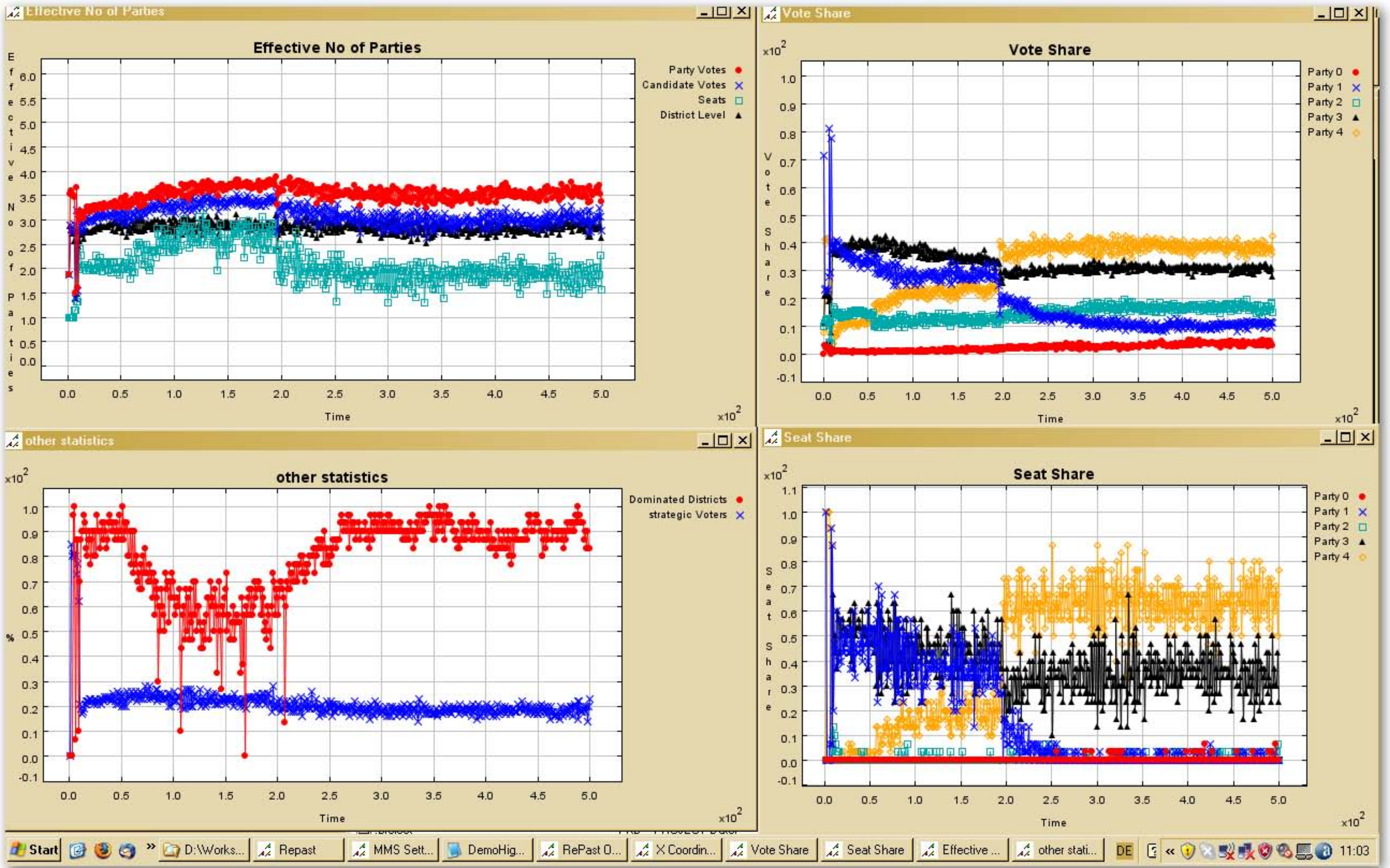
The data collected in this way are suited to the conventional statistical techniques since the frequentist view postulates that the analyzed data are only one realization of infinitely repeatable data collection processes. This is, however, hardly realistic for the data collected in conventional methods. The simulation technique is one of a few methods compatible with the frequentist assumption above. Furthermore, there is no risk of multicollinearity in multivariate analysis since the random draws of parameters were conducted independently of each other

[7]The model was programmed using Repast and is available from the author upon request. Repast (**Re**cursive **P**orus **A**gent **S**imulation **T**oolkit) is a broadly used, free and open-source Java-based toolkit for agent-based modeling and further simulation techniques.

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

[9]Using a desktop machine with Pentium 4 (2.80 GHz) processor one simulation run approximately takes four minutes in the batch mode without displaying graphics. The length of each individual run, however, varies depending on the number of parties and further parameters.

Figure 4: An example of a single simulation run



and there is no correlation between parameters.

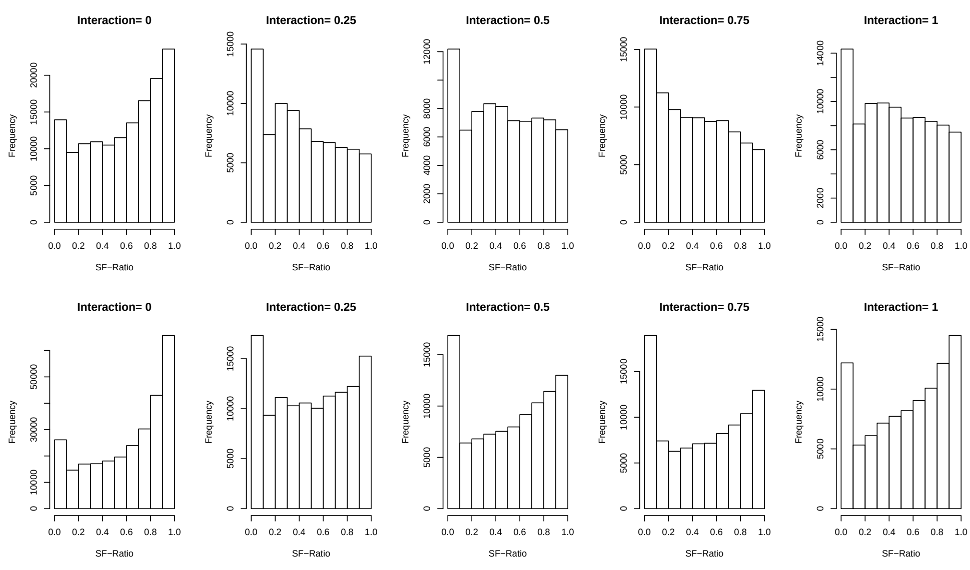
4.2 Implications of the computational model

Now, we are turning to the simulation results to draw implications from the model. As stated above, I expect a tendency toward Duvergerian equilibrium in dominated districts where the candidates of two main parties at national level finish the first and second. This tendency should be reinforced by using the national level results as heuristics in expectation formation. To see whether the computational model yielded the corresponding results, we begin with a descriptive investigation of simulation results. For this purpose, we first need to introduce the measure for Duvergerian and Non-Duvergerian equilibrium: SF ratio. This measure is suggested by Cox (1994, 1997) and computed as follows:

$$SF = \frac{v_{M+2}}{v_{M+1}} \tag{3}$$

where v_{M+1} denotes the vote share of the first loser and v_{M+2} that of the second loser. If Duvergerian equilibrium is operating v_{M+2} tends to be zero so that the SF ratio approaches to zero. Contrary, $v_{M+1} \approx v_{M+2}$ under Non-Duvergerian equilibrium. In this case, the SF ratio should be near unity.

Figure 5: SF ratio, linkage and interactive effects in expectation formation



Upper panels are distributions in dominated districts; lower panels are those in undominated districts.

Figure 5 presents distributions of SF ratios with various parameter combinations. The upper panels are the distributions of SF ratios in dominated districts and the lower panels are those in undominated districts. Accordingly, a bimodal distribution is replicated in dominated as well as undominated districts when voters form their expectation using their own district results (the left most panels). Contrary, the difference between dominated and undominated districts is clearer if there are interactive effects in expectation formation. When the expectation is formed using the national level PR results (at least partially) Non-Duvergerian equilibrium tends to be more observed in undominated districts than in dominated districts. This demonstrates that two kinds of effects are operating in undominated districts. One is that of interactive effects in expectation formation which reinforces candidates of top two parties at national level. On the other hand, there is also another effect in favor of top two candidates at district level. Since, per definition, the set of top two candidates do not correspond to the set of top two national parties in undominated districts, three or more candidates can enjoy the advantage of these effects. As a result, Non-Duvergerian equilibrium is more likely to be observed in undominated districts. Reversely, top two candidates in dominated districts also enjoy the effects of their own parties' result at national level. This makes difficult for the other candidates to contest with one of the top two candidates. Consequently, Duvergerian equilibrium is easier to be achieved.

Note that I do not speak of a causal relationship between districts dominance and existence of (Non-)Duvergerian equilibrium above. Both are simultaneous results of diverse parameters in the computational model including λ_3.

To confirm the visual observations above based on Figure 5, I estimate a statistical model using the simulated data. It is, however, tricky to observe multivariate relationships between the SF ratio and a number of independent variables. Here, the conventional regression analysis using ordinary least squares (OLS) is inappropriate due to some characteristics of the SF ratio. First, the range of the SF ratio is bounded between two extreme values 0 and 1. There is, however, no guarantee that predicted values yielded by OLS fall inside of these scale. Second, not all theoretical models discussed above expect non-unimodal symmetric distribution. In the context of this study, eventually expected bimodal distributions of SF ratios are not congruent with the assumption of OLS.

To cope with these problems, an alternative estimation model based on the beta distribution has been suggested by some researchers (Brehm and Gates, 1993; Paolino, 2001; Buckley, 2003). Accordingly, the dependent variable, here the SF ratio, is assumed to be generated according to the beta distribution. Paolino (2001) suggested an alternative modeling in which the expected value and dispersion of SF ratios are modeled.[10] I follow this approach:

$$\mathrm{SF}_{rck} \sim \mathcal{B}(\alpha_{rck}, \beta_{rck}) \quad (4)$$

$$\Gamma_{dt} = \gamma_{\alpha 0} + \gamma_{\alpha 1}\lambda_{3,r} + \gamma_{\alpha 2}\mathrm{dom}_{rck} + \gamma_{\alpha 3}\lambda_{3,r} \times \mathrm{dom}_{rck} \quad (5)$$

$$\Gamma'_{dt} = \gamma_{\beta 0} + \gamma_{\beta 1}\lambda_{3,r} + \gamma_{\beta 2}\mathrm{dom}_{rck} + \gamma_{\beta 3}\lambda_{3,r} \times \mathrm{dom}_{rck} \quad (6)$$

where,

SF_{rck} is SF ratio of district k at c'th cycle of r'th simulation run

dom_{rck} is one if district k at rc is dominated, otherwise zero;

$\lambda_{3,r}$ is the degree of interactive effects in expectation formation at election t.

Table 1 reports the estimation results.[11] According to the table, the use of PR tier information in expectation formation processes, that is, a nonzero value of λ_3 increases in general the likelihood of Duvergerian equilibrium and decreases that of Non-Duvergerian equilibrium. This is consistent with the theoretical expectation and the visual investigation of Figure 5 as well. The use of national level results as heuristics facilitates the coordination among voters and results in higher probability of Duvergerian equilibrium. The coefficient for dominated districts also has the positive sign. That is, if the district is contested among the candidates of the national level largest two parties, Duvergerian equilibrium is more likely and Non-Duvergerian takes place less likely. These two factors above, interactive effects in expectation formation and district dominance working in favor of Duvergerian equilibrium, are not only additive, but also multiplying. This is shown by the coefficients of interaction effects of both factors. Accordingly, the bimodality is weakened if both factors are present. Under higher-level interactive effects in expectation formation, we can expect more divergence concerning Duvergerian/Non-Duvergerian equilibria between dominated and undominated districts as we have already observed in Figure 5.

[10]See Appendix for more details.

[11]District results in which one candidate wins all the votes (about 3% of all district results) are excluded from the analysis since their SF ratios cannot be calculated. Furthermore, the log-likelihood returns an infinite value if $\mathrm{SF}_{rct} = 0$ or $\mathrm{SF}_{rct} = 1$. To avoid this, 0 is replaced with 0.5/400 and 1 with 399.5/400 respectively. Since one district has 400 voters the minimum value of nonzero SF-ratio is 1/199; the maximum value of SF-ratio which is smaller than one is 132/134. Therefore, these values cannot be reached by the value of the other SF ratios.

Table 1: Results of the beta maximum likelihood estimation

	E(SF)		Var(SF)	
	Γ_{dt}		Γ'_{dt}	
	Coeff.	SE	Coeff.	SE
Constant	0.184	0.003	0.241	0.002
λ_3	-0.391	0.004	0.124	0.004
dom	-0.203	0.003	0.076	0.003
$\lambda_3 \times$ dom	0.041	0.006	0.263	0.006

Log-likelihood: 120057.9; N= 1162911

The different distribution forms of SF ratios serve also as a kind of measure of λ_3 in empirical data analysis in the next section. That is, if one can observe a strong tendency toward Duvergerian equilibrium in dominated districts and a bimodal distribution in undominated districts, it can be inferred that voters take into account the PR tier results in forming expectations concerning plurality outcomes in their districts (a higher λ_3). Contrary, certain bimodality in both kinds of districts is interpreted as lower-level interactive effects between both tiers regarding the expectation formation process (a lower λ_3). This is an advantage of computational modeling approach. While this study is interested in λ_3, there is no empirical data available which measured it directly. Being confronted with this problem, the results of the computational model provided an alternative measure which can be applied to aggregate level data.[12]

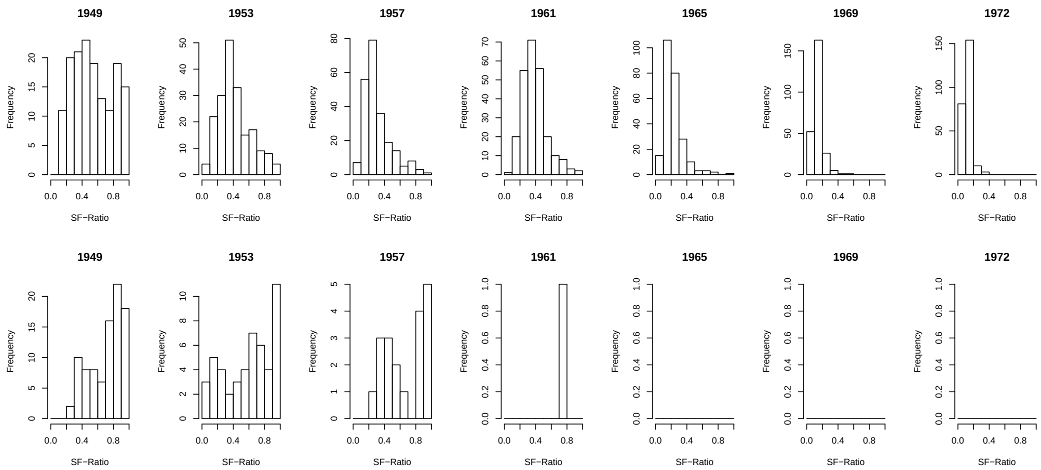
4.3 Integrating theoretical implications with empirical data via Bayesian statistics

We begin with the descriptive observation of SF ratios in dominated and undominated districts in West Germany. Figure 6 presents the distributions of SF ratios between 1949 and 1972. While SF ratios in dominated districts at 1949 election seems to be randomly distributed the distributions at 1953 election and thereafter reveals a trend towards Duvergerian equilibrium.

[12] A similar way of using computational models can be found in de Marchi (2005, Chapter 5). The author develops a computational model to validate his measure of the preference separability (for the preference separability see Lacy, 2001). Since the measure of preference separability is not available in conventional survey data, data are generated assuming different levels of separability using a computational model. Using the generated data, he developed a new alterative measure which can be applied to conventional survey data. The measure is compared with the full information measures of preference separability. After showing the consistency between the measurements, he applied his measure to survey data to observe the existence of the preference separability among the electorate in the US and Israel.

At 1972 election we can see a high concentration of SF ratio between 0.0 and 0.2. Undominated districts which existed until 1961 election reveals a reversed trend, i.e. a trend towards Non-Duvergerian equilibrium. This trend can be observed already at the first general election held in 1949. This means that most of non-SPD/CDU candidates who ranked second have never enjoyed a privileged status of visible candidate.

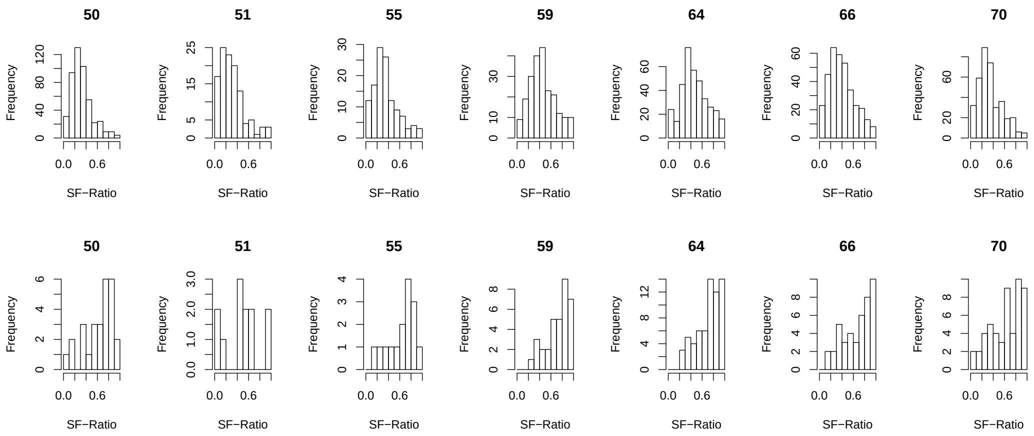
Figure 6: SF ratio in dominated and undominated districts (West Germany)



The trend that more Duvergerian equilibrium in dominated and more Non-Duvergerian equilibrium in undominated district is in line with the implication under high-level interactive effects in expectation formation obtained in the last Chapter (see Figure 5 on Page 14). To see that this West German trend is not necessarily the case in other countries, we can observe the corresponding figure for British elections. While the distributions at elections between 1950 and 1970 in dominated districts tend to have Duvergerian and undominated districts Non-Duvergerian equilibrium (Figure 7), no dynamic reinforcing trend can be observed as is the case in West Germany. That is, the distribution forms in 1950 and 1970 demonstrate no clear difference.

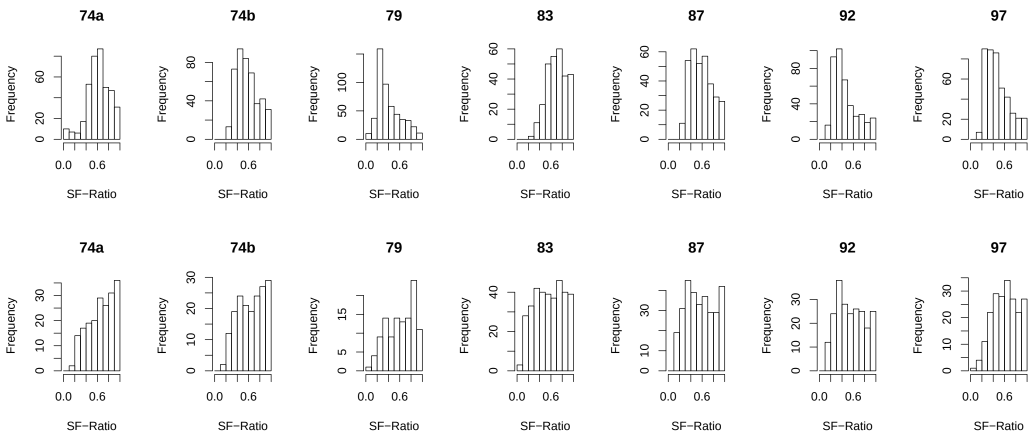
If we take a look at the distributions of SF ratios after 1974 (Figure 8), dominated districts do not always show a tendency toward Duvergerian equilibrium. Especially at the 1983 election, dominated districts also show a stronger tendency toward Non-Duvergerian equilibrium than undominated district. As Figure 1 (Page 5) shows, the percentage of dominated districts at this election was under 50% and at the lowest level after the War. At the national-level vote share, the third largest SDP-Liberal Alliance (25.4%) was very close behind the second largest

Figure 7: SF ratio in dominated and undominated districts (United Kingdom 1950-1970)



Upper panels are distributions in dominated districts; lower panels are those in undominated districts.

Figure 8: SF ratio in dominated and undominated districts (United Kingdom 1974-1997)



Upper panels are distributions in dominated districts; lower panels are those in undominated districts.

Labour Party (27.6%).[13] These clarify that neither linkage nor Duvergerian equilibrium was operating at the 1983 election. One possible explanation is a small-size split of the Labour party prior to the election. In 1981 some members of parliament left Labour and founded a new party, the Social Democratic Party (SDP). This new party agreed with the Liberals to form an electoral alliance at the 1983 election. Besides, participation of some senior Labour figures in the SDP caused a shift of many votes from Labour to the SDP-Liberal Alliance. One further favorable circumstance for the Alliance was redistricting prior to the election. These factors could disrupt voters' expectation formation concerning the outcome in their district which resulted in less Duvergerian equilibrium also in dominated districts.[14]

Now, this trend of beta distributions is modeled by using implications from the computational model. The statistical model utilized here has the same structure as that in the last section (see also Appendix). One problem in applying this model to the empirical data is that the use of national-level information in expectation formation ($\lambda_{3,t}$) is unknown. To solve this problem, I treat the values of λ_3 as latent variable and estimated them via Bayesian estimation process. Accordingly, we treat $\lambda_{3,t}$ for each year t as a random variable. Since its scale is limited between 0 and 1 the uniform distribution seems to be appropriate as prior distribution. Thus, I assume for each t:

$$\lambda_{3,t} \sim \text{Unif}(0, 1) \tag{7}$$

This kind of treatment of $\lambda_{3,t}$ have some advantages. First, λ_3 is assumed to be random, that is, variable among individual voters. Second, this randomness of λ_3 enables us to make statistical inference.

Furthermore, we have prior information concerning γ- and γ'-parameters obtained from the analysis in the last subsection (see Table 1 on Page 17). This information can be legitimately used as prior distribution to get the posterior distribution of parameters due to the same model structure based on the beta distribution. Among eight γ- and γ'-parameters, we are theoretically interested in three parameters: γ_1, γ_2 and γ'_3. Therefore, I use the simulation results for these three parameters while uninformative priors are used for the other parameters.

The posterior distribution is determined by the data as well as by the prior distribution using the variance/standard error as weighting factor. In this regard, the results from simulated data have very small standard error due to the large amount of data points. This could

[13] Despite its vote share, SPD-Liberal Alliance could obtain only 3.5% of seats in the parliament. This is the reason why the seat share of the two largest parties at 1983 election is quite high.

[14] Furthermore, dominated districts at both elections in 1974 demonstrate a certain tendency toward Non-Duvergerian equilibrium although it is weaker than that of undominated districts.

lead that the prior distributions fully dominate the posterior distributions and empirical data do not play any role in estimation. To avoid this, I re-estimate the model using randomly drawn 2000 and 4000 data points in the simulated data. The number of 2000 is chosen arbitrarily being oriented by the number of overall observations in the West German district data 1949-1972 (N=1722). Analogously, 4000 roughly correspond to the number of British overall districts between 1950 and 1987. [15]

Table 2, 3 and 4 summarize the prior information used in the following estimations.[16] Each prior distribution is assumed to be distributed normal with mean and standard deviation given in the tables. For uninformative priors, their normal distribution with standard deviation of 100 is so flat that it hardly influences the posterior distribution. Contrary, a prior distribution with standard deviation of e.g. 0.113 can have a certain impact upon the posterior distribution. However, this is not necessarily the case since it depends also upon the quality of empirical data (variance and the number of observations).

Table 2: Informative priors for West Germany

		expected value			dispersion	
		mean	sd		mean	sd
constant	γ_0	.000	100.000	γ'_0	.000	100.000
λ_3	γ_1	-.350	.113	γ'_1	.000	100.000
dom	γ_2	.203	.084	γ'_2	.000	100.000
$\lambda_3 \times$ dom	γ_3	.000	100.000	γ'_3	-.263	.147

Table 3: Informative priors for the United Kingdom

		expected value			dispersion	
		mean	sd		mean	sd
constant	γ_0	.000	100.000	γ'_0	.000	100.000
λ_3	γ_1	-.350	.050	γ'_1	.000	100.000
dom	γ_2	.203	.042	γ'_2	.000	100.000
$\lambda_3 \times$ dom	γ_3	.000	100.000	γ'_3	-.263	.072

Table 4: Non-informative priors

		expected value			dispersion	
		mean	sd		mean	sd
constant	γ_0	.000	100.000	γ'_0	.000	100.000
λ_3	γ_1	.000	100.000	γ'_1	.000	100.000
dom	γ_2	.000	100.000	γ'_2	.000	100.000
$\lambda_3 \times$ dom	γ_3	.000	100.000	γ'_3	.000	100.000

Table 5 presents estimation results with and without using the informative prior distribu-

[15]The estimation models without this kind of informative prior suffer the identifiability problem (Bafumi et al., 2005). To identify the model, $\lambda_{3,t}$ are rescaled so that the mean of $\lambda_{3,t}$ for each iteration equals 0.5.

[16]Estimation was conducted using OpenBUGS, a free open-source software for Bayesian inference. It was called from R using BRugs package. The posterior distributions are based on 5000 simulation runs after the burn-in period with 20000 runs.

tions based on the district-level election results between 1949 and 1973. In terms of the effect of the use of national-level information, dominance of district and their interaction, γ_1, γ_2 and γ'_3 are of our interest. All of these coefficients have the expected sign in both estimation models. That is, the use of national-level information as well as dominance of districts lowers SF ratios. And their interaction results in a lower variance of SF ratios. These effects seem to be not sensitive to the use of prior information, while there are some differences in the size of the effects.

Now, turning to the results estimated using the informative prior of Table 2, all of γ_1, γ_2 and γ'_3 show expected effects in line with the implication of the computational model. Here, one could criticize that the estimation results are very sensitive to the prior distributions. This is, however, only partly true. As Figure 9 visualizes, $\lambda_{3,t}$ in both estimations demonstrate a similar increasing trend. Two monotonically increasing trends are accompanied with a slight decrease of λ_3 between elections 1957 and 1961. While there is a large difference of the level of λ_3 at the first election in 1949, the assumption that λ_3 did not reach it maximum at 1961 or 1965 but continued to grow till election 1972. This parallel relationship of both trends can also be confirmed by Bayesian estimation. Since each $\lambda_{3,t}$ consists from estimates in a number of iterations we can easily compute the probability whether the use of national-level information increased between specific two elections. Comparisons between different pairs of elections confirm that both increasing trends are similar to each other despite their difference in their dispersion.

Figure 9: The degree of interactive effects (West Germany): Estimation results of λ_3 with 95% confidence interval

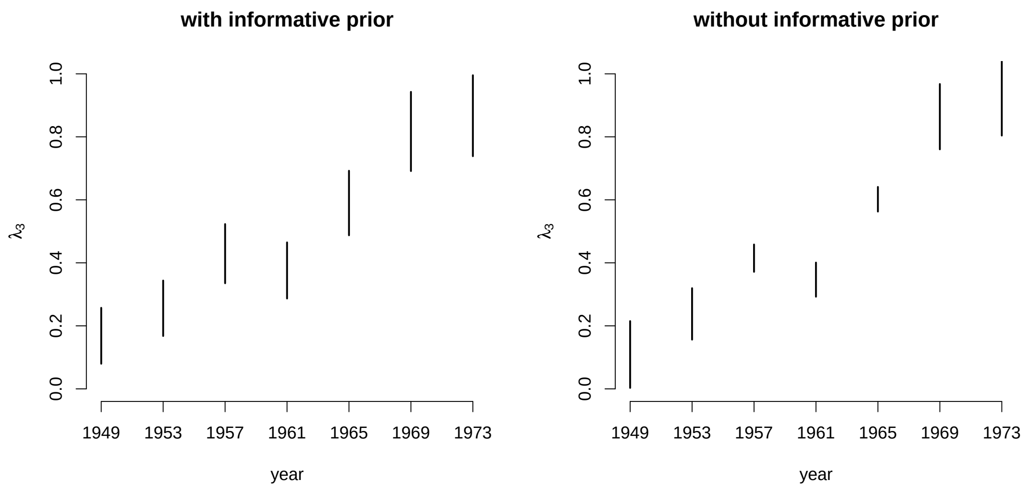


Table 5: Results of Bayesian estimation (West Germany)

	with informative prior				without informative prior			
	mean	sd	2.5%	97.5%	mean	sd	2.5%	97.5%
Effects on expected value								
γ_0	0.72	0.09	0.55	0.89	0.92	0.14	0.68	1.20
γ_1	-0.34	0.13	-0.68	-0.11	-2.20	0.71	-3.46	-0.92
γ_2	-0.21	0.09	-0.38	-0.02	-0.54	0.16	-0.81	-0.25
γ_3	-2.41	0.28	-3.04	-1.99	-1.97	1.23	-4.32	-0.35
Effects on dispersion								
γ_0	0.18	0.21	-0.23	0.48	1.23	0.14	0.96	1.51
γ_1	3.38	0.38	2.81	4.32	-2.20	1.25	-4.30	-0.06
γ_2	0.30	0.12	0.05	0.52	-0.60	0.30	-1.14	-0.15
γ_3	0.35	0.14	0.07	0.62	7.89	1.52	5.48	11.13
Estimated degree of interactive effects								
$\lambda_{3,1949}$	0.16	0.05	0.08	0.26	0.09	0.06	0.00	0.22
$\lambda_{3,1953}$	0.25	0.05	0.17	0.34	0.24	0.05	0.16	0.32
$\lambda_{3,1957}$	0.43	0.05	0.34	0.52	0.42	0.02	0.37	0.46
$\lambda_{3,1961}$	0.37	0.05	0.29	0.46	0.35	0.03	0.29	0.40
$\lambda_{3,1965}$	0.60	0.06	0.49	0.69	0.60	0.02	0.56	0.64
$\lambda_{3,1969}$	0.84	0.07	0.69	0.94	0.87	0.06	0.76	0.97
$\lambda_{3,1972}$	0.90	0.07	0.74	1.00	0.93	0.07	0.80	1.04
Model fit								
mse	0.03	0.00	0.03	0.03	0.03	0.00	0.03	0.03

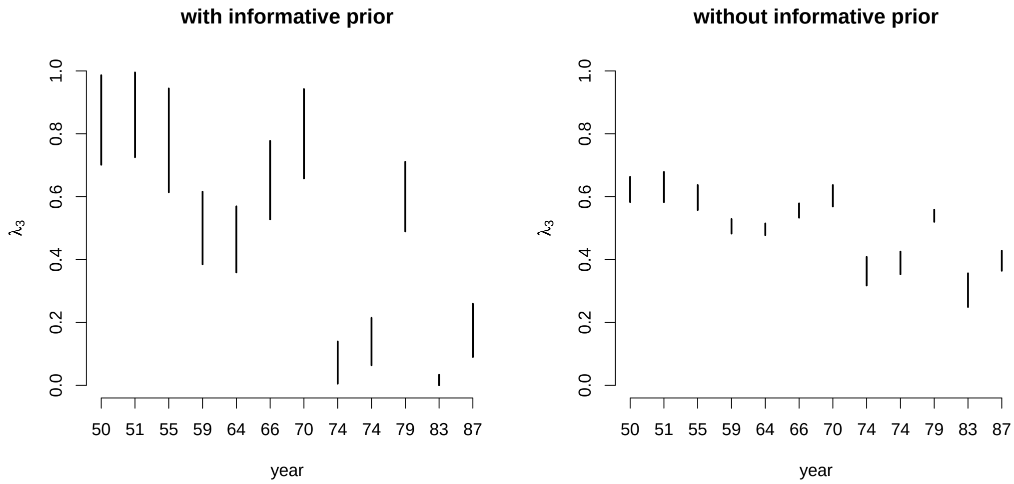
The analogous analysis using British data yielded the results in Table 6. The estimates of γ are more sensitive to the prior information than in the analysis of West-German data. Furthermore, the estimates of λ_3 show certain difference as well. Still, both estimates of the use of national-level information show similar downwards trend (see Figure 10).

If we compare the development of the use of national-level information estimated as $\lambda_{3,t}$, the difference between West Germany and the United Kingdom is apparent. While West German voters increasingly utilize the national-level information after the introduction of their mixed systems, British voters show no increase in using the national-level information. The opposite is the case: a long-termed decreasing trend is observed. Can West German voter's increasing use really attribute to the mixed electoral system? Analogous analysis of data from New Zealand and Japan shows a similar trend after the introduction of mixed systems in these countries (Shikano, 2007).

Table 6: Results of Bayesian estimation (the United Kingdom)

	with informative prior				without informative prior			
	mean	sd	2.5%	97.5%	mean	sd	2.5%	97.5%
Effects on expected value								
γ_0	0.62	0.03	0.56	0.67	0.87	0.26	0.31	1.30
γ_1	-0.32	0.05	-0.41	-0.22	-0.64	0.44	-1.33	0.21
γ_2	-0.04	0.03	-0.10	0.01	1.46	0.51	0.67	2.56
γ_3	-1.14	0.14	-1.47	-0.95	-2.60	0.55	-3.79	-1.68
Effects on dispersion								
γ'_0	1.12	0.03	1.05	1.17	0.87	0.07	0.75	1.01
γ'_1	-0.21	0.05	-0.30	-0.13	0.23	0.11	-0.02	0.43
γ'_2	0.35	0.03	0.30	0.42	0.81	0.10	0.61	0.99
γ'_3	0.21	0.04	0.15	0.30	-0.50	0.13	-0.70	-0.24
Estimated degree of interactive effects								
$\lambda_{3,1950}$	0.88	0.08	0.70	0.99	0.61	0.02	0.58	0.66
$\lambda_{3,1951}$	0.90	0.08	0.73	1.00	0.62	0.02	0.58	0.68
$\lambda_{3,1955}$	0.79	0.09	0.61	0.94	0.59	0.02	0.56	0.64
$\lambda_{3,1959}$	0.50	0.06	0.38	0.62	0.51	0.01	0.48	0.53
$\lambda_{3,1964}$	0.47	0.06	0.36	0.57	0.50	0.01	0.48	0.52
$\lambda_{3,1966}$	0.66	0.07	0.53	0.78	0.55	0.01	0.53	0.58
$\lambda_{3,1970}$	0.81	0.08	0.66	0.94	0.60	0.02	0.57	0.64
$\lambda_{3,1974}$	0.06	0.04	0.01	0.14	0.37	0.02	0.32	0.41
$\lambda_{3,1974}$	0.14	0.04	0.06	0.21	0.40	0.02	0.35	0.43
$\lambda_{3,1979}$	0.61	0.06	0.49	0.71	0.54	0.01	0.52	0.56
$\lambda_{3,1983}$	0.01	0.01	0.00	0.03	0.31	0.03	0.25	0.36
$\lambda_{3,1987}$	0.17	0.04	0.09	0.26	0.40	0.02	0.36	0.43
Model fit								
mse	0.05	0.00	0.05	0.05	0.05	0.00	0.05	0.05

Figure 10: The degree of interactive effects (Great Britatin): Estimation results of λ_3 with 95% confidence interval



5 Conclusion

Analyzing the effect of electoral systems on linkage of plurality races, this paper exemplify how to utilize computational modeling and Bayesian statistics to overcome a typical problem of comparative political studies.

The main message is that one needs a theoretical model to make inferences using empirical data with limited information. To this end, I suggest to build a formal model which is extended to consider further factors whereby a computational modeling technique is employed to derive implications of the model. The implications, which will be obtained as results of the computational model, will be then integrated into empirical model via Bayesian statistics. The methodological innovation of this paper lies in the last step since, to my knowledge; there has been no empirical analysis which incorporates results from computational modeling using Bayesian approach.

References

Bafumi, Joseph, Andrew Gelman, David K. Park and Noah Kaplan. 2005. "Practical issues in implementing and understanding Bayesian ideal point estimation." *Political Analysis* 13:171–87.

Black, Jerome H. 1978. "The multicandidate calculus of voting: Application to Canadian Federal Elections." *American Journal of Political Science* 22:609–38.

Brehm, John and Scott Gates. 1993. "Donut shops and speed traps: Evaluating models of supervision on police behavior." *American Journal of Political Science* 37(2):555–81.

Buckley, Jack. 2003. "Estimation of models with beta-distributed dependent variables: A replication and extention of Paolino (2001)." *Political Analysis* 11(1):1–12.

Cederman, Lars-Erik. 1997. *Emergent Actors in World Politics: How States and Nations Develop and Dissolve.* Princeton, NJ: Princeton University Press.

Chhibber, Pradeep K. and Ken Kollman. 2004. *The Formation of National Party Systems: Federalism and Party Competition in Canada, Great Britain, India, and the United States.* Princeton: Princeton University Press.

Chhibber, Pradeep and Ken Kollman. 1998. "Party Aggregation and the Number of Parties in India and the United States." *American Political Science Review* 92(2):329–42.

Clough, Emily. 2007. "Strategic voting under conditions of uncertainty: A re-evaluation of Duverger's Law." *British Journal of Political Science* 37:313–32.

Coleman, James S. 1986. "Social theory, social research, and a theory of action." *American Journal of Sociology* 91(6):1309–35.

Cox, Gary W. 1994. "Strategic voting equilibria under the Single Nontransferable Vote." *American Political Science Review* 88:608–21.

Cox, Gary W. 1997. *Making Votes Count: Strategic Coordination in the World's Electoral Systems*. Cambridge: Cambridge University Press.

de Marchi, Scott. 2005. *Computational and Mathematical Modeling in the Social Sciences*. Cambridge: Cambridge University Press.

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

EITM-Report. 2002. *The empirical implications of theoretical models (EITM) workshop: Report of the political science program, directorate for social, behavioral and economic sciences, National Science Foundation*.

Esser, Hartmut. 1999. *Soziologie. Spezielle Gundlagen. Band 1: Situationslogik und Handeln*. Frankfurt: Campus.

Fey, Mark. 1997. "Stability and coordination in Deuverger's law: A formal model of preelection polls and strategic voting." *American Political Science Review* 91(1):135–47.

Gill, Jeff. 2002. *Bayesian Methods: A Social and Behavioral Sciences Approach*. Boca Raton, FL: Chapman and Hall/CRC.

Jackman, Simon. 2000. "Estimation and inference via Bayesian simulation: An introduction to Markov Chain Monte Carlo." *American Journal of Political Science* 44(2):369–98.

Kollman, Ken, John Miller and Scott Page. 1992. "Adaptive parties and spatial elections." *American Political Science Review* 86:929–937.

Lacy, Dean. 2001. "A theory of nonseparable preferences in survey responses." *American Journal of Political Science* 45(2):239–58.

Laver, Michael. 2005. "Policy and the dynamics of political competition." *American Political Science Review* 99(2).

Leys, Colin. 1959. "Models, Theories, and the Theory of Political Parties." *Political Studies* 7:127–46.

McClelland, David C. 1967. *The achieving society.* New York: Free Press.

McKelvey, Richard D. and Peter C. Ordeshook. 1972. A general theory of the calculus of voting. In *Mathematical Applications in Political Science, VI*, ed. J. F. Herndon and J. L. Bernd. Charlottesville: University Press of Virginia pp. 32–78.

Morton, Rebecca B. 1999. *Methods and Models: A Guide to the Empirical Analysis of Formal Models in Political Science.* Cambridge: Cambridge University Press.

Myerson, Roger B. and Robert J. Weber. 1993. "A Theory of Voting Equilibria." *American Political Science Review* 87(1):102–14.

Paolino, Philip. 2001. "Maximum likelihood estimation of models with beta-distributed dependent variables." *Political Analysis* 9(4):325–46.

Sartori, Giovanni. 1968. Political development and political engineering. In *Public Policy*, ed. John D. Montgomery and Alvert O. Hirschman. Cambridge, Mass.: Harvard University Press pp. 261–98.

Shikano, Susumu. 2007. *Interactive Mechanism of Mixed-member Electoral Systems: A Theory-driven Comparative Analysis via Computational Modelling and Bayesian Statistics*. Habilitationsschrift an der Universität Mannheim.

Troizsch, Klaus G. 1997. Social science simulation - origins, prospects, purposes. In *Simulating Social Phenomena*, ed. Rosaria Conte, Rainer Hegselmann and Pietro Terno. Berlin: Springer pp. 41–54.

Western, Bruce and Simon Jackman. 1994. "Bayesian Inference for Comparative Research." *American Political Science Review* 88:412–23.

Appendix: the beta distribution model

The beta distribution is a function of two parameters α and β with $\alpha, \beta > 0$. The beta distribution has the following convenient characteristics: First, the beta distribution has a upper and lower limit. This is convenient since the SF ratio also has a limited scale between 0 and 1. Second, the beta distribution is enough flexible to model a uni- or bimodal distribution of SF ratios.

One caveat of the use of the beta distribution is difficulty to interpret both parameters. Consequently, if one models directly both parameters α and β, interpretation of the results cannot be straightforward. To this problem, Paolino (2001) suggested an alternative modeling in which the expected value and dispersion of the beta distribution are modeled. If we assume that SF ratios are distributed according to the beta distribution, one can model their expected values and dispersion as follows:

$$E(\text{SF}) = \frac{\exp(\Gamma X)}{1 + \exp(\Gamma X)} \tag{8}$$

$$Var(\text{SF}) = \frac{E(\text{SF})[1 - E(\text{SF})]}{\exp(\Gamma' X) + 1} \tag{9}$$

where Γ is the vector of parameters to be estimated and X is the vector of independent variables. α and β-parameters' relationships to the expected value variance are known as:

$$\alpha = E(\text{SF}) \left(\frac{E(\text{SF})(1 - E(\text{SF}))}{Var(\text{SF})} - 1 \right) \tag{10}$$

$$\beta = (1 - E(\text{SF})) \left(\frac{E(\text{SF})(1 - E(\text{SF}))}{Var(\text{SF})} - 1 \right) \tag{11}$$

Due to the intuitive interpretation of results, I take this modeling strategy using the following model specification:

$$\text{SF}_{rck} \sim \mathcal{B}(\alpha_{rck}, \beta_{rck}) \tag{12}$$

$$\Gamma_{dt} = \gamma_{\alpha 0} + \gamma_{\alpha 1}\lambda_{3,r} + \gamma_{\alpha 2}\text{dom}_{rck} + \gamma_{\alpha 3}\lambda_{3,r} \times \text{dom}_{rck} \tag{13}$$

$$\Gamma'_{dt} = \gamma_{\beta 0} + \gamma_{\beta 1}\lambda_{3,r} + \gamma_{\beta 2}\text{dom}_{rck} + \gamma_{\beta 3}\lambda_{3,r} \times \text{dom}_{rck} \tag{14}$$

where,

SF_{rck} is SF ratio of district k at c'th cycle of r'th simulation run

dom_{rck} is one if district k at rc is dominated, otherwise zero;

$\lambda_{3,r}$ is the degree of interactive effects in expectation formation at election t.

To get estimates of γ, one maximizes the following log-likelihood function (Paolino, 2001):

$$\begin{aligned} LL &= \sum_{r=1}^{1000} \sum_{c=1}^{40} \sum_{k=1}^{30} \{\ln\Gamma(\alpha+\beta) - [\ln\Gamma(\alpha) + \ln\Gamma(\beta)] \\ &\quad + (\alpha-1)\ln(\mathrm{SF}) + (\beta-1)\ln(1-\mathrm{SF})\} \end{aligned} \tag{15}$$