Hindawi
Complexity
Volume 2017, Article ID 7354642, 15 pages
https://doi.org/10.1155/2017/7354642

Research Article

Empirical Analysis and Agent-Based Modeling of the Lithuanian Parliamentary Elections

Aleksejus Kononovicius

Institute of Theoretical Physics and Astronomy, Vilnius University, Vilnius, Lithuania

Correspondence should be addressed to Aleksejus Kononovicius; aleksejus.kononovicius@tfai.vu.lt

Received 31 August 2017; Accepted 9 November 2017; Published 29 November 2017

Academic Editor: Gilberto C. Gonzalez-Parra

We analyze a parties' vote share distribution across the polling stations during the Lithuanian parliamentary elections of 1992, 2008, and 2012. We find that the distribution is rather well fitted by the Beta distribution. To reproduce this empirical observation, we propose a simple multistate agent-based model of the voting behavior. In the proposed model, agents change the party they vote for either idiosyncratically or due to a linear recruitment mechanism. We use the model to reproduce the vote share distribution observed during the election of 1992. We discuss model extensions needed to reproduce the vote share distribution observed during the other elections.

1. Introduction

While any individual vote is equally important to determine the outcome of an election, the probability for a single vote to decide the outcome is extremely small. As utility of casting a vote in this context seems to be small and as there is at least a minor associated cost, it seems that a rational choice would be simply not to vote. Yet this simplistic context may be further extended to provide sound reasoning for why people vote. Some argue that people vote to show support for the political system [1] or to avoid a risk of regret [2]; there might also be a social cost for abstention [3]. Some of the aforementioned works as well as numerous other earlier game-theoretic approaches, such as [4–6], had shown promise that game-theoretic voting models would soon provide rich and sophisticated explanation for the voting behavior. Yet further researches have shown that general game-theoretic models of the voting behavior with pure Nash equilibrium, and even mixed Nash equilibrium, might be impossible unless under certain specific conditions [7–9]. But people are rarely well informed and ideally rational, as they are not homo economicus nor are they Laplace's demons [10–12].

The above context provides good reasoning to consider the modeling of the voting behavior from the perspective of psychology [13–20]. Usually these models appear to be very sophisticated at least if compared to the game-theoretic models. They are hard to implement numerically and, most importantly, it is usually hard to provide clear insights into the modeled phenomena. Also usually these models involve a large number of parameters, which may lead to overfitting the data or different parameter sets providing similar results. Notably, recently, a psychologically motivated model was successfully used to predict the Polish election of 2015 [17, 19].

Another possible approach to the modeling of the voting behavior has its roots in statistical physics. This perspective could be neatly summarized by quoting Boltzmann's molecular chaos hypothesis [21]:

> *The molecules are like so many individuals, having the most various states of motion, and the properties of gases only remain unaltered because the number of these molecules which on the average have a given state of motion is constant.*

During the last three decades, physicists have approached social and economic systems from this perspective, looking for universal laws and important statistical patterns, while proposing simple theoretical models to explain the empirical observations. This effort by numerous more or less prominent physicists became what is now known as sociophysics and econophysics [22–27]. The opinion dynamics, and the voting behavior as a proxy of opinion, is still one of the major topics in sociophysics [25–30].

This paper contributes to the understanding and description of the voting behavior from a couple of different points of view. First of all, Lithuania is a young democratic nation and the analysis of the Lithuanian parliamentary elections' data sets seems interesting in the context of similar analyses carried out on the data sets gathered in the mature democratic nations, such as Brazil, England, Germany, France, Finland, Norway, or Switzerland [31–35]. In the political science and sociological literature, one would find numerous previous approaches to the Lithuanian parliamentary elections, for example, [36–40]. Yet most of these approaches had a quite different perspective: most of these papers discuss general electoral trends in the context of social, demographic, and economic changes. For this kind of discussion, highly aggregated (e.g., on a municipal district level) data sets prove to be sufficient, while in this paper we will consider the data on the smallest scale available (polling station level).

Another key contribution of this paper is a simple agent-based model, which is used to explain the statistical patterns uncovered during the empirical analysis. The proposed model is built upon a two-state herding model originally proposed by Kirman in [41]. In the recent years, the two-state herding model was quite frequently and rather successfully applied to reproduce the statistical patterns observed in the empirical data of the financial markets [42–48]. In this paper, we extend the two-state herding model to allow the agents to switch between more than two states. We discuss the similarity between the proposed model and the well-known Voter model [49–54].

Our approach is unique in a sense that we consider reproducing the parties' vote share distribution observed in the Lithuanian parliamentary elections. In the previous literature, there was only a single attempt to model, and predict, popular vote (aggregated vote share) in the Lithuanian parliamentary elections using regression model (see [55]). Numerous previous sociophysics papers have mainly ignored the vote share distribution, likely due to the belief that the vote share distribution reflects electoral sensitivity to the policies promoted by the parties and less due to the endogenous interactions between the voters (a similar argument is given in [31]). To some extent, this belief is supported by game-theoretic models (see [5, 6]). Notably, there were a couple of sociophysics papers considering two-state agent-based models of the voting behavior, for example, voting for or against certain proposals in a referendum [29]. It is interesting to note that, recently, the binary models considered in [29] were used to construct a simple financial market model [56], while we start from the financial market model [47] and move towards the model of the voting behavior. While the vote share distribution was mainly ignored in the previous sociophysics papers, the other statistical patterns arising during the many different elections were considered for the empirical analysis and modeling. A branching process model was proposed to reproduce the individual candidate's (in an open party list) vote share distribution [31]. A network model was used to explain how people decide whether to vote in the municipal elections [34]. A spatial diffusive model for the election turn-out was proposed in [33, 35]. One of more similar approaches was taken by [57], in which a generative model was proposed to reproduce the rank-size distribution of parties' vote share. Another similar approach, taken by [54], considered the vote share distribution observed in the elections of House of Representatives in Japan. The latter approach [54] also used a mean-field Voter model to explain the empirical observations.

This paper is organized as follows. In Section 2, we discuss the Lithuanian parliamentary election system and carry out the empirical analysis. Next, in Section 3, we briefly introduce the two-state herding model and extend it to account for the multiple states. Afterwards, in Section 4, we apply the extended model to reproduce the statistical patterns uncovered during the empirical analysis. Finally, we end the paper with a discussion (see Section 5).

2. Empirical Analysis of the Data from the Lithuanian Parliamentary Elections

Let us start by discussing the parliamentary voting system used in Lithuania. Lithuanian parliamentary elections are held every 4 years. During every election, all of the 141 parliamentary seats are distributed using two-tier voting system. Namely, 71 seats in the parliament are taken by elected district representatives (there are 71 electoral districts in total), while the other 70 seats are distributed according to the popular vote among the parties that received more than 5% of the popular vote. In other words, each individual voter is able to vote for a single candidate to represent his electoral district (the two-round voting system is used) and for an open party list (listing up to 5 individuals from that list). Every electoral district has multiple polling stations (their number varies over the years), which further subdivide the electoral districts. Every eligible voter is assigned to a single polling station based on location of their residence. Each of the polling stations may have widely different number of the assigned voters—some of the smallest polling stations have as few as 100 assigned voters, while the largest have up to 7000 assigned voters.

In this paper, we consider only votes cast for the open party lists in each of the local polling stations. We do not analyze ranking of the individuals on the party lists (similar analysis was carried out in, e.g., [31]), voting for the representative of electoral district (similar data was previously considered in, e.g., [54, 57]), nor turnout rates (modeling and analysis of which were previously considered in, e.g., [33, 35]). We ignore votes cast in the polling stations abroad or votes cast by post. In the analysis that follows, we consider only parties that were elected to the parliament (total vote share larger than 5%), while all other less successful parties were combined into a single party, which we have labeled as the "Other" party.

In this paper, we consider the three data sets from the Lithuanian parliamentary elections of 1992, 2008, and 2012. All of the original data sets were made publicly available by the Central Electoral Commission of the Republic of Lithuania (at https://www.rinkejopuslapis.lt/ataskaitu-formavimas). We have downloaded the original data sets from the website on August 31, 2016. During the preliminary phase of the empirical analysis, we have found some small inconsistencies within the original data. The original 1992 election data set had seven polling stations with incorrect total vote counts. We have identified three pairs of polling stations which were, most likely, swapped among themselves as the number of missing votes in one polling station matched the number of surplus votes in the other, while we have dealt with the remaining polling stations by simply adjusting the total vote count to match the sum of votes cast for each of the parties in that polling station. We have also found that data from 51 (out of 2034) polling stations is missing (the data was filled with zeros) from the original 2008 election data set. We have not identified any issues with the original 2012 election data set. These minor inconsistencies would not impact the overall result of any of the considered elections nor the results reported in this section. We have made the modified data sets available online at https://github.com/akononovicius/lithuanian-parliamentary-election-data.

In the analysis that follows, we consider the parties' vote share distribution across each polling station. The vote share, v_{ij}, is defined as total number of votes cast for the party V_{ij} divided by the total number of votes cast in that polling station:

$$v_{ij} = \frac{V_{ij}}{\sum_{k=1}^{K} V_{kj}}; \tag{1}$$

here index i varies over the parties (K is the total number of parties participating in the election) and index j varies over the polling stations. We consider the probability and rank-size distributions of v_{ij} across all of the polling stations during the same parliamentary election. The probability distribution is estimated using the standard probability density functions (abbr. PDFs). The rank-size distributions are often used if the data varies significantly in scale, for example, word occurrence frequency [58], earthquake magnitudes [59], city sizes [60], and cross country income distributions [61]. When using this technique, the original empirical data is sorted in descending order. Afterwards, the sorted data is plotted with the rank being the abscissa coordinate and the actual value being the ordinate coordinate. In our case, we sort the parties' vote shares, v_{ij}, for each party i separately to produce $\tilde{v}_{ik}$, for which

$$\tilde{v}_{i1} \geq \tilde{v}_{i2} \geq \cdots \geq \tilde{v}_{iM} \tag{2}$$

is true. In the above, M is the total number of polling stations, so that index k represents the rank. Note that, in this representation, the same polling station may be ranked differently for different parties (namely, k might be different for the same polling station for different i). Evidently, these two approaches, PDFs and rank-size distributions, are interrelated, but using both of them allows uncovering different statistical patterns.

2.1. The Parliamentary Election of 1992. The parliamentary election of 1992 was held in 2061 local polling stations. 17 parties competed in the parliamentary election, but only 4 of them were able to obtain more than 5% of popular vote. For the sake of simplicity, we will use the following abbreviations for these parties: SK, "SąjūdÂžio koalicija," LSDP, "Lietuvos socialdemokratų partija," LKDP, "Lietuvos KrikÂščionių demokratų partijos, Lietuvos politinių kalinių ir tremtinių sąjungos ir Lietuvos demokratų partijos jungtinis sąraÂšas," and LDDP, "Lietuvos demokratinė darbo partija." We have combined the other 13 parties to form the "Other" party (abbr. O) and considered the votes cast for the combined "Other" party alongside the votes cast for the 4 main parties.

As you can see from Figures 1 and 2 and Table 1, all of the parties with a notable exception of the "Other" party are very well fitted by assuming that data is distributed according to the Beta distribution, PDF of which is given by

$$p\left(v_{ij}\right) = \frac{\Gamma\left(\alpha\right)\Gamma\left(\beta\right)}{\Gamma\left(\alpha+\beta\right)} v_{ij}^{\alpha-1}\left(1-v_{ij}\right)^{\beta-1}. \tag{3}$$

The "Other" party stands out, because it includes "Lietuvos lenkų sąjunga" party (abbr. LLS; en. Association of Poles in Lithuania). The LLS party had heavily relied on the support of the ethnic minorities, which were spatially segregated. Namely, the representatives of ethnic minorities mostly live in larger cities and Vilnius County. The observed spatial segregation could easily cause the segregation observed in the voting data.

In Figure 3, we confirm this intuition by splitting the LLS party away from the "Other" party. After the split, the rank-size distribution of the "Other" party is well approximated by the Beta distribution with parameters $\alpha = 4.9$ and $\beta = 35.4$ ($R^2_{\mathrm{RS}} = 0.992$). To provide a good fit for the LLS rank-size distribution, we assume that underlying data is distributed according to a mixture of the two Beta distributions: one (95% of points) with parameters $\alpha_1 = 0.08$ and $\beta_1 = 10$ and the other (5% of points) with parameters $\alpha_2 = 1.22$ and $\beta_2 = 1.37$.

The parties' vote share rank-size distributions were previously considered in [57]. Unlike in this paper, Fenner and others assumed that the parties' vote share is distributed according to Weibull distribution; they have obtained rather good fits for the UK election data. Yet fits provided by the Beta distribution, are also rather good. We believe that Beta distribution is superior for this purpose from the theoretical point of view. Namely, Beta distribution has reasonable support, as probabilities are defined for $v \in [0; 1]$, while Weibull distribution needs to be arbitrary truncated, as probabilities are defined for $v \in [0; +\infty)$. Interestingly, Fenner and others also use a mixture distribution (of two Weibull distributions) to fit the UK election data. Similar observations were also made when studying Brazilian presidential election data [62]. In [63], it was noted that multiple different distributions, Weibull, log-normal, and normal, provide good fits for the distribution of religions' adherents. To discriminate between the possibilities, a deeper theoretical insight is needed.

2.2. The Parliamentary Election of 2008. The parliamentary election of 2008 was held in 2034 polling stations, yet we have only 1983 points in the data set as the data from 51 polling stations is missing. In this election, a slightly smaller number of parties had participated (16), but now 7 of them

Table 1: Parameters of the Beta distribution, α and β, used to fit the data in Figures 1 and 2 as well as wellness of fit for the PDFs, R^2_{PDF}, and the rank-size distributions, R^2_{RS}.

Party	α	β	R^2_{PDF}	R^2_{RS}
SK	3.9	16.6	0.956	0.994
LSDP	2.7	51	0.935	0.953
LKDP	2.2	16	0.926	0.995
LDDP	5.7	6.1	0.907	0.998
O	3	19.4	0.895	0.854

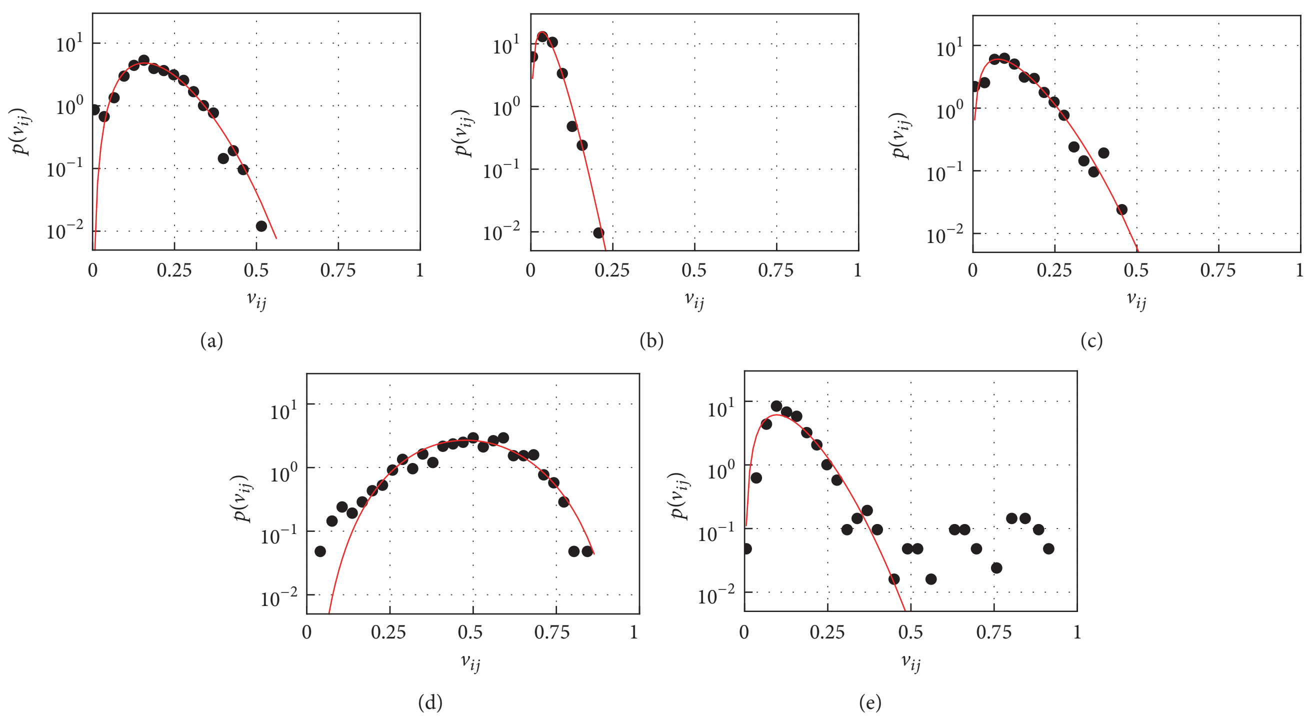

Figure 1: The vote share PDFs of the most successful parties during the 1992 election. The following parties were considered: SK (a), LSDP (b), LKDP (c), LDDP (d), and O (e). The empirical values are shown as black circles, while theoretical fits using the Beta distribution are shown as solid curves. The values of the Beta distribution parameters are given in Table 1.

were able to obtain more than 5% of popular vote. For the sake of simplicity, we will use the following abbreviations for them: LSDP, "Lietuvos socialdemokratų partija" (formed by LSDP and LDDP, which participated in the 1992 election), TS-LKD, "Tėvynės sąjunga—Lietuvos krikÂščionys demokratai" (could be considered to be a successor of the SK and LKDP, which participated in the 1992 election), TPP, "Tautos prisikėlimo partija," DP, "Koalicija Darbo partija + jaunimas," LRLS, "Lietuvos Respublikos liberalų sąjūdis," TT, "Partija Tvarka ir teisingumas," and LiCS, "Liberalų ir centro sąjunga." As previously all other parties (9 of them) were combined to form the "Other" party (abbr. O), we have considered the votes cast for the combined "Other" party alongside the votes cast for the 7 main parties.

As is evident from Figures 4 and 5 in the 2008 election, the vote share distributions of most of the parties are well fitted by a mixture of two Beta distributions. Although now there is no clear-cut explanation for this phenomenon, we would like to conjecture that this observation indicates that other spatial segregations (e.g., by income) of voters in Lithuania have started to play an important role. Namely, most of the parties could now be identified with specific socioeconomic classes; for example, some party starts favoring higher income voters (gaining the support in cities), consequently losing the support of poorer voters (losing the support in rural areas).

2.3. The Parliamentary Election of 2012. The parliamentary election of 2012 was held in 2017 polling stations (thus we have 2017 data points). 18 parties had participated in the election, while 7 of them were able to obtain more than 5% of popular vote. For the sake of simplicity, we will use the following abbreviations for them: LRLS, "Lietuvos Respublikos liberalų sąjūdis," DP, "Darbo partija," TS-LKD, "Tėvynės sąjunga—Lietuvos KrikÂščionys demokratai," DK, "Drąsos kelias politinė partija," LLRA, "Lietuvos lenkų rinkimų akcija," LSDP, "Lietuvos socialdemokratų partija,"

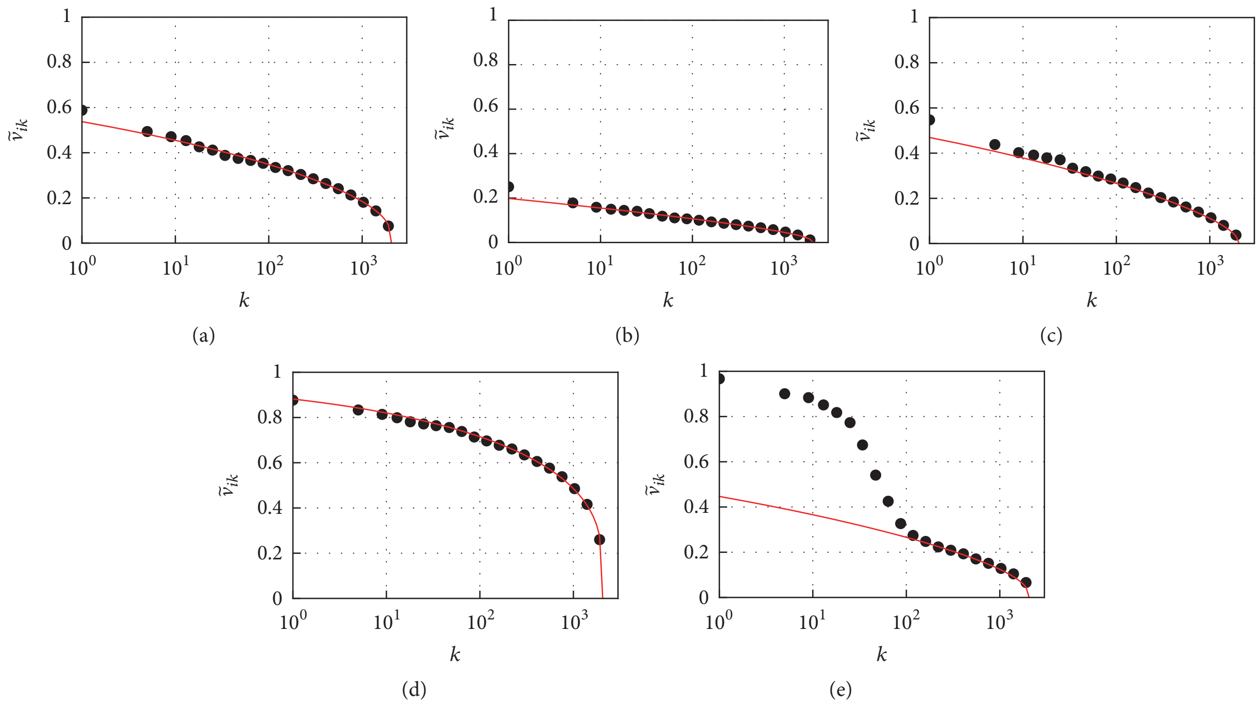

Figure 2: The rank-size distribution of the most successful parties during the 1992 election. The following parties were considered: SK (a), LSDP (b), LKDP (c), LDDP (d), and O (e). The empirical values are shown as black circles, while theoretical fits using the Beta distribution are shown as solid curves. The values of the Beta distribution parameters are given in Table 1.

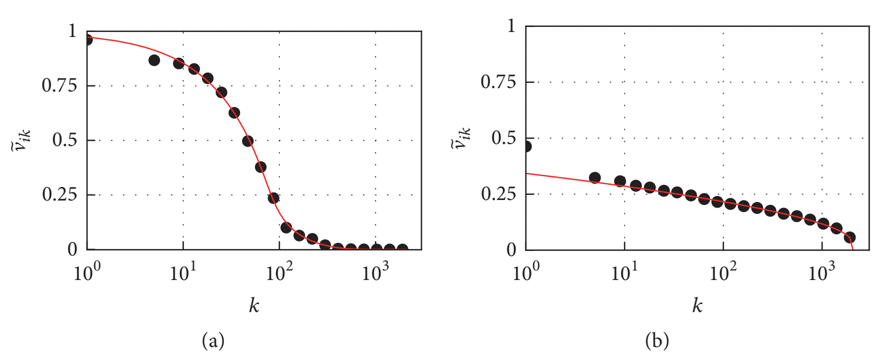

Figure 3: The rank-size distributions of the LLS party (a) and the other parties in the O party (b). The empirical values are shown as black circles, while solid curves provide fits by assuming that $v_{ij} \sim 0.95 \cdot \mathscr{B}e(0.08, 10) + 0.05 \cdot \mathscr{B}e(1.22, 1.37)$ ($R^2_{\mathrm{RS}} = 0.997$; (a)) and $v_{ij} \sim \mathscr{B}e(4.9, 35.4)$ ($R^2_{\mathrm{RS}} = 0.992$; (b)).

and TT, "Partija Tvarka ir teisingumas." Matching abbreviations indicate the same (or mostly the same) parties as in 2008 election. Once again the less-successful parties were combined to form the "Other" party (abbr. O). The votes cast for the combined "Other" party were analyzed alongside the votes cast for the 7 main parties.

Once again, as well as in the 2008 parliamentary election data set, it is evident that the vote share distributions of most of the parties in the 2012 parliamentary are also well described by a mixture of two Beta distributions (see Figures 6 and 7).

3. A Multistate Agent-Based Model of the Voting Behavior

In this section, we propose a simple multistate agent-based model, which is to describe the voting behavior within a small nonspecific geographic region covered by a single polling station. Unlike in some previous approaches [13–20], our aim is not to incorporate complex ideas from psychology but to reproduce the empirical parties' vote share distribution. It is known that an agent-based herding model proposed by

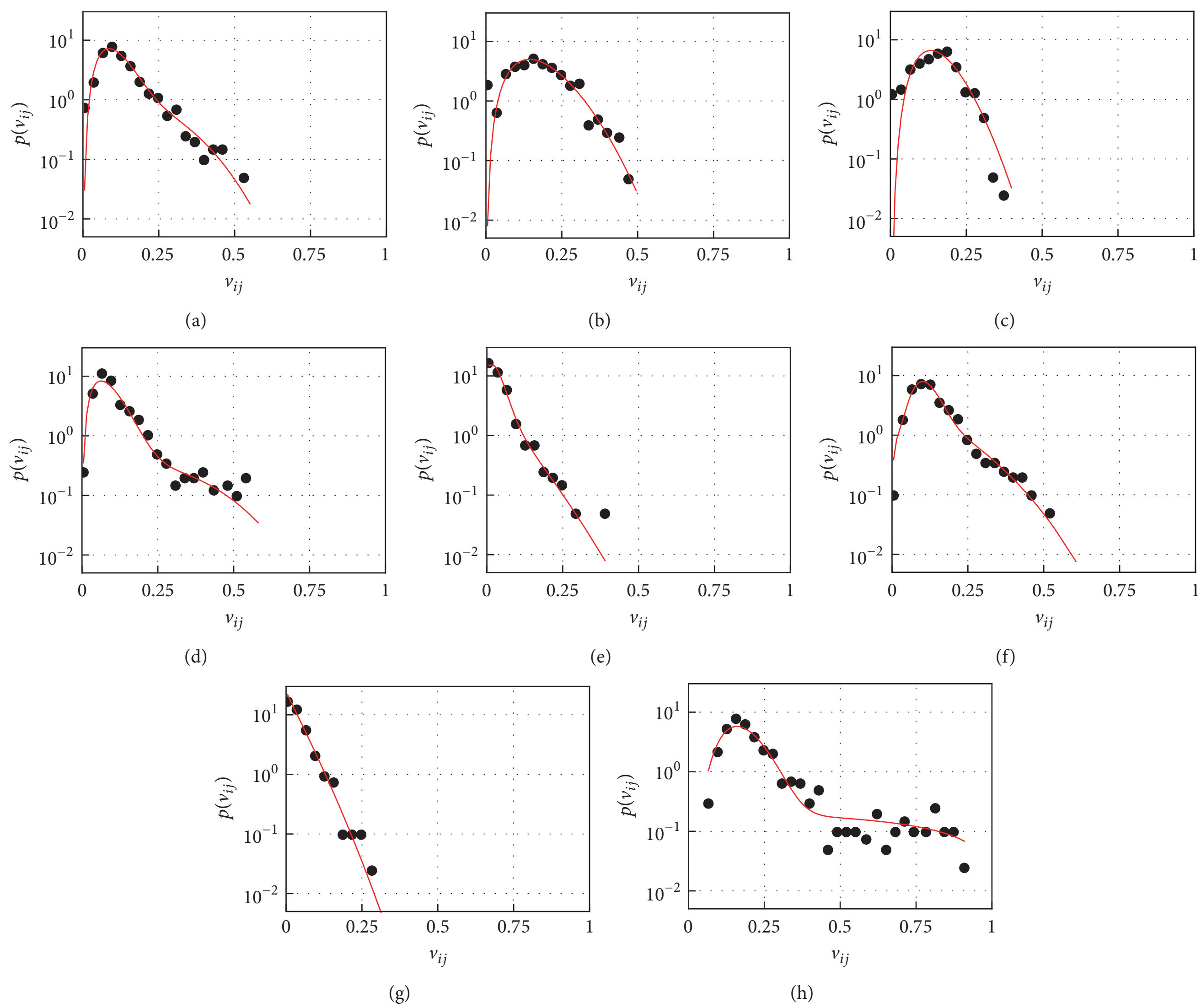

Figure 4: The vote share PDFs of the most successful parties during the 2008 election. The following parties were considered: LSDP (a), TS-LKD (b), TPP (c), DP (d), LRLS (e), TT (f), LiCS (g), and O (h). The empirical values are shown as black circles, while theoretical fits using a mixture of Beta distributions are shown as solid curves (the values of the parameters are given in Table 2).

Kirman, in [41], reproduces Beta distribution. So let us start by introducing Kirman's herding model.

Originally in [41] Kirman noted that biologists and economists observe similar behavioral patterns. Apparently both ants and people show interest in things that are more popular among their peers regardless of their objective properties [64–68]. In [41], a simple two-state model was proposed to explain these observations. In the contemporary interpretations of Kirman's model, the following mathematical form of the one-step transition probabilities is used (see [42–45]):

$$P(X \longrightarrow X+1) = (N-X)(\sigma_1 + hX)\Delta t,$$
$$P(X \longrightarrow X-1) = X[\sigma_2 + h(N-X)]\Delta t; \quad (4)$$

here N is the total number of agents acting in the system, X is the total number of agents occupying the first state (consequently there are $N-X$ agents occupying the second state), σ_i are the perceived attractiveness parameters (may differ for different states), h is a recruitment efficiency parameter, and Δt is a relatively short time step.

Let us now show that the two-state model produces Beta distribution. This can be done by using birth-death process formalism, which is well described in [69]. The dynamics of $x = X/N$ (let N be large) can be alternatively described by the following master equation:

$$\frac{\Delta\omega(x,t)}{\Delta t} = N^2 \left\{(\mathbf{E}-1)\left[\pi^-(x)\,\omega(x,t)\right] + \left(\mathbf{E}^{-1}-1\right)\left[\pi^+(x)\,\omega(x,t)\right]\right\}; \quad (5)$$

here $\omega(x,t)$ is time-dependent distribution, $\pi^{\pm}(x)$ are the transition rates per unit of time (defined as $\pi^{\pm}(x) = p(X \to X \pm 1)/N^2 h\Delta t$), and $\mathbf{E}$ and $\mathbf{E}^{-1}$ are the one-step increment and

Table 2: The parameters of a mixture of Beta distributions: here c is a weight of $\mathscr{B}e(\alpha_2, \beta_2)$, used to fit the data in Figures 4 and 5 as well as wellness of fit for the PDFs, R^2_{PDF}, and the rank-size distributions, R^2_{RS}.

Party	α_1	β_1	c	α_2	β_2	R^2_{PDF}	R^2_{RS}
LSDP	3.9	31.7	0.15	4.3	12.9	0.968	0.999
TS-LKD	3.7	16.8	0	—	—	0.915	0.999
TPP	5.1	27.8	0	—	—	0.884	0.992
DP	3	30.3	0.09	3.2	7.7	0.942	0.992
LRLS	2.7	67.9	0.54	0.6	12.6	0.986	0.994
TT	7.6	59.5	0.42	1.8	9.4	0.987	0.992
LiCS	0.98	23.5	0	—	—	0.955	0.993
O	6.6	30.4	0.15	1.2	1.6	0.944	0.995

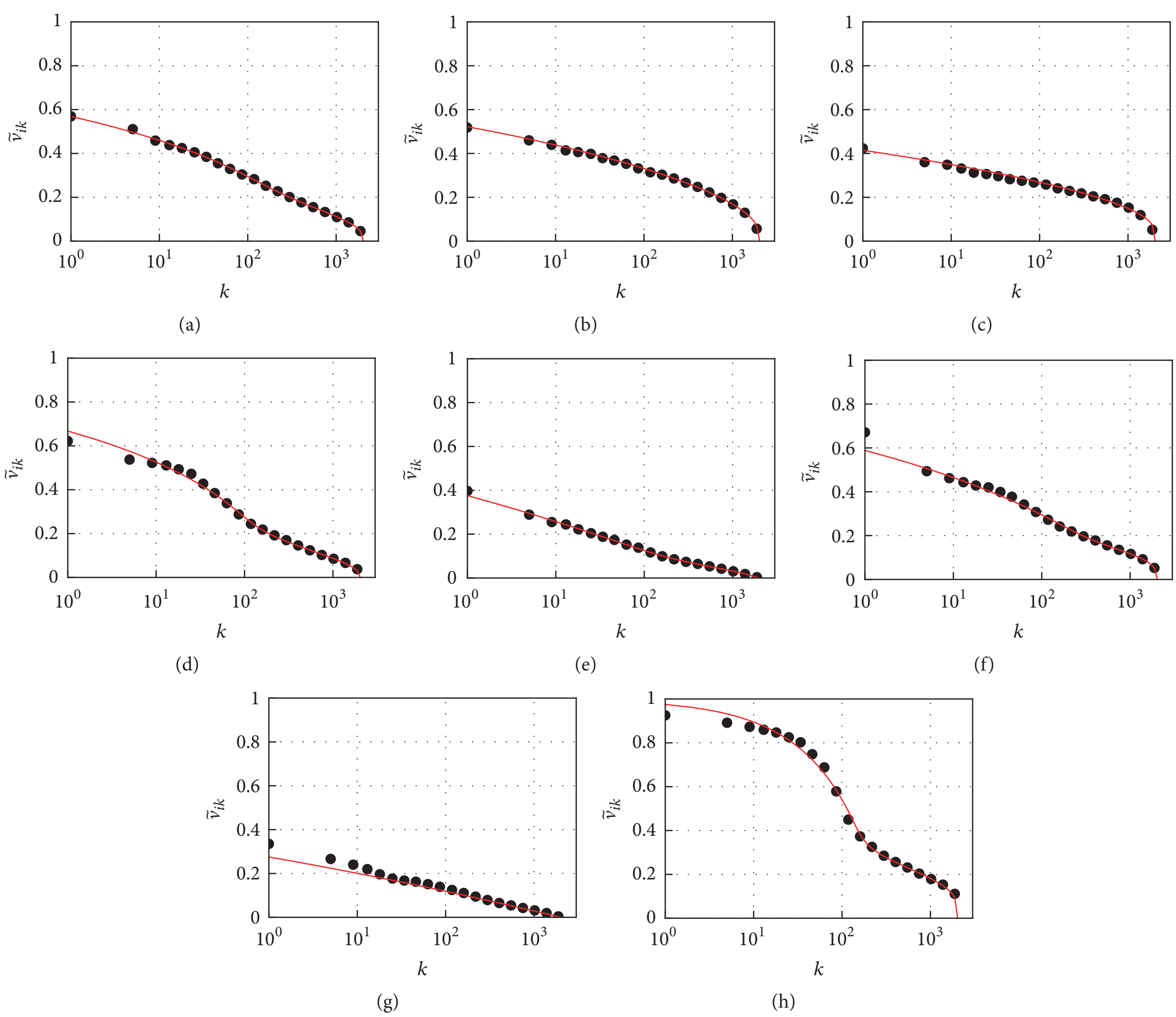

Figure 5: The rank-size distributions of the most successful parties during the 2008 election. The following parties were considered: LSDP (a), TS-LKD (b), TPP (c), DP (d), LRLS (e), TT (f), LiCS (g), and O (h). The empirical values are shown as black circles, while theoretical fits using a mixture of Beta distributions are shown as solid curves (the values of the parameters are given in Table 2).

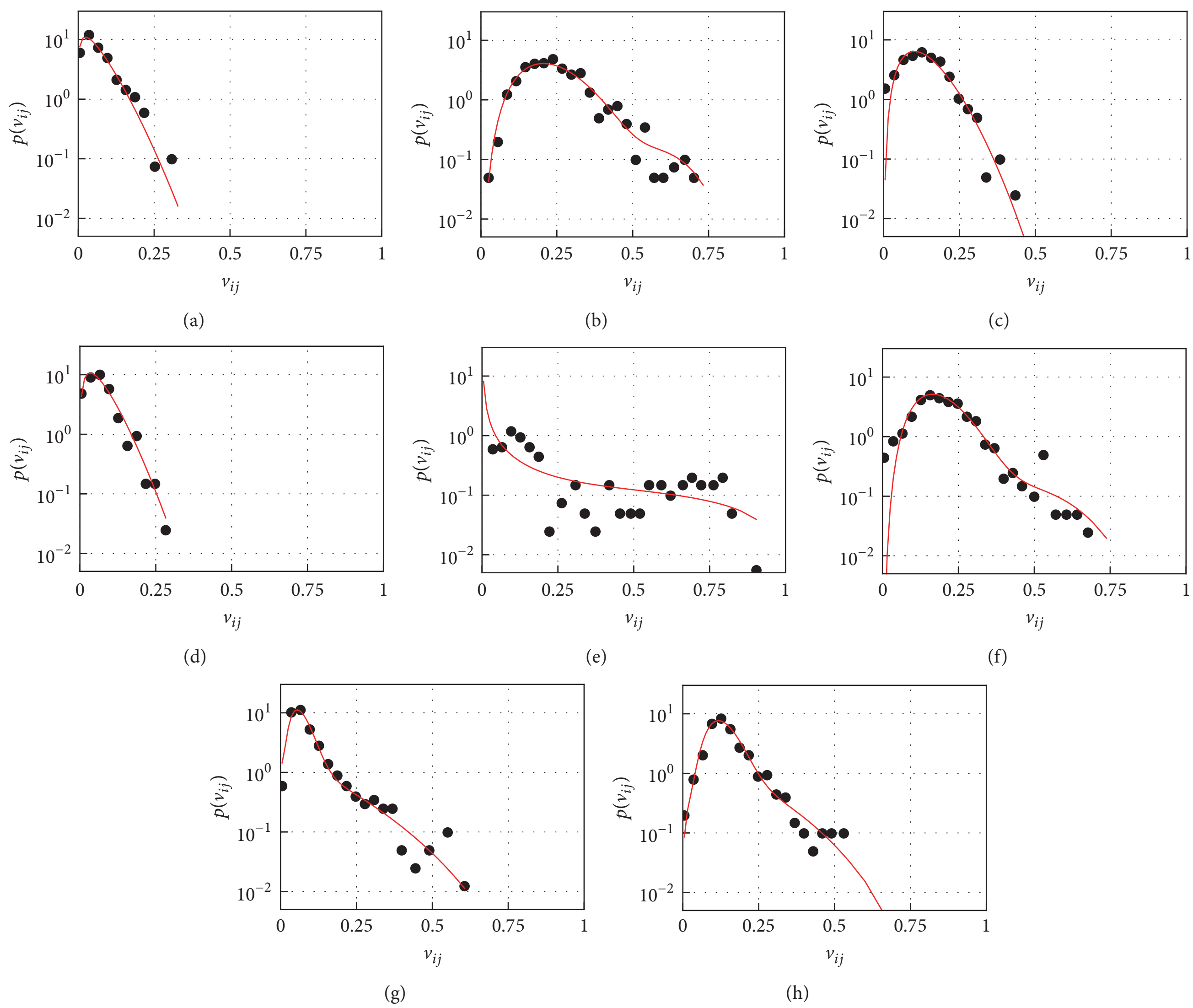

FIGURE 6: The vote share PDFs of the most successful parties during the 2012 election. The following parties were considered: LRLS (a), DP (b), TS-LKD (c), DK (d), LLRA (e), LSDP (f), TT (g), and O (h). The empirical values are shown as black circles, while theoretical fits using a mixture of Beta distributions are shown as solid curves (the values of parameters are given in Table 3).

decrement operators. Let us expand these operators in Taylor series up to the second-order term:

$$\begin{aligned} \mathbf{E}\left[f(x)\right] &= f(x+\Delta x) \\ &\approx f(x) + \Delta x \frac{d}{dx} f(x) + \frac{\Delta x^2}{2}\frac{d^2}{dx^2} f(x), \\ \mathbf{E}^{-1}\left[f(x)\right] &= f(x-\Delta x) \\ &\approx f(x) - \Delta x \frac{d}{dx} f(x) + \frac{\Delta x^2}{2}\frac{d^2}{dx^2} f(x); \end{aligned} \tag{6}$$

here $\Delta x = 1/N$. Putting these Taylor expansions back into the master equation as well as taking small time step limit yields the following Fokker-Planck equation:

$$\begin{aligned} \frac{\partial}{\partial t}\omega(x,t) &= -\frac{\partial}{\partial x}\left[A(x)\,\omega(x,t)\right] \\ &\quad + \frac{1}{2}\frac{\partial^2}{\partial x^2}\left[B(x)\,\omega(x,t)\right], \end{aligned}$$

$$\begin{aligned} A(x) &= N\left[\pi^{+}(x) - \pi^{-}(x)\right] = \varepsilon_1 (1-x) - \varepsilon_2 x, \\ B(x) &= \pi^{+}(x) + \pi^{-}(x) = 2x(1-x) + \mathcal{O}\left(N^{-1}\right), \end{aligned} \tag{7}$$

where $\varepsilon_i = \sigma_i/h$. Steady-state distribution of this Fokker-Planck equation can be obtained by solving

$$A(x)\,\omega_{\mathrm{st}}(x) - \frac{1}{2}\frac{\partial}{\partial x}\left[B(x)\,\omega_{\mathrm{st}}(x)\right] = 0. \tag{8}$$

In general case, the solution of this ordinary differential equation is given by

$$\omega_{\mathrm{st}}(x) = \frac{C_0}{B(x)} \exp\left[2\int^{x} \frac{A(u)}{B(u)}\,du\right], \tag{9}$$

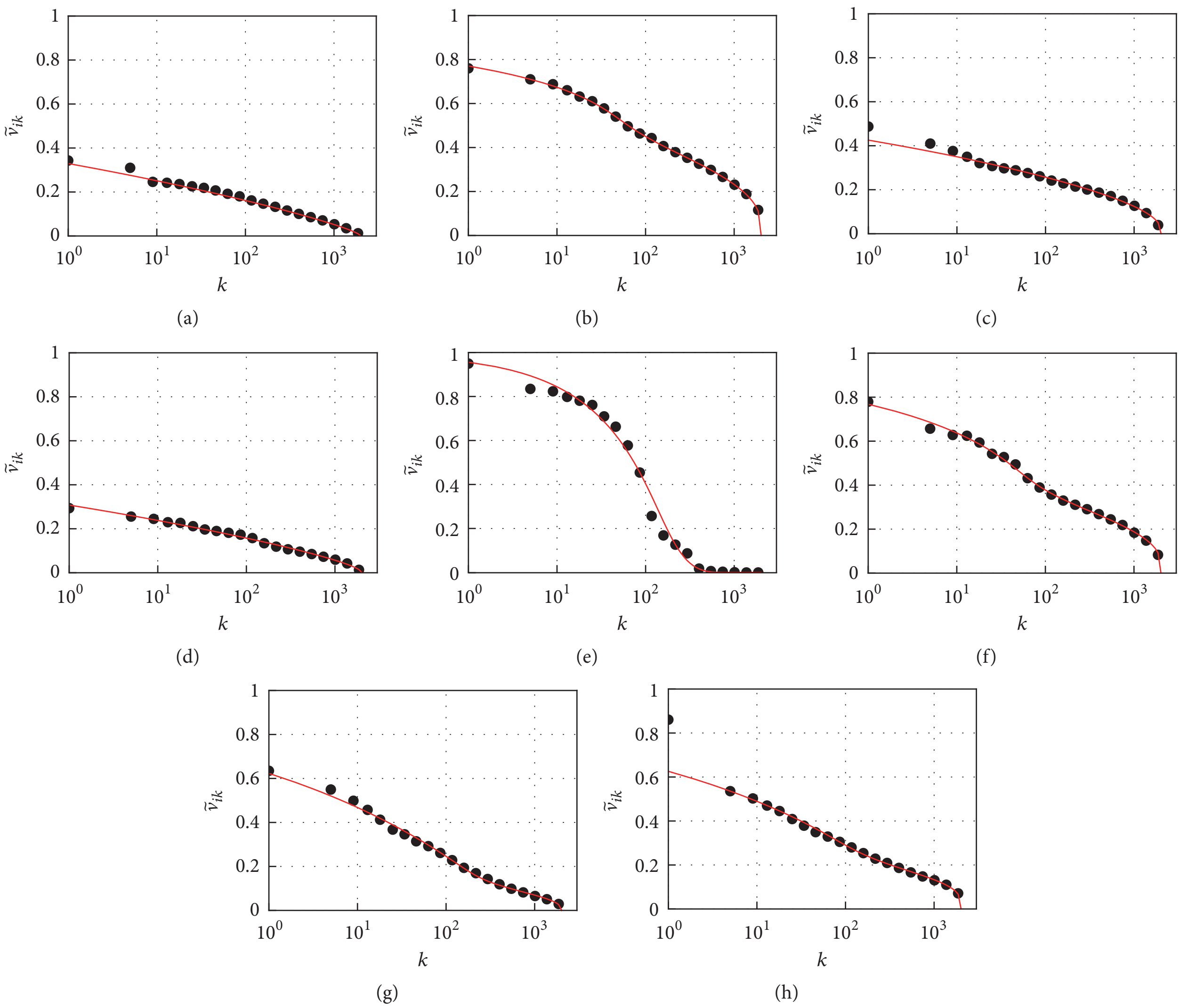

Figure 7: The rank-size distributions of the most successful parties during the 2012 election. The following parties were considered: LRLS (a), DP (b), TS-LKD (c), DK (d), LLRA (e), LSDP (f), TT (g), and O (h). The empirical values are shown as black circles, while theoretical fits using a mixture of Beta distributions are shown as solid curves (the values of parameters are given in Table 3).

where C_0 is normalization constant. In our specific case, we obtain a PDF for the Beta distribution, $\mathscr{B}e(\varepsilon_1, \varepsilon_2)$:

$$\omega_{\mathrm{st}}(x) = \frac{C_0}{2} x^{\varepsilon_1 - 1} (1 - x)^{\varepsilon_2 - 1}. \tag{10}$$

As in the typical parliamentary election, there are more than two competitors; we need to generalize the model to incorporate more than two states. From the conservation of the total number of agents N, we have

$$P(X_i \longrightarrow X_i \pm 1) = \sum_{j \neq i} P(X_i \longrightarrow X_i \pm 1,\ X_j \longrightarrow X_j \mp 1). \tag{11}$$

Assuming that the right-hand side probabilities have the same form as (4), we obtain

$$\begin{aligned} P(X_i \longrightarrow X_i + 1) &= \sum_{j \neq i} X_j \left(\sigma_{ji} + h_{ji} X_i\right) \Delta t, \\ P(X_i \longrightarrow X_i - 1) &= X_i \sum_{j \neq i} \left[\sigma_{ij} + h_{ij} X_j\right] \Delta t. \end{aligned} \tag{12}$$

These one-step transition probabilities for X_i depend not only on X_i (as in the two-state model) but also on all other X_j (here $j \neq i$). To circumvent this potentially cumbersome dependence, one needs to assume that $\sigma_{ij} = \sigma_j$ and $h_{ij} = h$ (where $j \neq i$). The first assumption, $\sigma_{ij} = \sigma_j$, means that the perceived attractiveness of any party does not depend on who is attracted to it. While the second assumption, $h_{ij} = h$, means that the recruitment mechanism is uniform (symmetric and independent of interacting agents). Note that these assumptions contrast with the assumptions underlying the bounded confidence model [15, 16]. Yet these assumptions are needed to ensure that $x_i = X_i/N$ is distributed according to the Beta distribution. After making these assumptions, we can further simplify the one-step transition probabilities:

$$P(X_i \longrightarrow X_i + 1) = (N - X_i)(\sigma_i + hX_i)\Delta t,$$
$$P(X_i \longrightarrow X_i - 1) = X_i(\sigma_{-i} + h[N - X_i])\Delta t; \quad (13)$$

here $\sigma_{-i} = \sum_{j \neq i} \sigma_j$ is the total attractiveness of all of the competitors of party i. By the analogy with the two-state model, it should be evident that

$$x_i \sim \mathscr{B}e(\varepsilon_i, \varepsilon_{-i}). \quad (14)$$

Unlike the two-state model, it seems impossible to provide a useful general aggregated macroscopic description, using the Fokker-Planck equation or a set of stochastic differential equations, of the generalized M-state model. In [70], a three-state model was considered and given an aggregated macroscopic description, by a system of two stochastic differential equations, yet it was possible only under specific conditions.

The one-step transition probabilities, while describing agent's behavior, are still aggregate description of individual agent-level dynamics. So some discussion on what do (13) represent is relevant. Selecting one random agent, per time step, and setting his switching probability to $\varepsilon_{-i}\Delta t_s$ gives us idiosyncratic behavior term, $X_i\varepsilon_{-i}\Delta t_s$, while selecting another random agent and, if both agents vote for the different parties, allowing the first agent to copy the second agent's voting preference gives us the recruitment term, $X_i(N - X_i)\Delta t_s$. This description of agent-level dynamics could be further generalized to allow the model to be run on the randomly generated networks [71, 72]. This agent-based algorithm might be seen to be a special case of the well-known Voter model [25, 49–51, 53, 54].

4. The Modeling of the Parliamentary Election

Now let us apply the proposed model to reproduce the empirical vote share PDFs and rank-size distributions, which were observed during the 1992 parliamentary election. Here we consider only the simplest case by ignoring the "Other" party and in this way removing distortions caused by the ethnic segregation (see the discussion in Section 2.1). We do not consider the segregated data, the full data of the 1992 parliamentary election, or the data of the 2008 and 2012 parliamentary elections, as to account for the vote segregation a more sophisticated approach is needed. Namely, in order to account for the full complexity of the empirical data, additional information, such as the spatial polling data or the sociodemographic data, would be needed, although, in general, one could try to infer the correct partition of the polling stations, where the vote share distribution of each party would be modeled using the proposed model using the same parameter set.

In Figures 8 and 9, we compare the vote share PDFs and the rank-size distributions numerically generated by the proposed model, (13), and the respective empirical vote share distributions of the 1992 parliamentary election. Yet we cannot use the previously empirically estimated Beta distribution parameters, by assuming $\alpha_i = \varepsilon_i$ and $\beta_i = \varepsilon_{-i}$, as model parameters, because an important model implication, $\sum_{l \neq i} \alpha_l = \beta_i$, does not hold for the empirical data. Yet one may obtain the parameter values by fitting the empirical data with the model implication in mind.

As can be seen from Figures 8 and 9 and Table 4, the proposed model excellently fits three of the four parties, while for the LSDP the fit is not as good as one might expect. Note that model overestimates the success of LSDP (the solid curve is above black circles for small k in subfigure (d)) and underestimates the electoral support of LDDP (the solid curve is below black circles for small k in subfigure (h)). Thus, it is likely that the LSDP had small perceived chance to win the 1992 election (their aggregated vote share was near 5%); thus voters who would actually consider voting for the LSDP cast their votes for the other left-wing party, which had better perceived chance at winning the 1992 election (aggregated vote share of LDDP was above 40%).

We can check this intuition by violating the assumption that the perceived attractiveness should not depend on the current state of the agent (agent's currently supported party); namely, instead of ε_i we now have ε_{ji}. Previously, this assumption was needed to ensure that vote share is distributed according to the Beta distribution. Let us introduce a single exception: if j corresponds to LSDP and i corresponds to LDDP, then ε_{ji} can be different in value from ε_i. This gives us the following matrix of ε_{ji} values (the numeric indices are assigned according to Table 4):

$$\varepsilon = \begin{pmatrix} 0 & 1.13 & 2.27 & 8.87 \\ 3.52 & 0 & 2.27 & 22 \\ 3.52 & 1.13 & 0 & 8.87 \\ 3.52 & 1.13 & 2.27 & 0 \end{pmatrix}. \quad (15)$$

Note that the diagonal elements of the matrix ε are set to zero, as it is not possible to switch to the state the agent is already in. As can be seen in Figures 10 and 11, the fit provided by the model has significantly improved for the LSDP by just making this small change.

5. Conclusions

In this paper, we have considered the parties' vote share PDFs and the rank-size distributions observed during the Lithuanian parliamentary elections. Namely, we have considered the 1992, 2008, and 2012 parliamentary elections' data sets. We have determined that the empirical vote share PDFs and the rank-size distributions are rather well fitted by assuming that the underlying distribution is the Beta distribution or a mixture of two Beta distributions. Reviewing literature, we have found that [54, 57, 62, 63] have reported somewhat similar results. In [57, 62], it was reported that the empirical data is rather well fitted by a mixture of Weibull distributions. In [63], it was noted that multiple different distributions, Weibull, log-normal, and normal, provide good fits for the distribution of religions' adherents. We argue that the Beta distribution is more suitable as it has correct

Table 3: The parameters of a mixture of Beta distributions: here c is a weight of $\mathcal{B}e(\alpha_2, \beta_2)$, used to fit the data in Figures 6 and 7 as well as wellness of fit for the PDFs, R^2_{PDF}, and the rank-size distributions, R^2_{RS}.

Party	α_1	β_1	c	α_2	β_2	R^2_{PDF}	R^2_{RS}
LRLS	1.5	22	0	—	—	0.969	0.994
DP	4.5	14.5	0.03	15.3	11.1	0.973	0.999
TS-LKD	3.4	22	0	—	—	0.938	0.982
DK	1.9	26	0	—	—	0.946	0.994
LLRA	0.06	1.9	0.05	2.1	1.9	0.857	0.990
LSDP	5	21.4	0.05	5.5	6.6	0.948	0.997
TT	4.6	62	0.29	1.1	6.7	0.963	0.995
O	7	45.8	0.23	2	8	0.966	0.970

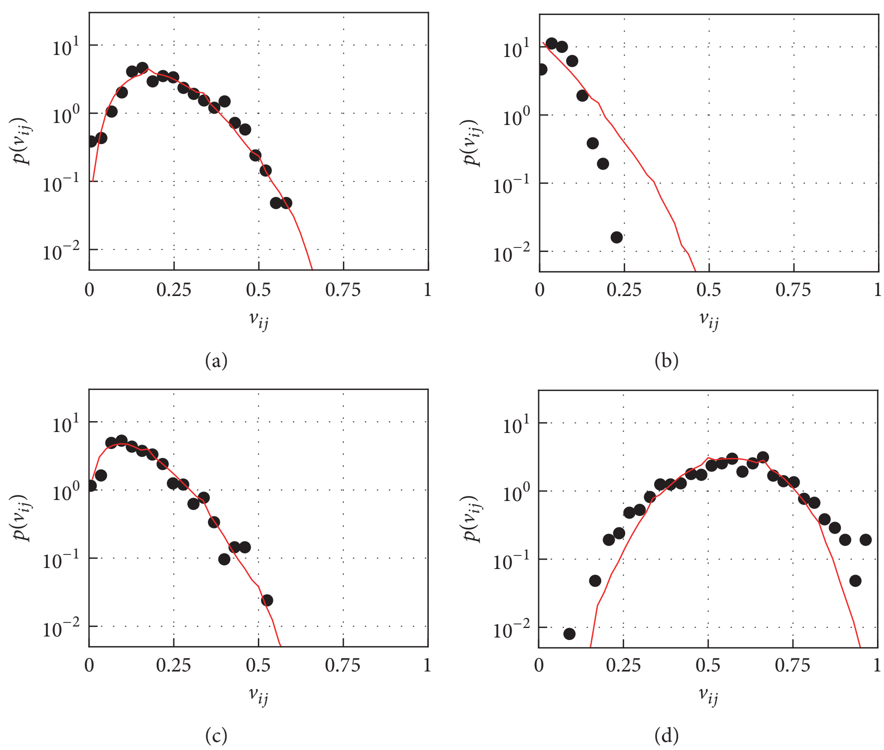

Figure 8: The vote share PDFs of the most successful parties during the 1992 election fitted by the proposed model. Only the following parties were considered: SK (a), LSDP (b), LKDP (c), and LDDP (d). The empirical values are shown as black circles, while solid curves represent numerical results obtained from the proposed model, driven by (13) (the same model run as in Figure 9). The parameters of the model are given in Table 4.

Table 4: The parameters of the proposed model used to reproduce the vote share PDFs and the rank-size distributions of the 1992 election as well as wellness of fit for the respective distributions.

Party	ε_i	R^2_{PDF}	R^2_{RS}
SK	3.52	0.951	0.997
LSDP	1.13	0.705	0.881
LKDP	2.27	0.952	0.998
LDDP	8.87	0.904	0.973

support (probabilities are defined for $v \in [0;1]$, although the other distributions could be arbitrary truncated) and it arises from a simple easily tractable agent-based model. From our empirical analysis, it follows that the mixture of Beta distributions is needed to fit the data if there is underlying spatial segregation of the electorate. In [54], it is also reported that the empirical data is rather well fitted by Beta distribution, which arises from a noisy Voter model. Yet [54] did not observe the vote share segregation pattern.

Having in mind the stark difference between the psychologically motivated models, such as bounded confidence model [15, 16], we would like to point out that the observed statistical patterns as well as the applicability of the model could arise due to numerous unrelated reasons. One of the alternative possibilities would be the people mobility patterns. In the proposed model, a single agent switching from supporting one party to supporting another party could

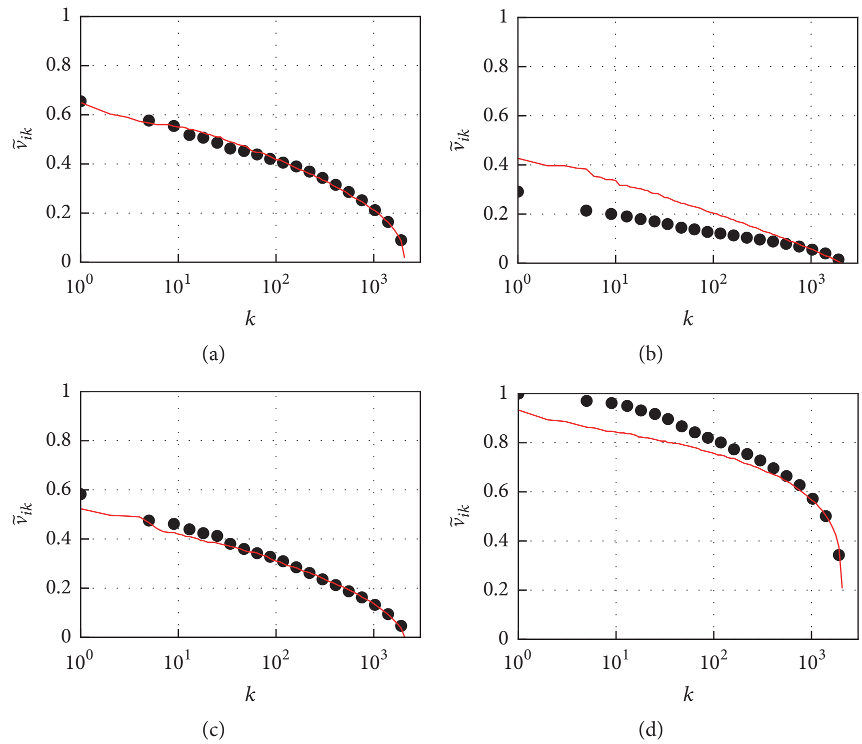

FIGURE 9: The rank-size distributions of the most successful parties during the 1992 election fitted by the proposed model. Only the following parties were considered: SK (a), LSDP (b), LKDP (c), and LDDP (d). The empirical values are shown as black circles, while solid curves represent numerical results obtained from the proposed model, driven by (13) (the same model run as in Figure 8). The parameters of the model are given in Table 4.

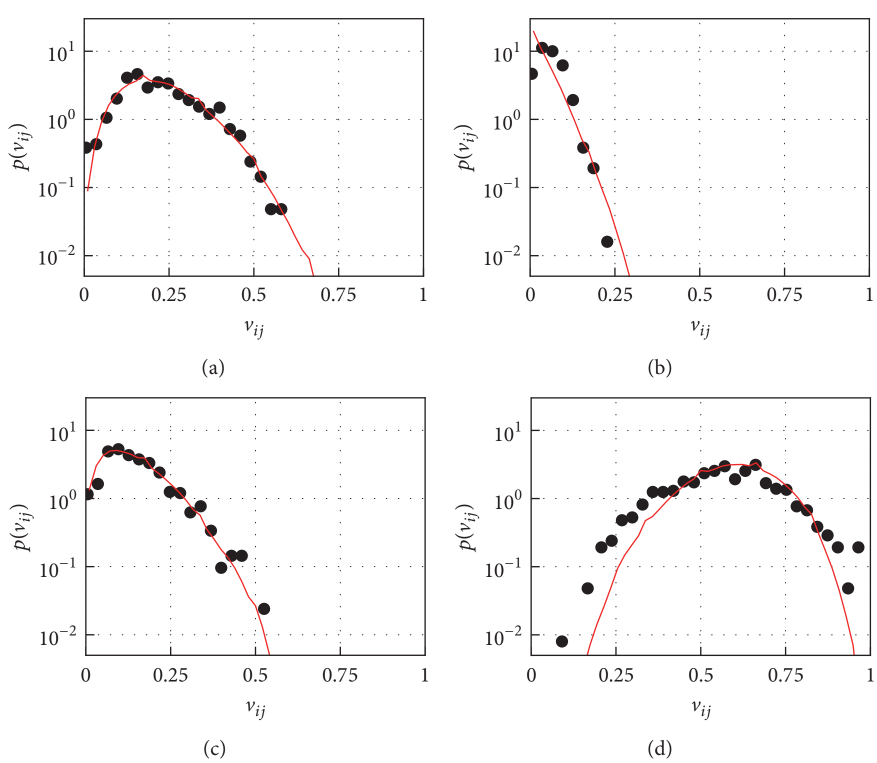

FIGURE 10: The vote share PDFs of the most popular parties during 1992 election fitted by the asymmetric model. Only the following parties were considered: SK (a), LSDP (b), LKDP (c), and LDDP (d). The empirical values are shown as black circles, while solid curves represent data numerically generated by the proposed model with attractiveness dependent on both current agent state and the perceived state. Model parameters are given by (15) (the same model run as in Figure 11).

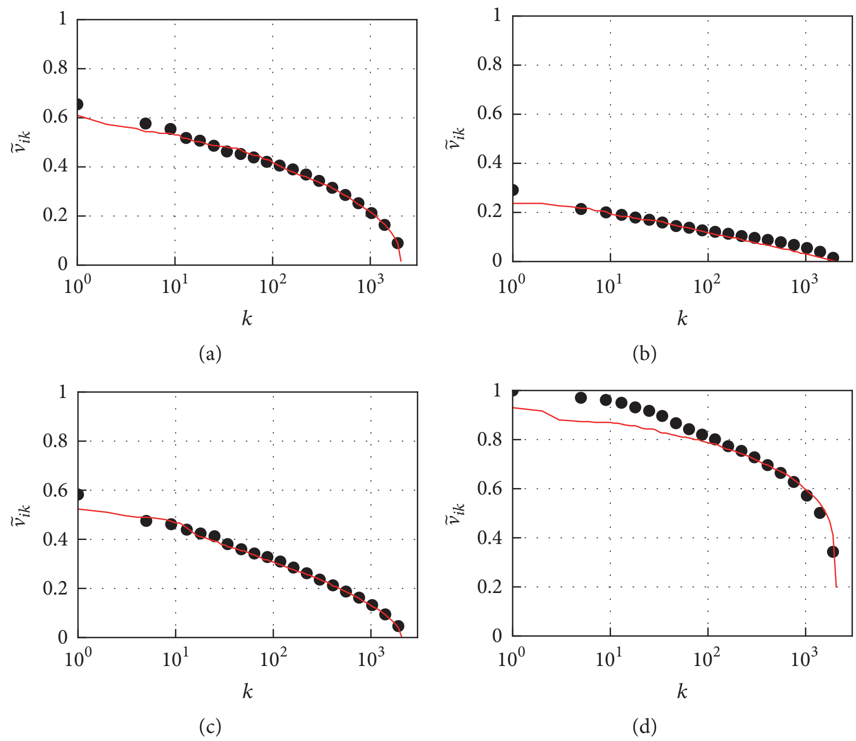

Figure 11: The rank-size distributions of the most popular parties during 1992 election fitted by the asymmetric model. Only the following parties were considered: SK (a), LSDP (b), LKDP (c), and LDDP (d). The empirical values are shown as black circles, while solid curves represent data numerically generated by the proposed model with attractiveness dependent on both current agent state and the perceived state. Model parameters are given by (15) (the same model run as in Figure 10).

also represent one agent moving away from the modeled geographic location, due to social or economic reasons, and another agent, holding different political views, moving in. A similar idea was raised in [53].

In the nearest future, we will consider spatial modeling of the Lithuanian parliamentary elections. Another possible approach, with forecasting possibility, could be considering a temporal regression model for the attractiveness parameters of the proposed model, ε_i, as well as the estimation of the agent interaction rates, h.

Conflicts of Interest

The author declares that there are no conflicts of interest regarding the publication of this article.

Acknowledgments

The author would like to thank Professor Ainė Ramonaitė for suggesting the idea to analyze the Lithuanian parliamentary election data and for numerous useful remarks. The author would also like to acknowledge valuable discussions and general interest in this venture shown by his colleagues at VU ITPA, Julius Ruseckas and Vygintas Gontis.

References

[1] W. H. Riker and P. C. Ordeshook, "A theory of the calculus of voting," *American Political Science Review*, vol. 62, no. 1, pp. 25–42, 1968.

[2] J. A. Ferejohn and M. P. Fiorina, "The paradox of not voting: a decision theoretic analysis," *American Political Science Review*, vol. 68, no. 2, pp. 525–536, 1974.

[3] C. J. Uhlaner, "Rational turnout: the neglected role of groups," *American Journal of Political Science*, vol. 33, no. 2, pp. 390–422, 1989.

[4] H. Hotelling, "Stability in competition," *The Economic Journal*, vol. 39, no. 153, pp. 41–57, 1929.

[5] A. Downs, "An economic theory of political action in a democracy," *Journal of Political Economy*, vol. 65, no. 2, pp. 135–150, 1957.

[6] D. Black, *The Theory of Committees and Elections*, Springer, Dordrecht, the Netherlands, 1958.

[7] C. R. Plott, "A notion of equilibrium and its possibility under majority rule," *American Economic Review*, vol. 57, no. 4, pp. 787–806, 1967.

[8] R. D. McKelvey, "Intransitivities in multidimensional voting models and some implications for agenda control," *Journal of Economic Theory*, vol. 12, no. 3, pp. 472–482, 1976.

[9] S. Ansolabehere and J. M. Snyder Jr., "Valence politics and equilibrium in spatial election models," *Public Choice*, vol. 103, no. 3/4, pp. 327–336, 2000.

[10] W. B. Arthur, "Inductive reasoning and bounded rationality," *American Economic Review*, vol. 84, pp. 406–411, 1994.

[11] G. Akerlof and J. Shiller, *Animal Spirits: How Human Psychology Drives the Economy, and Why It Matters for Global Capitalism*, Princeton University Press, Princeton, NJ, USA, 2009.

[12] D. Helbing and A. Kirman, "Rethinking economics using complexity theory," *Real-Wold Economic Review*, no. 64, pp. 23–51, 2013.

[13] A. Nowak, J. Szamrej, and B. Latané, "From private attitude to public opinion: a dynamic theory of social impact," *Psychological Review*, vol. 97, no. 3, pp. 362–376, 1990.

[14] R. Axelrod, "The dissemination of culture: a model with local convergence and global polarization," *Journal of Conflict Resolution*, vol. 41, no. 2, pp. 203–226, 1997.

[15] R. Hegselmann and U. Krause, "Opinion dynamics and bounded confidence: models, analysis and simulation," *Journal of Artificial Societies and Social Simulation*, vol. 5, no. 3, article 2, 2002.

[16] G. Deffuant, "Comparing extremism propagation patterns in continuous opinion models," *Journal of Artificial Societies and Social Simulation*, vol. 9, no. 3, article 8, 2006.

[17] P. Sobkowicz, "Discrete model of opinion changes using knowledge and emotions as control variables," *PLoS ONE*, vol. 7, no. 9, Article ID e44489, 2012.

[18] D. G. Lilleker, "Modelling political cognition," in *Political Communication and Cognition*, Political Campaigning and Communication, pp. 198–205, Palgrave Macmillan, London, UK, 2014.

[19] P. Sobkowicz, "Quantitative agent based model of opinion dynamics: polish elections of 2015," *PLoS ONE*, vol. 11, no. 5, Article ID e0155098, 2016.

[20] P. Duggins, "A psychologically-motivated model of opinion change with applications to american politics," *JASSS*, vol. 20, no. 1, article 13, 2017.

[21] P. Ball, "The physical modelling of society: a historical perspective," *Physica A: Statistical Mechanics and its Applications*, vol. 314, no. 1-4, pp. 1–14, 2002.

[22] S. Galam, "Majority rule, hierarchical structures, and democratic totalitarianism: a statistical approach," *Journal of Mathematical Psychology*, vol. 30, no. 4, pp. 426–434, 1986.

[23] R. N. Mantegna and H. E. Stanley, *Introduction to Econophysics: Correlations and Complexity in Finance*, Cambridge University Press, Cambridge, UK, 2000.

[24] B. K. Chakrabarti, A. Chakraborti, and A. Chatterjee, *Econophysics and Sociophysics: Trends and Perspectives*, John Wiley & Sons, Hoboken, NJ, USA, 2006.

[25] C. Castellano, S. Fortunato, and V. Loreto, "Statistical physics of social dynamics," *Reviews of Modern Physics*, vol. 81, no. 2, pp. 591–646, 2009.

[26] D. Stauffer, "A biased review of sociophysics," *Journal of Statistical Physics*, vol. 151, no. 1-2, pp. 9–20, 2013.

[27] F. Abergel, H. Aoyama, B. Chakrabarti et al., *Econophysics and Sociophysics: Recent Progress and Future Directions*, Springer, Berlin, Germany, 2017.

[28] K. Sznajd-Weron, "Sznajd model and its applications," *Acta Physica Polonica B*, vol. 36, no. 8, pp. 2537–2547, 2005.

[29] S. Galam, "Sociophysics: a review of galam models," *International Journal of Modern Physics C*, vol. 19, no. 3, pp. 409–440, 2008.

[30] P. Nyczka and K. Sznajd-Weron, "Anticonformity or independence?—insights from statistical physics," *Journal of Statistical Physics*, vol. 151, no. 1-2, pp. 174–202, 2013.

[31] S. Fortunato and C. Castellano, "Scaling and universality in proportional elections," *Physical Review Letters*, vol. 99, no. 13, Article ID 138701, 2007.

[32] C. A. Andresen, H. F. Hansen, A. Hansen, G. L. Vasconcelos, and J. S. Andrade Jr., "Correlations between political party size and voter memory: a statistical analysis of opinion polls," *International Journal of Modern Physics C*, vol. 19, no. 11, pp. 1647–1657, 2008.

[33] C. Borghesi and J.-P. Bouchaud, "Spatial correlations in vote statistics: a diffusive field model for decision-making," *The European Physical Journal B*, vol. 75, no. 3, pp. 395–404, 2010.

[34] M. C. Mantovani, H. V. Ribeiro, M. V. Moro, S. Picoli, and R. S. Mendes, "Scaling laws and universality in the choice of election candidates," *EPL*, vol. 96, no. 4, 2011.

[35] C. Borghesi, J.-C. Raynal, and J.-P. Bouchaud, "Election turnout statistics in many countries: similarities, differences, and a diffusive field model for decision-making," *PLoS ONE*, vol. 7, no. 5, Article ID e36289, 2012.

[36] A. Krupavicius, "The lithuanian parliamentary elections of 1996," *Electoral Studies*, vol. 16, no. 4, pp. 541–575, 1997.

[37] M. Degutis, "How Lithuanian voters decide: reasons behind the party choice," *Lithuanian Political Science Yearbook*, pp. 69–111, 2000.

[38] M. Kreuzer and V. Pettai, "Patterns of political instability: affiliation patterns of politicians and voters in post-communist Estonia, Latvia, and Lithuania," *Studies in Comparative International Development*, vol. 38, no. 2, pp. 76–98, 2003.

[39] M. Jurkynas, "Emerging cleavages in new democracies: the case of Lithuania," *Journal of Baltic Studies*, vol. 35, no. 3, pp. 278–296, 2004.

[40] A. Ramonaite, "The development of the Lithuanian party system: from stability to perturbation," in *Post-Communist EU Member States: Parties and Party Systems*, pp. 69–88, Routledge, Abingdon, UK, 2006.

[41] A. Kirman, "Ants, rationality, and recruitment," *The Quarterly Journal of Economics*, vol. 108, no. 1, pp. 137–156, 1993.

[42] S. Alfarano, T. Lux, and F. Wagner, "Estimation of agent-based models: the case of an asymmetric herding model," *Computational Economics*, vol. 26, no. 1, pp. 19–49, 2005.

[43] S. Alfarano, T. Lux, and F. Wagner, "Time variation of higher moments in a financial market with heterogeneous agents: an analytical approach," *Journal of Economic Dynamics and Control*, vol. 32, no. 1, pp. 101–136, 2008.

[44] V. Alfi, M. Cristelli, L. Pietronero, and A. Zaccaria, "Minimal agent based model for financial markets I: origin and self-organization of stylized facts," *The European Physical Journal B*, vol. 67, no. 3, pp. 385–397, 2009.

[45] A. Kononovicius and V. Gontis, "Agent based reasoning for the non-linear stochastic models of long-range memory," *Physica A: Statistical Mechanics and its Applications*, vol. 391, no. 4, pp. 1309–1314, 2012.

[46] S. Alfarano, M. Milaković, and M. Raddant, "A note on institutional hierarchy and volatility in financial markets," *The European Journal of Finance*, vol. 19, no. 6, pp. 449–465, 2013.

[47] V. Gontis and A. Kononovicius, "Consentaneous agent-based and stochastic model of the financial markets," *PLoS ONE*, vol. 9, no. 7, Article ID e102201, 2014.

[48] V. Gontis, S. Havlin, A. Kononovicius, B. Podobnik, and H. E. Stanley, "Stochastic model of financial markets reproducing scaling and memory in volatility return intervals," *Physica A: Statistical Mechanics and its Applications*, vol. 462, pp. 1091–1102, 2016.

[49] P. Clifford and A. Sudbury, "A model for spatial conflict," *Biometrika*, vol. 60, pp. 581–588, 1973.

[50] T. Liggett, *Stochastic Interacting Systems: Contact, Voter, and Exclusion Processes*, Springer, Berlin, Germany, 1999.

[51] S. Galam, B. Chopard, and M. Droz, "Killer geometries in competing species dynamics," *Physica A: Statistical Mechanics and its Applications*, vol. 314, no. 1-4, pp. 256–263, 2002.

[52] C. Castellano, D. Vilone, and A. Vespignani, "Incomplete ordering of the voter model on small-world networks," *EPL*, vol. 63, no. 1, pp. 153–158, 2003.

[53] J. Fernández-Gracia, K. Suchecki, J. J. Ramasco, M. San Miguel, and V. M. Eguíluz, "Is the voter model a model for voters?" *Physical Review Letters*, vol. 112, Article ID 158701, 2014.

[54] F. Sano, M. Hisakado, and S. Mori, "Mean field voter model of election to the house of representatives in japan," in *Proceedings of the Asia-Pacific Econophysics Conference—Big Data Analysis and Modeling toward Super Smart Society (APEC-SSS '16)*, Tokyo, Japan, August 2016.

[55] M. Jastramskis, "Forecasting the results of Seimas elections: model of government parties' votes," *Parlamento studijos*, vol. 11, 2011.

[56] S. Galam, "The invisible hand and the rational agent are behind bubbles and crashes," *Chaos, Solitons & Fractals*, vol. 88, pp. 209–217, 2016.

[57] T. Fenner, M. Levene, and G. Loizou, "A multiplicative process for generating the rank-order distribution of UK election results," *Quality & Quantity*, pp. 1–11, 2017.

[58] G. K. Zipf, *Human Behavior and the Principle of Least Effort*, Addison-Wesley, Boston, Mass, USA, 1949.

[59] D. Sornette, L. Knopoff, Y. Y. Kagan, and C. Vanneste, "Rank-ordering statistics of extreme events: application to the distribution of large earthquakes," *Journal of Geophysical Research: Solid Earth*, vol. 101, no. B6, pp. 13883–13893, 1996.

[60] X. Gabaix, "Zipf's law for cities: an explanation," *The Quarterly Journal of Economics*, vol. 114, no. 3, pp. 739–767, 1999.

[61] J. Shao, P. C. Ivanov, B. Uroević, H. E. Stanley, and B. Podobnik, "Zipf rank approach and cross-country convergence of incomes," *EPL*, vol. 94, no. 4, Article ID 48001, 2011.

[62] R. F. da Paz, R. S. Ehlers, and J. L. Bazán, "A weibull mixture model for the votes of a Brazilian political party," in *Interdisciplinary Bayesian Statistics*, vol. 118 of *Springer Proceedings in Mathematics & Statistics*, pp. 229–241, Springer International Publishing, Berlin, Germany, 2015.

[63] M. Ausloos and F. Petroni, "Statistical dynamics of religions and adherents," *EPL*, vol. 77, no. 3, Article ID 38002, 2007.

[64] F. Bass, "A new product growth model for consumer durables," *Management Science*, vol. 15, pp. 215–227, 1969.

[65] J. M. Pasteels, J. L. Deneubourg, and S. Goss, "Self-organization mechanisms in ant societies (I): trail recruitment to newly discovered food sources," in *From Individual to Collective Behaviour in Social Insects*, J. M. Pasteels and and J. L. Deneubourg, Eds., pp. 155–175, Birkhauser, Basel, Switzerland, 1987.

[66] J. M. Pasteels, J. L. Deneubourg, and S. Goss, "Self-organization mechanisms in ant societies (II): learning in foraging and division of labor," in *From Individual to Collective Behaviour in Social Insects*, J. M. Pasteels and and J. L. Deneubourg, Eds., 196, p. 177, Birkhauser, Basel, Switzerland, 1987.

[67] G. S. Becker, "A note on restaurant pricing and other examples of social influences on price," *Journal of Political Economy*, vol. 99, no. 5, pp. 1109–1116, 1991.

[68] A. Ishii, H. Arakaki, N. Matsuda et al., "The 'hit' phenomenon: a mathematical model of human dynamics interactions as a stochastic process," *New Journal of Physics*, vol. 14, Article ID 063018, 2012.

[69] N. G. van Kampen, *Stochastic Process in Physics and Chemistry*, 2007.

[70] A. Kononovicius and V. Gontis, "Three-state herding model of the financial markets," *EPL*, vol. 101, no. 2, Article ID 28001, 2013.

[71] S. Alfarano and M. Milaković, "Network structure and N-dependence in agent-based herding models," *Journal of Economic Dynamics & Control*, vol. 33, no. 1, pp. 78–92, 2009.

[72] A. Kononovicius and J. Ruseckas, "Continuous transition from the extensive to the non-extensive statistics in an agent-based herding model," *The European Physical Journal B*, vol. 87, no. 8, article 169, 2014.