CARESS Working Paper 98-10

Ergodicity and Clustering in Opinion Formation[*][†]

Antonella Ianni[‡]
University of Southampton (U.K.),
University of Pennsylvania (U.S.A.)

Valentina Corradi
University of Pennsylvania (U.S.A.)

April, 27, 1998

Abstract

We study a simple model of pre-electoral opinion formation that posits that interaction between neighbouring voters leads to bandwagons in the dynamics of the individual process, as well as in that of the aggregate process. We show that in different speci cations of the model, there is a tendency for the process to show consensus, i.e. to approach a con guration of homogeneous support for one candidate, out of the two who run the electoral campaign. We point out that the process displays the feature that, after long time spans, a sequence of states occur which, when viewed locally, remain almost stationary and are characterized by large clusters of individuals of the same opinion.

[*]**JEL:** D72, D82, C44. **Keywords:** Ergodicity, Clustering, Local interaction, Voter model.

[†]We have bene ted from fruitful comments from L. Anderlini, D. Cass, S. Coate, N. Persico, M.O. Ravn, L. Samuelson, and all seminar participants at the University of Pennsylvania. The rst author wishes to thank the University of Southampton for granting a research leave and the University of Pennsylvania for the generous hospitality.

[‡]Address for Correspondence: A. Ianni, Department of Economics, University of Southampton, Southampton SO14 2AY, U.K.. e-mails: `a.ianni@soton.ac.uk, aianni@econ.sas.upenn.edu`

> Public opinion, if we wish to see it as it is, should be regarded as an organic process, and not merely as a state of agreement about some question of the day.
>
> Cooley (1918) (p. 378)

1 Introduction

Almost a century has passed since Cooley's words and though scholars, observers and analysts do not seem to agree as to what public opinion actually is, there seems to be a widespread consensus that it should be studied as an "interactive, multidimensional, continuously changing phenomenon whose diverse aspects form causally interrelated patternings" (Crespi (1997), p. ix[1]). Two particular aspects are often emphasized in the sociological literature. The first is the fact that individuals faced with different choices as to whom - or what - to support show a tendency to be influenced by the opinion of some collective majority (*mutual awareness*, as defined in Crespi (1997)). The second is that environmental conditions that are specific to each agent seem to matter in determining the outcome of individual choices (*situational correlates of opinion*, as in Crespi (1997)). These features of the public opinion process seem to be well documented in terms of experimental and empirical evidence.

Issues related to the process of public opinion formation are not tangential to economic theory. Public opinion plays a key role in shaping animal spirits, expectations, voting decisions, patterns of consumer and producer behaviour, as well as dynamics of adoptions of different technologies and innovation. In a nutshell, as common sense has it, public opinion may and does affect daily life, as well as business cycles. Indeed a relevant body of modern economic theory, in particular in the field of Finance (Banerjee (1992), Birkhandani, Hirshleifer and Welch (1992), Lee (1993), Devenow and Welch (1996) among others) and Industrial Organization (Farrel and Saloner (1985), Katz and Shapiro (1985), Kandori and Rob (1998) among others), has often dealt with issues arising from network externalities and band-wagon effects, where the formalizations explicitly postulate a dependence of the level of utility or of profit of a single individual on the proportion of other individuals who think and act in an analog fashion. From the point of view of pure theory a great deal of attention has been devoted to the analysis of simple models that admit multiple equilibria and in particular coordination games, where the underlying incentive structure supports herding as equilibrium behaviour (Vega-Redondo (1995) and therein references).

This paper formalizes and studies a simple process of private and public opinion formation. The main aim is to account for an individual process by which each agent forms an opinion, as well as an aggregate process that shapes (and in many respects de nes) public opinion in its collective dimension. As a metaphor that helps us describing these processes, we think of a model of pre-electoral opinion formation, where individuals repeatedly form their own opinion as to which, out of two, candidate to vote for, at the time when elections come.

As there are very many voters in the population, voting decisions are almost by de nition deprived of any strategic element and each voter sees the future electoral outcome as an unknown state of the world. Nevertheless, in order to account for an explicitly interactive element, we postulate that preference are such that each voter has an incentive to conform to what (s)he perceives to be the winning side of the elections. The motivation we have in mind is in terms of *band-wagon* effect (voters favour the party that is doing well in the polls), or *projection* effect (voters tend to project their intended vote onto their election outcome expectations). Since the theoretical studies of Simon (1954) and Baumol (1957), empirical work carried out in the UK and in the USA seems to provide evidence for these hypothesis (see McAlister and Studlar (1991), Zuckerman A.S., Valentino and Zuckerman E.W. (1994) and therein references).

We formally think in terms of a side-payment that each voter will receive if (s)he votes for the candidate who wins the elections. This raises an incentive for an individual to gather some information on the current state of public opinion, as this would determine, through a simple majority voting, the electoral outcome. To capture the fact that choices are often determined by the features of the environment where interaction takes place, we postulate that each voter can only observe the opinions adopted within the small subset of her or his neighbours, colleagues, friends or relatives, and for modeling purposes we endow each voter with a speci c location on an appropriately characterized topological structure. Models of interactive behaviour with a local connotation have been extensively analyzed in the modern literature on learning and evolution (Blume (1993), Ellison (1993) and (1995), Anderlini and Ianni (1996), Morris (1997a) and (1997b), Eshel, Samuelson and Shaked (1998) and Ianni (1998) among others). We see this process of local interaction as being a minor component of each voter s daily life; as such it takes place repeatedly over time and it may lead to different decisions.

We refer to the process of public opinion formation as to the dynamic process generated by the aggregate of all the individual decisions, and we provide a stochastic formulation. We are interested in analyzing the properties of the time evolution of this process. In particular, we address two complementary issues. The rst relates to the asymptotic properties of the dynamics. We ask what is likely to happen after many time periods have elapsed, and whether we are able to provide unique predictions as to the limit outcome of the process of public opinion formation. The second focuses on the behaviour of the process along the dynamics and relies on the explicit characterization of the process of cluster formation.

We believe that the analysis of the rate at which local areas of consensus grow over time is important in order to understand the dynamics of the process towards its long run behaviour. This feature is not peculiar to this model, but rather it is a common feature of many economic settings where multiple equilibria may arise. Indeed, it is a stylized fact that several economic and social variables show a high degree of local homogeneity and persistent cross-sectional variance, that is only partly explained by fundamental differences in economic conditions. This is the case, for example, for crime (where what is puzzling is not the overall level of criminal activity, but rather its high variance across time and space), or the persistence of income inequalities (ghetto formation and poverty traps), or the co-existence of different, maybe rival, techniques within an industry of identical rms (where standard economic theory would view the adoption of new technologies as essentially monotonic), or phenomena of price dispersion and tax dispersion (even in the absence of heterogeneities among agents). Last, but not least, a quick glance at the distribution of votes over geographical areas show large areas of consensus, certainly in the UK, in the USA and in many other European countries. The methodology we use allows to characterize situations where different opinions co-exist in the population in terms of clusters that are *almost stationary*, i.e. that vary very slowly over time. Hence, although these con gurations cannot be observed in any steady state of the process we study, still they can be a persistent feature of the dynamics along its evolution[2].

Although the aim of the paper and the questions we address are different, this paper also contributes to the recent literature on learning and evolution in interactive settings in two respects. First, the speci cation of the process of private opinion formation we model (based on Bayesian updating, given a sample of observations) produces a dynamics entirely analog to the speci cation of noisy best-reply dynamics used for example in Blume (1993), McKelvey and Palfrey (1995) and Ianni (1998). This paper departs from those speci cations in that we do not postulate any mistake on the part of individuals, as the probabilistic component that drives the dynamics stems entirely from the fact that the information available at the time when choices are to be made is limited. Second, the aggregate dynamics we study could be applied to an explicitly interactive setting, where randomly drawn couples of players repeatedly play a one-shot (2-by-2) coordination game. If this line were to be pursued, the results we obtain here (namely Theorem 2 and Theorem 3) would complement what is already known in the literature as to the asymptotic properties of myopic best-reply dynamics, with more information as to the time evolution of the process towards its steady states.

The paper is organized as follows. Section **??** describes the details of the process of private and public opinion formation. The individual process is formalized in terms of an estimate of the current public opinion, on the basis of which a voter forms her or his private opinion. The collective process relies on two main elements. First, it is assumed that opinions are formed repeatedly over time in a sequential manner (where only one voter at a time can revise or formulate an opinion). Second, as observations consists of other voters opinions, the distribution from which observations are drawn at random is endogeneized in terms of a simple statistic of the voter s neighbourhood. Section **??** analyses the properties of the dynamic process of public opinion formation. As anticipated, the study relies on the characterization of the long-run properties of the dynamics, as well as that of its short-term (or nite time) features. The results show that these two aspects are complementary and provide a better understanding of the process itself. Although decisions are highly decentralized, as the modelled incentive structure is such that each voter has an incentive to conform to what (s)he perceives as being the current collective opinion, the aggregate process displays an analog feature. By pursuing a space-time analysis (i.e. by relating the two dimensions, time and space, over which our process is de ned) we study the process of *cluster formation*. In Section **??** we provide some heuristic considerations as to the implications that these ndings would have within a more general model that allows for strategic behaviour on the part of the two candidates. Finally Section **??** concludes and the Appendix contains the technical proofs, as well as a Remark on the speci c characterization we use.

2 The model

The model formalizes in a simple way the process of pre-electoral public opinion formation. Elections are going to be held at a future date. Two candidates, A and B, run the elections and the winner will be decided through simple majority voting. To focus the model on the behaviour of the public, we disregard completely any strategic element on the part of the candidates. For the purposes of our analysis, each of them has some well de ned electoral plan, the implementation of which will affect each voter s utility, after the elections are run and the winner is decided.

In the model there are countably many identical voters[3] that formulate their opinion as to which candidate to support when the elections will be held. Voters behave in an identical manner, though, as we shall see, asymmetries might arise due to differences in the information they possess. The next two sessions describe the process of opinion formation on the part of a single voter and the process of public opinion formation respectively.

2.1 Private Opinion Formation

As there are many voters in the population, voting decisions on the part of each single voter are almost by de nition deprived of any strategic content. Only the pivotal voter will eventually determine the outcome of the elections and, for each voter, the probability of being pivotal is negligible[4]. Hence, undergraduate micro-economics textbook wisdom has it, a rational voter does not exist[5]. The model we formalize takes into account the fact that each voter cannot marginally determine the electoral outcome, but still accounts for speci c preferences over the two candidates. In particular, we shall think of each possible electoral outcome as an unknown state of nature for the single voter. Preferences, formalized by a utility function, depend on the state of nature, and expected utility considerations determine the process of opinion formation, and ultimately, the outcome of the elections.

We are going to describe the way in which, given current public opinion, a voter formulates his or her own. For the voter, the outcome of this process will be an opinion, a or b, which would correspond to a vote (for A or for B respectively) if elections were to be held at the same point in time. Voters may be forgetful and go through this process repeatedly in their electoral life.

Ingredients of this process are: two exogenous states of nature, labelled A and B corresponding to the event candidate A wins the elections and candidate B wins the elections , and a utility function that depends on the outcome of the election and on the vote chosen. The idea we want to pursue is that, although the outcome of the elections is exogenous to the voter, utility depends on the vote itself. The simplest way to formalize this is to think in terms of a side payments, denoted by $\varepsilon_A > 0$ and $\varepsilon_B > 0$, that a voter gets if (s)he has voted for candidate A or B respectively and if that candidate wins the elections. We further postulate that the utility function is quasi-linear in this latter argument:

	A wins	B wins
vote a	$U(A) + \varepsilon_A$	$U(B)$
vote b	$U(A)$	$U(B) + \varepsilon_B$

Given any probability distribution, P, over the state space $\{A \text{ wins}, B \text{ wins}\}$, it is easy to notice that $E_P U(a) > E_P U(b)$ if $\Pr(A \text{ wins}) > \varepsilon_B(\varepsilon_A + \varepsilon_B)^{-1} \equiv \varepsilon$. As, in order to win the elections, candidate A must have the support of at least half of the population, if we let $\nu_A \in [0, 1]$ denote the fraction of the electorate who is currently supporting candidate A, the above inequality can be restated as $\Pr(\nu_A > 1/2) > \varepsilon$.

If, whenever forming an opinion, the voter knew what was the current public opinion (i.e. if (s)he knew exactly ν_A), then (s)he would form an opinion consistently with the above inequality. However, we assume that this information is not readily available, and the need for some inference on the part of the voter arises. We formalize this process of inference as follows.

We assume the voter has flat priors over $\nu_A \in [0,1]$ which is the fraction of voters in the population who currently support candidate A, and updates this priors after having observed a sample of observations. We take priors to be given by a Beta distribution with equal parameters, $Be(1,1)$. Each observation consists of a randomly chosen other voter in the population, the opinion of whom is observed. Each observation comes from a Binomial distribution, $Bi(p)$, where $0 \leq p \leq 1$ is the parameter and all observations are i.i.d.. We denote the density of the probability distribution that generates observations by $f_{Bi}(r \mid n, p)$ where r is the number of opinions a in a sample of n observations.

If a voter has observed a sample of r opinions a in a sample of n observations, then (s)he updates her prior $Be(1,1)$ to the posterior $Be(1+r, 1+n-r)$, with density $f_{Be}(z \mid 1+r, 1+n-r)$ and mean $(1+r)(n+2)^{-1}$. Given this posterior, the voter would choose opinion a if the $\Pr_{f_{Be}}[z > 1/2 \mid r, n] > \varepsilon$, opinion b otherwise. As observations come from the above distribution, with parameter p, we can calculate the probability of the voter choosing opinion a, given n observations as:

$$\Pr[a \mid n, p, \varepsilon] = \sum_{r=0}^{n} \binom{n}{r} p^r (1-p)^{n-r} \; 1(\{\frac{(n+1)!}{r!(n-r)!} \int_{1/2}^{1} z^r (1-z)^{n-r} dz > \varepsilon\}) \qquad (1)$$

where $1(\{\cdot\})$ is an indicator function that takes value of one whenever $\{\cdot\}$ is true. This quantity depends in a non trivial way on n (the number of observations), on p (the parameter that determines the probability of observing an a) and on ε (the threshold value defined by the voter's preferences).

In the Sections that follow we shall formalize a dynamic process of public opinion formation where infinitely many voters repeatedly form their opinion, in the way we set out in this Section, in a setting where information is highly decentralized. This will naturally introduce a specific form of asymmetry among voters: although they are identical in terms of preferences, as well as in the way they gather and process information, what will determine their information set (i.e. p in the above formulation) will be voter specific and correlated across different voters. In order to achieve this aim, we proceed as follows.

First, we take the parameter ε to be a half. This practically means that the side payments a voter would receive by the winning candidate are identical (though the utility achievable in each state of the world can of course be different) and creates an incentive for the voter to form an opinion in favor of the candidate that is supported by the simple majority of the electorate[6]. Furthermore, in order to focus on two parameters (instead of three, as in (**??**)) we assume that each voter, given a sample of observations, draws one realization at random from the probability distribution de ned by her updated posterior. Hence, the probability with which (s)he will choose opinion a is given by $\Pr_{f_{Be}}[z > 1/2 \mid r, n]$. In this case the r.h.s of (**??**) can be re-written as:

$$\Pr[a \mid n, p] = \sum_{r=0}^{n} \binom{n}{r} p^r (1-p)^{n-r} \, \frac{(n+1)!}{r!(n-r)!} \int_{1/2}^{1} z^r (1-z)^{n-r} dz \tag{2}$$

It is important to notice that, unlike (**??**), (**??**) formalizes a process that is not entirely consistent with expected utility maximization, in that the behavior of the voter, given a sample of observations, is only de ned probabilistically. The motivation we provide for (**??**) relies on the considerations that follow.

Note that (**??**) depends on n (number of observations) and p (the parameter that determines the probability of observing an a). In order to understand the behaviour of this probability, we study its dependence from each of the two arguments separately.

For a xed $n = \overline{n}$ the above probability is a continuous function of $p \in (0, 1)$, and we denote it by $\Pr[a \mid \overline{n}, p]$. It is not difficult to show that this is increasing in p, formalizing the fact that the more likely observations as are, the higher is the probability that the voter will adopt opinion a.

To see the way in which the above probability changes as the number of observation increases, suppose p was actually the true proportion of A s supporters in the population, i.e. $p = \nu_A$. Let $s^*(p)$ denote the probability with which an expected utility maximizer voter should choose opinion a, i.e. $s^*(p) = 0$ for $p < 1/2$, $s^*(p) = 1$ for $p > 1/2$ and, conventionally, set $s^*(p) = 1/2$ for $p = 1/2$. Then numeric computations show that, for the number of observations becoming very large and for each given $p = \overline{p}$, $\lim_{n\to\infty} \Pr[a \mid n, \overline{p}] = s^*(\overline{p})$. In other words, for n large, $\Pr[a \mid n, \overline{p}]$ is essentially described by $s^*(\overline{p})$, although it remains differentiable, since only in the limit, for $n \to \infty$, its image is restricted to the values $0, \frac{1}{2}, 1$. However, for small values of n, the probability with which the voter would formulate opinion a can be substantially different from what (s)he would do, had (s)he perfect information about the aggregate. This difference, namely $\mid \Pr[a \mid n, \overline{p}] - s^*(\overline{p}) \mid$, is a bias due to the fact that information is limited.

We shall be interested in formalizing the process of opinion formation for very small values of n. Intuitively, if a voter observes only a few observations and draws inference on the aggregate, than such inference will only be partially correct. As priors are uninformative, the voter s opinion formation process will be strongly in uenced by what (s)he observes. Such in uence is actually so strong to resemble pure imitative behaviour, in the sense that a voter who goes through the above reasoning *de facto* behaves in a way that reproduces the frequencies that (s)he observes. Speci cally, let $v(p) = p$ for $p \in [0, 1]$ be the function that describes the very simple behaviour of a voter who observes a randomly drawn observation from a binomial $Bi(p)$ and imitates such observation. Then for $n \leq 4$ the error we incur by approximating $\Pr[a \mid n, p]$ with $v(p)$ is bounded above by $(2)^{n+1}$. The next picture plots the two functions for $n = 4$.

Uninformed Voter and Linear Voter: Probability of choosing a, as a function of p.

Although we shall not make use of this approximation in what follows, we consider it to be an insightful relation between purely imitative behaviour on the one hand and uninformed behaviour on the other. The above considerations provide the motivation for the behavioral speci cations we introduce here and to which we shall refer later.

De nition 1 (Uninformed voter) *Given n (number of observations being sampled) and p (parameter of the binomial distribution from which observations are* *drawn) an **uninformed voter** chooses opinion a with probability:*

$$\Pr{}^U[a \mid n, p] = \sum_{r=0}^{n} \binom{n}{r} p^r (1-p)^{n-r} \frac{(n+1)!}{r!(n-r)!} \int_{1/2}^{1} z^r (1-z)^{n-r} dz \qquad p \in \{0, \frac{1}{n}, \frac{2}{n} ..., 1\} \tag{3}$$

De nition 2 (Linear voter) *Given n (number of observations being sampled) and p (parameter of the binomial distribution from which observations are drawn) a **linear voter** chooses opinion a with probability:*

$$\Pr{}^L[a \mid n, p] = p \qquad p \in \{0, \frac{1}{n}, \frac{2}{n} ..., 1\} \tag{4}$$

Both speci cations involve a process of sampling of n opinions among other voters: in the rst, an opinion is formed by updating priors on the basis of the observations, in the second (a voter has no prior to update and) the opinion that is observed is blindly imitated. Before we proceed, we stress that a key difference between the two speci cations is the fact that for $p = \{0, 1\}$ the two processes are the most different: the linear voter can only imitate opinions that are observed, while the uninformed voter who observes all identical opinions can still not imitate what (s)he sees on the ground of the uniform prior.

2.2 Public Opinion Formation

As anticipated, we are going to model a dynamic process of public opinion formation, in which we take the behavioral speci cations introduced in the previous Subsection as primitives. In order to account for different plausible information structures and information transmission among agents, we will analyze different speci cations of the model.

The general notation of the model we study has individual $x \in S$ choosing opinion $\eta(x) \in \{i = 0, 1\}$, where the set of opinion has been re-labeled for notational convenience and opinion a is now re-labeled as 1. A con guration of opinions in the population will be denoted by $\eta \in \{0, 1\}^S$. We model the dynamics of the process where at each point in time at most one individual changes opinion. To this aim, we assume that each individual may choose a new opinion at a random exponential time, with mean one. Whenever individual x is to form a new opinion, (s)he will do so according to De nition **??** or **??**, according to the model we are studying. Within the same model, all individuals form their opinion in exactly the same way.

We shall endogeneize the parameter p (that, we recall from the previous Section, is the parameter that determines the probability of observing opinion 1) as $p_t(x, \eta) \equiv p(x, \eta_t)$, meaning that such probability depends on the agent s identity, as well as on the current con guration of opinions in the population, but is homogeneous over time.

The speci cation of the model provides each agent with a spatial location on a d-dimensional lattice Z^d, and postulates that (s)he can only observe the opinions adopted within the set of agents that live in her/his vicinity. Formally, we take $S = Z^d$ and de ne the set of x s nearest neighbours as $\{y :\| y - x \|= 1\}$, i.e. the set of $2d$ agents who live at Euclidean distance one from agent x. We assume that each voter is equally likely to observe any of the opinions adopted among her nearest neighbors. As a result $p(x, \eta) = (2d)^{-1} \sum_{\{y:\|y-x\|=1\}} \eta(y)$.

In general, we shall denote by η_t the process at time t (η_t is clearly an element of the state space $\{0, 1\}^S$) and we are interested in characterizing the time evolution of the stochastic process η_t and its asymptotic properties as $t \to \infty$. We denote any probability distribution over the state space by μ_t, and the initial distribution by μ_0. A degenerate probability distributions that has pointmass on the con gurations where all individuals adopt exactly the same opinion i (that is con guration η_i where $\eta(x) = i$ for all x in S) is denoted by μ_i. Given μ_0, we let $\mathcal{L}(\eta_t^{\mu_0})$ be the law of $\eta_t^{\mu_0}$, and we write $\lim_{t\to\infty} \mathcal{L}(\eta_t^{\mu_0}) = \mathcal{L}(\eta_\infty^{\mu_0})$ to mean that $\mathcal{L}(\eta_t^{\mu_0})$ is weakly convergent. We also denote by $\mathfrak{I}$ the set of invariant measures for η_t and $\mathfrak{I}_e \subset \mathfrak{I}$ the set its extreme points. We shall de ne the process η_t to be ergodic if and only if $\mathfrak{I}$ is a singleton; in this case the above limit will not depend on the initial condition, in the sense that $\lim_{t\to\infty} \mathcal{L}(\eta_t^{\mu_0}) = \mathcal{L}(\eta_\infty)$ for any μ_0.

The following De nitions summarize the details of the processes of public opinion formation that we study.

De nition 3 (Uninformed voters) *Consider a population of S voters. At a random exponential time t, with mean one, voter $x \in S$ chooses opinion* 1 *with probability* $\Pr^U[1 \mid n, p(x, \eta_t)]$ *(given by equation (**??**)), where $n < \infty$ and $p(x, \eta) \in$*

$\{0, \frac{1}{n}, \frac{2}{n} ..., 1\}$. *Then* $\eta_t^U(n)$ *de nes the process of aggregate opinion formation for a population of* uninformed voters*, when sampling* n *observations.*

De nition 4 (Linear voters) *Consider a population of* S *voters. At a random exponential time* t*, with mean one, voter* $x \in S$ *chooses opinion* 1 *with probability* $\Pr{}^L[1 \mid n, p(x, \eta_t)] = p(x, \eta_t)$ *(given by equation (??)), where* $n < \infty$ *and* $p(x, \eta) \in \{0, \frac{1}{n}, \frac{2}{n} ..., 1\}$. *Then* η_t^L *de nes the process of aggregate opinion formation for a population of* linear voters.

Before we proceed to state the main results of the paper, it is worth noticing that some feature of the processes we shall analyze are relatively intuitive. First, since whenever called to form an opinion, a voter samples current observations, and on the basis of these (s)he decides, the aggregate process clearly satis es some Markovian properties. Second, as pointed out earlier, the *uninformed voter* who samples nitely many observations can choose each of the opinions with strictly positive probability. This suggests that this process may be ergodic, in that all possible con gurations of opinions in the population could be visited in nitely often by the process. In this case a question of interest is the explicit characterization of its limit behavior. This feature is clearly not shared by the *linear voter* process, as only opinions that are observed can be adopted with positive probability. It should be clear from the speci cation that, for this latter process, the set of probability measures that have pointmass one on state η_i where everybody adopts opinion i, $\{\mu_i \; i \in \{0, 1\}\}$, are invariant. Although the lack of ergodicity does not allow for unique predictions as for the limit behavior of this process, it is interesting to analyze what happens along the dynamics of the process, in terms of the process of cluster formation.

3 Main Results

We are now ready to state the main results of the paper. The Theorems that follow study the asymptotic behaviour of the two processes described in the previous Sections. The results are obtained for a speci c characterization of the set of all voters, that are located on a one-dimensional lattice and only observe a sample of opinions from their nearest neighbors. This speci cation is chosen mostly for convenience, as it allows for an intuitive characterization of the space-time analysis we address towards the end of this Section. Further comments on this particular issue are given in a Remark in the Appendix.

Theorem 1 shows that when voters are *uninformed*, the process of public opinion formation is ergodic, in that it admits a unique limit distribution that is fully characterized.

Theorem 1 *Consider η_t^U as in De nition* **??** *and suppose that*

$$p(x, \eta_t) = \frac{1}{2} \sum_{\{y: \|y-x\|=1\}} \eta_t(y)$$

For $S = Z^1$, then:

1. *for any $n < \infty$, $\eta_t^U(n)$ is ergodic (in the sense that $\mathfrak{I}$ is a singleton) and*
2. *for any initial μ_0, $\mathcal{L}(\eta_t^{\mu_0}) \xrightarrow{p} \mathcal{L}(\eta_\infty) \equiv \mu_\infty$, given by:*

$$\mu_\infty^\sigma(\eta) = K \exp[\sum_x \sum_{\{y: \|y-x\|=1\}} \sigma(2\eta(x) - 1)(2\eta(y) - 1)] \tag{5}$$

where K is such that $\sum_\eta \mu_\infty(\eta) = 1$ and $\sigma = \frac{1}{4} \log(2^{n+1} - 1) < \infty$.

Proof. See Appendix.

The above Theorem provides a characterization of the limit behaviour of the dynamics η_t^U: no matter where the process starts, the probability with which each con guration could be observed asymptotically is given by the above limiting distribution. As the latter has full support, each of the possible con gurations of opinions in the population can be observed in the limit. However, it is clear from the above formulation that some con gurations are more likely to be observed than others. In particular, as the sum of which in the square brackets of (**??**) is taken over all *couples* of nearest neighbours, and as the addendum is equal to one if and only if $\eta(x) = \eta(y)$, the two con gurations which are more likely to be observed are those where every voter chooses exactly the same opinion, i.e. η_0 and η_1. Since (**??**) is continuous in the parameter[7] σ, this proves the next Corollary.

Corollary 1 *Under the assumptions of Theorem* **??**,

$$\text{for all } \sigma \ \frac{\mu_\infty^\sigma(\eta_0)}{\mu_\infty^\sigma(\eta_1)} = 1 \qquad \text{and} \qquad \lim_{\sigma \to \infty} \frac{\mu_\infty^\sigma(\eta)}{\mu_\infty^\sigma(\eta_i)} = 0$$

where $\{\eta_i, i \in \{0,1\}\}$ *with* $\eta_0 \equiv \{\eta \in \{0,1\}^S : \eta(x) = 0 \quad \forall x \in S\}$, $\eta_1 \equiv \{\eta \in \{0,1\}^S : \eta(x) = 1 \quad \forall x \in S\}$ *and* $\eta \notin \{\eta_i, i \in \{0,1\}\}$.

The interpretation of the above Corollary in our model is the following. Since $\sigma = \frac{1}{4}\log(2^{n+1} - 1)$, taking a limit for $\sigma \to \infty$ means studying what happens when an uninformed voter samples an in nite number of observations (n). Observations are opinions randomly gathered in the neighbourhood and the interpretation we have in mind is that each of these (even though they may come from the same neighbour) provides the voter with some further information about the state of the system at the time at which (s)he forms an opinion. In other words, a voter who repeatedly talks to the same neighbours, updates her beliefs at each time. Ergodicity breaks down only in the limit, as the transition probabilities of which in De nition **??** become discontinuous in the parameter p, as n grows to in nity. It is clear that Corollary **??** only relies on a comparative static exercise over the limit distributions of a sequence of identical processes, that differ only in the parameter σ and, as such, it does not provide a full understanding of the dynamics at any nite point in time.

The message that the above result conveys is that, asymptotically, we are more likely to observe a con guration of homogeneous opinions in the population. Due to the underlying symmetries that the process satis es, this is perhaps not surprising. Given these asymptotics (and given that elections typically take place at a nite, rather than in nite, time), what we would like to know in more detail what happens along the dynamics of the process. Loosely speaking, we would like to know to what extent the logic we followed in Theorem **??** (where we rst looked at the asymptotics for $t \to \infty$, and then at the limit for $\sigma \to \infty$) can be reversed by rst looking at a process where a voter knows exactly what the opinions in her neighbourhood is (i.e. in the limit for $\sigma \to \infty$) and second at what happens along the dynamics of this process (i.e. asymptotically for $t \to \infty$). By this doing we can gather some further understanding of the way the process evolves, when the behaviour of voters is not driven by lack of information about the current con guration of opinions in the neighbourhood. The Theorem that follows establishes that this line can be pursued and that the resulting process is exactly the process η_t^L, for which asymptotic behaviour is characterized.

Theorem 2 *Consider $\eta_t^U(\sigma)$ and let $S = Z^1$. Then:*

1. for any given t, $\eta_t^U(\infty)$ corresponds to η_t^L (as in De nition **??**).

2. Consider η_t^L, and for $0 \leq \theta \leq 1$, let μ_θ be the product measure with density θ, i.e. $\mu_\theta\{\eta(x) = 1\} = \theta$ for all $x \in S$. Then:

$$\mathfrak{I}_e = \{\mu_0, \mu_1\} \text{ and } \mathcal{L}(\eta_t^{\mu_\theta}) \rightarrow (1 - \theta)\mu_0 + \theta\mu_1$$

Proof. See Appendix.

The above Theorem shows that the process η_t^L can be interpreted as a limiting case of the process η_t^U and that this latter process admits no extreme invariant measures other than $\{\mu_0, \mu_1\}$. Part (2.) is relevant for our purposes because it shows that the extreme invariant measures of η_t^L are exactly those to which $\eta_t^U(\sigma)$ collapses for $\sigma \rightarrow \infty$. If this was not the case the invariant measures of η_t^L could be of a different nature than $\eta_t^U(\infty)$; they could for example identify con gurations where different opinions coexist in equilibrium, and this (from Corollary **??**) would not be consistent with the process η_t^U.

The second reason why the above result is important is that it shows that along the dynamics, the process shows *consensus*, in that if we look at any possible couple of voters, x and y in S, the probability that they choose different opinions approaches zero asymptotically:

$$\lim_{t \rightarrow \infty} \Pr[\eta_t(x) \neq \eta_t(y)] = 0 \text{ for all } x \text{ and } y \text{ in } S$$

Clearly, for any $0 < \theta < 1$, each single voter may change her or his opinion in nitely many times (as $\lim_{t \rightarrow \infty} \eta_t(x)$ does not necessarily exist). However, as a result of the above considerations, the observed frequencies of individuals choosing the same opinion grows over time. Our aim is now to characterize more in detail how this occurs.

As anticipated, we intend to study the properties of the dynamics of our process for any nite t. As our process is de ned in the two dimensions of time and space, we shall nd it useful to relate these two dimensions in a space-time analysis. In particular, we aim at characterizing a *clustering* process, by relying on the local speci cation of the model. With the term cluster we mean a connected group of individuals holding the same opinion, that is the length of a segment with all connected individuals of the same opinion. In order to see how the size of a cluster increases with time, we shall later express the length of a cluster as a function of t. Formally, given a con guration, η, we de ne a *cluster* as the connected components of $\{x : \eta(x) = 0\}$ or $\{x : \eta(x) = 1\}$ and the mean cluster size of η (around the origin) as:

$$C(\eta) = \lim_{l \to \infty} \frac{2l}{\text{number of clusters of } \eta \text{ in } [-l, l]}$$

whenever this limit exists.

The Theorem that follows states that, if the initial distribution is a product measure, and if we re-scale the length of a cluster as $l = \sqrt{t}$, then the mean cluster size around the origin converges in distribution, as $t \to \infty$. Its limit depends on the initial distribution and it is possible to provide bounds, expressed as a function of this probability.

Theorem 3 (Bramson and Griffeath (1980)) *Consider η_t^L (as in De nition* **??***) and for $0 < \theta < 1$, let μ_θ be the product measure*[8] *with density θ, i.e. $\mu_\theta\{\eta(x) = 1\} = \theta$ for all $x \in S = Z^1$. Then, for any initial distribution μ_θ :*

$$\sqrt{\pi}\left(\frac{1}{2\theta(1-\theta)}\right) \leq \lim_{t \to \infty} E[\frac{C^{\mu_\theta}(\eta_t)}{\sqrt{t}}] \leq 2\left(\frac{\theta^2 + (1-\theta)^2}{\theta(1-\theta)}\right)\sqrt{\pi}$$

Proof. Theorem 7 in Bramson and Griffeath (1980), p. 211.

An immediate corollary of the above Theorem is the fact that, as t gets large, the largest segment containing all individuals choosing the same opinion, has side of probability order $\sqrt{t}$. For t very large, such cluster tends to be *almost stationary*, in the sense that the rate at which it changes is slower than the rate at which time changes[9].

The Theorem provides a numerical lower and an upper bound for the expected mean cluster size. To interpret this estimate, consider a process that starts with an initial distribution where each voter chooses opinion 1 with probability, say, $\theta = \frac{1}{2}$. As choices are initially independent, clearly, at time zero, the probability of observing a cluster of $k = 100$ voters with the same opinion is 2^{-100}. As the process evolves, however, choices show a certain amount of spatial correlation. For $t \to \infty$ the mean cluster size, re-scaled by $\sqrt{t}$, will converge to a limit that lies between $2\sqrt{\pi} = 3.5449$ and $4\sqrt{\pi} = 7.0898$. Hence a cluster of $k = 100$ voters could be approximately observed as early as after $t = 198.94$ [10], and is on average not going to vary until $t = 795.78$, as the gure that follows illustrates:

Limiting mean cluster size for $\theta = \frac{1}{2}$.

In other words, in order to observe the cluster size to double (say from $k = 100$ to $k = 200$ in the above picture), the process needs to go through four times as many periods (say from $t \sim 200$ to $t \sim 800$).

Simple calculus shows that the lower and the upper bound of the (limiting) mean cluster size are convex in θ and symmetric around $\theta = \frac{1}{2}$. Hence for $\theta \neq \frac{1}{2}$ a cluster of a given mean size is likely to be observed earlier than if θ was $\frac{1}{2}$ and is likely to persist for a relatively longer spell of time. Hence, conditional on a candidate winning the elections, his or her support in terms of absolute number of votes grows at rate $\sqrt{t}$. The higher is θ, the lower is the number of time periods that are necessary to achieve a given minimum expected cluster size of votes in her or his favour, and the longer is the spell of time within which his or her electoral support is going to remain almost stationary.

4 Insights for further research

As the dynamics we studied are speci ed over time and over space, natural questions to be addressed relate to the optimal *spatial* allocation of funding in an electoral campaign (i.e. among different districts or different states), as well as to the optimal *timing* of such allocation (i.e. between the time when the elections are called and the time just before the elections are actually held). Although a formal treatment of these interesting questions warrants future research, in what follows we elaborate on the insights that the model we studied in this paper provides.

The first thing that all specifications of our model show is that the spatial distribution of votes matters in the long run, as well as in the short run. In particular, simply by looking at the limit distribution for the ergodic process generated by the dynamics of the *uninformed voters* model, as in Theorem 1, it is easy to see that the limit probability of each configuration depends on the opinions chosen in its connected components, and not on the frequency with which opinions are adopted in the population. For example, consider two configurations, η_A and η_B, identical at all sites apart from the sites $\{x-2, x-1, x, x+1\}$ which are as follows:

$$\begin{array}{llllll} \eta_A: & .... & \eta(x-2)=1 & \eta(x-1)=0 & \eta(x)=1 & \eta(x+1)=0 & .... \\ \eta_B: & .... & \eta(x-2)=1 & \eta(x-1)=1 & \eta(x)=0 & \eta(x+1)=0 & .... \end{array}$$

From Theorem 1 we infer that the limit probabilities of these configurations (where the frequencies of 1s is exactly the same) are respectively:

$$\mu_\infty^\sigma(\eta_A) \propto \exp[-6\sigma]$$
$$\mu_\infty^\sigma(\eta_B) \propto \exp[2\sigma]$$

Configuration η_B is given higher probability, as more coordinates agree with their neighbouring coordinates. These considerations clearly relate to the long-run distribution of the process, but the insight applies to the short-run, as can be seen by looking at the dynamics of the specification of the model in terms of *linear voters*, to which we focus next.

Much of the descriptive and normative literature on elections in political science identifies at least two alternative basic rules that a candidate may follow when deciding where to allocate resources (in terms of money, as well as time spent campaigning) among different constituencies or states. The first posits that a candidate should allocate campaign resources roughly in proportion to the electoral votes of each state (Brams and Davis (1974)). The second suggests that candidates should mostly be concerned with the likelihood that resources can swing a state from one candidate to another, and by this advocates a competitive allocation of resources to be directed to the marginal states (Colantoni et al. (1974)). With some heroic simplifications, we can translate these two alternatives into the set-up of our model, by asking the following question: suppose a candidate had the possibility to buy one vote (i.e. to buy the support of one voter), would (s)he rather do so *within* a cluster of voters who support the other candidate, or exactly at the *border* of a cluster? It turns out that our model suggests that the best alternative is this latter possibility. To see this, consider the following con guration, η, that has a border at $x = 0$, in that $\eta(x-1) \neq \eta(x)$:

$$.... \quad \eta(x-2) = 1 \quad \eta(x-1) = 1 \quad \eta(x) = 0 \quad \eta(x+1) = 0 \quad \eta(x+2) = 0 \quad ....$$

Suppose, for simplicity, that the process is started deterministically at con guration η. In this case the duality equation (**??**) (see the proof of Theorem 2 in the Appendix) states that the probability that starting from con guration η, the voter at site x supports candidate 1 is: $E^\eta \eta_t(x) = \sum_y p_t(x,y)\eta(y)$, which, applied to the subset $\{x-1, x, x+1\}$, becomes:

$$E^\eta[\eta_t(x-1)+\eta_t(x)+\eta_t(x+1)] = \sum_y p_t(x-1,y)\eta(y) + \sum_y p_t(x,y)\eta(y) + \sum_y p_t(x+1,y)\eta(y)$$

The above probabilities are given explicitly in equation (**??**), and it is not difficult to see that, for any nite t, since $p_t(x, x+j) = p_t(x, x-j)$ for any $j \geq 1$ and since $p^{(0)}(x, x+1) = p^{(0)}(x, x-1) = \frac{1}{2}$:

$$\frac{1}{2} \geq p_t(x, x+1) - p_t(x, x+j) > 0 \qquad \forall j > 1$$

formalizing the fact that a voter s opinion is more strongly affected by the opinions held in the neighbourhood than by opinions held further away.

If we take into account of this fact, and we denote $p_t(x, x+1)$ as p, we can re-write the above equation as:

$$\begin{aligned} E^\eta[\eta_t(x-1) + \eta_t(x) + \eta_t(x+1)] &\approx p[\eta(x-2) + \eta(x-1) + 2\eta(x) + \eta(x+1) + \eta(x+2)] \\ &= p[1 + 1 + 2\eta(x) + \eta(x+1) + \eta(x+2)] \end{aligned}$$

Hence, by buying the vote of voter x, candidate 1 increases the probability that at time t voters in $\{x-1, x, x+1\}$ support her or him by twice as much as (s)he would do by buying the vote of voter $x+1$ or voter $x+2$. This is because by moving the border of a cluster by one voter, the candidate guarantees stability of the area inside the cluster, that being inward looking is not so exposed to sudden swings in opinions.

A further insight that the model provides relates to the optimal timing of resource allocations in an electoral campaign. As we showed before, in the *linear voter* model the process is path-dependent, as its long run behaviour depends crucially on the initial distribution. This determines the basins of attraction of the two limit distributions (that, we recall, show consensus), as well as the lower and upper bound of the expected minimum cluster size. Hence the model suggests that what happens at the very beginning of an electoral campaign has a very strong effect on its later developments, and raises the incentive for a candidate to invest campaign resources on whatever is deemed to have any power to affect the initial distribution. Along the dynamics, clusters emerge and are almost stationary when viewed locally. Clearly, as the model is non-ergodic, there is no guarantee that once a minimum cluster size is reached, electoral support for a candidate will continue to grow unboundedly. Hence, if a candidate could gather some information about the current distribution of potential votes (for example through an electoral poll) and if this was favourable to her or him, then delaying the date of the elections could have a detrimental effect on the outcome. As clusters grow at rate $\sqrt{t}$ the model also seems to suggest that a linear allocation of funding over time during an electoral campaign might be sub-optimal, as the returns in terms of electoral support are decreasing over time[11].

5 Concluding remarks

This paper studied a dynamic model of pre-electoral public opinion formation. We treated public opinion as a continuously changing process and we analyzed the emergence of interactive patterns of behavior.

The model involves a countable population of individuals that repeatedly choose to support one of two candidates. Each individual has a well de ned preference structure over the nal electoral outcome that formalizes an incentive to conform to the opinion held by a perceived majority. In the rst speci cation of the model agents update their beliefs over the current distribution of opinions by sampling a number of observations within their neighbours. In the second, voters simply follow a linear rule, by choosing opinion 1 with probability α if $\alpha\%$ of their neighbours hold opinion 1 (and viceversa).

We analyzed the dynamics of the public opinion process by addressing two related questions. The rst relates to the asymptotics of the process. For certain speci cations we showed that the process is ergodic, while for others the process admits two extreme invariant measures, where one candidate is supported by the whole population. The second question explicitly focused on the dynamics itself, by pursuing a space-time analysis. It turned out that the process displays the feature that, after long time spans, a sequence of states occur which, when viewed locally, remain almost stationary and are characterized by large clusters of individuals of the same opinion.

Finally, we provided some heuristic considerations on the implications that these ndings could have within a more general model that allows for strategic behaviour on the part of the two candidates.

Appendix

Proof of Theorem 1

(1) (ergodicity)

We interpret the process η^U as a system of interactive, nearest neighbours, particles on a one-dimensional lattice, Z^1. We rst show that the process is *attractive* (or monotonic) in that coordinates tend to agree with neighbouring coordinates. It is known (see, for example, Liggett (1985), Theorem 3.14, p.152) that a sufficient condition for an attractive system with a countable state-space to be ergodic, is that the transition probabilities that generate the process be strictly positive. This is the logic we follow.

We introduce the following partial order on $\{0,1\}^{Z^1}$. We say that, for $\eta, \zeta \in \{0,1\}^{Z^1}$, $\eta \leq \zeta$ if $\eta(x) \leq \zeta(x)$ for all $x \in Z^1$. Then a process is de ned to be *attractive* if, whenever $\eta \leq \zeta$:

$$\begin{aligned} c(x,\eta) &\leq c(x,\zeta) \text{ if } \eta(x)=\zeta(x)=0 \\ c(x,\eta) &\geq c(x,\zeta) \text{ if } \eta(x)=\zeta(x)=1 \end{aligned}$$

where $c(x,\cdot)$ are the ip rates that generate the dynamics (i.e. $c(x,\cdot)$ is the probability with which coordinate x ips, in state $\cdot$).

In order to check this condition, and for later purposes, we re-write the transition probabilities of which in equation (**??**) by substituting $\sigma = \frac{1}{4}\log(2^{n+1}-1)$:

$$\Pr{}^U[1 \mid n, p(x,\eta)] \equiv \Pr{}^U[1 \mid \sigma, p(x,\eta)] = \tag{6}$$

$$= \frac{1}{1+\exp[-4\sigma(2p(x,\eta)-1)]} \tag{7}$$

where we recall $p(x,\eta) = \frac{1}{2}\sum_{y:\|y-x\|=1}\eta(y)$ and takes values in $\{0, 1/2, 1\}$. For example, if $p(x,\eta)=0$ the above equation states that the probability that opinion 1 is chosen is given by $[1+\exp[4\sigma]]^{-1} = [1+\exp[\log[2^{n+1}-1]]]^{-1} = (2^{n+1})^{-1}$.

Hence, the ip rates can be written as:

$$\begin{aligned} \Pr{}^U[1 \mid \sigma,\, p(x,\eta),\, \eta(x)=0] &= \frac{1}{1+\exp[-4\sigma(2p(x,\eta)-1)]} \\ \Pr{}^U[0 \mid \sigma,\, p(x,\eta),\, \eta(x)=1] &= \frac{1}{1+\exp[+4\sigma(2p(x,\eta)-1)]} \end{aligned}$$

As clearly $p(x,\eta) \leq p(x,\zeta)$ whenever $\eta \leq \zeta$, our process is attractive. It is also clear from equation (**??**) with $n < \infty$, or equivalently from the above speci cations, for $\sigma < \infty$, that transition probabilities are strictly positive. Hence $\mathfrak{I}$ is a singleton. ■

(2) (characterization of the limit distribution)

In order to characterize the unique invariant measure of the process, we establish a relation between the process η_t^U and a class of stochastic processes known as Ising models (Liggett (1985), Chapter IV provides details on Ising models).

It is easy to see that the transition probabilities of which in equation (**??**) correspond exactly to the ip rates of a stochastic Ising model, with nearest neighbour interactions, relative to the following potential:

$$J_R = \begin{cases} \sigma \text{ if } R = \{x,y\} \text{ and } y : \| y - x \| = 1 \\ 0 \text{ otherwise} \end{cases} \tag{8}$$

Hence the unique limit distribution of the process is the Gibbs state corresponding to the above potential:

$$\mu_\infty^\sigma(\eta) = K \exp[\sum_x \sum_{\{y:|y-x|=1\}} \sigma(2\eta(x)-1)(2\eta(y)-1)]$$

where $K = \{\sum_\eta \exp[\sum_x \sum_{\{y:|y-x|=1\}} \sigma(2\eta(x)-1)(2\eta(y)-1)]\}^{-1}$ (see for example equation (1.4) on p. 180 in Liggett (1985)).

To see this, it suffices to notice that the above measure is *reversible*, in that:

$$\Pr{}^U[1 \mid \sigma,\, p(x,\eta),\, \eta(x)=0]\mu_\infty^\sigma(\eta_{x=0}) = \Pr{}^U[0 \mid \sigma,\, p(x,\eta),\, \eta(x)=1]\mu_\infty^\sigma(\eta_{x=1})$$

where the two con gurations $\eta_{x=0}$ and $\eta_{x=1}$ differ only in the coordinate x (i.e. $\eta_{x=0}(x) = 0, \eta_{x=1}(x) = 1$ and $\eta_{x=0}(y) = \eta_{x=1}(y)$ for all $y \neq x$):

$$\begin{aligned} \frac{\mu_\infty^\sigma(\eta_{x=1})}{\mu_\infty^\sigma(\eta_{x=0})} &= \exp[2\sigma \sum_{\{y:\|y-x\|=1\}} (2\eta(y)-1)] \\ &= \frac{1}{1+\exp[-2\sigma \sum_{\{y:\|y-x\|=1\}}(2\eta(y)-1)]} \cdot \\ &\quad \cdot [\frac{1}{1+\exp[2\sigma \sum_{\{y:\|y-x\|=1\}}(2\eta(y)-1)]}]^{-1} \\ &= \frac{\Pr{}^U[1 \mid \sigma,\, p(x,\eta),\, \eta(x)=0]}{\Pr{}^U[0 \mid \sigma,\, p(x,\eta),\, \eta(x)=1]} \end{aligned}$$

Clearly any reversible measure is also an invariant measure (i.e. $\mu_\infty^\sigma \in \mathfrak{I}$). From part (1) we know that the process is ergodic, i.e. $\mathfrak{I}$ is a singleton. Hence the assert follows. ■

Proof of Theorem ??

1. Recall that the process $\eta_t^U(\infty)$ is de ned by the transition probabilities of which in (**??**). Hence, we only need to show that $\lim_{\sigma\to\infty} \Pr^U[1 \mid \sigma, p(x,\eta)] = p(x,\eta)$ for $p(x,\eta) \in \{0, \frac{1}{2}, 1\}$. This is clear by looking at (**??**) and taking such limit, for any given $p(x,\eta)$. ■

2. Clearly, for the process η_t^L, $\mathfrak{I} \supseteq \mathfrak{I}_e \supseteq \{\mu_0, \mu_1\}$, as by simple inspection of the transition probabilities that de ne the process (namely $p(x,\eta)$ for $p \in \{0, \dfrac{1}{2}, 1\}$) it is clear that any state for which $\eta(x) = \eta(y)$ for all x, y in S is stationary for the process. Hence, the result relies on the proof that these are the *only* two extreme invariant measures (i.e. $\mathfrak{I}_e \subseteq \{\mu_0, \mu_1\}$), so that, as $\mathfrak{I}$ is a convex set, any other invariant measure is fully characterized. Furthermore, one needs to show that the domains of attraction of each extreme invariant measure, depend on the stochastic initial condition given by the product measure μ_θ, as in $\mathcal{L}(\eta_t^{\mu_\theta}) \to (1-\theta)\mu_0 + \theta\mu_1$.

Results along these lines are well known in the statistical literature on the Voter s model in the case Z^1 and can be found in Liggett (1985), Section 1 and 3, Chapter V or in Bramson and Griffeath (1980). As the logic of the proofs is interesting in its own right, we sketch the proof in what follows.

The process η_t^L (shortened to η_t in what follows) can be studied in terms of its dual process in terms of coalescing random walks. The duality relation transforms questions about η_t in questions concerning the cardinality of the coalescing random walk system.

We rst show that such duality can be used, by checking the conditions of which in equation. (4.3) (p. 158) in Liggett (1985). To this aim, note that the ip rates for the process η_t^L can be written as:

$$\begin{aligned} c(x,\eta) &= \eta(x) + p(x,\eta)(1-2\eta(x)) \\ &= (1-\eta(x)) + (2\eta(x)-1) \sum_{\{y:\|y-x\|=1\}} \frac{1}{2}(1-\eta(y)) \end{aligned}$$

as $p(x,\eta) = \sum_{\{y:\|y-x\|=1\}} \frac{1}{2}\eta(y)$. These coincide with equation. (4.3) (p. 158) in Liggett (1985), once we take $c(x) = 1$, $A = \{y\}$ and $p(x,A) = p(x,y) = \frac{1}{2}$ if $y :\| y - x \|= 1$ and zero otherwise.

The dual process is a system of countably many continuous time, symmetric random walks that jump after an exponential mean-1 holding time, with probabilities $p(x, x+1) = p(x, x-1) = \frac{1}{2}$. Whenever two random walks meet (i.e. if one jumps to a site that is already occupied), then they coalesce, i.e. they merge into one. In particular, any such random walk de nes a continuous time Markov chain, $X(t)$, with transition probabilities:

$$p_t(x,y) = e^{-t} \sum_{n=0}^{\infty} \frac{t^n}{n!} p^{(n)}(x,y) \tag{9}$$

where $p^{(n)}(x,y)$ are the n-step transition probabilities associated with $p(x,y)$. Any system of nitely many independent copies of $X(t)$, where any two merge whenever they meet, de nes a system of nitely many coalescing Markov chains over the state space of all nite subsets of S.

We denote by A_t the system of coalescing random walks at time t, that started at time zero in the nite subset $A \subset S$. For any such subset A, let:

$$g_t(A) = \Pr{}^A[\mid A_t \mid < \mid A \mid \text{ for some } t \geq 0]$$

where $\mid \cdot \mid$ denotes the cardinality of a set. This represents a measure of how far apart the single processes are. Clearly, for any t, $g_t(A) = 0$ when $\mid A \mid= 1$, as a single recurrent random walk is never going to die. If $\mid A \mid= 2$, $g_t(A) \rightarrow_{t\rightarrow\infty} 1$, meaning that two recurrent random walks will tend to meet and coalesce, as time grows, and possibly only asymptotically. In order to shorten an otherwise very long proof, we shall however assume that $g_{t^*}(A) = 1$ when $\mid A \mid= 2$ for some $t^* < \infty$.

Let $A = \{x \in S : \eta(x) = 1 \quad \text{for all } x \in A\}$ and, for μ being a probability measure on $\{0,1\}^S$, let $\mu(A) = \mu\{\eta : \eta(x) = 1 \text{ for all } x \in A\}$. Then the duality equation can be stated as follows (see equation. 1.7, p. 230 in Liggett (1985)):

$$\mu_t(A) = E^A \mu(A_t) \tag{10}$$

where $\mu_t(A)$ is the probability that the process η_t has $\eta_t(x) = 1$ for all $x \in A$ and $E^A\mu(A_t)$ is the probability that $\mid A_t \mid$ random walks, started at A, are still alive at time t.

By using this duality relation, we now show that, given a product measure μ^θ, $\mathcal{L}(\eta_t^{\mu_\theta}) \rightarrow (1-\theta)\mu_0 + \theta\mu_1$.

To characterize the basins of attraction of $\{\eta_0, \eta_1\}$, suppose the process η_t is started (stochastically) with product measure μ_θ. If τ is the rst time that $| A_t |= 1$ (which is nite with probability one by our assumption that $g_{t^*}(A) = 1$ when $| A |= 2$) the duality equation (**??**) implies that:

$$\lim_{t\to\infty} E^A \mu(A_t) = E^A[\lim_{t\to\infty} E^{A_\tau} \mu(A_t)]$$

Applying this again to $A = \{x\}$ we obtain:

$$\lim_{t\to\infty} \sum_y p_t(x, y)\mu(\{y\}) = \theta \text{ for all } x \in S$$

But, by part (b) of Theorem 1.9 in Liggett (1985) (p. 231), this is a necessary and sufficient condition for $\mathcal{L}(\eta_t^{\mu_\theta}) \rightarrow (1-\theta)\mu_0 + \theta\mu_1$ to be true. Hence the assert follows. ■

Remark

The speci cation of the model we use throughout the paper allows for a countable set of individuals located on a one-dimensional lattice. One may wonder whether the results are peculiar to this characterization, other things being equal.

Most of the results would also hold for a nite (hence countable) set of individuals. In this case the dynamics of the uninformed voter model and that of the linear voter model would de ne a continuous time Markov chain over a nite state-space. The analog of Theorem 1 (and its Corollary) would hold in any dimension, on the grounds that the chain would be ergodic (whereas for a countable state-space, it is known that in a d-dimensional setting with $d \geq 2$ the process may admit multiple invariant distributions). As for the analog of Theorem 2, independently of the dimension, the process η_t^L would converge with probability one to $\{\eta_i, i \in \{0, 1\}\}$, as the chain would be absorbing, with these con gurations as the only absorbing states. As a result, for any given initial condition, the process would get trapped (in nite time) in a con guration that shows consensus. Unlike in the case of a countable population, each voter could change her or his opinion only nitely many times. We conjecture that also the domains of attraction of the two stationary measures, μ_0 and μ_1, would be exactly the same. However, the technique used in the proof we provide, as well as that used to proof Theorem 3 (Bramson and Griffeath (1980)) heavily relies on the duality in terms of coalescing random walks and it is known that these behave differently for $d \leq 2$ and for $d > 2$. In particular, as a finite system will be trapped with probability one, the analysis of the clustering process could not be carried out. An interesting question to address in this case is the study of the way absorption times vary, for the number of individuals growing large (as in Cox (1989)). In Corradi and Ianni (1998b) we study the details of the clustering process in a 2-dimensional model related to those we analyzed in this paper.

Notes

[1]The wording is by I. Crespi, the former president of the American Association for Public Opinion Research.

[2]In Corradi and Ianni (1998a) we further investigate the relation between this clustering process and stationary co-existence of different opinions in a model analog to the one we study in this paper, for different speci cations of the dynamics.

[3]As we shall discuss later, the focus on a countable (not necessarily nite) population of individuals is motivated by the fact that we want to analyze explicitly the process of cluster formation. To this aim, we shall assume that the set of individuals is located on a one dimensional lattice. Consistently, the frequency of opinions in the population is de ned as the limit of its natural restriction to $[-l, l]$, as $l \to \infty$. For the purposes of the exposition, we assume that this limit exists when describing the process of Private Opinion Formation. As it will become clear, such assumption will be trivially satis ed by our process of Public Opinion Formation, as interactions will only have nite range.

[4]Riker and Ordershook (1968) estimated such probability for the USA as being 10^{-8}.

[5]An interesting philosophical discussion of *rational* voting decisions is provided in Meehl (1977).

[6]This assumption is only introduced for technical convenience. It is consistent with the idea of symmetric (equilibrium) behaviour, on the part of the two candidates, if side-payments where decided strategically. As we formally show in Corradi and Ianni (1998a), the dynamics of the public opinion process in the case where the side-payments are asymmetric are far less articulate than otherwise.

[7]More precisely, it varies upper hemicontinuously with σ in the weak convergence topology.

[8]The result is actually true for any *n-fold mixing measure* as de ned in Bramson and Griffeath (1980).

[9]In Ellison (1993) the author studies the rates of convergence of best-reply dynamics for an underlying coordination game, repeatedly played by couples of players drawn at random from a nite population. Our model differs from the cited paper in a number of respects. First, the speci cation of the dynamics that Ellison (1993) studies is perturbed by mistakes (that take the form of a binomial distribution that assigns small, though strictly positive, probability, uncorrelated across players and over time, to actions that are not best-replies to the current con guration of play). This is substantially different from the way we model the individual process of opinion formation (that in the speci cation in terms of uninformed voters could be motivated in terms of mistakes that *do depend* on expected payoffs). Second, the dynamics of Ellison s (1993) are de ned over a nite state-space and modeled as nite, discrete time, regular Markov chains, whereas our dynamics de ne a Markovian process over a countable state-space, that is ergodic if voters are uninformed , but is path-dependent if voters are linear . Lastly, the cited paper compares the speed of convergence of

transition probabilities to their limit values, in a model with local interaction and in an (analog) model with global interaction. All speci cations of our model rely on a local characterization of the way in which interaction takes place.

[10]Of course the quality of the approximation improves with t.

[11]However, considerations of this sort require an explicit consideration of the strategic interaction between the two candidates, which at present is not part of the model.

REFERENCES

ANDERLINI, L. AND A. IANNI (1996), Path Dependence and Learning From Neighbours," *Games and Economic Behavior,* **13**, 141-177.

BANERJEE, A.V. (1992), A Simple Model of Herd Behavior," *Quarterly Journal of Economics,* **107**, 797-817.

BAUMOL, W.J. (1957), Interactions between Successive Polling Results and Voting Intentions," *Public Opinion Quarterly,* **21**, 318-323.

BIRKHCHANDANI, S., D. HIRSHLEIFER AND I. WELCH (1992), A Theory of Fads, Fashions, Custom, and Cultural Change as Information Cascades," *Journal of Political Economy,* **100**, 992-1026.

BLUME, L.E. (1993), The Statistical Mechanics of Strategic Interaction," *Games and Economic Behaviour,* **5**, 387-424.

BOGART, L. (1985), *Polls and the Awareness of Public Opinion,* L. Erlbaum Ass.

BRAMS, S.J. AND M.D. DAVIS (1974), The 3/2 s Rule in Presidential Campaigning," *The American Political Science Review,* **68**, 113-134.

BRAMSON M. AND D. GRIFFEATH (1980), Clustering and Dispersion Rates for Some Interacting Particle Systems on Z^1," *Annals of Probability,* **8**, 183-213.

COLANTONI, C.S., T.J. LEVESQUE AND P.C. ORDESHOOK (1974), Campaign Resource Allocation Under the Electoral College," *The American Political Science Review,* **69**, 141-154.

COOLEY, C.H. (1918), *Social Process,* Charles Scribner and Son.

CORRADI, V. AND A. IANNI (1998a), Consensus and Co-existence in a Simple Model of Opinion Fromation," *mimeo,* University of Pennsylvania.

CORRADI, V. AND A. IANNI (1998b), A Note on the Dynamic Properties of Imitative Behaviour," *mimeo,* University of Pennsylvania.

COX, J.T. (1989), Coalescing Random Walks and Voter Model Consensus Times on the Torus in Z^d," *Annals of Probability,* **17**, 1333-1366.

CRESPI, I. (1997), *The Public Opinion Process,* L.E.A. Publishers.

DEVENOW, A. AND I. WELCH (1996), Rational Herding in Financial Economics," *European Economic Reviw,* **40**, 603-615.

ELLISON, G. (1993), Learning, Local Interaction, and Coordination," *Econometrica* **61**, 1047-1071.

ELLISON, G. (1995), Basins of Attraction, Long Run Equilibria and the Speed of Step-by-Step Evolution," *mimeo,* MIT.

ESHEL, I., SAMUELSON L. AND A. SHAKED (1998), Altruists, Egoists, and Hooligans in a Local Interaction Model," *The American Economic Review* **88**, 157-179.

FARRELL, J. AND G. SALONER (1985), Standardization, Compatibility, and Innovation," *Rand Journal of Economics* **16**, 70-83.

IANNI, A. (1998), Learning Correlated Equilibria in Potential Games," CARESS, 98-05.

KANDORI, M. AND R. ROB (1998), Bandwagon Effects and Long-Run Technology Choice," *Games and Economic Behavior* **22**, 30-60.

KATZ, M. AND C. SHAPIRO (1985), Network Externalities, Competition and Compatibility," *American Economic Review* **75**, 424-440.

LEE, I.H. (1993), On the Convergence of Informational Cascades," *Journal of Economic Theory* **61**, 395-411.

LIGGETT, T.M. (1985), *Interacting Particle Systems,* Springer-Verlag.

MCALISTER I AND D.T. STUDLAR (1991), Bandwagon, Underdog, or Projection? Opinion Polls and Electoral Choice in Britain 1979-1987," *The Journal of Politics,* **53**, 720-741.