Physics Letters A 374 (2010) 3197–3200

Contents lists available at ScienceDirect

Physics Letters A

www.elsevier.com/locate/pla

On the empirical relevance of the transient in opinion models

Sven Banisch [a,c,*], Tanya Araújo [b,c]

[a] Mathematical Physics, Physics Department, Bielefeld University, 33501 Bielefeld, Germany
[b] Research Unit on Complexity in Economics (UECE), ISEG, TULisbon, 1249-078 Lisbon, Portugal
[c] Institute for Complexity Science (ICC), 1249-078 Lisbon, Portugal

ARTICLE INFO

Article history:
Received 25 March 2010
Received in revised form 28 May 2010
Accepted 29 May 2010
Available online 2 June 2010
Communicated by C.R. Doering

Keywords:
Opinion dynamics
Empirical confirmation
Voting behaviour
Transient dynamics

ABSTRACT

While the number and variety of models to explain opinion exchange dynamics is huge, attempts to justify the model results using empirical data are relatively rare. As linking to real data is essential for establishing model credibility, this Letter develops an empirical confirmation experiment by which an opinion model is related to real election data. The model is based on a representation of opinions as a vector of k bits. Individuals interact according to the principle that similarity leads to interaction and interaction leads to still more similarity. In the comparison to real data we concentrate on the transient opinion profiles that form during the dynamic process. An artificial election procedure is introduced which allows to relate transient opinion configurations to the electoral performance of candidates for which data are available. The election procedure based on the well-established principle of proximity voting is repeatedly performed during the transient period and remarkable statistical agreement with the empirical data is observed.

1. Introduction

Using physical tools in the analysis of social collective phenomena can help uncovering invisible structures, patterns and mechanisms at work in real-world social systems. The work of Fortunato and Castellano in 2007 [1] is at the leading edge of this endeavour. Their analysis of the electoral performance of candidates in various proportional elections revealed a universal voting pattern, which was shown to be independent of the characteristics of the voting population, being instead, a consequence of the elementary interactions.

Empirical data coming from electoral contexts provide one of the most relevant accounts of preference distributions in existing societies. An opinion model with empirical relevance should match these accounts of real-world preference distributions.

In a recent paper [2], we analysed the interplay of opinion dynamics and communication networks. Using a bit-string model it was shown that non-trivial social structures emerge from simple rules for individual communication. Here, using the same abstract bit-string model, we show that the universal scaling function found in Ref. [1] is reproduced when artificial elections are run on the transient opinion profiles. Such an empirical confirmation further increases our confidence on the model capabilities to capture and to reproduce some fundamental aspects of real-world dynamics of opinion exchanges.

2. Opinion models and election data

Attempts to compare model results to empirical data are relatively rare in opinion dynamics [3]. There are however some studies with a reference to real data which mostly use election results in the comparison (see [4, Section III.H] for an overview). Galam, for instance, has published on votes and politics, including the French elections [5]. A more quantitative approach to election data was initiated by a statistical analysis of the 1998 Brazilian elections by Filho et al. [6]. Their study revealed that the distribution of votes among candidates ($P(v)$) follows a hyperbolic law (i.e., $P(v) \propto \frac{1}{v}$) in a range of two orders of magnitude. Similar patterns were found for the Indian elections [7]. However, due to party commitment or strategic voting behaviour a universal scaling could not be expected [1,4].

A different scenario characterises the so-called proportional elections, where each party competes with an open list of candidates for multiple seats in the parliament. In Ref. [1], the statistical analysis of proportional elections in Italy (1958, 1972, 1987), Poland (2005) and Finland (2003) revealed that “the distribution of the number of votes received by the candidates is a universal scaling function, identical in different countries and years” (p. 1). This remarkable result is obtained by a re-scaling of the vote numbers v by the number of candidates of the same party Q and the total number of votes received by this party N. The distribution

* Corresponding author at: Mathematical Physics, Physics Department, Bielefeld University, Postfach 100131, 33501 Bielefeld, Germany.
E-mail addresses: sven.banisch@universecity.de (S. Banisch), tanya@iseg.utl.pt (T. Araújo).

0375-9601/$ – see front matter
doi:10.1016/j.physleta.2010.05.071

of the function $F(\frac{vQ}{N})$ is the same for all the elections considered and a log-normal fit is shown to approximate the data quite well.

Opinion studies referring to these new empirical insights either concentrate on adaptations of the Sznajd model [8,7,9] or on very simple models of opinion spread in different network topologies [10,1]. Using a Sznajd model variant[1] where the opinion states directly account for the preference for one out of a set of candidates, Bernardes et al. [8] show that a microscopic opinion model reproduces the characteristic $\frac{1}{v}$-pattern of the 1998 Brazilian elections. In this approach, there is first a stage to construct an adequate initial condition, in which different candidates have different initial chances of being voted, and secondly a stage in which the usual Sznajd process is performed in order to represent the electoral campaign. The latter dynamical process is stopped at some (arbitrary) iteration number and the respective transient state is used in the comparison to real-world results.

Subsequent studies [7,9] basically use the same mechanisms and analyse the effects of different network structures on the distribution of votes. The actual problem with the approach due to Bernardes and colleagues [8] is the termination of the Sznajd process after a "certain carefully chosen time" [4, p. 612]. No reasonable argument is presented for the choice of this iteration number. Furthermore, it is not entirely clear how much of the similarity is due to the quite complicated construction of the initial condition.

An alternative opinion model capable of reproducing the voting pattern of the Brazilian and the Indian elections was proposed by Travieso and Costa in 2006 [10]. Voters are treated as the nodes of a network. Initially, some of these nodes are assigned to a favourite candidate and all the others are treated as undecided. Then, decided nodes are chosen randomly and all their undecided neighbours are associated to the respective candidate. Already decided nodes change the candidate preference with a given switching probability. In some sense, this model is similar to the set-up stage of the initial conditions in Ref. [8]. This simple model is run on Erdõs–Rényi and Barabási networks and it succeeds in reproducing the pattern in the first but not in the latter case. In Ref. [1], a similar model of opinion spread on treelike graphs is used to explain the universal pattern found for proportional elections.

In what follows, an alternative microscopic explanation is provided. Briefly, the essential lines used in previous attempts to empirically confirm opinion models are the following: (i) the models attempt to reproduce universal patterns, which are normally expressed in terms of scaling laws; (ii) in explaining these universal patterns, the system underlying topology is frequently called into place and (iii) the dynamical process comprise three different periods in time: (1) setting up initial conditions, (2) the process final (steady) state, and (3) an intermediate time interval lying in between the (1) and (2).

3. Method

3.1. The model

In our recently introduced model of opinion exchange (Ref. [2]), opinions are represented as a series of k bits, accounting for the positions concerning k different issues in the agents mind. This is similar to the well-known model of cultural dissemination introduced by Axelrod [15,16]. In the beginning of the simulation N agents are generated and a random bit-string is assigned to them. In the iteration process, two agents meet at random. The two players (i, j) are willing to communicate about an issue (one element of the bit-string), only if the number of unequal bits is below or equal to a similarity threshold d_I (i.e., $d(x_i, x_j) \leqslant d_I$). The result of successful communication is that the agent chosen first (x_i) adopts the opinion of the other (x_j) by flipping one of the unequal bits. The conceptual idea behind this is that provided that the views of two individuals are close enough, similarity leads to interaction and interaction leads to still more similarity. These dynamic rules are summarised in the following steps:

1. an initial random set-up of N bit-strings of length k according to the uniform distribution;
2. a dynamic process which iterates:
   (a) random choice of two agent strings x_i, x_j,
   (b) compute the Hamming distance $d(x_i, x_j)$ and if $d(x_i, x_j) \leqslant d_I$ flip one of the unequal bits chosen at random for x_i (opinion exchange);
3. the termination of this process as soon no more exchange is possible.

By applying the rules repeatedly, the process converges to a stable opinion profile in which every two agents either agree in all the issues or their disagreement is larger than d_I. Depending on d_I, different behaviour of the population is observed: low values result in a state of highly fragmented opinions and higher values yield consensus. A precise study of the opinion distribution in the frozen state is presented in Ref. [2]. In the present work, we concentrate on the opinion profiles before freezing in a stable configuration. Model parameters are chosen in order to eventually lead to a global (quasi-)consensus profile, while requiring a relatively long time to reach the absorbing state (i.e., $k = 20$, $d_I = 5$ and $200 \leqslant N \leqslant 4000$).

3.2. Artificial elections and the transient

The dynamical evolution of the preferences is characterised by three different eras. In the first period, called the *burn-in phase* (this terminology follows the work of Laver and Sergenti [17]), preference patterns which do not deviate significantly from the random initial case are observed. The period after the simulation "burnt-in", we refer to as *transient phase*. The opinion structure is somewhere in between randomness and order, and the main hypothesis made in this Letter is that preference distributions comparing to real-world preference profiles have emerged. The third and *final dynamic era* is characterised by a relatively fast convergence to a stable profile with all the agents in the same state. Fig. 1 shows the dynamical evolution of the relative support provided for five issues.

The question for the empirical confirmation experiment is whether the preference distribution observed in the transient of the model is realistic. In order to relate opinion profiles to the electoral performance of candidates (for which the data are available [18]) an artificial election procedure is introduced. The election procedure is based on the well-established principle of proximity voting which assumes that a voter chooses the candidate which is closest to her/him. Proximity voting was first proposed back in 1929 by Hotelling [19] in the context of economic competition and later (in 1957) applied to the problem of candidate positioning by Downs [20].

Initially, Q random bit-strings are generated and taken to account for the policy propositions of Q different candidates. Then, a proximity voting election procedure is performed. The implementation of such procedure is based on the work of Araújo et al. in two different contexts: the one where consumers are driven by

[1] See Ref. [11] for the original Sznajd model and [12,14] for later analyses of the model. See also [13] for an early analysis of the voter model and [5] for a recent review of Galam's majority rule models.