Universality in systems with group-outcome decision making

Christian Borghesi 1*, Laura Hernández 1, Rémi Louf 2, Fabrice Caparros 3

1 Laboratoire de Physique Théorique et Modélisation, UMR-8089 CNRS-Université Cergy Pontoise, France
2 Institut de Physique Théorique, CEA-Saclay, France
3 D´epartement G´eographie, Centre Universitaire de Formation et de Recherche de Mayotte, France

Abstract

Elections constitute a paradigm of decision making problems that have puzzled experts of different disciplines for decades. We study two decision making problems, where groups make decisions that only impact themselves as a group. In both studied cases, participation to local elections and the number of democratic representatives at different scales (from local to national), we observe a universal scaling with the constituency size. These results may be interpreted as constituencies having a hierarchical structure, where each group of \(N\) agents, at each level of the hierarchy, is divided in \(N^{1/3}\) subgroups. Following this interpretation, a phenomenological model of vote participation, where abstention is related to the perceived link of an agent to the rest of the constituency, reproduces quantitatively the observed data.

Keywords: universal scaling | democratic decision making | group-size effects | empirical study of turnout.

Introduction

Elections, appear to be an excellent tool to study decision making. Unlike other examples of opinion dynamics, where data come from declarative opinion polls that may be affected by different biases, elections are an occasion where objective and reliable data about a significant number of people are gathered. Their results play the role of observed quantities in natural phenomena.

Traditionally, studies of opinion dynamics in the physics community are posed as a direct problem, this is: assumptions on the way people interact with each other are made and a model of opinion spreading based on them, is built [1, 2]. These models have successfully described the qualitative evolution of opinion at the macro-level.

More recently, the community has undertaken to study inverse problems in opinion dynamics. Several authors [3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14] have been looking more closely at data in the hope of finding regularities –often called stylized facts– before trying to find the underlying mechanisms. For example, it has been found that the distribution of the number of votes received by candidates is a universal scaling function, suggesting that the global opinion form in a branching process [5]. Also, spatial and temporal regularities observed in the results of French elections [8] and the strong spatial correlations observed in the turnout rate of elections in many different countries [11], may be interpreted in terms of the existence of a ‘cultural field’ whose dynamics obeys a two-dimensional random equation.

The observed regularities seem to indicate so far that the processes at stake in opinion formation are independent from details regarding the social, cultural and economical background of the voters. Therefore, they might be explained by the nature of the particular interactions ruling the opinion dynamics or, as has been recently reported, to an underlying structure of the opinion [14].

In this work we first present results concerning the study of the participation to local elections (e.g. for the Mayor). We have gathered and analysed data from 21 elections in 10 different countries (see Tab. [1]) with different social, economical and cultural characteristics. We measure a common behaviour of the participation with respect to the constituency size in almost all cases, which we interpret as constituencies behaving as if they were divided into subgroups, whose size is universally related to the constituency size. In order to further investigate this claim we gathered and analysed data about the number of representatives to different chambers (from local to national level) in different countries. Again, these numbers exhibit a common behaviour with respect to the constituency size. Finally, we present a phenomenological model based on these observations, which brings a new outlook on the reasons behind abstention in local elections. More generally, our results show that the structure and the cohesion of a group become important when the group makes a decision that only involves it as a group.

*christian.borghesi@u-cergy.fr
Local elections analysis

Local elections refer to municipal districts which go from small villages of hundreds of inhabitants to large cities of hundred of thousand inhabitants.\(^1\)

The elections chosen for this study verify the following properties: (i) voting is not compulsory, (ii) there are enough available data so as to assure good statistics (typically over 2000 municipalities), (iii) the elections only concern local issues, (iv) the number of registered voters is well known.

For each municipality and each election, we denote by \(N\) the total number of registered voters, \(N_+\) the total turnout and \(N_\text{-}\) the total number of abstentions (\(N = N_+ + N_\text{-}\)). We define the logarithmic turnout rate (LTR) as \(\tau = \ln\left(\frac{N_+}{N}\right)\), a symmetric and unbounded variable, that enlightens statistical regularities appearing over different elections and countries [8, 11, 16, 17].

We first study the dependence of the average LTR –the first moment of \(P(\tau|N)\)– with the size of the considered municipality (Fig. 1A). Each point corresponds to the average value of \(\tau\) over municipalities of about the same size (see SI section S1.1). The first striking result is that fitting these data assuming a logarithmic dependency, we find:

\[
\langle \tau \rangle (N) \approx \text{cst} - \alpha \ln(N).
\]

Moreover, the slope \(\alpha\) seems to be the same for most of the studied elections, with \(\alpha \approx 0.31\) on average (Tab. S1). This behaviour holds for elections that have a turnout rate ranging from 18% to 88%, implying that \(\alpha\) is independent of the global turnout rate of the country.

Interestingly, the variance of \(P(\tau|N)\), \(\sigma_\tau^2(N)\), also exhibits regularities even though the curves are noisier than for the first moment (Fig. 1B). We fit these curves assuming a power-law dependency and we find:

\[
\sigma_\tau^2(N) \approx \frac{\text{cst}}{N^\beta},
\]

where the constant, \(\text{cst}\), is of the order of unity. Again, the exponent \(\beta\) is found to take similar values for most of the studied elections, with \(\beta \approx 0.3\) on average.

Although there are a few exceptions to this regular behaviour (in Israel, for instance, where --interestingly-- Arab and Jewish localities exhibit different behaviours (Fig. S4)), these results are robust: they hold for a large range of municipality sizes and they depend neither on the rural or urban character of the considered municipalities, nor on the particular type of ballot or the region the municipality belongs to (see SI section S2).

It is important to note that these results contrast with previous studies of the LTR in non-local elections for the same constituencies [11], which showed a radically different behaviour. Indeed, it has been found that the variations of \(\langle \tau \rangle\), and \(\sigma_\tau^2(N)\), depend in this case on the country being considered, and on the particular election (Figs. S1 and S2). Hence we conclude that the aforementioned regularities arise when the issue of the election only concerns the interests of the group that is asked to vote.

This not only confirms the well known fact that the electoral participation decreases with the constituency size, but also it shows that, in the case of local elections, some properties of this decrease are universal.

We would like to further understand how the constituency size influences the turnout rate. To that purpose, we first note that Eq. 2 indicates that the turnout rate cannot be understood as a process where the voters act independently (as we would then have \(\sigma^2 \approx 1/N\)). On the contrary, it is compatible with a subdivision of the constituency of size \(N\) into \(n_g\) \(\approx N^2\) independent groups (see SI section S4).\(^2\) In order to further investigate this fact, we study how the number of democratic representatives of a given constituency scales with its size. The intuition is that, if significant, this group structure must be mirrored in the political organisation at different scales.

---

\(^1\)Data may be downloaded from [13], see SI section S1.1 for details.

\(^2\)As \(\alpha \approx \beta\), from Eq. 11 \(d(\tau) \approx -\frac{\text{d}n_g}{n_g}\).
Figure 1: LTR, \( \tau \), as a function of the municipality size, \( N \), for different elections (for the sake of clarity we only plot one election for each country, the remaining ones may be found in Figs. S1 and S2). An election is identified by the country and the corresponding year. Average value (A), \( \langle \tau \rangle(N) \), and variance (B), \( \sigma^2(N) \) of the conditional distribution \( P(\tau|N) \). Only municipalities with Jewish majority are shown in (A). Due to poor statistics, \( \sigma^2(N) \) is shown neither for Israeli nor Polish elections.

Number of democratic representatives

We gathered and analysed data regarding the number \( n_r \) of representatives of a constituency, for different countries and at different scales (municipality, state, country). We only consider elections where the representatives are directly elected by the voters, and for states or countries with bicameral legislatures, only the Lower House (or House of representatives) is considered (see SI section S1.2 for details).

In particular, we focus on the dependence of \( n_r \) on the size \( N \) of the corresponding constituency (Fig. 2A-C) and we find that, disregarding of the considered scale, \( n_r \) scales as (Tab. S2):

\[
n_r \sim c N^{\gamma}, \quad \gamma \approx 1/3, \tag{3}
\]

with \( c \) being a constant of the order of the unity.

As for the local elections, these results are robust: they do not depend on the type of chosen representative assembly (see SI section S3.1 the example for Czech Republic). Interestingly, we also gathered data regarding the evolution of the number of congressmen in the United States (which shows a large increase in population along with a stable democratic system) with regards to the evolution of the population of the country over time (Fig. 2D) and show that it follows the same trend until the beginning of the twentieth century, when the number of congressmen last changed. (See also SI section S3.2 for a short historical investigation of non-present “representatives”).

This scaling can be easily retrieved with the following heuristic in the manner of the least difficulty principle [18]. A good representative democracy must meet two competing requirements. First, it must maximise the representativeness, that is to say, the number \( n_r \) of representatives should be such that the average number of citizens per representative \( N_1 = N/n_r \) is as small as possible. On the other hand, it must be efficient: consensus should be reached after a reasonable number of interactions among the \( n_r \) representatives. Roughly, if one considers all pairwise interactions, the corresponding term is of the order of \( n_r^2 \). Therefore, a ‘good’ representative regime is such that the function: \( F = N_1 + A n_r^2 \), where \( A \) is a measure of the relative importance of representativeness and efficiency, is minimum. This leads immediately to \( n_r \approx N^{1/3} \).

Both, these results and those concerning the participation to local elections hint at a self-organisation of constituencies in a hierarchical structure appearing at different scales. At the first level of the hierarchy, the whole constituency is divided in \( N^{1/3} \) subgroups of size \( N_1 \), which are themselves divided in \( N_1^{1/3} \) subgroups and so on until the subgroups only contain one citizen (Fig. S10).

We apply this idea to build a phenomenological model for the participation to local elections, without making any hypothesis on the exact nature or the origin of these groups, which are beyond the scope of this study.

Phenomenological model for local elections

We propose a very simple, phenomenological model that reproduces the results observed for the LTR in local elections, assuming the peculiar organisation of constituencies explained above. The intention \( C_i \) of the voter \( i \)

---

\footnote{With \( n_r = N^{1/3} \), \( N_1 \) is the geometric mean of \( N \) and \( n_r \).}
is assumed to rely on three different components. First, an idiosyncratic part reflecting the intrinsic tendency of agents to vote or not, related to their cultural, social, etc. background. We assume that agents belonging to the same group share the same idiosyncrasy. In this way, the groups at level $\ell + 1$ in the hierarchy, issued from the same parent group at level $\ell$, share some common trait while having some variability which accounts for the differences emerging at smaller scales. We thus write for the idiosyncrasy of a group at level $\ell + 1$, $H_{\ell+1} = H_{\ell} + \gamma_{\ell+1}$, where $\gamma_{\ell}$ is random variable of zero mean and whose variance decreases with $\ell$. Second, we assume that the agent $i$’s intention to vote also depends on his perceived link to the whole constituency. In other words, his intention to vote on a collective issue depends on how strong the feeling to belong to that constituency is. We represent this component by the number of levels $d(i)$ that separates a citizen from the higher level, the whole constituency. Finally, the average intention to vote of the global population is represented by a global field $F_{\text{nat}}$ that is a constant over all municipalities, for each election. Thus, $C_i = H(i) - d(i) + F_{\text{nat}}$, and following [2, 8, 11], the decision to vote (1) or not to vote (0) is given by $S_i = \Theta(C_i)$, where $\Theta$ is the Heaviside function. For a particular choice of level-dependence of the group idiosyncrasies (see SI section S4), the model reproduces remarkably the results obtained empirically for $\langle \tau \rangle$ and $\sigma^2(\tau)$ (Fig. 1).

Discussion

We have found a universal behaviour of the participation to a decision making problem when the $N$ agents have to decide about an issue that only concerns themselves as a group. The LTR to local elections exhibits in most of the cases a universal behaviour, while the LTR to national/European elections registered for the same set of municipalities does not. This difference shows that agents in each municipality behave as if their group structure and cohesion became important when the object of the election only concerns them as a group. The observed behaviour hints at a universal subdivision of the constituencies in $n_g \approx N^{1/3}$ subgroups. We also find a similar universal behaviour for the number of democratic representatives from municipal to national scales, as a function of the size $N$ of the constituency that has elected them. These results agree with the study of the candidature process [10].

A very simple phenomenological model, based on the existence of groups at all scales, with the corresponding relation $n_g \approx N^{1/3}$ at each scale, allows to retrieve the observed behaviour of LTR. This idea of a hierarchy of group-idiosyncrasies is also found in [14]. As in this work we calculate average values of the LTR over many (typically 100) municipalities of similar size, $N$, our results do not imply that other factors might influence the participation rate of a given municipality as has been largely studied in [19, 20].
As these results seem to be related to the representation that $N$ agents have about themselves as a group, one could ask whether they reflect some intrinsic property of the whole group of agents, common to the different analysed situations—as has also been observed in anthropological studies [21]—, or on the contrary, they reflect the way in which the agents (here humans) use to deal with group classification—as has been observed also in taxonomic and soil classifications [22, 23].

Acknowledgments CB thanks Brigitte Hazart—who sent by herself French élections municipales data, which triggered this study—, Eva Jakubcová and Peter Kurz for help in gathering data; Yonathan Anson, Yves Couder, Stamatios Nicolis, Alexandra Reisigl-Borgesi, Giulia Sandri, Lionel Tabourier and Gérard Toulouse for enlightening discussions; and the Région IdF for partially supporting this work through a DIM post-doctoral grant.

References

[1] Castellano C, Fortunato S, Loreto V (2009) Statistical physics of social dynamics. Rev Mod Phys 81:591-646.
[2] Bouchaud J-P (2013) Crises and Collective Socio-Economic Phenomena: Simple Models and Challenges. J Stat Phys 151:567-606.
[3] Costa Filho RN, Almeida MP, Andrade Jr. JS, Moreira JE (1999) Scaling behavior in a proportional voting process. Phys Rev E 60:1067-1068.
[4] Lyra ML, Costa UMS, Costa Filho RN, Andrade JS (2003) Generalized Zipf’s law in proportional processes. Europhys Lett 62:131-137.
[5] Fortunato S, Castellano C (2007) Scaling and universality in proportional elections. Phys Rev Lett 99:138701.
[6] Hernández-Saldaña H (2009) On the corporate votes and their relation with daisy models. Physica A 388:2699-2704.
[7] Araripe LE, Costa Filho RN (2009) Role of parties in the vote distribution of proportional elections. Physica A 388:4167-4170.
[8] Borgesi C, Bouchaud JP (2010) Spatial correlations in vote statistics: a diffusive field model for decision-making. Eur Phys J B 75:395-404.
[9] Araújo NAM, Andrade Jr. JS, Herrmann HJ (2010) Tactical Voting in Plurality Elections. PLoS ONE 5(9):e12446.
[10] Mantovani MC, Ribeiro HV, Moro MV, Picoli Jr. S, Mendes RS (2011) Scaling laws and universality in the choice of election candidates. Europhys Lett 96:48001-48005.
[11] Borgesi 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(5): e36289.
[12] Chatterjee A, Mitrovic M, Fortunato S (2013) Universality in voting behavior: an empirical analysis. Sci. Rep. 3:1049.
[13] Valori L, Picciolo F, Allandsdottir A, Garlaschelli D (2012) Reconciling long-term cultural diversity and short-term collective social behavior. Proc Natl Acad Sci USA 109:1068-1073.
[14] Klimek P, Yegorov Y, Hanel R, Thurner S (2012) Statistical detection of systematic election irregularities. Proc Natl Acad Sci USA 109:16469–16473.
[15] Franklin MN, Van Der Eijk C, Evans D, Fotos M, Hirczy De Minio W, Marsh M, and Wessels B (2004) Voter Turnout and the Dynamics of Electoral Competition in Established Democracies since 1945. (Cambridge University Press).
[16] Geys B (2006) Explaining voter turnout: A review of aggregate-level research. Electoral Studies 25:637-663.
[17] Kline MA, Boyd R (2010) Population size predicts technological complexity in Oceania. Proc R Soc London Ser B 277:2559-2564.
[18] Caretta Cartozo C, Garlaschelli D, Ricotta C, Barthélémy M, Caldarelli G (2008) Quantifying the taxonomic diversity in real species communities. J Phys A 41:224012.
[19] Ibanez JJ, Arnold RW, Ahrens RJ (2009) The fractal mind of pedologists (soil taxonomists and soil surveyors). Ecol Complex 6:286-293.
Supporting Information

Universality in systems with group-outcome decision making

Christian Borghesi, Laura Hernández, Rémi Louf, Fabrice Caparros

S1 Materials and methods: Data collection

Data used in this paper can directly be downloaded from [1].

S1.1 Local elections study

The studied countries are listed on table 1, each election is referred to by its abbreviation and year. Data has been taken from Austria [2], Costa Rica [3], Czech Republic [4], France [5], Israel [6], Poland [7], Portugal [8], Romania [9], Slovakia [10] and Spain [11].

When dealing with real data some compromise is sometimes needed in order to better fulfil the four conditions listed in the article. For instance, in France and Romania, where the local election may require a second round, only the first round is studied because it occurs in every municipality of the country. On the other hand, in Spain we have only analysed data from the provinces (comunidades autónomas) that vote exclusively for local representatives on the election date: Andalucía, Cataluña and Galicia. In Polish elections, the first round is excluded because it concerns local and regional elections and we have kept the second round which concerns only local elections. These considerations lead sometimes to a lower statistics, e.g. the polish second round takes place in ≈ 700 and 800 municipalities in the two analysed elections. Finally the case of Israel is shown, despite its poor statistics, because it may help to understand the observed regularities as discussed in section S2.

Some remarks on the notation used to describe the analysed data:

• In Austria, though the election date may change from one province to another, we have labelled the elections in all the provinces with a single date.

• Some Romanian electors, not registered in the lista electorala permanenta, are able to vote. We do not consider those electors.

• In one third of the municipalities of Czech Republic, the local election takes place on the same day as Senate election, and in Romania, the municipal election and the election concerning representatives of județ (which are a constituted by a group of municipalities, except for large cities, where the corresponding județ has the same size as the city) are simultaneous. Nevertheless we have integrated them in the general group (above the horizontal line of table 1), because we observe that this part of mixed data does not completely destroy the universal behaviour of LTR.

Each point of Fig. 1 of the article is obtained as the average over 100 municipalities of approximatively the same size, \( N \), with the exception of France, Poland and Israel where each point represents the average over 200, 50 and 20 municipalities, respectively. Apart from the two extremal bins, (largest and smallest municipalities) each municipality is taken into account twice, i.e. belongs to two neighbouring bins.

In order to get rid of electoral errors, we decide to exclude the extreme values of the LTR. Let us consider for instance some 100 municipalities of size \( \approx N \) (as in Fig. 1), and of LTR \( \tau_i \) \( (i = 1, 2, \ldots, 100) \). We denote by \( \mu \) and \( \sigma \) the raw average and standard deviation of \( \tau \) over these 100 municipalities. Then, the final average value \( \langle \tau \rangle (N) \) and of the variance \( \sigma^2 (N) \) only take the municipalities \( i \) such that \(|\tau_i - \mu| < 5\sigma\) into account.
S1.2 Study of the number of democratic representatives

We collect data of the number of representatives in democratic chambers at different scales: national, regional/state, municipal as a function of the total number of registered voters $N$ or, when unknown, the corresponding total population.

For “representative” we understand either a particular (elected) individual that represents a given group of people that has voted for him (as in the case where distinct constituencies exist and one particular representative talks in the name of each constituency), or the situation where a given number of representatives is attributed to the whole population. This approach neither distinguishes between councillors and deputies, nor between representatives directly elected from people constituency from those elected from their political parties.

- **Municipal Councils**
  French Local councillors (conseillers municipaux) in 2008, are taken from [12]. The number of representatives of municipalities is directly given for some particular important towns like Paris, Lyon, Marseille, etc., and for the remaining municipalities, it is given as a function of population range. The population of these towns comes from a 2011 national estimation [13]. Data from the smallest municipalities is not taken into account in order to avoid a cutoff effect because when the population is less than 100 inhabitants, the number of councillors is fixed by law, to 9. (Indeed some municipalities have less than 20 inhabitants). For the last Czech and Slovakian local elections, the total number of elected deputies/councillors for each municipality, as a function of the $N$ registered voters, is directly obtained from the corresponding official electoral dataset.

- **Regional Councils**
  The analysed cases are: 16 German (Ge) Landesparlamente [14], 15 Italian (It) Consigli regionali [15] (in this case the 5 Autonomous regions with special statute are not considered for the sake of consistency), 17 Spanish Parlamentos autónomicos, 26 French Parlements régionaux, 16 Polish voivodeship sejmiks (sejmiki wojewódzkie), 13 Czech Regional Councils (Zastupitelstva krajù) [16], 41 Romanian County Councils (Consiliile Judec˘atelor) and Bucharest Municipal General Council [17], 27 Brazilian (Br) Assembleias Legislativas Estaduais, 30 State Legislative Assemblies in India, 56 State Lower Houses in USA, 8 Austrian Landtage and the Municipal Assembly of Vienna, 8 Regional Self-Governments in Slovakia [18] and 2 Assembleias Legislativas regionais (for Açores and Madeira) in Portugal [19]. In general, data corresponding to the last election is given. The number $N$ of registered voters is known in every region, with the exception of India and U.S.A. (where the total population is taken instead). Brazilian regions are analysed with respect both to $N$ and to the population in order to check for robustness.

- **National chambers of Deputies**
  Data concerning 181 democratic countries has been gathered using the sources in [20, 21], where we have taken whenever possible, the information corresponding to the latest election. We do not consider National Constituent Assemblies. We take into account the total number $N$ of registered voters per country. In some few cases this number is estimated, using the same source [20, 21].

- **Evolution in time of the number of deputies in the U.S. House of Representatives**
  The evolution in time of total number of deputies in the House of Representatives of the U.S.A. is the only empiric case where two important properties meet: (1) stability of the democratic regime over a long period, and (2) a significant change of the total population during this period (here the population has increased around two decades from 1790 up to 2000). Data is taken from [22, 23] or directly from [24], where details on the considered population, etc., are also given.

  Finally, general widespread data like the area and the nominal gross domestic product (GDP) per country used here, correspond to 2011/2012.

---

4Here by the word “Regional” we mean an intermediate political organization between municipalities and the whole nation. These political units have different names and sizes in different countries, like “States” in USA, or “Régions” in France.
Figure S1 compares the behavior of the average value of LTR, $\langle \tau \rangle (N)$, in local elections, as a function of $N$ with its behavior for national elections in different countries (data taken mainly from [25]). It is shown that the described universal behavior does not hold for national elections, where the shape of the curves changes from one country to another and, may also change from one election to another in the same country. It is worthwhile noticing that the curves for different local elections in the same country are almost coincident.

Table S1 summarizes the results of $\langle \tau \rangle (N)$ and $\sigma^2(\tau)$ for the elections listed above the horizontal line of Table 1.

Let us consider the cases where the described universal behavior does not hold (lower part of table 1). In Portuguese elections, a cutoff is imposed to large cities: they are fractioned into freguesias of $N \leq 40000$. In this way, as large cities are fractioned, small or average constituencies are created, thus artificially modifying the
Figure S2: Variance of the LTR, $\sigma^2_\tau$, as a function of the municipality size, $N$, for local and global elections. An election is identified by the country and the corresponding year and its type $m$: Municipal election, E: European election, P: President election, D: Deputies election, R: Referendum. Variances have been calculated over the same number of municipalities as in Fig. 1. The Polish case is not shown because the very low statistics make the variance extremely noisy.

...number of the municipalities of a given size. This situation mainly concerns the province of Lisboa. Nevertheless there are other provinces (regiões) where there are very few municipalities that reach threshold of $N = 40000$, and still these provinces (North and Centro) do not follow the universal behaviour previously described. At the same time other provinces, like Alentejo and Algarve or the two islands Açores and Madeira follow the behaviour described in the article. Aggregating data at the level of the country (in order to have enough statistics), the $\sigma^2_\tau(N)$ behaves in agreement with Eq. (2), with the exponent $\beta \approx 0.34$. We have neither an organisational nor a sociological or historical explanation for single case of Spanish elections (1987) which is an exception to the observed universal behaviour. Here again, we observe that $\sigma^2_\tau(N)$ behaves in agreement with Eq. (2), with the exponent $\beta \approx 0.37$. These results are shown in figure S3.

The anomalous behaviour of Israeli elections is particularly interesting. When plotting data of all the municipalities of the country, the behaviour described by Eq. (1) is not observed, but as it is shown in figure S4, when splitting data into municipalities with a majority of Jewish population on one side and with a majority of Arab population on the other, one retrieves the universal behaviour in the first case but not in the second. This result reinforces the idea that the observed regularities reveal some property of the concerned group.

We have also studied the robustness of the observed behaviour when data is grouped according to the
Table S1: Parameters $\alpha$ of $\langle \tau \rangle (N)$ and $\beta$ of $\sigma_\tau^2(N)$ from fits to Eqs. (1) and (2), respectively, for all the elections in 8 out of the 9 countries shown in the upper part of table 1. $\alpha$ and $\beta$ are only measured in the range where Eqs. (1) and (2) hold (see Figs. 1, S1 and S2). Concerning Israeli municipalities with a Jewish majority, Eq. (1) holds with $\alpha \approx 0.32$ (see Fig. S4). $\star$: recall Polish elections have a very poor statistic.

|     | At  | CR  | Cz  | Fr  | Pl  | Ro  | Sk  | Sp  | average |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|---------|
| $\alpha \approx$ | 0.36 | 0.26 | 0.27 | 0.29 | 0.37 | 0.31 | 0.32 | 0.28 | 0.31    |
| $\beta \approx$  | 0.26 | 0.22 | 0.44 | 0.33 | 0.17$^*$ | 0.26 | 0.39 | 0.34 | 0.30    |

Figure S3: Exceptions to the universal behaviour of LTR; in all the panels we plot the 2001 local French election (Fr-2001) as a reference. (A and B) $\langle \tau \rangle (N)$ and $\sigma_\tau^2(N)$ respectively, for local elections in Portugal (Pt) and Spain (Sp). (C) The Portuguese 2009 local election, different curves correspond to to municipalities of different regions.

Figure S4: Exceptions to the universal behaviour of LTR. Scatter plot of the local elections in Israel; in red municipalities with a majority of Arab population, in black, municipalities with a majority of Jewish population. The straight line corresponds to the fit $\langle \tau \rangle (N) \approx 3.6 - 0.32 \ln(N)$. 

10
different regions or provinces inside a country. The results are shown in figure S5 for the French case (France has much more municipalities than other countries studied here, which allows a better statistical analysis). Data split into regions follow the same behaviour as when the whole country is considered.

It is worthwhile noticing that the results depend neither on the voting rules nor on the rural/urban character of the municipalities. For instance, in France the voting rule is different according to the population of the municipality, the limit being 3500 inhabitants. Figure S5 shows that the slope of the curve is the same in both cases. Moreover, both curves fit the same line, thus indicating that the total turnout rate describes a property of that particular electoral event. On the other hand, using Romanian data we are able to separate rural from urban municipalities. In both cases the described behaviour remains, as shown in Fig. S6B.

In order to check the robustness of our measures, we have evaluated the average of LTR (i.e. \( \langle \tau \rangle (N) \)) over municipalities with approximatively the same size \( N \), in a different manner. Instead of averaging \( \tau \) over, for instance, 100 municipalities of size \( N \approx N \); each one with a number of registered voters \( N_i \) \((i = 1, 2, ..., 100)\), a number of voters \( N_{i,+} \) which leads to its LTR \( \tau_i = \ln \left( \frac{N_{i,+}}{N_i} \right) \). According to what precedes, \( \langle \tau \rangle (N) = 1/100 \sum_{i=1}^{100} \tau_i \). The other manner we have used to define a kind of average LTR over these 100 municipalities is \( \ln \left( \frac{\sum_{i=1}^{100} N_{i,+}}{\sum_{i=1}^{100} (N_i - N_{i,+})} \right) \).

\[5\] Let us consider 100 municipalities, \( i \), of size \( N \approx N \); each one with a number of registered voters \( N_i \) \((i = 1, 2, ..., 100)\), a number of voters \( N_{i,+} \) which leads to its LTR \( \tau_i = \ln \left( \frac{N_{i,+}}{N_i} \right) \). According to what precedes, \( \langle \tau \rangle (N) = 1/100 \sum_{i=1}^{100} \tau_i \). The other manner we have used to define a kind of average LTR over these 100 municipalities is \( \ln \left( \frac{\sum_{i=1}^{100} N_{i,+}}{\sum_{i=1}^{100} (N_i - N_{i,+})} \right) \).
S3 Supplementary text: Number of democratic representatives

S3.1 Modern representative democratic systems

Table S2 gives the power-law fitting parameters of the number of democratic representatives, \( n_r \), with respect to the corresponding number of registered voters, \( N \).

| Scale        | Country   | \( n_r \)  | \( N \)  |
|--------------|-----------|------------|--------|
| Municipal    | Cz        | 1.9 \( N^{0.46} \) |        |
| Regional     | Ge (16)   | 0.57 \( N^{0.43} \) |        |
|              | Fr (26)   | 0.26 \( N^{0.39} \) |        |
|              | Br (27)   | 0.19 \( N^{0.35} \) |        |
| National     | all (181) | 0.49 \( N^{0.37} \) |        |
|              | Africa (50)| 0.39 \( N^{0.39} \) |        |

Table S2: Power-law fits of \( n_r \) with respect to \( N \) at municipal, regional and national scales (for French municipal representatives, data is given in terms of the population, \( \text{pop} \), instead). At regional and national scales, the total number of regions or countries is written in brackets.

It is interesting to note that the observed behaviour does not depend on the particular type of democratic chamber even at a given scale. For instance in Czech Republic, different democratic assemblies exists at municipal scale: Municipal Council, Town Council, Statutory Town Council, Prague City Assembly, City Part/District Council, Market Town Council. The results for these assemblies are shown in figure S7.

To further investigate how \( n_r \) behaves as a function of other extensive variables characterising the studied population, we have plotted the \( n_r \) of national assemblies as a function of surface of the country \( S \), and the nominal Gross Domestic Product (GDP). In table S3 we give the results of the corresponding power-law fit for the case of national parliaments. For all the cases the best fits are obtained as a function of \( N \).

In order to study large regional parliaments we need to consider very large countries. The results for Brazil, India and USA are shown in figure S8. For these countries the number of registered voters is not always known, so we plot \( n_r \) as a function of the population. The behaviour is qualitatively the same in the case of India and Brazil (except for the lower cutoff in the latter) but quite different for USA regions. This anomaly may be related to historical particularities that have differently fixed the number of local representatives in different regions. Such details cannot be described by this coarse-grained approach. A saturation effect similar to the one observed at national scale, for the US House of Representatives after 1910 (see Fig. 2D), could account for the anomalous behaviour observed at regional scale in Fig. S8.

Figure S7: Number of councillors in different Czech assemblies at municipal scale. Lines correspond to the power-law fits in the legend.

Figure S8: Number of councillors in different regional parliaments in large countries as a function of the total population. Lines correspond to the power-law fits in the legend.
Table S3: Power-law fits of the total number of deputies with respect to an extensive quantity of the considered country. Parameters of the fit are written on the left, while the goodness of fit $R^2$ is on the right. The extensive quantities are the number of registered voters ($N$), the total area ($S$ in km$^2$) and the Nominal Gross domestic product ($GDP$, in billion US $). The total number of considered countries per continent or for the world is written in brackets.

S3.2 A short historical investigation in ancient cities

We have investigated the number of “representatives” in ancient cities which had a procedure of election to choose their government, in order to check if they also verified the behaviour observed for present societies.

We show examples presenting the following characteristics: (1) they correspond to different historical periods, and (2) only assemblies or citizens who directly voted or acclaimed, etc., their “representatives” are considered, in order to compare with our nowadays study. Unfortunately, most of the cases do not correspond to this criterion because either there were intermediate structures between the citizenry and those who held the legitimate authority (e.g. Geneva in the 17th had the Petit Conseil, $\approx 25$ members, the Conseil des deux-cents, $\approx 200$-300 members, and the Conseil général, $\approx 1500$ members) or the “body of electors” was composed by members of different qualities (e.g. the popolo grasso and the popolo minuto in Florence in the 14th century).

In Sparta, the citizens who corresponded to the hoplites elected five ephors for one year (27, pp. 161-210). To have an approach as precise as possible of the number of Spartan citizens, we chose the 480BC period, because it was the peak of the Sparta’s military and political power, with 8000 hoplites (28, VII, 234), and 371BC when Sparta entered its decline with the defeat in Leuctra (29, VI, 4-15) where there were only 800 hoplites (27, p. 269).

For the Athenian democracy, in the 5th century BC, we used the Dictionary of Antiquity by the University of Oxford (20). Thus, we could identify the population (30, pp. 802-803) and with more specific entries about ekklesia (30, p. 348) or the boule (30, p. 154) we could clarify the overall organization of the Athens Constitution 31. On the Pnyx which was the official place where the citizens gathered, there could not be more than 6000 citizens because it was the maximum number of people who could really gather there to elect 10 strategists for 1 year.

Medieval Geneva had a population of about 5000 inhabitants in the 14th and 15th centuries (32) and all the male inhabitants could participate in the political life. Thus, about 2000 active citizens can be expected to have elected 4 syndics for 1 year (33 pp.13-14).

As for the colony of Virginia, in 1619 was established a House of Burgesses composed of 22 members directly elected by all the free men over 17, protestants and owners (34 35). As the population was of 1277 people in 1624 (36), we can assume that the active citizens who participated in the elections of their representatives were about 600.

The last example comes from the dialogue/utopia Laws by Plato –and particularly in books V and VI. The city of Magnesia should have 37 guardians of the laws from the around 15000 citizens (see 27 section Annexes I, from which we have taken a low number of the estimated number of citizens).

Fig. S9 plots the number of “representatives” with respect to the number of citizens, $N$. It appears that regularities observed in present data (i.e. $n_r \approx N^{1/3}$) do not hold at all for these non recent cities –maybe apart from the Laws by Plato. This is probably due to the fact that, though those ancient cities elected their “representatives”, the nature of their politic regime was different from our representative democracy.

Figure S9: Total number of “representatives” with regard to the corresponding total number of citizens, $N$, from non-recent cities and from a dialogue/utopia written in Antic Greece. Councillors from Czech Republic in present municipalities (like in Fig. 2A) are also plotted as a guide view.
The main idea leading to the phenomenological model described in the article is quite simple: it is based on the hypothesis that \( N \) agents divide into \( n_g \approx N^{1/3} \) subgroups according to some shared common trait. This division in groups may go on at all scales, reflecting a hierarchy of differences among agents (as the subdivisions go on, agents inside the same group are more and more similar among themselves). Moreover, we assume that the closer the agent feels to the whole population, the larger his tendency is to participate to the election. This feeling (in other words, the strength of his link to the whole population) is taken into account by the number of levels of subdivisions a given agent has undergone in the hierarchy.

Let a random variable \( x \) be the sum of two independent terms, one characterising the group \( \langle \zeta \rangle \), i.i.d. with variance \( \sigma^2_{\zeta} \), and another representing a variability inside a group, described by a random variable \( \epsilon \) that may fluctuate around the average value of the group, with variance \( \sigma^2_{\epsilon} \). Hence an agent \( i \) which belongs to the group \( g_i \) is characterised by a value: \( x_i = \zeta_{g_i} + \epsilon_i \). Assuming that there are \( N \) agents distributed in \( n_g \) groups of approximatively the same size, the variance of \( S = \frac{1}{N} \sum_{i=1}^{N} x_i \) (which could represent the turnout rate of a municipal election) is equal to \( \frac{1}{n_g} \sigma^2_{\zeta} + \frac{1}{n_g} \sigma^2_{\epsilon} \). If \( n_g \approx N^{1/3} \) it yields, when \( N \gg 1 \) and when variances \( \sigma^2_{\zeta} \) and \( \sigma^2_{\epsilon} \) are of the same order of magnitude, that the variance of \( S \) behaves as \( \text{Var}(S) \approx \frac{1}{n_g} \sigma^2_{\zeta} \). This is the dependency that we observe in data with \( n_g \approx N^{1/3} \).

Due to the scale invariance of \( \langle \tau \rangle \) and \( \sigma^2_{\tau} \) we may create a hierarchy assuming that \( N \) individuals form \( n_1 \) groups (with \( n_1 \approx N^{1/3} \)), of approx \( N_l \) individuals each, and each of these groups subdivide again in approximatively \( n_2 \) subgroups (with \( n_2 \approx N^{1/3} \)) with \( N_2 \) individuals in each subgroup, and so on as is depicted in Fig. S10.

As the number of subgroups at each level fluctuates (see for example at level \( l = 1 \), the group of \( N_1 \) individuals subdivides into \( n_2 = 3 \) subgroups but the neighbouring group of \( N_1' \) individuals subdivides into \( n_2' = 2 \) subgroups), the length of all the branches (a branch stops when the group is too small, namely, one single individual) is not the same. The total distance of an agent with respect to the original population of \( N \) agents is \( d(i) = l_f(i) \) where \( l_f(i) \) is the final level of agent \( i \). For instance, the agent represented by the grey-filled box in Fig. S10 (in this case, the size of its ‘subgroup’ is \( N_4 = 1 \)), is at a distance \( d(i) = 4 \) from the whole group.

We call the group-idiomsyncrasy, \( H_\ell \), the property of the group shared by all the agents belonging to it. Then the idiomsyncrasy of one of the subgroups at level \( \ell + 1 \) is: \( H_{\ell+1} = H_\ell + \gamma_{\ell+1} \), where \( \gamma_{\ell+1} \) is a random variable of zero mean and a variance proportional to \( 1/2^{\ell+1} \). The decrease in the variance stands for the fact that differences become smaller at a smaller scale. The decrease law is chosen for simplicity considerations\(^6\).

As indicated in the article, we assume that the overall intention to vote, \( C_i \), of an agent \( i \), is given by three terms: (i) his group-idiomsyncrasy, (ii) his degree of cohesion with the whole population of \( N \) agents and (iii) a national voting field.\(^7\) Thus:

\[
C_i = H(i) - d(i) + F_{\text{nat}}. \tag{S1}
\]

The final binary decision of voting, \( (S_i = 0, 1) \), of the agent \( i \) can be modelled by \( S_i = \Theta(C_i) \), where \( \Theta \)

\(^6\)In this case, the variance of \( H_\ell \) is proportional to \( 1/2 + 1/4 + 1/8 + \cdots + 1/2^\ell \) which is \( \approx 1 \) for \( \ell \gg 1 \).

\(^7\)It is possible to add an other local ‘diffusive field’\(^\text{\footnote{\textit{\cite{1,2}}}}\)\(^\text{\footnote{\textit{\cite{5,6}}}}\), which is not relevant in this study.
means the Heaviside function. This choice is based on a threshold model of decision-making, see e.g. [25, 38] and [39] for a review.

As the national average of the LTR does not play any role in the observed universal behaviour of local elections, Eq. (S1) depends on the variance of the group-idiosyncratic term, $H(i)$. It can easily be seen that the mathematical expression of the LTR, $\tau$, obtained from the data, is particularly convenient for idiosyncrasies distributed according to a logistic distribution of zero mean. Let us recall that the probability that an agent, taken at random, will vote is $p = \frac{N_+}{N}$, for large $N$. Then, from the definition $\tau = \ln\left(\frac{N+}{N-N_-}\right)$, one gets: $p = \frac{1}{1+e^{-\tau}}$. On the other hand, let $H_\ell$, in Eq. (S1), be a random variable distributed accordingly to a logistic distribution with zero mean and a scale parameter $s$, and let also, for simplicity, $d(i) = 1 \forall i$. Then $p = P(C_i > 0) = \frac{1}{1+e^{-(F-d)/s}}$ in the large $N$ limit. This yields $\tau = (F - \overline{d})/s$. $\gamma$ can also be taken, for convenient reasons, from a logistic distribution without loss of generality. The only parameter is thus $s$, the scale parameter of the logistic distribution, which is taken $s \approx 1$. More precisely, according to the chosen decrease law for the variance of the idiosyncrasy as the level of the hierarchy increases, the scale parameter at the level $\ell \geq 1$ of the distribution of $\gamma_\ell$ is $s/\sqrt{2^{\ell}}$.

We have performed simulations over 200 realisations of the group distribution, each leading to a different realisation of the hierarchy depicted in Fig. S10 for a given $N$. For each realisation and for each group at the level $\ell$, the size of subgroups at level $\ell + 1$ are such that $N_{\ell+1} = N_\ell^{2/3} (1 + \eta)$. The noise, $\eta$, is taken from a Gaussian distribution with zero mean, and standard deviation equal to $0.2$. This small irrelevant value of standard deviation only allows to avoid step-like curves. We have calculated, using Eq. (S1) and the voting rule, the corresponding $\langle \tau \rangle$ and $\sigma_\tau^2$. The results are shown by the continuous line of Fig. 1 in the article. Simulated $\langle \tau \rangle(N)$ approximately decays logarithmically, with the slope $\approx 0.31$ in agreement with observed curves. The model imposes that the simulated $\sigma_\tau^2(N)$ behaves as $1/N^\beta$ (with $\beta \approx 0.3$) but interestingly, also the pre-factor is very near to the one obtained from observed data.

We have checked that these results are robust when varying the national field, $F_{\text{nat}}$, in the same range than the $\langle \tau \rangle(N)$ obtained from empirical curves (i.e. when the turnout rate spread from $\approx 20\%$ up to $\approx 90\%$). Moreover the robustness of the model has been tested with respect to different parameters. First, the scale parameter $0.75 \lesssim s \lesssim 1.15$ provide results in agreement with the empiric regularities. Next, instead of directly introducing the noise in each subgroup (such that $N_{\ell+1} = N_\ell^{2/3} (1 + \eta)$), we have directly injected the noise in the number of groups. More precisely a group with $N_\ell$ agents at the level $\ell$ is directly split into $A N_\delta$ subgroups. Simulations retrieve empiric regularities when $0.28 \lesssim \delta \lesssim 0.35$ with $A = 1$, and $0.5 \lesssim A \lesssim 2$ with $\delta = 1/3$. 


References

[1] http://www.u-crete.fr/fr/laboratoires/labo-lptm/donnees-de-recherche.html
[2] http://www.bmi.gv.at/cms/BML_wahlen/
[3] http://www.consulta.tse.go.cr/estadisticas_elecciones.htm
[4] http://www.volby.cz
[5] http://www.interieur.gouv.fr/Elections/Les-resultats
[6] http://www.moin.gov.il/Subjects/Bchirot/Pages/resultmekomi.aspx (Thanks to Prof. Yonathan (Jon) Anson).
[7] http://pkw.gov.pl/wyniki-wyborow-i-referendow/wybory-i-referenda.html
[8] http://eleicoes.cne.pt/
[9] http://www.beclocale2012.ro/rezultate.html
[10] http://portal.statistics.sk/showdoc.do?docid=5673
[11] http://infoelectoral.mir.es/min/
[12] http://www.politiquemania.com/nombre-conseillers-municipaux-par-commune.html
[13] http://professionnels.ign.fr/rcg
[14] http://www.wahlrecht.de/ergebnisse/index.htm
[15] http://elezionistorico.interno.it/
[16] http://www.volby.cz/index_en.htm
[17] http://www.beclocale2012.ro/
[18] http://portal.statistics.sk/showdoc.do?docid=3090
[19] http://www.cne.pt/
[20] http://www.ipu.org/parline/
[21] http://africanelections.tripod.com/
[22] http://www.thirty-thousand.org/pages/analyses.htm
[23] https://pantherfile.uwm.edu/margo/www/apport/datasets.htm
[24] www.uwm.edu/margo/apport/apportionment1.pdf
[25] 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 7(5): e36289 (2012).
[26] C. Borghesi, J. Chiche and J.-P. Nadal, *Between Order and Disorder: A Weak Law on Recent Electoral Behavior among Urban Voters?*. PLoS ONE 7(7): e39916 (2012).
[27] Edmond Levy, *Sparte, Histoire politique et sociale jusqu’à la conquête romaine* (Ed. Seuil, 2003).
[28] Hérodote, *L’enquête* (Translation by A. Barguet, Ed. Gallimard, 1990).
[29] Xénophon, *Helléniques* (Translation by E. Talbot, Ed. Gallimard, 2012).
[30] M. C. Howatson, University of Oxford, *Dictionnaire de l’Antiquité* (Collection Bouquin, 1993).
[31] Aristote, *Constitution d’Athènes* (Translation by G. Mathieu, B. Haussoulier, and C. Mossé, Ed. Les Belles-Lettres, 1996).
[32] *Dictionnaire historique de la Suisse*, www.hls-dhs-dss.ch/textes/f/F2903.php
[33] L. Binz, *Brève histoire de Genève* (Chancellerie de Genève, 2000).
[34] http://www.49online.org/webpages/nschumacher/index.cfm?subpage=506084
[35] http://www.statemaster.com/encyclopedia/House-of-Burgesses
[36] http://en.wikipedia.org/wiki/History_of_Virginia
[37] Platon, *Les Lois* (Translation by L. Brisson and J.-F. Pradeau, Ed. Flammarion, 2006).
[38] C. Borghesi and J.-P. Bouchaud, *Spatial correlations in vote statistics: a diffusive field model for decision-making*. Eur. Phys. J. B 75, 395-404 (2010).
[39] J.-P. Bouchaud, *Crises and Collective Socio-Economic Phenomena: Simple Models and Challenges*. J. Stat. Phys. 151, 567-606 (2013).