ARTICLE

Universality of neutral models: decision process in politics

Volker Hösel[1], Johannes Müller[1,2] & Aurelien Tellier[3]

ABSTRACT   Political elections exemplify complex decision processes in human populations. Data of proportional elections show a striking feature at different spatial scales, across years, and for several democracies: when ranking the parties according to their number of voters, the amount of votes grows exponentially with the party's rank. We develop a mechanistic mathematical model of birth and death of parties and voter grouping based only on word of mouth and not on political contents, close to neutral models used in evolutionary biology (Ewens sampling formula), or Hubbell's model of species biodiversity. Data and model agree strikingly well. The model explains, for instance, the steady loss of big-tent parties in France and Germany by the increasing number of parties standing for elections. A cannibalism effect (parties/candidates at a given rank systematically withdraw votes from others) can be identified. The interpretation and consequences of the rational or lack thereof of voters' choices for modern democracies are discussed.

[1] Centre for Mathematical Sciences, Technical University Munich, Munich, Germany. [2] Institute of Computational Biology, Helmholtz Zentrum München, Munich, Germany. [3] Section of Population Genetics, Center of Life and Food Sciences Weihenstephan, Technical University Munich, Munich, Germany. Correspondence and requests for materials should be addressed to A.T. (email: tellier@wzw.tum.de)

Introduction

Elections are at the heart of modern democracies. In our understanding of elections, mature citizens consider the challenges the society is faced with, and decide rationally about the best approach for the coming years. Finding any systematic statistical pattern in election results is thus not expected, and is even distinctively incompatible with our basic idea of democracy. It is more than surprising that several studies reveal patterns repeating themselves over a whole range of democracies and elections, namely when correlating the relative number of votes per candidate (or per given party) (Chatterjee et al., 2013; Fortunato and Castellano, 2007; Palombi and Toti, 2015) with respect to spatial correlations (Borghesi et al., 2012; Borghesi and Bouchaud, 2010; Fernandez-Gracia et al., 2014) or properties such as voter participation or hierarchical structures (Borghesi et al., 2013; Hernández-Saldana, 2009). In contrast, several papers address the dependency of the candidates on the number of voters, in case of many candidates (Araripe and Costa Filho, 2009; Costa Filho et al., 2003, 1999). It is interesting here, that either candidates are only compared within a party, or the candidates of all parties are pooled. That is, these papers do not focus on the selection of alternatives, but on overall statistics. Perhaps the most striking finding in these studies is that the double logarithmic plot appears to be linear, indicating a power law relation. In contrast to these papers, we investigate proportional elections with only few candidates/parties, where each candidate/party represents a different choice. If the parties are ranked according to the number of votes, the logarithm of the number of votes appear to depend linearly on the rank. This observation indicates an exponential growth of the number of voters with the rank of the parties. We find this relationship in proportional elections in Germany since 1949, the first round of presidential elections in France (2005, 2012 and 2017), recent elections in the Netherlands (1972–2017), and the Republican primaries in the USA (2016) (Fig. 1a–d and Supplementary Information (SI) Figures). Such a repeated pattern, at several spatial scale (for Germany and France results over large towns, regional unit or whole country) hints to an underlying dynamical process that seems to be incompatible with our idea of elections. Particularly in view of recent developments in western democracies it is necessary to identify the underlying mechanisms and to give appropriate interpretations of these observations.

Much academic work and media analyses are done to predict and better understand the results of elections, that is the way voters decide rationally or irrationally, and how parties position themselves to maximize their success (Schofield, 2007; Schofield et al., 2010). Most models in this field focus on the weight of political opinion of voters and the (demographic, socio-economic) targets of parties. The present paper is based on an alternative argument, namely a 'neutral dynamics', for which the content of the opinion (e.g., left- or right wing, liberal or conservative, …) does not play a role. Neutral models appeared to be surprisingly successful in several branches of biology, also there partially contradicting the valid doctrine of non-neutrality. In population genomics, for instance, the use of neutral models has been particularly successful since its introduction by Wright, Fisher, Haldane, Hardy, Weinberg and Kimura (see textbooks such as Charlesworth and Charlesworth (2010) for details). At the first glance, these models contradict our idea of evolution driven by the 'survival of the fittest'. Neutral models allow to test rigorously for the occurrence of natural selection versus the effect of past demography in shaping nucleotide diversity across the genome (e.g., Nielsen et al. (2017); Tellier et al. (2011), review in Stephan 2016). These models facilitate the connection between the observed patterns in data and the underlying driving evolutionary or ecological mechanisms. We are interested in the

following in developing a 'neutral model' for human choice and complex decision processes using as a case study the voting behavior of individuals during elections.

There exist few examples for this kind of dynamics such as the voter model (Liggett, 2012) in which voters copy the opinion of randomly chosen neighboring voters. The noisy voter model (Fernandez-Gracia et al., 2014; Granovsky and Madras, 1995), voter model with zealots (Braha and de Aguiar, 2017; de Aguiar and Bar-Yam, 2011; Palombi and Toti, 2015) or Sznajd models (Gonzalez et al., 2004; Sznajd-Weron and Sznajd, 2000) are also popular versions of (nonlinear) voter models. In a similar spirit to the present paper, Fortunato and Castellano (2007) are able to explain a unimodal distribution appearing in some election data (Italy, Poland, and Finland) by a linear branching process. Interestingly, as noted in Braha and de Aguiar (2017) the voter models and most of its variants used are mathematically identical with variants of the Moran model, a basic neutral model in population genomics. However, we still miss a persuading, simple, and robust mechanism that allows explaining the observations in multi-party systems (Braha and de Aguiar, 2017 have built a two-party model).

In order to give a persuading explanation of our empirical findings, we develop a model based on a stochastic process to capture the outcome of voters' decision for a multi-party system. The model resembles the famous voter model (Granovsky and Madras, 1995; Liggett, 2012) but is based on two well established models of population genetics: the infinite allele model (Durrett, 2008) and the Ewens sampling formula (Durrett, 2008; Ewens, 1972; Kingman, 1978). In particular, the foundation and liquidation of groups (or parties) are an integral part of the model, which we therefore call 'voter model with party dynamics'. Despite the models relative simplicity, the analysis of data in a quantitative manner was possible. We indeed recover the log-linear structure observed in the order statistics of election results. Based on this model we are able to identify and describe the cannibalism effect. Moreover, we suggest inferring the number of parties present in the parliament based on the number of parties that stand for election in case a threshold is implemented in the electoral process. We defer all mathematical work to the supplementary information (SI), and focus below on the model approach and results in order to target on the consequences and conclusions of our work for modern democracies.

Results

Statistical patterns in election data. Data for elections, presented in Fig. 1a–d, Table 1, and SI section 4, reveal a striking statistical pattern if the parties are ranked according to the number of votes. These plots are also known as Whittacker plots (where the sequence of ranks in these plots classically is reversed). Whittacker plots are a rather universal tool to reveal structures in species-abundance data (Saeedghalati et al., 2017). The logarithm of votes appears to depend linearly on the rank of a party. Strikingly, similar observations can be found for species abundance data in ecosystems and for allele frequencies in population genetic data. The quality of the fit to a linear correlation is measured by R^2, the fraction of variability in the data explained by the model (Fig. 1e and SI Figure 6). Strikingly, the R^2 values observed in elections in France, Germany, the Netherlands, and the USA republican primaries, are almost always above 0.9. This log-linear structure is conserved over several years and over magnitudes of organizational units (cities, states, country). In Germany, even the demographic and societal changes due to the German Reunification did not apparently affect this data structure. The only case for which this linear relation is only weakly
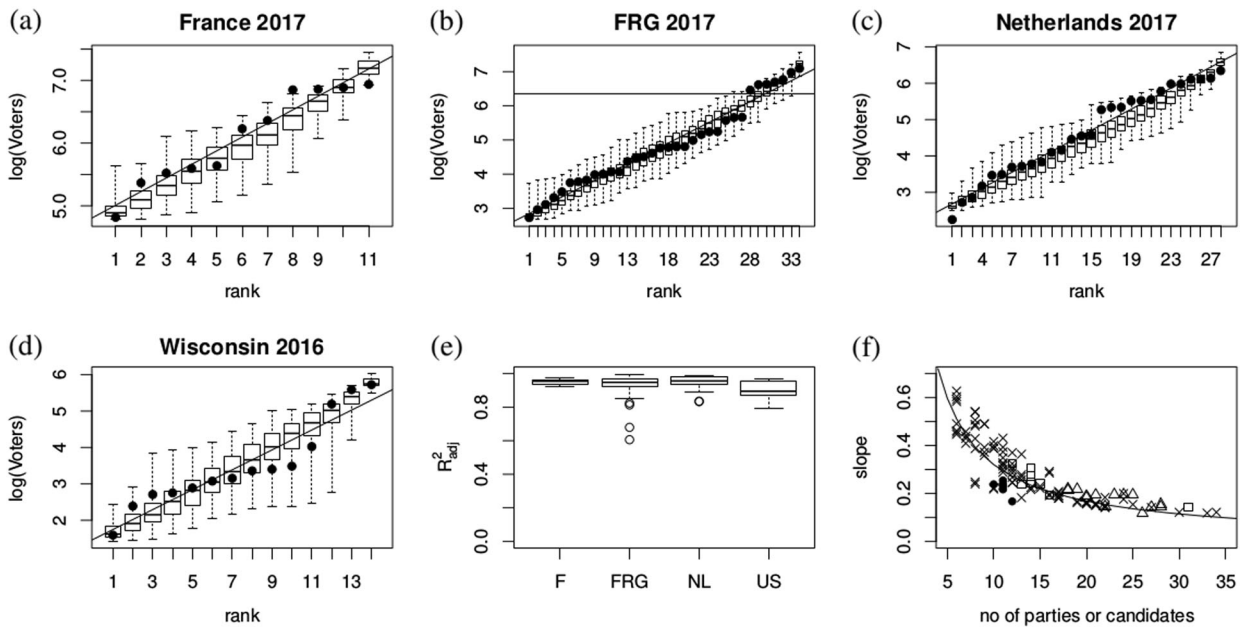
Fig. 1 Occurrence of the log-linear law in election results and model fitting. **a–d** The logarithm of the number of voters over the rank of a party (data as filled circles) for different countries, together with a linear fit (solid line), indication of the 5% threshold for the FRG (horizontal, dashed line) and results of 100 runs of the stochastic model (boxplot). Filled circles are the actual data of the election indicated. **e** Boxplot of R^2 for the linear fit of the data for the four countries. **f** The slope of the linear fit for the log-rank data over the number of parties (France: bullet, FRG: cross—excluding the atypical election from 1949, NL: triangle, US: square), together with the model prediction (solid line)

Table 1 Unit of organization (city, state, or country), number of parties, number of active voters, and parameter of the fit of the linear model for the FRG election in the year 2013

Unit	#parties	#voters	intercept	slope	R^2_{adj}
Stuttgart	20	284541	1.651	0.156	0.9438
Munich	20	650216	1.933	0.165	0.9889
Baden Wuerttemberg	20	5642019	2.953	0.156	0.9759
Bavaria	20	6580755	3.009	0.161	0.9826
Fed. Rep. Germany	30	43726856	3.146	0.121	0.9473

observed is the election in the Federal Republic of Germany (FRG) in 1949 (SI, Figure 6b). This election was the first one after the Second World War, and the elective system was changed afterwards, thereby possibly explaining the somewhat larger deviance.

This structure in the data, namely the strong log-linear correlation between votes and the rank, is rather unexpected and the underlying mechanism is so far unexplained in the literature. When inspecting the log-linear relation over organizational units of different size (cities, states, country), we find the slope to be relatively similar and only the intercept depends on the size of the considered unit (Table 1). Obviously, the mechanism creating this structure scales with the size of the population. This scale invariance in the population size is a strong requirement for the model we introduce next.

Model overview and analytical results. We consider a population of voters. Each voter supports a party, meaning he/she plans to vote for this party (see also SI, Fig. 1). We aim to model the dynamics of the affiliation of voter to party and the time evolution of the number of parties. The dynamics is not based on

political aims and competition of ideas, but driven by the word of mouth according to the following rules: a voter either just copies the opinion of another voter, or creates a new party. At each time step, a randomly chosen person rethinks his/her opinion. This person may decide to keep his/her opinion with probability v, and if not there are two possibilities: 1) with a given small probability $(1-v)u$, he/she founds a new party, or 2) he/she selects randomly one individual in the population, and supports the party of that individual (with probability $(1-v)(1-u)$). When the last supporter of a party leaves, that party is eliminated. In this way we obtain a voter model that allows for the formation/destruction (or birth and death) of parties. We repeat the process until an equilibrium in the party structure is reached, that is when the number of parties reaches a stable value. This model is thus a birth and death process analogous to the well-known infinite allele Moran model in population genetics. The distribution of party size and number is described by the Ewens Sampling formula (Durrett, 2008; Ewens, 1972).

In our model, the equilibrium solution, i.e., the invariant measure of this stochastic process, is not directly the result obtained at the elections. Analysis reveals that the model described so far predicts the appearance of many parties with only very few supporters. This situation is not realistic. Small groups are less likely to bear the effort required to be a party in country-wide elections, so that parties with a subcritical number of supporters would not stand for election. The relative critical size z (minimal size possible/population size) is introduced as the last parameter of the model. As a consequence, the model behavior scales with the population size. We introduce this cut-off in the model only at the time the actual election takes place. That is, the invariant measure of the model is determined without consideration of a minimal party size. At the time of the election, subcritical parties are eliminated, and their supporters are distributed to supercritical parties. Here, each supercritical party has a probability proportional to its size to attract a supporter of a subcritical party.

As the infinite allele Moran model and the Ewens sampling formula are well understood, our analysis (see SI section 1 for an overview, and SI section 2 for the demonstrations) can be based on those results. In particular, we can condition on the number of parties that stand for election K (SI, sections 2.1 and 2.2). It turns out that the invariant measure is in this case determined by three parameters: the number of voters n, the running number of parties K, and the relative critical party size z; it is important to note that the parameters u and v do not influence the distribution any more if we condition on K. Moreover, it is straight forward to characterize the invariant measure of the model. Determine K identical and independent realizations $X_1, \dots, X_K$ of integer-valued random variables with $P(X = j) = \frac{c}{j}$ for $zn \le j \le n$ and $P(X = j) = 0$ else, where $c^{-1} = \sum \frac{1}{j}$.

In order to address the log-linear structure, we order the realizations according to the natural order. We name the ith ordered random variable $X_{(i)}$, such that $X_{(1)} \le X_{(2)} \le \dots \le X_{(K)}$. If the numbers of voters are given beforehand, we condition on $\sum_{i=1}^{K} X_{(i)} = n$. We demonstrate the exponential growth of the number of voters with the rank of a party, namely (SI, section 2.3):

Theorem: Let $\Theta = K/\ln(1/z)$. For $i \in \{1,\dots,K-1\}$ we find

$$\lim_{n \to \infty} E\left(X_{(i+1)}/X_{(i)}\right) = G(\Theta, z) = \int_{z}^{1} \left(\frac{1}{z} - \frac{1}{x}\right) \left(\frac{\ln(1/x)}{\ln(1/z)}\right)^{\ln(z)\Theta - 1} dx.$$

Note that $\Theta = k/ln(1/z)$ parallels the classical Watterson estimator for the rescaled mutation rate Θ (Watterson, 1975). The theorem indicates that $E(X_{(i+1)}/X_{(i)})$ is independent of i. In this sense, we find a log-linear dependency in the rank data: In case of deterministic variables, a constant quotient $X_{(i+1)}/X_{(i)} = G(\Theta,z)$ immediately implies that the logarithm depends linearly on i, $\log(X_{(i+1)}) = \log(X_{(i)}) + \log(G(\Theta,z))$, and hence $\log(X_{(i+1)}) = (i-1)\log(G(\Theta,z)) + \log(X_{(i)})$. As $X_{(i)}$ are random variables and the logarithm is a non-linear function, this relation does not hold exactly but only approximately (which can

concluded by the delta-method, Roussas, 2015). The slope mainly depends on the lumped parameter Θ and only in a less degree on z in the sense that $G(\Theta,z) \to \Theta/(\Theta-1)$ for $z \to 0$ and $\Theta > 1$ (see SI, proposition 2.8). Note that in the theorem above, we did not condition on the total number of voters. The conditioned rank statistics does not exactly show the exact log-linear relation, but numerical experiments indicate that the deviance of the log-linear relation is negligible in the parameter range of interest (SI, section 2.4).

Results of the model fit to data
We compare data of several elections with our model and find surprisingly a very good agreement (Fig. 1a–d, and SI section 5 with data of elections and boxplots of 100 realizations of our model). Two parameters (K, n) can be readily read off the data, while for the parameter z we developed an estimator (Methods and SI, section 2.3.4). The log-rank data are well approximated by a linear fit, that is, by slope and intercept of a linear function. Additionally, we have a mathematical model for the data with the three parameters K, n, and z. This situation allows in particular investigating the dependence of the slope on the parameters. We find that n has asymptotically, for $n \to \infty$, no influence and can be neglected and z has only weak influence (SI, 2.3.3 and SI, 3). For the consideration of the slope we fix z as 0.3 per mille. Only K has a reasonable influence. If we compare the empirical slope in the data (for all four democracies considered) in dependence on the number of parties/candidates with the model prediction, we find a good agreement (Fig. 1 f). We used the Akaike information criterion (AIC) to compare the model developed here with a phenomenological regression model (SI, section 4). The AIC indicates that the regression model performs slightly worse than the mechanistic model, mainly as the number of parameters in the regression is larger. There is, of course, additionally the point that a mechanistic model is less arbitrary than a generic statistical model, and is therefore to prefer. All in all, this analysis indicates
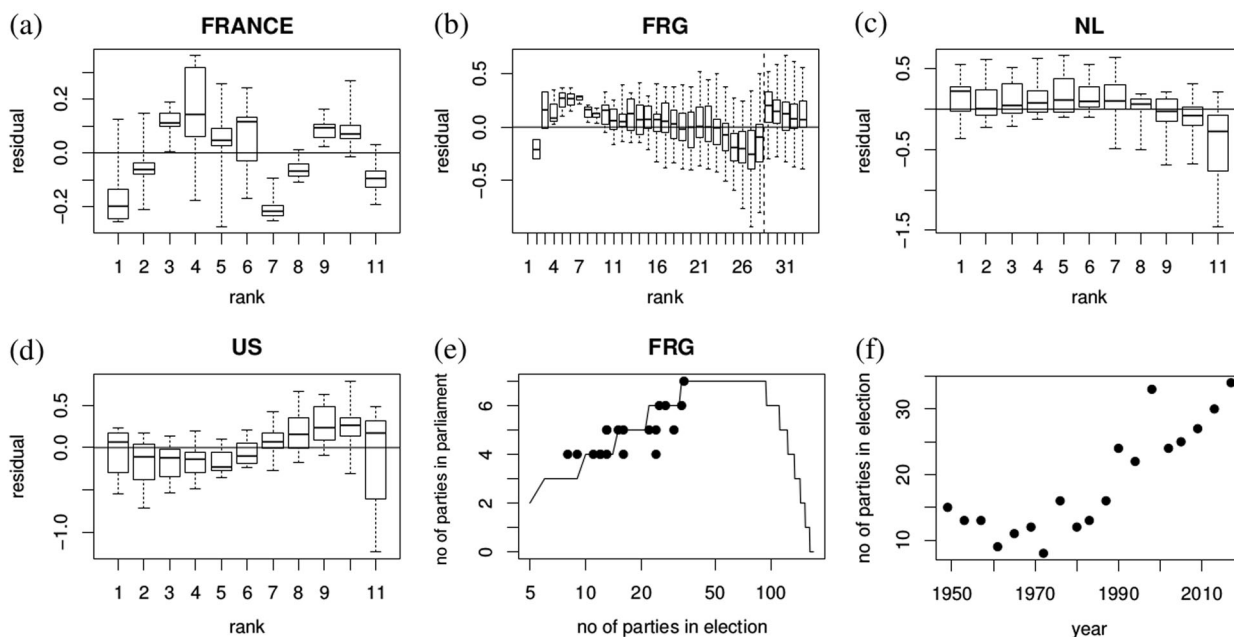

Fig. 2 a–d Boxplot of residuals. For France, NL, and US, only the 11 strongest candidates/parties are shown (rank 11: party/candidate with most votes). For the FRG, data are adapted such that the 5% threshold is always located between rank 28 and 29 (dashed vertical line). Namely, if the 5% threshold is for example located between rank 20 and 21, we add 8 to all ranks. **e** For FRG, number of parties in the parliament over number of parties in the election (solid lines for model prediction, circles for data; we only count parties over the 5% threshold, not parties exclusively present due to direct mandates). **f** Number of parties at the level of the FRG running for second votes over year of this election

that the model explains well the most important structures of the data.

Fine structure: cannibalism effect. The residuals of the linear fits show—for all four democracies we consider—characteristic patterns (Fig. 2a–d). The linear model does not completely meet the data structure. Depending on the rank, the number of votes for a candidate or party is systematically under- or over-estimated. The neutral model developed above does not show such a systematic bias (SI, Figure 6). The reason of any systematic bias in the deviance found in the data is not covered by our model assumptions.

We suggest that this deviation from the correlation is due to strategic thinking by the voters. The reasoning depends on the details of the electoral system, and therefore also the patterns of the residuals are characteristic for the system at hand. Perhaps most striking is the result for the FRG (Fig. 2b), which has a 5% threshold. If we compare election results with those predicted by the neutral model (that is, by the linear fit), we find that parties above the 5% threshold receive more, and parties just below the threshold too few data. We claim that voters who originally would want to vote for parties just below the threshold may expect that their vote would eventually not count (as the party would not be present in the parliament). Therefore, these voters change their mind and vote for a party above the threshold. The bigger parties thus cannibalize the slightly smaller ones and evaluations of detailed polls confirm this interpretation empirically (Meffert et al., 2011). Note that this effect is different to the key ingredient of our model, that tiny proto-parties (relative size smaller z) will not stand for election.

The elections in the Netherlands (Fig. 2c) are similar to that in the FRG, only the 5% threshold is missing. Obviously, this fact reverses the effect: voters who originally wanted to vote for large parties assume that these parties will be strong anyway. Therefore, some of these voters change their mind and choose smaller parties in order to support an aspect in politics that would not this prominent otherwise.

The elections in France and the US (Figs. 2a, d) do appear quite similar in this context, as only the one or two most successful candidates win (or go through to another round). Somehow, the favorite is rather well met by the linear models, but the second and third strongest candidates seem to withdraw votes from the weaker candidates (with a similar reasoning as in the FRG). We explain this observation by strategic thinking (voting) which is known in presidential elections (Bouton, 2013).

In all cases, some candidates/parties cannibalize other ones, depending on the details of the election at hand. In all cases, we expect a similar motivation causing the patterns observed: voters aim to maximize their impact. The resulting strategies and hence also the patterns observed depend on the given electoral system.

Number of parties in the parliament for proportional elections with threshold. A further prediction can be used to support the validity of the model. For proportional elections with a threshold (FRG with 5% threshold), we can calculate the expected number of parties present in the parliament given the number of parties that stand for election: We can predict the slope s of the log-linear relation from the number of parties that stand for election K, and we know the total number of voters n. Consequently, we can compute the intercept a of the log-linear relation and are able to predict the expected outcome of the election. Approximately, the size of the ith group x_i satisfies

$$x_i = 10^{a+si}, \quad \sum_{i=1}^{K} x_i = n \quad \Rightarrow \quad x_i = n \frac{10^s - 1}{10^{s(K+1)} - 1} 10^{si}.$$

This relation allows to compute the number of parties that are present in the parliament as a result of the elective process (that is parties reaching the 5% threshold, $\#\{i|x_i/n > 0.05\}$). Recall that the slope of the log-linear regression mainly depends on the number of parties that stand for election, so that we can draw the expected number of parties in the parliament (Fig. 2e).

Our predictions are compared with the election data (Fig. 2e). We find a good agreement. The number of parties present in the parliament cannot grow to arbitrary high numbers, but we expect that at the present time we already reached this maximum in Germany.

Methods

Data. Presidential elections in France: We considered the first round. Any voter has one vote. If no candidate reaches the absolute majority (which is mostly the case), one of the two most successful candidates will become president in a second ballot.

Federal Republic of Germany (FRG). The elections for the Bundestag have been considered. Each voter has two votes (one for parties, one for direct candidates). Only the votes for parties have been considered. The FRG has a 5% threshold: parties that collect more than 5% of the votes are present in the parliament. The other way to enter the parliament, namely to win a direct mandate via the first vote; is ignored in the present considerations.

The Netherlands. The elections for the House of Representatives have been considered, where any citizen has exactly one vote. The Netherlands have no threshold implemented.

United States of America (USA). The republican, presidential primaries are taken into account, where the republican candidate for the presidential election is selected. Depending on the state, primaries are closed (only registered members of the party are allowed to vote), semi-closed (also unaffiliated citizens are allowed to vote) or open (any citizen may vote). We did not take this difference into account. In these elections, citizen determine how delegates of the state will vote (details are highly state-dependent). Only primaries with at least 10 candidates were considered.

Simulations. Simulating the model is equivalent to drawing realizations from the invariant measure of the model. Recall that we characterized the invariant measure above in the section 'Model overview and analytical results' as follows: Let z be the relative minimal party size, n the number of voters, and K the number of parties that stand for election. Let furthermore $X_1, \ldots, X_K$ i.i.d integer-valued random variables with $P(X=j) = \frac{c}{j}$ for $zn \leq j \leq n$ and $P(X=j) = 0$ else, where $c^{-1} = \sum \frac{1}{j}$.

We condition the realization on $\sum_{i=1}^{K} X_i = n$, and order $X_1, \ldots, X_K$ according to the natural order. The ordered X_i are denoted by $X_{(1)}, \ldots, X_{(K)}$, such that $X_{(1)} \leq \ldots \leq X_{(K)}$. The challenge here is to condition the random variables $X_{(i)}$ on the total number of voters, $\sum_{i=1}^{K} X_i = n$. Let $\widetilde{X}_{(i)}$ denote conditioned realizations. We utilize the fact that $X_{(i)}/n$ tend to a truncated Dirichlet distribution with parameters $\alpha_i = 0$ (SI, section 2.4.1). Note that the Dirichlet distribution requires $\alpha_i > 0$, as integrability is lost otherwise at $x_i = 0$. In our case, the condition $x_i \geq z$ allows for zero parameters. This result indicates, that realizations of $X_{(i)}$ scale approximately linearly with the total voter number n. If n is large, we can draw the $X_{(i)}$ for some different $\widehat{n}$ and rescale the $X_{(i)}$ afterwards to obtain (approximate) realization for the desired n. Depending on z and K, the number of realizations that need to be dismissed to condition on $\widehat{n}$ (and therewith the efficiency of the simulation algorithm) depends on the choice of $\widehat{n}$.

In order to draw samples we select $\widehat{n} = 10^{5+m}$ for an integer m, such that $zn > 5$, draw realizations of $X_{(i)}$ for voter size $\widehat{n}$ until the condition $\sum_{i=1}^{K} X_{(i)} = \widehat{n}$ is approximately satisfied (error less than $\pm 2.5\%$). The resulting realizations are ordered and scaled by $n/\widehat{n}$.

Parameter estimation. The only parameter to estimate is z. We use a heuristic estimation, based on the analytic results for $X_{(i)}$ (SI, section 2.3.4). We find that

$$E\left(\ln\left(X_{(1)}\right)\right) - \ln(n) = -K \int_{z}^{1} \frac{1}{x} \left(\frac{\ln(1/x)}{\ln(1/z)}\right)^{K} dx$$

We replace $X_{(1)}$ by the voter number of the smallest party as predicted by a linear fit to the data (we use the fit to remove some of the variability that is inherent to extreme value statistics) and solve the equation above for z. Simulations of the model based on this estimator yield a striking agreement with the data, though the estimator is only heuristic.

Discussion

The universality of neutral models. We start this study by highlighting the statistical structure of electoral data in elections in France (presidential elections), Germany, the Netherlands, and the USA (Republican primaries) when more than six parties or candidates are present. We find a log-linear relation for voter number and rank of a party. This observation motivates us to build a stochastic mechanistic model. This model indeed explains the pattern in the statistics of the data. It has some intuitive ingredients, as the importance of the word-of-mouth in opinion dynamics, which we assume here to be the mechanism underlying our 'neutral' model of decision process at the population level.

In effect, each individual chooses initially a group, and the dynamics starts, i.e., voters can change their mind and move to another group or found a new one; groups vanish if the last member leaves. Groups can be parties or candidates (for the US and French elections). The groups do not exhibit in our model any specific characteristic (ideas, socio-economic policies, sociological identity,…), and therefore the moving from one group to another occurs at random (with a given probability). Our model is thus a 'neutral' model of voter choice analogous to population genetics model of neutral alleles which frequency varies as a function of genetic drift in a population. The predictions of our model are in line with the overall statistical features of the data. We identify the cannibalism effect that promotes candidates or parties on certain ranks on the costs of candidates or parties on other ranks, depending on the details of the election at hand. Finally we provide a relation between the expected number of parties entering the parliament and the running number of parties for an election with threshold.

The present model approach resembles the role of Hubbell's neutral theory (Hubbell, 2001, 2005) for ecology: The species-abundance curve in ecology is comparable with the log-rank statistics of party size in our case. Hubbell's theory translates the neutral infinite allele model into the ecological setting, seemingly contradicting the classical ecological approach in that every species has a well define niche within an ecosystem. In doing so, Hubbell's work triggered a lot of discussions (Clark, 2009), but finally his theory and further refinements proved to provide a useful tool to analyze ecosystems (Matthews and Whittaker, 2014). Similarly, following Kimura's work in population genomics (Kimura, 1983, 1964) numerous debates and controversies did and still do occur on the role of neutral versus selection in the evolutionary process (Jensen et al., 2018; Kern and Hahn, 2018) while neutral population genetics model have become standards

in population genomics (Charlesworth and Charlesworth, 2010; Jensen et al., 2018; Nielsen et al., 2017; Tellier et al., 2011). We expect in the same way the neutral theory developed here and elsewhere to be useful in politics, though it disregards the classical description of party systems, e.g., by specifying conservative and liberal parties. Our model is also a more general tool to analyze human decision processes or choice behavior at the population level.

Note that our model, as well Hubbell's and that of genetics of a single population, neglects spatial structure and are concerned by one population of freely interacting agents. On the one hand, this assumption is for mathematical convenience, justified *a posteriori* as the model meets the overall structure of the data. On the other hand, we argue that modern media and social networks decrease the importance of spatial structure (the paper by Braha and de Aguiar (2017) gives empirical support for this hypothesis). More fundamentally, fine scale spatial structure within a population (country) may not be very important because the absorption probabilities for the classical voter model on a finite graph does not depend on the graph structure, but follows that of a homogeneous Moran model (Lieberman et al., 2005).

Loss of big tent parties and the number of running parties. In recent years, e.g., in France and in the FRG, the large big-tent parties steadily loose votes, and more parties enter the parliament. In view of our analysis, we identify the reason that contributes to or even primarily causes this effect as the steady increase of running and founded parties since approximately 1975 (Fig. 2f). We suggest that the marginalization of big tent parties is the consequence of a long-term process, and not only a recent trend. The increase in the total number of parties is due to a fragmentation of the political spectrum into specialized parties, focusing on specific issues (e.g. environmental problems, women equality, elderly peoples' place in the society, development of internet and information technology, promoting animal welfare, …). As these parties persist (often on a low level of votes, though), this yields the decrease in voters of large parties. The marginalization of big-tent parties is not necessarily based on a societal shift towards political extremes.

Cannibalism effect distorts election results. Our findings suggest that the actual voting behavior is influenced by two different processes. First of all, voters select a party/candidate. Here, the word of mouth is central. This is the more important process. However, before the citizen actually ballot, a second process of game-theory and strategical thinking influences the decision. This strategic thinking can be recognized most clearly in the FRG, where parties above the threshold withdraw votes from parties just below the threshold. Voters want to avoid to 'waste' their votes in supporting parties that will not be part of the parliament. Note that voters have some pre-knowledge due to election predictions and polls about which party will be in the parliament. We assume that such predictions are often quite reliable at that point (on election day). In the Netherlands, where also the parliament is elected (without threshold), votes are withdrawn from the largest parties (most likely since citizen assume that these parties are strong anyway), but parties in the middle range are promoted. A similar effect takes place in presidential elections in France and pre-elections in the US. The favorite is basically set, and does not much deviate from the neutral model. The second and third candidates, however, receive more votes than expected by the neutral model. Voters try to push promising candidates towards the favorite candidate, at the cost of the less promising candidates. In any case, strategic thinking targets on an increase of the importance of one's own vote. Instead to vote for a party/

candidate where either success is expected anyway or no success at all is expected, a fraction of the votes are re-directed to change the effective outcome. This process may become a self-fulfilling prophecy. Some parties/candidates may fail since voters expect them to fail. On the other hand, in particular in the elections with threshold (FRG), a new party will be able to enter the parliament if (1) it embraces a new hot topic, but also (2) if there the number of parties in total is large enough so that novel ones can enter the parliament (Fig. 2e).

Global social interactions drives the dynamics. Our model presents a voter as an individual primarily driven by social interactions and influenced mainly by the opinions of his/her fellow citizens, and not by rational considerations about the direction the society should take in the next years. This theoretical framework has been modeled before as the voter model (Granovsky and Madras, 1995; Liggett, 2012) or the Sznajd model (Sznajd-Weron and Sznajd, 2000). A new aspect in the present study is the insight that the party (birth and death) dynamics and the truncation of small groups play a central role. This view doubtless capture one but not all important aspects in the dynamics of opinions. With the advent of social media, we predict that the word of mouth becomes increasingly important and generalized, and the opportunity for voters to change their mind (the probability v) becomes larger. There is some indication for this fact already presented in the study by Braha and de Aguiar (2017), where it is shown that the variability in US presidential elections is rather stable since 1920 till 2000, and steadily increasing afterwards. The results of future elections and voter distributions can be studied using our model to test this hypothesis. Particularly, the birth and death process of parties is expected to accelerate.

Implications for modern democracies. We now interpret further the results of this study. In present times, democracy is criticized, and alternatives are proposed from an adaptation of the electoral process, e.g. by incorporation of random elements and chance (Van Reybrouck, 2016), up to a fundamental change of democracy to epistocracy (Brennan, 2017). Proponents of these ideas could be tempted, in order to support their point of view, to use our observation that elections have an inherited structure which is not in line with a rational decision process. However, we propose three other interpretations which could be more appropriate and fruitful.

First, the ultimate decision process, and the shaping of governments and ruling majorities, takes place in the competition of large parties for voters. Most part of the data discussed here is concerned with smaller parties that only represent a relatively small part of the population (note that we take the logarithm of the voter numbers). This effect is best observed in the data of the presidential elections in France 2017 (see SI, section 4). Basically four candidates, representing drastically different political ideas, did compete and attracted fairly similar amount of votes. Nevertheless, when including all candidates' results, the data show the log-linear structure discussed in the present paper. The impact of rational, open, and critical decisions could be hidden in (logarithmically small) perturbations.

Second, the model opens the door for problem driven decision processes. A central, new aspect in our model is the dynamics of party creation/destruction. This process is formulated in a neutral way, as parties are formed at a certain probability, and destroyed, if—by chance—the last supporter leaves. However, when inspecting the political aims of new parties, it appears often that their focus is driven by actual political issues. In other words, elections and the parliamentary system are crucially influenced by

the question of whether new topics are handled adequately by big-tent parties. If not, the citizens may feel that new parties and movements are necessary to handle them. Our results indicate that new, even small, parties are able to influence the complete political landscape and are important for the distribution of votes.

Third, our model cannot disentangle between two components of the switch probability v: (1) the decision of voters per se, or (2) the adequacy between voters' need and party programs. In other words, our results can indicate some elements of randomness in the way that voters may switch their opinion and/or that parties change policies between elections. The feedback loop between voters need and priorities and parties' policies being based upon a complex multi-factorial set of ideas filtered by rationale and irrational decisions (at the individual and at the party level), which may generate the neutral distribution observed in our results. This point of view is in agreement with previous suggestions (e.g. Costa Filho et al., 2003).

The aim of such study is not a complete description of the democratic process, but the insight into some basic mechanisms acting in the system, such as the rationale or randomness of voters. We believe that our results reveal fundamental trends such as the steady increase of the number of political parties, and the cannibalism effect that influences the political geometry. The democratic system, which in essence should be based on equal opportunities and individuals' participation to the society rather than only rely on elections, is here not questioned.

References

Araripe L, Costa Filho R (2009) Role of parties in the vote distribution of proportional elections. Phys Stat Mech Its Appl 388:4167–4170

Borghesi C, Bouchaud J-P (2010) Spatial correlations in vote statistics: a diffusive field model for decision-making. Eur Phys J B 75:395–404

Borghesi C, Hernández L, Louf R, Caparros F (2013) Universal size effects for populations in group-outcome decision-making problems. Phys Rev E 88:062813

Borghesi C, Raynal J-C, Bouchaud J-P (2012) Election turnout statistics in many countries: similarities, differences, and a diffusive field model for decision-making. PLoS ONE 7:e36289

Bouton L (2013) A theory of strategic voting in runoff elections. Am Econ Rev 103:1248–88

Braha D, de Aguiar MAM (2017) Voting contagion: Modeling and analysis of a century of U.S. presidential elections. PLoS ONE 12:e0177970

Brennan J (2017) Against democracy. Princeton University Press, Princeton, NJ, USA

Charlesworth B, Charlesworth D (2010) Elements of evolutionary genetics. Roberts & Company Publishers, Greenwood Village, CO, USA

Chatterjee A, Mitrovic M, Fortunato S (2013) Universality in voting behavior: an empirical analysis. Sci Rep 3:1049

Clark JS (2009) Beyond neutral science. Trends Ecol Evol 24:8–15

Costa Filho R, Almeida M, Moreira J, Andrade Jr J (2003) Brazilian elections: voting for a scaling democracy. Phys Stat Mech Its Appl 322:698–700

Costa Filho RN, Almeida MP, Andrade JSJ, Moreira JE (1999) Scaling behavior in a proportional voting process. Phys Rev E 60:1067–1068

de Aguiar MAM, Bar-Yam Y (2011) Moran model as a dynamical process on networks and its implications for neutral speciation. Phys Rev E 84:031901

Durrett R (2008) Probability models for DNA sequence evolution. Springer, New York, NY, USA

Ewens WJ (1972) The sampling theory of selectively neutral alleles. Theor Popul Biol 3:87–112

Fernandez-Gracia J, Suchecki K, Ramasco JJ, San Miguel M, Eguiluz VM (2014) Is the voter model a model for voters? Phys Rev Lett 112:158701

Fortunato S, Castellano C (2007) Scaling and universality in proportional elections. Phys Rev Lett 99:138701

Gonzalez M, Sousa A, Herrmann H (2004) Opinion formation on a deterministic pseudo-fractal network. Int J Mod Phys C 15:45–57

Granovsky BL, Madras N (1995) The noisy voter model. Stoch Process Their Appl 55:23–43

Hernández-Saldana H (2009) On the corporate votes and their relation with daisy models. Phys Stat Mech Its Appl 388:2699–2704

Hubbell S (2001) The unified neutral theory of biodiversity and biogeography. Princeton University Press, Princeton, NJ

Hubbell SP (2005) Neutral theory in community ecology and the hypothesis of functional equivalence. Funct Ecol 19:166–172

Jensen JD, Payseur BA, Stephan W, Aquadro CF, Lynch M, Charlesworth D, Charlesworth B (2018) The importance of the Neutral Theory in 1968 and 50 years on: a response to Kern and Hahn 2018. Evolution 73:111–114

Kern AD, Hahn MW (2018) The neutral theory in light of natural selection. Mol Biol Evol 35:1366–1371

Kimura M (1983) The neutral theory of molecular evolution. Cambridge University Press, Cambridge, UK

Kimura M (1964) Diffusion models in population genetics. J Appl Probab 1:177–232

Kingman JFC (1978) Random partitions in population genetics. Proc R Soc Lond A 361:1–20

Lieberman E, Hauert C, Nowak MA (2005) Evolutionary dynamics on graphs. Nature 433:312–316

Liggett TM (2012) Interacting particle systems. Springer, Berlin, Germany

Matthews TJ, Whittaker RJ (2014) Neutral theory and the species abundance distribution: recent developments and prospects for unifying niche and neutral perspectives. Ecol Evol 4:2263–2277

Meffert MF, Huber S, Gschwend T, Pappi FU (2011) More than wishful thinking: causes and consequences of voters' electoral expectations about parties and coalitions. Elect Stud 30:804–815

Nielsen R, Akey JM, Jakobsson M, Pritchard JK, Tishkoff S, Willerslev E (2017) Tracing the peopling of the world through genomics. Nature 541:302–310

Palombi F, Toti S (2015) Voting behavior in proportional elections from agent–based models. Phys Procedia 62:42–47

Roussas GG (2015) An introduction to probability and statistical inference. Elsevier, Amsterdam, Netherlands

Saeedghalati M, Farahpour F, Budeus B, Lange A, Westendorf AM, Seifert M, Küppers R, Hoffmann D (2017) Quantitative Comparison of Abundance Structures of Generalized Communities: From B-Cell Receptor Repertoires to Microbiomes. PLoS Comput Biol 13:e1005362

Schofield N (2007) The mean voter theorem: necessary and sufficient conditions for convergent equilibrium. Rev Econ Stud 74:965–980

Schofield N, Claassen C, Ozdemir U, Zakharov A (2010) Application of a theorem in stochastic models of elections. Int J Math Math Sci. 2010: ID 562813

Stephan W (2016) Signatures of positive selection: from selective sweeps at individual loci to subtle allele frequency changes in polygenic adaptation. Mol Ecol 25:79–88

Sznajd-Weron K, Sznajd J (2000) Opinion evolution in closed community. Int J Mod Phys C 11:1157–1165

Tellier A, Laurent SJY, Lainer H, Pavlidis P, Stephan W (2011) Inference of seed bank parameters in two wild tomato species using ecological and genetic data. Proc Natl Acad Sci U S A 108:17052–17057

Van Reybrouck D (2016) Tegen verkiezingen. Bezige Bij bv, Uitgeverij De. Amsterdam, Netherlands

Watterson GA (1975) On the number of segregating sites in genetical models without recombination. Theor Popul Biol 7:256–276

Acknowledgements

Additional information

Supplementary information The online version of this article (https://doi.org/10.1057/s41599-019-0222-0) contains supplementary material, which is available to authorized users.

Competing interests: The authors declare no competing interests.

Reprints and permission information is available online at http://www.nature.com/reprints

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.