Consensus, Contagion and Clustering in a Space-Time Model of Public Opinion Formation[*][†]

Antonella Ianni[‡]

University of Venice (Italy), E.U.I. and University of Southampton (U.K.),

Valentina Corradi
University of Warwick (U.K.)

Abstract

We study a simple model of public opinion formation that posits that interaction between neighbouring agents leads to bandwagons in the dynamics of individual opinions, as well as in that of the aggregate process. We show that in different specifications of the model, there is a tendency for the process to show consensus on one of the two competing opinions.

We show how a publicly available poll of current public opinion may lead to a form of contagion, by which public opinion tends to agree with the poll. We point out that, in the absence of a poll, 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.

The running metaphor we use is that of a model of pre-electoral public opinion formation, with two candidates running. We provide some heuristic considerations on the implication that these findings could have in terms of space-time allocation of fundings in an electoral campaign.

1 Introduction

In order to understand the emergence and the dynamic stability of social norms and patterns of behaviour in society, a growing body of recent literature analyzes interdependent settings where the communication structure within a large population of agents is highly decentralized. Most of this literature focuses on situations where the relative advantage of choosing one, out of finitely many actions, is increasing in the probability with which other agents do so, thus formalizing an explicit payoff externality. This incentive structure raises an issue of coordination, that even in the absence of any explicit coordination device, may be achieved by society over time when individuals are randomly matched and interaction takes place repeatedly. The results to date indicate that, whenever a risk-dominant (as introduced by Harsanyi and Selten (1988)) equilibrium exists, this is robust to stochastic perturbations that may come from postulated mistakes or mutations, or from randomness in the initial condition. An implication of these findings is that the risk-dominant action spreads in the population in a form of *contagion*. In particular, in a locally interactive setting, starting from a small *cluster* of agents playing the risk-dominant action, myopic best reply dynamics ensure that this action will, eventually and with probability one, be adopted by all other agents in the population.

This paper analyzes a simple stochastic dynamic process of public opinion formation and it studies the way in which this evolves over space and over time. The setting to which our results apply is basically the following. There is a large population of agents who repeatedly choose one out of two possible actions. Returns to each action depend on the state of the world, and this is not known with certainty. Agents derive a posterior probability on the basis of a symmetric binary signal they receive and/or by observing a sample of actions taken within a subset of other agents (their *neighbours*). Observed actions are informative, since signals are (*informational externality*). While signals are generated by the same probability distribution, observed actions are not, and this introduces a degree of informational heterogeneity across agents.

Specifically, the metaphor we use to describe the model is that of a process of pre-electoral opinion formation, where individuals repeatedly form their own opinion as to which, out of two, candidates to vote for, as the time when elections come. We postulate that the way in which opinions are chosen reflects an informational externality that leads agent to conform to other individual choices. Since the population is large and communication among individuals is highly decentralized, other individuals' opinions are not perfectly observed. Choices are hence made on the ba-

sis of a sample of observations that each agent gathers each time a choice is to be made. Heterogeneity among agent may hence arise endogenously from the different information available to each. Additionally, individuals may rely on some publicly available statistics of the current public opinion. We refer to the process of public opinion as to the dynamic process generated by the collection of all individual opinions and we are interested in analyzing the properties of this dynamics. In particular, we address two complementary issues. The first relates to the asymptotic properties of the dynamics. Starting from an initial random configuration of opinions in the population, will the process of public opinion formation eventually show *consensus* in the aggregate, or will different opinions *co-exist* indefinitely? Under which condition will one opinion propagate in the population in a form of *contagion*? The second focuses on the behaviour of the process along the dynamics and relies on the spatial characterization of the process of *clustering*, by which consensus is achieved. In our model, it is the aggregate of all individual opinions that shapes and defines public opinion in its collective dimension, and in particular, electoral support for each of the candidates. In this sense, and as we discuss towards the end of the paper, our results provide a basis for analyzing issues of space and time allocation of funding in an electoral campaign.

Issues related to the process of public opinion formation are in fact 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. The process of public opinion has also been extensively studied in fields other than economics. Our model provides a formalization of two aspects that 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, as well as empirical evidence.

The paper is organized as follows. Section 2 describes the details of the process of private and public opinion formation. The probability with which a voter chooses an opinion depends on a sample of observations of other voters' opinions, as well as on a publicly available electoral poll. The collective processes rely 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, the distribution from which observations are drawn at random is endogeneized in terms of a simple statistic of the opinions adopted in a voter's neighbourhood. Section 3 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 process, as well as those of its dynamics. The results show that these two aspects are complementary and provide a better understanding of the process itself. By pursuing a space-time analysis (i.e. by relating the two dimensions, time and space, over which our process is defined) we also study the process of *cluster formation.* In Section 4 we provide some heuristic considerations as to the implications that these findings would have within a more general model that allows for strategic behaviour on the part of the two candidates. Finally Section 5 concludes and the Appendix contains the technical proofs.

2   The model

A single round of elections is going to be held at a future date. Two candidates, 0 and 1, run the elections and the winner will be decided through simple majority voting[1]. 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 defined 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 is a population of voters that formulate their opinion as to which candidate to support when the elections will be held. Voters behave in an identical manner, although, as we shall see, asymmetries might arise due to differences in the information they possess. In our setting, voting decisions on the part of each single voter are almost by definition deprived of any strategic content. The model we formalize in fact postulates that each possible electoral outcome is 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, 0 or 1, which would correspond to a vote (for 0 or for 1 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 as $w = \{0, 1\}$ corresponding to the event "candidate 0 wins the elections" and "candidate 1 wins the elections"; two possible ballots, labelled $v = \{0, 1\}$ and a utility function that depends separably on the outcome of the election and on the ballot chosen, $U(w, v)$. The idea we want to pursue is that, although the outcome of the elections is exogenous to the voter, utility also depends on the vote itself. Specifically, and for $\varepsilon > 0$, we assume that the utility of each voter is given by:

$$U(w, v) = \begin{cases} 0 & v \neq w \\ \varepsilon & v = w \end{cases}$$

This formalizes the idea that voters strictly prefer to vote for the winning candidate[2]. Given these preferences, and given that the outcome of the elections depends on the distribution of ballots in the population, a voter needs to assess the probability with which each candidate is going to win the elections. Given this assessment, a voter will conform to what (s)he perceives to be the opinion of the majority of voters in the population.

2.1 Private Opinion Formation

Voters' decisions clearly depend on the assessment of the probability with which each candidate will win the elections. We postulate that the latter is made on the basis of the inference drawn from some privately gathered observations of other voters' opinions, as well as on the basis of a publicly available electoral poll. We think of the poll as a signal on the uncertain state of the world. If we let $p \in [0, 1]$ be a parameter that defines the probability with which a voter privately observes an opinion in favour of candidate 1 in a sample of n observations, $\pi \in [0, 1]$ the assessment of candidate 1's percentage votes produced by a public electoral poll, and $\alpha \in [0, 1]$ the relative weight of privately and publicly gathered information in the voter's assessment, the probability with which the voter formulates an opinion that favours candidate 1 is modeled as:

$$\Pr\left[1 \mid \alpha, p, n, \pi\right] \equiv \alpha \Pr\left[1 \mid p, n\right] + (1 - \alpha)\Pr[1 \mid \pi]$$

Throughout the paper we shall take α to be given exogenously and identical for all voters in the populations. The electoral poll that produces π is clearly endogenous, but is not modeled explicitly. The idea we have in mind is that a single poll, in terms

of some sample statistics, is produced at a given point in time and it is freely available to all voters. Most of the Section that follows is devoted to the endogeneization of the parameter p, that will account for heterogeneities across voters that arise from a highly decentralized information structure in the population. For the sake of the exposition, we take α to be one in the rest of this sub-Section and we formalize the process of inference that leads to the assessment of the probability with which either candidate will win the elections as follows.

We assume that each voter has flat priors over $\nu \in [0,1]$ which is the fraction of voters in the population who currently support candidate 1, 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. Thus, each observation comes from a Binomial distribution, $Bi(p)$, where $0 \leq p \leq 1$. 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 1 in a sample of n observations.

If a voter has observed a sample of r opinions 1 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 1 if the $\Pr_{f_{Be}}[z > 1/2 \mid r,n] > 0.5$, and opinion 0 otherwise. Hence, we can calculate the ex-ante probability of choosing opinion 1, given p and $n \geq 1$ observations[3]. We find it convenient to take n to be an odd number and to write this probability, for $n = 2m+1$ as:

$$\Pr[1 \mid p, m] = \sum_{r=m+1}^{2m+1} \binom{2m+1}{r} p^r (1-p)^{2m+1-r} \quad m \geq 0 \tag{1}$$

This quantity depends on m (the number of observations) and on p (the parameter that determines the probability of observing a 1). It is not difficult to see that, for any fixed $m = \overline{m} < \infty$ the above probability is a continuous and increasing function of $p \in [0,1]$, formalizing the fact that the more likely observations 1s are, the higher is the probability that the voter will adopt opinion 1. It is also clear from (1) that, for any m, $\Pr^c[1 \mid p=0, m] = 0$, $\Pr^c[1 \mid p=0.5, m] = 0.5$ and $\Pr^c[1 \mid p=1, m] = 1$. However, for any $p \in (0, 0.5) \cup (0.5, 1)$, since observations are realizations of the random variable $Bi(p)$, the voter's behaviour is only described probabilistically in the sense that $0 < \Pr^c[1 \mid p, m] < 1$.

It is worth noticing that, for $m = 0$, the above probability is linear in p, since $\Pr^c[1 \mid p, m=0] = p$. This resembles pure imitative behaviour on the part of a voter,

who only samples one observation and blindly imitates it. To see the way in which the above probability changes as the number of observation increases, suppose p was actually the true proportion of 1's supporters in the population. Let $s^*(p)$ denote the probability with which an expected utility maximizer voter should choose opinion 1, i.e. $s^*(p) = 0$ for $p < 0.5$, $s^*(p) = 1$ for $p > 0.5$ and, conventionally, set $s^*(p) = 0.5$ for $p = 0.5$. Then[4], for the number of observations becoming very large and for each given $p = \bar{p}$, $\lim_{m \to \infty} \Pr^c[1 \mid \bar{p}, m] = s^*(\bar{p})$. In other words, for m large, $\Pr^c[1 \mid \bar{p}, m]$ is essentially described by $s^*(\bar{p})$, although it remains differentiable, since only in the limit, for $m \to \infty$, its image is restricted to the values $\{0, 0.5, 1\}$.

In order to account for the fact that the information gathered by a voter, or his or her behaviour may be not be as systematic as prescribed by equation (1), we shall also analyze a noisy version of it. Specifically, we shall assume that, given a sample of observations, a *noisy conformist voter* draws one realization at random from the probability distribution defined by her updated posterior. In this case the probability with which (s)he will choose opinion 1 is given by $\Pr_{f_{Be}}[z > 1/2 \mid r, n]$ to produce[5]:

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

(2)

where we take $n = 2m + 1$ for $m \geq 0$. Again, for any fixed $m = \bar{m} < \infty$ the above probability is a continuous and increasing function of $p \in [0, 1]$, formalizing the fact that the more likely observations 1s are, the higher is the probability that the voter will adopt opinion 1. However, unlike (1), (2) reflects the fact that behaviour is noisy since, for any m, $\Pr^{nc}[1 \mid p = 0, m] > 0$, $\Pr^{nc}[1 \mid p = 0.5, m] = 0.5$ and $\Pr^{nc}[1 \mid p = 1, m] < 1$. This noise tends to disappear as the number of observations increases, since equation (2) behaves exactly in the same fashion as equation (1) did[6]: for each given $p = \bar{p}$, $\lim_{m \to \infty} \Pr^{nc}[1 \mid \bar{p}, m] = s^*(\bar{p})$ and, for m large, it is essentially described by $s^*(\bar{p})$.

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 specifications introduced in the previous Section as primitives. In order to account for different plausible information structures and information transmission among agents, we shall introduce a specific type of heterogeneity among the voters in the population.

The general notation of the model we study has individual $x \in S$ choosing ballot $v(x) \in \{0, 1\}$. We shall assume throughout that S is countable. A configuration of opinions in the population will be denoted by $v \in \{0, 1\}^S$. 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 time runs continuously and each individual may choose a new opinion at a random exponential time, with mean one. The time-dependence of all variables will be denoted by sub-t, but sometime dropped to lighten notation. Whenever individual x is to form a new opinion, (s)he will do so in the way described in the previous Section.

The dynamic aspect of the model is determined by the endogeneization of the parameter p (that, we recall from the previous Section, is a voter's probability of observing an opinion in favour of candidate 1). Specifically, we assume that at each point in time, $p_t \equiv p(x, v_t)$, meaning that such probability depends on the agent's identity, as well as on the current configuration of opinions in the population, but is homogeneous over time. In order to capture the amount of correlation that seems to be particularly pertinent in processes of opinion formation, we proceed as follows. We provide 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 $S \subseteq Z^d$ and define 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 $p(x, v) = (2d)^{-1} \sum_{\{y : \|y-x\|=1\}} v(y)$ and the parameters that determine the probability of observing opinions are (spatially) correlated among voters who are neighbours. Since the set of neighbours is finite, p only takes values in $\{0, (2d)^{-1}, 2(2d)^{-1}, ..., 2d(2d)^{-1}\}$.

As we anticipated, although clearly endogenous, the process that produces the electoral poll π is not modeled explicitly. The idea we have in mind is that at any point in time a voter is to form an opinion, (s)he may take into account the assessment of the percentage votes of candidate 1 produced by a freely and publicly available electoral poll. We assume that each voter treats such public information in exactly the same stochastic way a conformist voter does. Hence, from equation (1) (for $p \equiv \pi$ and $m = 0$) the assessed probability of voting for candidate 1 given the poll is given by:

$$\Pr[1 \mid \pi] \equiv \Pr {}^c[1 \mid \pi, 0] = \pi$$

where $\pi \in [0,1]$.

We also envisage a second way in which a voter may handle the information coming from an electoral poll: this postulates that, given the poll, the voter chooses the ballot that favours the winning candidate and tosses a coin otherwise. As a result, $\Pr[1 \mid \pi] = 1(\{\pi > 0.5\})$ $\Pr[0 \mid \pi] = 1(\{\pi < 0.5\})$ and $\Pr[1 \mid \pi] = \Pr[0 \mid \pi]$ if $\pi = 0.5$. Although these two specifications may be given different behavioural motivation, since we only treat π parametrically (in that we do not model explicitly the way in which the poll is produced), this latter case is *de facto* a particular case of the former, obtained for $\pi \in \{0,1\}$.

In order to make the model tractable, we require the process that produces the poll not to show any time-dependence, and we assume that the last poll is available at some finite time T. Furthermore, although subject to potential sample bias, we require the electoral poll not to mis-represent the current state of public opinion, in the sense that, for all T, if $v_T(x) = 1$ (vs. $v_T(x) = 0$) for all $x \in S$, then $\pi_T = 1$ (vs. $\pi_T = 0$).

The following Definition summarizes the details of the processes of public opinion formation that we study:

Definition 1 (Public Opinion) *Consider a population of voters denoted by $S \subseteq Z^d$. For any $t \geq 0$, let $\alpha \in [0,1]$, $\pi_T \in [0,1]$, $m \geq 0$ and $p_t = (2d)^{-1} \sum_{\{y : \|y-x\|=1\}} v_t(y) \in \{0, (2d)^{-1}, 2(2d)^{-1}, ..., 2d(2d)^{-1}\}$. At each random exponential time t, with mean one, voter x chooses ballot 1 at rate[7]:*

$$\Pr[1 \mid \alpha, p_t, m, \pi_T] \equiv \alpha \Pr[1 \mid p_t, m] + (1-\alpha) \Pr[1 \mid \pi_T]$$

where:

either

$\Pr[1 \mid p, m] \equiv \Pr^c[1 \mid p, m]$ *if all voters are* conformist *and behave according to equation (1). In this case, $v^c(m)$ defines the process of public opinion formation for a population of conformist voters.*

or

$\Pr[1 \mid p, m] \equiv \Pr^{nc}[1 \mid p, m]$ *if all voters are* noisy conformist *and behave according to equation (2). In this case, $v^{nc}(m)$ defines the process of public opinion formation for a population of noisy conformist voters.*

In general, we shall denote by v_t the process at time t. In all specifications, v_t are clearly an element of the state space $\{0,1\}^S$. We are interested in characterizing the

evolution of the stochastic process v_t over time and over space and we shall mostly focus on two extreme cases. The first is when a poll is available at time $T = 0$, it uniformly determines the initial opinion of each voter and it is then ignored by the voters at any subsequent time at which the opinion is revised. We motivate this poll in terms of an opening speech that candidates publicly give or some initial electoral poll that is made available only at the beginning of the electoral campaign. In the parametrization we adopt, this corresponds to the case when $\alpha = 0$ at $t = 0$ and $\alpha = 1$ for all $t > 0$. The second case we shall analyze is instead when an electoral poll is available at some finite time $T \geq 0$ and has a long lasting effect on the process of public opinion in that it is taken into account by each voter at any subsequent time $t \geq T$ at which opinions are to be revised. In this case we take $\alpha < 1$ for all $t \geq T$ and we shall study the dynamics of the process of public opinion for $t \geq T$.

3 Main Results

The first question we address relates to the asymptotics of the processes we study, i.e. their limit behaviour for $t \to \infty$.

Before we state the main results, 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 of public opinion clearly satisfies some Markovian properties. Second, as pointed out earlier, the *noisy conformist voter* who samples finitely many observations can choose each of the opinions with strictly positive probability. This suggests that this process may be ergodic, in that all possible configurations of opinions in the population could be visited infinitely often by the process. This feature is clearly not shared by the *conformist voter* process, as only opinions that are observed can be adopted with positive probability. It should be clear from the specification that, for this latter process, the set of probability measures that have pointmass one on a state where everybody adopts the same opinion are invariant.

Clearly, in the special case of $\alpha = 0$ for any t, full weight is given to the publicly available poll. Since any private information is disregarded by a voter, any potential asymmetries that arise from the local nature of information gathering vanish. As a result, whenever $\pi_T \in (0, 1)$, the behaviour of each single voter is Markovian over $\{0, 1\}$ and the asymptotic behaviour of the process of public opinion formation is described by the product measure of mutually singular countable Markov chains.

Hence the last electoral poll that becomes available fully characterizes the process of public opinion formation.

Interesting questions arise for $\alpha \in (0, 1]$. In this case the stochastic process of public opinion formation is still Markovian over its state-space (since any transition probability depends only on the current configuration of opinions and, possibly, on the currently available electoral poll), although the behaviour of a single voter no longer is (due to the local nature of the interaction).

Some of the results that follow are obtained for a specific 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 specification is chosen mostly for convenience, as it allows for sharper results and for an intuitive characterization of the space-time analysis we address towards the end of this Section. The basic properties of the processes in this case are the following. Since $d = 1$, private opinions are formulated by voters on the basis of the opinions chosen by the two adjacent voters. As a result the parameter that produces observations only takes values $p \in \{0, 1/2, 1\}$. Both specifications of the model involve a process of sampling of m opinions among their nearest neighbours, but while, for any value of m, a conformist voter *de facto* simply imitates the opinion that are observed, a noisy conformist voter who observes all identical opinions can still not imitate what (s)he sees on the ground of the uniform prior.

In the statement of the results we use the following further notation. We shall denote any probability distribution over the state space by μ_t, and the initial distribution by μ_0. Degenerate probability distributions that have pointmass on the configurations where all individuals adopt exactly the same ballot (that is configurations $v^{\mathbf{0}}$ where $v(x) = 0$ for all x in S and configuration $v^{\mathbf{1}}$ where $v(x) = 1$ for all x in S) are denoted by $\mu^{\mathbf{0}}$ and $\mu^{\mathbf{1}}$ respectively. Given μ_0, we let $\mu_t^{\mu_0}$ be the law of $v_t^{\mu_0}$, and we write $\lim_{t\to\infty} \mu_t^{\mu_0} = \mu_\infty^{\mu_0}$ to mean that $\mu_t^{\mu_0}$ is weakly convergent. We also denote by $\mathfrak{I}$ the set of invariant measures (i.e. a measure that is stationary over time) for v_t and $\mathfrak{I}_e \subset \mathfrak{I}$ the set its extreme points. We shall define the process v_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 $\lim_{t\to\infty} \mu_t^{\mu_0} = \mu_\infty$ for any μ_0.

A first understanding of the properties of the process is obtained by studying the process $v^{nc}(m)$ for any value of $m < \infty$ and is summarized in the following Theorem. In words, the Theorem asserts that in the absence of any electoral poll, when public opinion is only driven by private information on the part of each voter, the process

admits an invariant measure that can be fully characterized. Specifically, this invariant measure posits higher limit probability to configurations of public opinion where opinions among voters tend to be spatially homogenous. If and whenever the process is ergodic, the above invariant measure is unique, meaning that independently of any initial condition (that may or may not involve an initial electoral poll), the process will converge to it over time.

Theorem 1 *Consider $v_t^{nc}(m)$ as in Definition 1 for $m < \infty$.*

1. *If $\alpha = 1$ for all $t \geq 0$ and for all $0 < d < \infty$, the following measure is invariant for the process:*

$$\mu_\infty^{(m)}(v) = K \exp[\sum_x \sum_{\{y:\|y-x\|=1\}} \sigma^{(m)}(p(x,v))(2v(x)-1)(2v(y)-1)] \quad (3)$$

where K is such that $\sum_v \mu_\infty^{(m)}(v) = 1$ and $\sigma^{(m)}(p(x,v))$ is detailed in the proof. If $d = 1$, i.e. for $S = Z^1$, then $\sigma^{(m)}(p(x,v)) \equiv \sigma(m) = \frac{1}{4}\log(2^{2(m+1)}-1)$.

2. *If $\mid S \mid < \infty$ and/or $d = 1$, for all $\alpha \in (0,1]$ and for all $\pi_T \in [0,1]$, the process is ergodic.*

   Hence if $\alpha = 1$ for all $t > 0$, the measure in equation (3) is the unique invariant measure for the process and

$$\lim_{t \to \infty} \mu_t^{(m),\mu_0} = \mu_\infty^{(m)} \text{ for any initial distribution } \mu_0.$$

Proof. See Appendix. ∎

The above Theorem provides a characterization of the limit behaviour of the dynamics $v_t^{nc}(m)$. Part 2. states that one of the following two conditions guarantees that the process does not show any path dependence. The first requires the population of voters to be finite (i.e. for $\mid S \mid < \infty$). In this case the process can be regarded as a finite state Markov chain over the space of all possible configurations of opinions. Given that each voter only samples a finite number of observations (i.e. $m < \infty$), as previously noticed, any opinion can be adopted by each voter with strictly positive probability. This guarantees that all possible transitions among different configurations can occur with strictly positive probability and as a result, initial conditions become less and less important along the dynamics. The second instead requires voters to be located on a one-dimensional lattice (i.e. for $d = 1$). This allows to

12

formalize some monotonicity properties of the process, stated in terms of attractiveness, that rely on the fact that coordinates, i.e. voters, tend to agree in opinion with their nearest neighbours. It is known that attractive processes with strictly positive transition rates are ergodic.

Whenever ergodicity is guaranteed, then no matter where the process starts, the probability with which each configuration could be observed asymptotically is given by the above limiting distribution. Notice that, since this is true for any initial condition, this is also true for a specific initial condition where, for example, each voter initially chooses opinion 1 with probability $\pi \in [0,1]$, where π is an initial electoral poll. This formally corresponds to the case where $\mu_0 \equiv \mu_\pi$, that is a product measure with density π (such that $\mu_\pi\{v(x) = 1\} = \pi$ for all $x \in S$). Hence, if voters' behaviour is noisy, an electoral poll available at time $T = 0$ does not have any effect on the asymptotic properties of the process. In an analogous manner, a poll available at a finite point in time $0 < T < \infty$ that is taken into account (with weight $0 < \alpha \le 1$) by all voters at any time $t > T$ does not affect the ergodicity properties of the process. However, as detailed in the proof, if some strictly positive weight $(0 < \alpha < 1)$ is given to an informative poll in the voters' assessments, the reversibility properties upon which the full characterization relies fail to hold.

As the limit distribution (3) has full support, each of the possible configurations of opinions in the population can be observed in the limit. However, it is clear from the above formulation that some configurations are more likely to be observed than others. To see this, consider the one-dimensional case, where the function σ is a mere re-parametrization of the model. As the sum of which in the square brackets of (3) is taken over all *couples* of nearest neighbours, and as the addendum is equal to one if and only if $v(x) = v(y)$, the two configurations which are more likely to be observed are those where every voter chooses exactly the same opinion, i.e. $v^{\mathbf{0}}$ and $v^{\mathbf{1}}$. Since, in the one dimensional case, (3) is continuous in the parameter[8] m, this proves the next Corollary.

Corollary 2 *Consider $v_t^{nc}(m)$ as in Definition 1. Assume $d = 1$ and $\alpha = 1$ for all $t > 0$. Then*

$$for\ all\ m, \qquad \frac{\mu_\infty^{(m)}(v^{\mathbf{0}})}{\mu_\infty^{(m)}(v^{\mathbf{1}})} = 1 \qquad and \qquad \lim_{m \to \infty} \frac{\mu_\infty^{(m)}(v)}{\mu_\infty^{(m)}(v^i)} = 0$$

where $v^i = \{v^{\mathbf{0}}, v^{\mathbf{1}}\}$ and $v \ne v^i$.

The interpretation of the above Corollary in our model is the following. Taking

a limit for $m \to \infty$ means studying what happens when a voter samples an infinite number of observations. Observations are opinions randomly gathered in the neighbourhood and the interpretation we have in mind is that each of these provides the voter with some further information about the state of the system at the time at which (s)he forms an opinion. Ergodicity breaks down only in the limit, as the transition probabilities of which in Definition 2 become discontinuous in the parameter p, as m grows to infinity. The message that the above result conveys is that, asymptotically, we are more likely to observe a configuration of homogeneous opinions in the population, due to the underlying symmetries that the process satisfies[9].

Given these asymptotics (and given that elections typically take place at a finite, rather than infinite, time), what we would like to know in more detail is what happens along the dynamics of the process. In particular we would like to 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 configuration of opinions in the neighbourhood. As we shall see, it is exactly in this setting that the importance of an electoral poll in shaping the dynamics becomes relevant.

The Theorem that follows establishes that this line can be pursued and that the resulting process is exactly the process $v^c(m)$, for which the asymptotic behaviour is fully characterized.

Theorem 3 *Consider $v^{nc}(m)$ and $v^c(m)$ as in Definition (1).*

1. *For all d, for all α and for all π_T, $v^c(\infty) = v^{nc}(\infty)$ as in Definition (1).*

   Suppose $d = 1$ and, for $\pi_0 \in [0,1]$, let $\mu_0 \equiv \mu_{\pi_0}$ be the product measure with density π_0, i.e. $\mu_{\pi_0}\{v(x) = 1\} = \pi_0$ for all $x \in S$. Suppose further that, at $t = 0$, $\alpha = 0$. Then, for all m, the process $v^c(m)$ is such that:

2. *If $\alpha = 1$ for all $t > 0$, then:*

$$\Im_e = \{\mu^{\mathbf{0}}, \mu^{\mathbf{1}}\} \text{ and } \lim_{t \to \infty} \mu_t^{\mu_{\pi_0}} = (1 - \pi_0)\mu^{\mathbf{0}} + \pi_0\mu^{\mathbf{1}}$$

3. *If $\alpha \in (0,1)$ for all $t > 0$, then the process is ergodic.*

   If $\pi_T = 1$ (vs. $\pi_T = 0$), then for $t > T$:

$$\Im = \{\mu^{\mathbf{1}}\} \text{ and } \lim_{t \to \infty} \mu_t^{\mu_T} = \mu^{\mathbf{1}}$$
$$(vs. \ \Im = \{\mu^{\mathbf{0}}\} \text{ and } \lim_{t \to \infty} \mu_t^{\mu_T} = \mu^{\mathbf{0}}).$$

14

Proof. See Appendix. ∎

Part 1. shows that the process v_t^c can be interpreted as a limiting case of the process v_t^{nc}. This is simply due to the fact that, as we already mentioned in Section 2, noisy conformist voters approximate conformist voters for the number of observations growing large.

Part 2. states some relevant properties of the public opinion process when voters are conformist and only an initial electoral poll is made available. In this case the process is clearly path-dependent, since both configurations where all voters adopt the same opinion are invariant for the process. Moreover, these are the *only* two invariant measures for the process. This is relevant for our purposes because it shows that the extreme invariant measures of v^c are exactly those to which v^{nc} collapses for $m \to \infty$.

The second reason why Part 2. is relevant is that it fully identifies the basins of attraction of each of the two invariant measures in terms of the initial electoral poll. The initial public assessment of the probability with which a candidate will win the election fully identifies the set of initial conditions that lead the process of public opinion to asymptotically show uniform agreement on that candidate.

The third reason why Part 2. is interesting 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:

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

Clearly, for any $\pi_0 \in (0,1)$, each single voter may change her or his opinion infinitely many times (as $\lim_{t \to \infty} v_t(x)$ does not necessarily exist). However, as a result of the above considerations, the observed frequencies of individuals choosing the same opinion grows, in probability, over time.

Part 3. describes the effects of an electoral poll available at some finite point in time T, and accounted for by voters from then on (since $\alpha \in (0,1)$), on the dynamics of the process of public opinion. If the poll is treated stochastically by each single voter, this introduces a further probabilistic component in voters' behaviour, since, for example, a conformist voter who is surrounded by opinion 1s, could still choose opinion 0 on the grounds that candidate 0 is favoured by the poll. However, as shown in the proof, this aggregate process is still attractive, since the probability of choosing

an opinion is increasing in the number of neighbours doing so. Hence ergodicity is guaranteed by the same argument used in the proof of Theorem 1.

Interesting cases arise when $\pi_T = \{0, 1\}$. In this case, as asserted by Part 3., the behaviour of the process of public opinion formation is substantially altered, in that the state that shows consensus and agrees with the last available electoral poll becomes a trap for the process. This amounts to saying that starting from any time-T distribution, μ_T, the $\lim_{t \to \infty} v_t^{\mu_T}(x)$ exists, meaning that each voter may change opinion only finitely many times, and will eventually choose the ballot that favours the winning candidate, assessed by the electoral poll. This leads to a form of *contagion* in the aggregate, as:

$$\Pr[\lim_{t \to \infty} v_t(x) = \lim_{t \to \infty} v_t(y) = \pi_T \text{ for all } x \text{ and } y \text{ in } S \mid \pi_T \in \{0, 1\}, t \geq T] = 1$$

meaning that if the last poll available identifies candidate 1 as the winning candidate, $\Pr[\lim_{t \to \infty} v_t(x) = 1 \text{ for all } x \text{ in } S] = 1$.

Our aim is now to characterize more in detail how this consensus (in the case of a poll only available as initial condition) or contagion (when a poll is available and accounted for by voters at any subsequent time) occurs along the dynamics.

As our processes are defined in the two dimensions of time and space, we shall find 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 specification 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 configuration, v, we define a *cluster* as the connected components of $\{x : v(x) = 0\}$ or $\{x : v(x) = 1\}$; the size of a cluster of ones in a segment of side l around the origin as:

$$\mid v_l \mid = \mid \{x : v(x) = 1; x \in [-l, l]\} \mid$$

and the mean cluster size of v around the origin as:

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

whenever this limit exists.

Given the asymptotics described in Theorem 3, we already know that the mean cluster size tends to grow indefinitely. The result that follows provides some information as to the rate of growth of the mean cluster size in the two cases studied in Part

2. (*consensus*) and Part 3. (*contagion*) of Theorem 3. In particular it shows that the rate of growth of the mean cluster size of the process of public opinion when a poll is available at some finite time T and accounted for by the voters at any subsequent time is higher than the rate of growth of the mean cluster size in the case where a poll is only available as initial condition.

Remark 1

a. *Under the same assumptions as in Part 2. of Theorem 3, and for $\pi_0 \in (0,1)$, the mean cluster size, $C^{\mu_{\pi_0}}(v_t)$, grows in probability at rate $\sqrt{t}$, in the sense that:*

$$\frac{C^{\mu_{\pi_0}}(v_t)}{t^{1/2}} \to_p K$$

where K is a positive constant depending on π_0.

b. *Under the same assumptions as in Part 3. of Theorem 3, for $\pi_0 \in (0,1)$ and $\pi_T = 1$, the size of a cluster of ones around the origin, grows in probability at least at rate t, in the sense that, for $t > T > 0$:*

$$\frac{\mid v_t \mid}{t^\gamma} \to_p \infty$$

for $\gamma < 1$.

Proof.

a. Under the assumptions of Part 2. of Theorem 3, our model reproduces the dynamics of a model known in the probabilistic literature as the Voter's model. On these grounds, statement *a.* is proved in Bramson and Griffeath (1980). In fact, Theorem 7, p. 211 of that paper also provides the following estimate for the lower and upper bound of the limit expected value of the above quantity:

$$\sqrt{\pi}\left(\frac{1}{2\pi_0(1-\pi_0)}\right) \leq \lim_{t\to\infty} E[\frac{C^{\mu_{\pi_0}}(v_t)}{t^{1/2}}] \leq 2\left(\frac{\pi_0{}^2 + (1-\pi_0)^2}{\pi_0(1-\pi_0)}\right)\sqrt{\pi}$$

where π_0 is the initial poll and (unfortunate notation) $\pi = 3.1416$.

b. The second statement is proved in the Appendix. ∎

The second statement in the above Remark asserts that whenever the available poll is taken into account by voters at any time an opinion is to be formulated (and hence the opinion that favours the winning candidate as assessed by the poll propagates in a form of *contagion*) the mean cluster size grows at least at a linear rate. This is in essence due to the fact that, beyond having an initial effect, the

17

poll also directly affects the rate at which each voter adopts an opinion in support of the candidate favoured by the poll. This determines a fundamental asymmetry in the rate of growth of the cluster size. The intuition for why this is true is the following. Suppose the poll favours candidate 1 and consider a cluster of opinion 1s. What determines the rate of growth of such cluster is what happens at the left and right border of this cluster. The right bordering 1 and her right bordering 0 face an identical distribution of opinion within their respective neighbourhood and, in the absence of a poll, would be equally likely to change their opinion. However, since the poll favours candidate 1 and since this is given a positive weight $(1 - \alpha > 0)$ in the opinion formation process, the rate at which the right neighbouring 0 flips to 1 is higher than the rate at which the right bordering 1 flips to 0. Hence, in expected terms, the cluster is going to increase.

The above reasoning fails to hold in the case where the poll has only an initial effect on the process of public opinion, since the borders of a cluster behave in a perfectly symmetric way. In this case, whenever the model shows *consensus*, the first statement of the above Remark asserts that, as t gets large, the largest segment containing all individuals choosing the same opinion, has side of probability order $\sqrt{t}$. As a result, 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. This amounts to saying that, although the process we analyze does not admit any stationary distribution where both opinions *co-exist* indefinitely (this is ruled out by Part 2. of Theorem 3), any such configuration can indeed be observed along the dynamics, and when viewed locally, remains almost stationary.

4 Some remarks on the allocation of funding in an electoral campaign

As the dynamics we studied are specified 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 dis-

tribution 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 *noisy conformist 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, in a one-dimensional setting, consider two configurations, v_A^{nc} and v_B^{nc}, identical at all sites apart from the sites $\{x-2, x-1, x, x+1\}$ which are as follows:

$$v_A: \quad .... \quad v(x-2) = 1 \quad v(x-1) = 0 \quad v(x) = 1 \quad v(x+1) = 0 \quad ....$$
$$v_B: \quad .... \quad v(x-2) = 1 \quad v(x-1) = 1 \quad v(x) = 0 \quad v(x+1) = 0 \quad ....$$

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

$$\mu_\infty^\sigma(v_A) \propto \exp[-6\sigma]$$
$$\mu_\infty^\sigma(v_B) \propto \exp[2\sigma]$$

Configuration v_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 its dynamics as well, as can be seen by looking at the dynamics of the specification of the model in terms of *conformist 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, even in the absence of a poll, our model suggests that the best alternative is this latter possibility. To see this, consider the following configuration, v, that has a

border at x, in that $v(x-1) \neq v(x)$:

$$\ldots \quad v(x-2) = 1 \quad v(x-1) = 1 \quad v(x) = 0 \quad v(x+1) = 0 \quad v(x+2) = 0 \quad \ldots$$

Suppose, for simplicity, that the process is started deterministically at configuration v. In this case the duality equation (9) (see the proof of Theorem 2 in the Appendix) states that the probability that starting from configuration v, the voter at site x supports candidate 1 is: $E^v v_t(x) = \sum_y p_t(x,y) v(y)$, which, applied to the subset $\{x-1, x, x+1\}$, becomes:

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

The above probabilities are given explicitly in equation (8), and it is not difficult to see that, for any finite 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:

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

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.

The importance of electoral poll, or analog quantifiable messages that candidates may send to the electorate, are made quite clear in the main results of this paper. If voters behaviour is affected by some noise (in the specification of the model in terms of *noisy conformist voters* and in the results of Theorem 1), the effect of an electoral poll is somewhat limited, since the noisy component of private information gathering de facto determines the asymptotic properties of the process, and these are only partially

affected by a poll. However, if and when voters' behaviour is not noisy, or in any case when such noise disappears in the limit (as in the specification of the model in terms of *conformist voters*), the importance of a poll becomes paramount. Even if such message is only taken into account at an early stage of the electoral campaign and disregarded by voters forever after (as in Theorem 3, Part 2.), due to the underlying monotonicity properties of the process of public opinion, the poll 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 (Remark 1, statement a.). 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. Loosely speaking, a very good opening speech in an electoral campaign, or some primary results, have a long lasting effect on the process of public opinion: although they do not determine the final outcome (since the process is path-dependent), they directly affect the probability with which a candidate achieves uniform support in the electorate.

If the poll is repeatedly taken into account by voters in their opinion formation process, then it not only singles out one configuration as the only asymptotic outcome (Theorem 3, Part 3.), but it also increases the rate at which support grows in the population (Remark 1) from $\sqrt{t}$ to at least t. This clearly emphasizes the importance of the last electoral poll in an electoral campaign and formalizes an incentive, on the part of candidates, to invest resources in producing a last electoral poll, as close as possible to the date of the elections.

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 *conformist voter* model when only an initial poll is made available, the process is path-dependent, as its long run behaviour depends crucially on the initial distribution. Along the dynamics, in the absence of any further poll, clusters emerge and are almost stationary when viewed locally, since their rate of growth is of probability order $\sqrt{t}$. It is instructive to interpret the numerical lower and an upper bound available for the expected mean cluster size. To this aim, consider a process that starts with an initial distribution where each voter chooses opinion 1 with probability, say, $\pi = 0.5$. 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 $t = 198.94$, and is on average not going to vary until $t = 795.78$. In other words, in order to observe the cluster size to double (say from $k = 100$ to $k = 200$), 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 π and symmetric around $\pi = 0.5$. Hence for $\pi \neq 0.5$ a cluster of a given mean size is likely to be observed earlier than if π was 0.5 and is likely to 'persist' for a relatively longer spell of time. Hence, conditional on a candidate winning the elections, the higher is π, 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. Hence, if a candidate could gather some information about the current distribution of potential votes and if this was favourable to her or him, then delaying the date of the elections could have a detrimental effect on the outcome.

These last considerations seem to suggest that a linear allocation of funding over time during an electoral campaign might not be fully and always optimal, since the returns in terms of growing support in the electorate are determined by the properties of the dynamics of the process of public opinion and these may be endogenously affected by candidates. It is however clear that considerations of this sort require an explicit account of the strategic interaction between the two candidates, which at present is not part of the model.

An important assumption that we have maintained throughout all of this paper is that voters are homogenous in their behaviour and the only form of heterogeneity in the opinions that are chosen stems endogenously from the configuration of other agents' opinions at the time choices are to be made. One important extension one may consider is to allow for modelled heterogeneities among voters, other than those stemming from the local nature of information. This is particularly interesting in the light of a recent line of research in the field of Political Economy that focuses on the social effects of preference falsification (T. Kuran (1997) provides a very insightful study[10]), where the opinion reported in public may not reflect true preferences due to social pressures or peer considerations. Our model may capture some aspects pertinent to this approach, once we allow for heterogeneous preferences among voters. We outline below two ways in which this may occur.

In the first case we assume conformist voters, located on a one-dimensional lattice sampling opinions among their nearest neighbours and no electoral poll. The form of heterogeneity we consider relates to voters' behaviour when facing an equal distribution of opinions within their neighbours: while a conformist voter as in our Definition ?? would toss a fair coin, now a type-1 voter chooses opinion 1, while a type-0 voter chooses opinion 0. This formalizes the idea that peer pressure are strong enough to fully determine opinions for a voter who is surrounded by all neighbours choosing the same opinion, but that in the absence of a strict majority within the neighbourhood, a voter's type determines one's choice. This seemingly innocent asymmetry in behaviour alters substantially the asymptotics of the process of public opinion, in that although consensus may still obtain, whenever both types exist in the population of voters, infinitely many configurations where both opinions co-exist may also be absorbing for the dynamics. To see this notice that a border between a cluster of at least two ones and a cluster of at least two zeros, where bordering voters are a type-0 and a type-1 respectively, is stable (in that no voter would flip). Hence the process admits *infinitely* many possible absorbing states where both opinions *co-exist*.

In the second case conformist voters are heterogenous in terms of the poll they account for in determining their opinion: suppose $0 < \alpha < 1$ and that type-1 voter receive $\pi = 1$, type-0 voters receive $\pi = 0$. In fact, we may take $\pi = \{1, 0\}$ to represent a voter true preferences and α to measure the weight given to social pressures that may lead to preference falsification. Looking at the implied flip rates, it is not difficult to see that a type-1 (vs. type-0) voter would choose opinion 1 (vs. opinion 0) with probability one if and only if (s)he is surrounded by all neighbours choosing opinion 1 (vs. opinion 0). In all other cases the probability of choosing an opinion is strictly between zero and one and is increasing in the number of neighbours choosing the same opinion. This means that, for example, a type-1 voter surrounded by all zeros flips to opinion zero at rate α. As a result, whenever both types exist in the population, the system admits *no* absorbing states. We conjecture that since this process is attractive, it may still display *clustering*. It is however not clear how the invariant measures could be characterized since reversibility fails to hold in this case.

5   Conclusions and some related issues

This paper analyses a simple model of public opinion formation that posits that interaction between neighbouring agents leads to bandwagons in the dynamics of

individual opinions, as well as in that of the aggregate process. Bandwagons emerge due to the local nature of information gathering and the potential heterogeneity in behaviour that this entails. We show however that in different specifications of the model, the process tends asymptotically to show consensus on one of the two competing opinions, meaning that initial correlation of opinions among agents tends to vanish over time. We consider the effects on the process of opinion formation of a publicly available poll and show that this may lead to a form of contagion, by which public opinion tends to agree with the poll. In the absence of a poll, the process displays the feature that, after long time spans, a sequence of states occur which remain almost stationary and, when viewed locally, are characterized by large clusters of individuals who hold the same opinion.

Throughout the paper we used the metaphor of a model of pre-electoral public opinion formation, with two candidates running. This allowed us to provide some heuristic considerations related to the space-time allocation of funding in an electoral campaign. As previously pointed out the model we study shares some analogies with the recent literature on learning and evolution in games. As a concluding remark we briefly elaborate on these latter points.

Consider the following 2-by-2 symmetric coordination game, where actions are labelled as 0 and 1 and payoffs are such that $a > c$ and $d > b$:

	0	1
0	a, a	b, c
1	c, b	d, d

By standard arguments, the game admits two strict Nash equilibria in pure strategies $\{1, 1\}$ and $\{0, 0\}$ and a Nash equilibrium in mixed strategies where each player plays action 1 with probability $\delta \equiv (a - c + d - b)^{-1}(d - b)$.

The behavioural specifications we introduced in Section 2.1, motivated in terms of binomial sampling with uniform priors, provide smoothed approximation for the best-reply correspondence induced by the above game[11]. To see this, let p be the probability with which action 1 is chosen, $BR(p)$ be the best reply correspondence and $\overline{BR}(p, m)$ and $\overline{\overline{BR}}(p, m)$, where m odd is the sample size, be obtained in analogy

to equations (1) and (2) respectively:

$$BR(p) = \begin{cases} 0 & p < \delta \\ [0,1] & p = \delta \\ 1 & p > \delta \end{cases}$$

$$\overline{BR}(p,m) = \sum_{r=0}^{m} \binom{m}{r} p^r (1-p)^{m-r} \, 1(\{\frac{(m+1)!}{r!(m-r)!} \int_{\delta}^{1} z^r (1-z)^{m-r} dz > \delta\})$$

$$\overline{\overline{BR}}(p,m) = \sum_{r=0}^{m} \binom{m}{r} p^r (1-p)^{m-r} \frac{(m+1)!}{r!(m-r)!} \int_{\delta}^{1} z^r (1-z)^{m-r} dz$$

These constitute approximations of $BR(p)$ since straight-forward adaptations of the results contained in the proof of Part 1 of Theorem 3, would show that $\lim_{m\to\infty} | \overline{\overline{BR}}(p,m) - \overline{BR}(p,m) | = 0$ and $\lim_{m\to\infty} \overline{BR}(p,m) = BR(p)$.

A locally interactive specification of the model is obtained by postulating that a player can only observe a sample of actions chosen within her neighbourhood. With the notation used throughout this paper, this amounts to saying that the values of p are defined as $p(x,v) = (2d)^{-1} \sum_{\{y:\|y-x\|=1\}} v(y)$ (where $v(y)$ now denotes the action chosen by player y, who is a nearest neighbour of player x) and are restricted to $\{0, (2d)^{-1}, 2(2d)^{-1}, ..., 1\}$.

The results we obtain in the paper refer to the case where $\delta = 0.5$. In the above parametrization of the game, this is the case when $a - c = d - b$, meaning that, since the two pure strategy Nash equilibria are risk-equivalent, the game does not admit any risk-dominant equilibrium as defined in Harsanyi and Selten (1988). The interpretation of the results of our paper are as follows.

Theorem 1 in the case of $\alpha = 1$ and $d = 1$ provides a result analog to Blume (1993). In fact an alternative route to prove this result is by establishing a relation between the process $v^{nc}(m)$ for m finite, and a class of stochastic processes known as Ising models. This could be based on the fact that the transition probabilities of which in equation (2) correspond exactly to the flip rates of a stochastic Ising model, with nearest neighbour interactions, relative to the following potential:

$$J_R = \begin{cases} \sigma(m) = \frac{1}{4} \log(2^{2(m+1)} - 1) \text{ if } R = \{x,y\} \text{ and } y : | x - y |= 1 \\ 0 \text{ otherwise} \end{cases} \tag{4}$$

For $d > 1$, it is worth noticing that, for any finite m and over values of $p \notin \{0, 1/2, 1\}$, $\overline{\overline{BR}}(p,m)$ does not coincide with the specification used in Blume (1993), given by $\widehat{BR}(p,m) = [1 + \exp[m\{(c - a)p + (d - b)(1 - p)\}]]^{-1}$ applied to a 2-dimensional

lattice. The logic followed in Corollary 2 is entirely analog to that used in equilibrium selection in games. The result says that, whenever a risk-dominant equilibrium does not exist, as the noise becomes infinitesimally small (i.e. for $m \to \infty$) the unique limit distribution of the process posits positive probability only on the two configurations that reproduce, in terms of observed frequencies, the two pure strategy Nash equilibria of the underlying game.

Theorem 3 for $\alpha = 1$ and $d = 1$ may be interpreted to describe a coordination game played in a one-dimensional locally interactive system, by countably many myopically best-responding players who, in case of indifference, choose each action with probability $1/2$ and where initial conditions are random (as in Lee and Valentinyi (2000)), i.e. where initially every player chooses action 1 with probability π_0. This result says that, whenever a risk-dominant equilibrium does not exist, a) the only absorbing states of the process are the two configurations that reproduce, the two pure strategy Nash equilibria of the underlying game and b) the initial condition fully identifies the basins of attraction of the two steady states. As a result, the process shows *consensus*, meaning that the probability with which all players adopt the same action tends to one over time.

Remark 1, Part a. characterizes the speed with which consensus is achieved as $\sqrt{t}$. This means that, although both actions cannot *co-exist* indefinitely in our model (in the sense used for example in Morris (2000)), the process is characterized by large clusters of adjacent agents who adopt the same action that persist over time, and when viewed locally look as if they were stationary. The analysis pursued here has some analogies, at least in its motivation, with Ellison (1993) and Ellison (2000), where 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 finite population. Our model differs from the cited paper in a number of respects. First, the specification of the dynamics that Ellison studies is perturbed by mistakes (that take the form of a multinomial distribution that assigns small, though strictly positive, probability, uncorrelated across players and over time, to actions that are not best-replies to the current configuration of play). This is substantially different from the way we model individual behaviour (that in the specification given by $\overline{\overline{BR}}(p, m)$ could be motivated in terms of mistakes that *do depend* on expected payoffs as in $\widehat{BR}(p, m)$). Second, Ellison's dynamics are defined over a finite state-space and modeled as finite, discrete time, regular Markov chains, whereas our dynamics define a Markovian process over a countable state-space, that is ergodic if players adopt

26

$\overline{\overline{BR}}(p,m)$, but is path-dependent if players adopt $\overline{BR}(p,m)$. 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 and shows that , whenever a risk-dominant equilibrium exists, the speed of convergence is higher in a locally interactive setting. All specifications of our model rely on a local characterization of the way in which interaction takes place and the results we obtain here rely on an explicit relation between the two dimensions over which our processes are defined (i.e. time and space) in a process that may show path-dependency. Since in a non-local specification of our model (where players are equally likely to interact with any other player in the population) consensus would grow at least at a linear rate, our results imply that, whenever a risk-dominant equilibrium does not exist, the speed of convergence is actually lower in a locally interactive system.

Appendix

Proof of Theorem 1

1. (characterization of the invariant measure) We start by looking at the one-dimensional case, i.e. for $d = 1$ and we claim that the following measure is invariant for the process $v^{nc}(m)$ for any $m < \infty$:

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

where $K = \{\sum_v \exp[\sum_x \sum_{\{y:\|y-x\|=1\}} \sigma(2v(x)-1)(2v(y)-1)]\}^{-1}$ and $\sigma = \frac{1}{4}\log(2^{2(m+1)} - 1)$.

We re-write the transition probabilities of which in equation (2) by substituting $\sigma = \frac{1}{4}\log(2^{2(m+1)} - 1)$:

$$\Pr{}^{nc}[1 \mid m, p(x, v)] \equiv \Pr{}^{nc}[1 \mid \sigma, p(x, v)] = \tag{5}$$

$$= \frac{1}{1 + \exp[-4\sigma(2p(x, v) - 1)]} \tag{6}$$

where we recall $p(x, v) = \frac{1}{2}\sum_{y:\|y-x\|=1} v(y)$ and, since $S = Z^1$, it takes values in $\{0, 1/2, 1\}$. For example, if $p(x, v) = 0$ the above equation states that the probability that opinion 1 is chosen is given by $[1 + \exp[4\sigma]]^{-1} = [1 + \exp[\log[2^{2(m+1)} - 1]]]^{-1} = (2^{2(m+1)})^{-1}$, which corresponds exactly to (2) for $p = 0$.

To prove the assert, it suffices to notice that the above measure is reversible, in that:

$$\Pr{}^{nc}[1 \mid \sigma, p(x, v), v(x) = 0]\mu^\sigma_\infty(v_{x=0}) = \Pr{}^{nc}[0 \mid \sigma, p(x, v), v(x) = 1]\mu^\sigma_\infty(v_{x=1})$$

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

$$
\begin{aligned}
\frac{\mu^\sigma_\infty(v_{x=1})}{\mu^\sigma_\infty(v_{x=0})} &= \exp[2\sigma \sum_{\{y:\|y-x\|=1\}} (2v(y) - 1)] \\
&= \frac{1}{1 + \exp[-2\sigma \sum_{\{y:\|y-x\|=1\}}(2v(y) - 1)]} \cdot \\
&\quad \cdot [\frac{1}{1 + \exp[2\sigma \sum_{\{y:\|y-x\|=1\}}(2v(y) - 1)]}]^{-1} \\
&= \frac{\Pr{}^{nc}[1 \mid \sigma, p(x, v), v(x) = 0]}{\Pr{}^{nc}[0 \mid \sigma, p(x, v), v(x) = 1]}
\end{aligned}
$$

The same logic applies to the case where $1 < d < \infty$. In this case, for any given $m < \infty$ and for $p \equiv p(x, v) = (2d)^{-1}\sum_{\{y:\|y-x\|=1\}} v(y)$ which takes values in

28

$\{0, (2d)^{-1}, 2(2d)^{-1}, ..., 1\}$, reversibility requires:

$$\frac{\mu_\infty^{(m)}(v_{x=1})}{\mu_\infty^{(m)}(v_{x=0})} = \exp[4\sigma^{(m)}(p)d(2p-1)]$$

$$= \frac{1}{1 + \exp[-2\sigma^{(m)}(p)d\sum_{\{y:\|y-x\|=1\}}(2v(y)-1)]} \cdot$$

$$\cdot [\frac{1}{1 + \exp[2\sigma^{(m)}(p)d\sum_{\{y:\|y-x\|=1\}}(2v(y)-1)]}]^{-1}$$

$$= \frac{\Pr^{nc}[1 \mid p, m, v(x)=0]}{\Pr^{nc}[0 \mid p, m, v(x)=1]} \equiv \frac{\Pr^{nc}[1 \mid p, m]}{1 - \Pr^{nc}[1 \mid p, m]}$$

where $\Pr^{nc}[1 \mid p, m]$ is as in equation (2). Hence $\sigma^{(m)}(p)$ solves the following equation:

$$\frac{\Pr^{nc}[1 \mid p, m]}{1 - \Pr^{nc}[1 \mid p, m]} = \exp[4\sigma^{(m)}(p)d(2p-1)]$$

This equation has a unique solution for any $p \neq 0.5$ given by:

$$\sigma^{(m)}(p) = \frac{\log[\Pr^{nc}[1 \mid p, m] - \log[1 - \Pr^{nc}[1 \mid p, m]]}{4d(2p-1)}$$

If $p = 0.5$, since $\Pr^{nc}[1 \mid p = 0.5, m] = 0.5$ for any m, any finite value of σ satisfies the above equation. It can be shown that $\sigma^{(m)}(p)$ is symmetric, in the sense that $\sigma^{(m)}(p) = \sigma^{(m)}(1-p)$. Hence, for any given d, $\sigma^{(m)}(p)$ is fully characterized by d values. In fact, for $d = 1$, its domain is restricted to $\{0, 0.5, 1\}$ and its co-domain is fully characterized by the parameter $\sigma = \frac{1}{4}\log(2^{2(m+1)} - 1)$, as used in the first part of this proof. ∎

2. (ergodicity) We interpret the process v^{nc} as a system of interactive, nearest neighbours, particles on the state space S.

We look first at the case where $\mid S \mid \equiv S < \infty$. For convenience, and in order to assume away bordering conditions (where, since there are only finitely many voters, a voter would be surrounded by only d neighbours, as opposed to $2d$), we think of the lattice Z^d as folded to form the torus $\Lambda(S) = Z^d \cap [-S/2, S/2]^d$ for $S = 2, 4, ...$

The process v^{nc} moves on the finite state space of all configurations $v \in \{0, 1\}^{\Lambda(S)}$. In the model, at any point in time, at most one voter may choose to revise her or his opinion. When (s)he does so, (s)he behaves according to equation (1), which we recall only depends on m, $p(x, v)$, α and π and are homogeneous over time. Hence the dynamics is generated by the following flip rates, $c(x, v)$ that define the probability with which coordinate x flips, from $v(x)$ to $1 - v(x)$, when the process is in state v:

$$c^{nc}(x, v, \alpha, \pi, m) = v(x) + (1 - 2v(x))[\alpha \Pr^{nc}[1 \mid p(x, v), m] + (1 - \alpha)\pi]$$

where $v(x) = \{0, 1\}$ and $\Pr^{nc}[1 \mid p, m]$ is as in equation (2).

It can easily be checked that for any value of $(\alpha, \pi) \in (0, 1] \times [0, 1]$, since $0 < \Pr^{nc}[1 \mid p, m] < 1$ for $m < \infty$, these flip rates are strictly positive. Hence, transition probabilities are strictly positive from each state to all, and only, the states that differ from that state by at most one coordinate. Hence we may regard the process as a finite-state Markov chain, and conclude that, since starting from one state, the process can reach any other state in at most $S < \infty$ steps, the process is ergodic.

Whenever $S = Z^d$ is countable, but possibly infinite, so is the state space of the process v^{nc} and hence the above logic does not hold. We proceed as follows. We first show that the process is *attractive* (or monotonic) in that coordinates tend to agree with neighbouring coordinates. We then use a result stating that, in Z^1, 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 (as we already know they are).

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 defined to be *attractive* if, whenever $\eta \leq \zeta$ flip rates satisfy the following:

$$c(x, \eta, \alpha, \pi, m) \leq c(x, \zeta, \alpha, \pi, m) \quad \text{if } \eta(x) = \zeta(x) = 0$$
$$c(x, \eta, \alpha, \pi, m) \geq c(x, \zeta, \alpha, \pi, m) \quad \text{if } \eta(x) = \zeta(x) = 1$$

Since for any $\eta \leq \zeta$, $p(x, \eta) \leq p(x, \zeta)$, also $\Pr^{nc}[1 \mid p(x, \eta), m] \leq \Pr^{nc}[1 \mid p(x, \zeta), m]$ for any m. Hence, the process is attractive. For example, for $d = \alpha = 1$, flip rates expressed as a function of $\sigma < \infty$ are:

$$c^{nc}(x, v, \sigma) \equiv \begin{cases} \Pr^{nc}[1 \mid \sigma, p(x, v), v(x) = 0] = [1 + \exp[-4\sigma(2p(x, v) - 1)]]^{-1} & \text{if } v(x) = 0 \\ \Pr^{nc}[0 \mid \sigma, p(x, v), v(x) = 1] = [1 + \exp[4\sigma(2p(x, v) - 1)]]^{-1} & \text{if } v(x) = 1 \end{cases}$$

or:

$$c^{nc}(x, v, \sigma) = \frac{1}{1 + \exp[-4\sigma(1 - 2v(x))(2p(x, v) - 1)]}$$

and it can easily be checked that attractivity is guaranteed.

As proved in Gray (1982) (and reported, for example, in Liggett (1985), as Theorem 3.14, p.152), this is a sufficient condition for ergodicity. Hence, the set of invariant measures, $\mathfrak{I}$, for the process $v^{nc}(m)$, with $m < \infty$, is a singleton.

As a result, for $\alpha_t \equiv \alpha = 1$ for all $t > 0$, the only such measure is the one identified in Part 1. of this proof and the last statement of Part 2. of the Theorem follows.

Notice that for any $\alpha_t \equiv \alpha < 1$ for all $t > 0$, the reversibility properties we used in Part 1. of the proof only hold for $\pi = 0.5$ (since in any other case $\alpha \Pr^{nc}[1 \mid p = 0, m] + (1-\alpha)\pi + [\alpha \Pr^{nc}[1 \mid p = 1, m] + (1-\alpha)\pi] \neq 1$, thus formalizing an asymmetry in the flip rates). ■

Proof of Theorem 3

1. Recall that the processes $v^{nc}(m)$ and $v^c(m)$ are ultimately defined by the transition probabilities of which in (2) and (1) respectively. Hence, we only need to show that $\lim_{m\to\infty} \Pr^{nc}[1 \mid m, p] = \lim_{m\to\infty} \Pr^c[1 \mid m, p]$ for any given $p \equiv p(x, v) \equiv i/2d$ for $i \in \{0, 1, ..., 2d\}$.

For $d = 1$, this is trivial, since $\Pr^{nc}[1 \mid m, p = 0] = 2^{-2(m+1)}$, $\Pr^{nc}[1 \mid m, p = 0.5] = 0.5$ and $\Pr^{nc}[1 \mid m, p = 1] = 1 - 2^{-2(m+1)}$ and, over $\{0, \frac{1}{2}, 1\}$ and for all m, $\Pr^c[1 \mid m, p = 0] = 0$, $\Pr^c[1 \mid m, p = 0.5] = 0.5$ and $\Pr^c[1 \mid m, p = 1] = 1$.

For $d > 1$, we show that convergence obtains over all values of p:

$$
\begin{aligned}
&\mid \Pr^{nc}[1 \mid m, p] - \Pr^c[1 \mid m, p] \mid \leq \\
&\leq \; \mid \sum_{r=0}^{2m+1} \binom{2m+1}{r} p^r (1-p)^{2m+1-r} - \sum_{r=m+1}^{2m+1} \binom{2m+1}{r} p^r (1-p)^{2m+1-r} \mid = \\
&= \sum_{r=0}^{m} \binom{2m+1}{r} p^r (1-p)^{2m+1-r} \leq \frac{1}{2m}
\end{aligned}
$$

which goes to 0 for $m \to \infty$.

To characterize this limit, notice that, $\Pr^c[1 \mid m, p]$ is symmetric around $p = 0.5$, in that $\Pr^c[1 \mid m, p] = 1 - \Pr^c[1 \mid m, 1 - p]$. Hence it suffices to show that, for all $0 < p < 0.5$, $\lim_{m\to\infty} \Pr^c[1 \mid m, p] = 0$. To this aim, notice that, for $0 < p < 0.5$, this is a sum of m decreasing terms. Hence:

$$
\begin{aligned}
0 &< \Pr^c[1 \mid m, p] \leq m \binom{2m+1}{m+1} p^{m+1}(1-p)^m \\
0 &< \lim_{m\to\infty} \Pr^c[1 \mid m, p] \leq \lim_{m\to\infty} m \binom{2m+1}{m+1} p^{m+1}(1-p)^m = 0
\end{aligned}
$$

thus concluding the proof. ■

2. Since $\alpha = 0$ at $t = 0$ the initial condition fo the process is given by the product measure μ_{π_0}. Since $\alpha = 1$ for all $t > 0$ flip rates for this process are:

$$
c^c(x, v) = v(x) + \Pr^c[1 \mid m, p(x, v)](1 - 2v(x))
$$

for $v(x) \in \{0, 1\}$ and $\Pr^c[1 \mid m, p(x, v)]$ as in equation (1). Since $d = 1$, $p(x, v) = (2)^{-1} \sum_{\{y: \|y - x\| = 1\}} v(y) \in \{0, 1/2, 1\}$ and over these values $\Pr^c[1 \mid m, p(x, v)] \equiv$

$p(x, v)$ for all $m \geq 0$. As a result:

$$c^c(x, v) = v(x) + p(x, v)(1 - 2v(x)) \tag{7}$$

By simple inspection of the flip rates that define the process it is clear that any state for which $v(x) = v(y)$ for all x, y in S is stationary for the process. Clearly, for this process v^c, $\mathfrak{I} \supseteq \mathfrak{I}_e \supseteq \{\mu^{\mathbf{0}}, \mu^{\mathbf{1}}\}$. Hence, the result relies on the proof that these are the *only* two extreme invariant measures (i.e. $\mathfrak{I}_e \subseteq \{\mu^{\mathbf{0}}, \mu^{\mathbf{1}}\}$), 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_{π_0}, and $\lim_{t \to \infty} \mu_t^{\pi_0} = (1 - \pi_0)\mu^{\mathbf{0}} + \pi_0 \mu^{\mathbf{1}}$.

We make use of results that are well known in the statistical literature on the Voter's model (Liggett (1985), Section 1 and 3, Chapter V or in Bramson and Griffeath (1980)) that our model reproduces for this specification of the parameters. In the Voter's model, a voter at $x \in Z^d$ changes his opinion at an exponential rate (with mean one) proportional to the number of $2d$ nearest neighbours with the opposite opinion. If $2d$ neighbours disagree with the person at x, the flip rate is 1. It can be seen by equation (7) that this is exactly the dynamics of our model.

As the logic of the proofs is interesting in its own right, we sketch the proof in what follows.

The process v^c can be studied in terms of its dual process in terms of coalescing random walks. The duality relation transforms questions about v^c in questions concerning the cardinality of the coalescing random walk system.

We first 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, at any $t > 0$, the flip rates of equation (7) can be written as:

$$c^c(x, v) = (1 - v(x)) + (2v(x) - 1) \sum_{\{y : \|y-x\|=1\}} \frac{1}{2}(1 - v(y))$$

These coincide with equation. (4.3) (p. 158) in Liggett (1985), once we take $c(x) = 1$, $A = \{y\}$ and $p(x, A) = p(x, y) = 1/2$ 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 $p(x, x + 1) = p(x, x - 1) = 1/2$. 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 defines a continuous time Markov chain, $X(t)$, with transition probabilities:

$$p_t(x, y) = e^{-t} \sum_{n=0}^{\infty} \frac{t^n}{n!} p^{(n)}(x, y) \tag{8}$$

where $p^{(n)}(x, y)$ are the n-step transition probabilities associated with $p(x, y)$. Any system of finitely many independent copies of $X(t)$, where any two merge whenever they meet, defines a system of finitely many coalescing Markov chains over the state space of all finite subsets of $S = Z^1$.

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

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

This represents a measure of how far apart the single processes are. Clearly, for any t, $g_t(A) = 0$ when $|A| = 1$, as a single recurrent random walk is never going to die. If $|A| = 2$, $g_t(A) \rightarrow_{t \to \infty} 1$, 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 $|A| = 2$ for some $t^* < \infty$.

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

$$\mu_t(A) = E^A \mu(A_t) \tag{9}$$

where $\mu_t(A)$ is the probability that the process v_t has $v_t(x) = 1$ for all $x \in A$ and $E^A \mu(A_t)$ is the probability that $|A_t|$ 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, $\lim_{t \to \infty} \mu_t^\theta = (1 - \theta)\mu^{\mathbf{0}} + \theta \mu^{\mathbf{1}}$.

To characterize the basins of attraction of $\{v^{\mathbf{0}}, v^{\mathbf{1}}\}$, suppose the process v^c is started (stochastically) with product measure μ_θ. If τ is the first time that $|A_t| = 1$ (which is finite with probability one by our assumption that $g_{t^*}(A) = 1$ when $|A| = 2$) the duality equation (9) 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 $\lim_{t \to \infty} \mu_t^\theta \to (1 - \theta)\mu^0 + \theta\mu^1$ to be true. Hence, for $\theta = \pi_0$ the assert follows. ∎

3. We follow the same logic we used in Part 2. of the proof of Theorem 1 for the case of $S = Z^1$. Flip rates for this model can be written as:

$$c^c(x, v, \alpha, \pi, m) = v(x) + (1 - 2v(x))[\alpha \Pr{}^c[1 \mid p(x, v), m] + (1 - \alpha)\pi] \qquad (10)$$

for $v(x) = \{0, 1\}$ where $\Pr{}^c[1 \mid p(x, v), m]$ is given in equation (1). Attractivity is, again, guaranteed by the fact that, for any m, $\Pr{}^c[1 \mid p, m]$ is increasing in p. Hence the process is ergodic if these flip rates are strictly positive.

Recall that $\alpha \in (0, 1)$ by assumption. Since, for all m, $\Pr{}^c[1 \mid p = 0, m] = 0$ and $\Pr{}^c[1 \mid p = 1, m] = 1$, it is clear that, for any $\pi \in (0, 1)$, $0 < c^c(x, v, \alpha, \pi, m) < 1$, which guarantees ergodicity in this case. For $\pi \in \{0, 1\}$ ergodicity is proven next.

We prove the statement for $\pi = 1$. (The proof for $\pi = 0$ is entirely analog). In this case the configuration $v^1 = \{v \in \{0, 1\}^{Z^1} : v(x) = 1\}$ is absorbing, since, form the flip rates of which in (10), no voter would chance opinion. Moreover, this would be the only absorbing state, since in any other configuration some voters, for whom $p(x, v) < 1$, could flip with positive probability. Since the state-space of the process v^c on Z^1 is only countable, ergodicity might still fail to hold[12].

Let S_N be finite sets that increase to S, such that $\lim_{N \to \infty} S_N = S$. Define the following flip rates:

$$c_i^{c,N}(x, v, \alpha, \pi, m) = \begin{cases} c^c(x, v(x)^i, \alpha, \pi, m) & \text{if } x \in S_N \\ 0 & \text{if } x \notin S_N \quad \text{and } v(x) = i \\ 1 & \text{if } x \notin S_N \quad \text{and } v(x) \neq i \end{cases}$$

with $v(x)^i = v(x)$ for $x \in S_N$, and $v(x) = i$ for $x \notin S_N$, $i \in \{0, 1\}$.

Denote the process with flip rates $c_i^{c,N}(x, v, \alpha, \pi, m)$ by $S_{i,N}(t)$, where $S_{i,N}(t)$ is equal to the original process for $x \in S_N$, and characterized by all coordinates set equal to i for $x \notin S_N$. Let $\mu^0 S_{0,N}(t)$ be the law of the process characterized by flip rates $c_0^{c,N}(x, v, \alpha, \pi, m)$ when the initial distribution is given by all 0 at time 0 and let $\mu^1 S_{1,N}(t)$ be the law of the process characterized by flip rates $c_1^{c,N}(x, v, \alpha, \pi, m)$ when

the initial distribution is given by all 1 at time 0. As $c^c(x, v(x), \alpha, \pi, m)$ is attractive, by Theorem 2.7 in Liggett (1985), also $c_i^{c,N}(x, v, \alpha, \pi, m)$ are attractive and

$$\mu^0 S_{0,N}(t) \leq \mu^\theta S(t) \leq \mu^1 S_{1,N}(t)$$

for $\theta \in (0, 1)$, and

$$\lim_{N \to \infty} \lim_{t \to \infty} \mu^0 S_{0,N}(t) = \lim_{t \to \infty} \mu^0 S(t)$$
$$\lim_{N \to \infty} \lim_{t \to \infty} \mu^1 S_{1,N}(t) = \lim_{t \to \infty} \mu^1 S(t)$$

Now $\lim_{t \to \infty} \mu^0 S_{0,N}(t) = \lim_{t \to \infty} \mu^1 S_{1,N} = \mu^{1,N}$, that is, as $t \to \infty$, independently of the initial distribution, the process restricted on S_N converges to a configuration all ones. In fact $S_{i,N}(t)$ is a finite Markov chain over S_N, and as there is a unique absorbing state $(v_N^1 \equiv \{v(x) = 1 \text{ for all } x \in S_N\})$ we know that the unique ergodic distribution posits pointmass one on this state. As $\lim_{N \to \infty} S_N = S$, il follows that

$$\lim_{N \to \infty} \lim_{t \to \infty} \mu^0 S_{0,N}(t) = \lim_{N \to \infty} \lim_{t \to \infty} \mu^1 S_{1,N}(t) = \lim_{N \to \infty} \mu^{1,N} = \mu^1$$

The desired result then follows. ∎

Proof of Remark 1

a. We have already shown that, under the assumptions of Theorem 3, Part 2. our model reproduces the dynamics of the Voter's model. Theorem 7 (p.211) in Bramson and Griffeath (1980) requires the initial condition to be a product measure (as such translation invariant) and in our model μ_{π_0} is so by definition.

b. We prove the statement for $\pi_T = 1$. In this case we know (Theorem 3, Part 3.) that, starting from any time T distribution, the system converges to v^1. We here characterize the minimum rate at which this occurs. First notice that, since $\alpha \in (0, 1)$, $\pi_T = 1$ and $d = 1$, flip rates are given by:

$$c^c(x, v, \alpha, \pi, m) = v(x) + (1 - 2v(x))[\alpha p(x, v) + (1 - \alpha)]$$

Hence, starting from v^0 (where $p(x, v) = 0$ for all x), ones are produced by the poll at rate $(1 - \alpha) > 0$.

Suppose at some time $t > T = 0$, $v(x) = 1$ and $v(y) = 0$ for all $y \neq x$. The minimum rate at which this one at x grows into a cluster of two adjacent ones is computed as follows. Within a small time interval, since at most one voter can change opinion, three things can happen:

a) $v(x) = 1$ flips to $v(x) = 0$, $v(x - 1) = v(x + 1) = 0$. This occurs at rate α.

b) $v(x) = 1$, $v(x-1) = 0$ flips to $v(x-1) = 1$, $v(x+1) = 0$. This occurs at rate $\alpha/2 + (1 - \alpha)$.

c) $v(x) = 1$, $v(x+1) = 0$ flips to $v(x+1) = 1$, $v(x-1) = 0$. This occurs at rate $\alpha/2 + (1 - \alpha)$.

Under a) the cluster disappears; under b) or c) the cluster grows by one unit. It can easily be checked that these are also the (minimum) rates at which a cluster of at least two adjacent ones grows by one unit. Hence, between $T = 0$ and t, with probability one, the cluster size is such that:

$$| v_t | \geq \int_0^t [2(\frac{\alpha}{2} + 1 - \alpha) - \alpha] dt = 2(1 - \alpha)t$$

As a result, since $\alpha \in (0, 1)$, for $t > T$:

$$\lim_{t \to \infty} \frac{| v_t |}{2t^\gamma} \geq (1 - \alpha) \lim_{t \to \infty} t^{1-\gamma}$$

which is equal to ∞ for $\gamma < 1$. ∎

Notes

[1] Since many specifications of our model allow for a countable population of agents, *simple majority* is to be intended as the limit of its natural restriction to $[-N, N]$ as $N \to \infty$ whenever this limit exists. As it will become clear, this will not play a key role in our results.

[2] The simplest way to motivate this is to think in terms of a side payments, denoted by ε, that a voter gets if (s)he has voted for the candidate who wins the elections and to further postulate that the utility function is quasi-linear in this latter argument. The relevance of this assumption within the framework of a standard herding model has been recently emphasized by Collander (2002).

[3] This is given by $\Pr\left[1 \mid p, n\right] = \sum_{r=0}^{n} \binom{n}{r} p^r (1-p)^{n-r} \mathbf{1}(\{\frac{(n+1)!}{r!(n-r)!} \int_{1/2}^{1} z^r (1-z)^{n-r} dz > 0.5\})$ where $\mathbf{1}(\{\cdot\})$ is an indicator function that takes value of one whenever $\{\cdot\}$ is true.

[4] This statement is proved in the proof of Part 1. of Theorem 3.

[5] See footnote 3.

[6] This statement is proved in the proof of Part 1. of Theorem 3.

[7] Specifically, we mean that, within a small time interval dt:

$$\Pr[v_{t+dt}(x) = 1 \mid \alpha, p_t, m, \pi_T] = \Pr[1 \mid \alpha, p_t, m, \pi_T] dt + o(t)$$

Also, the assumption that opinions are revised at random times ensures that the probability that more than one voter revises opinion at the same time is negligible.

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

[9] It is clear that Corollary 2 only relies on a comparative static exercise over the limit distributions of a sequence of identical processes, that differ only in the parameter m and, as such, it does not provide a full understanding of the dynamics. Loosely speaking, what we do next is to reverse the logic we followed in Corollary 2 (where we first looked at the asymptotics for $t \to \infty$, and then at the limit for $m \to \infty$) by first looking at a process where a voter knows exactly what the opinions in her neighbourhood is (i.e. in the limit for $m \to \infty$) and second at what happens along the dynamics of this process (i.e. asymptotically for $t \to \infty$).

[10] We are grateful to an anonymous referee of a previous version of this paper and to A. Hamlin for drawing this book to our attention.

[11] The term smooth refers to the fact that the best-reply correspondence are approximated by continuous functions. For an overview of analog approximations, see for example Fudenberg and Levine (1998), Chapter 4 and therein references.

[12] The intuition why this could be so, is as follows. Suppose the process starts at $v^{\mathbf{0}} = \{v \in \{0,1\}^{Z^1} : v(x) = 0\}$. An occasional $v(x) = 1$ appears and, when it does so, it may grow into a block of 1s. But if $(1-\alpha)$ is small, the length of the block of 1s surrounded by 0s may grow at a negative

rate. The process is described approximately by a countable positive recurrent Markov chain over the number of 0s, $(0, 1, 2, ...)$, absorbed at 0 after a time with finite expected value. If $(1 - \alpha)$ (the rate of production of 1s) is small relative to this expected time, then one may expect the limit distribution for $t \to \infty$, to be different from $\mu^{\mathbf{1}}$. Hence the process would not be ergodic. Hence we need to prove that ergodicity holds for any value of $\alpha \in (0, 1)$.

REFERENCES

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

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

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.

CHEW, M. S. Y. (2000), "Comunication and Coordination in Social Networks," *The Review of Economic Studies,* **67**, 1-16.

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.

COLLANDER S. (2002), "Bandwagons and Momentum in Sequential Voting," *mimeo,* Kellog Graduate School of Management.

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

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

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

ELLISON, G. (2000), "Basins of Attraction, Long-Run Stochastic Stability, and the Speed of Step-by-Step Evolution," *The Review of Economic Studies,* **67**, 17-45.

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.

FUDENBERG, D. AND D.K. LEVINE (1998), *The Theory of Learning in Games,* MIT Press.

HARSANYI, J. AND R. SELTEN (1988), *A General Theory of Equilibrium Selection in Games,* MIT Press.

HEATH, A. AND J. EVANS (1994), "Tactical Voting: Concepts, Measurements and Findings," *British Journal of Political Science,* **21**, 557-561.

IANNI, A. (2001), "Learning Correlated Equilibria in Population Games," *MAthematical Social Sciences,,* **42**, 271-294.

IANNI, A. AND V. CORRADI (1998), "Ergodicity and Clustering in Opinion Formation," CARESS, 98-10.

KANDORI, M., G. MAILATH AND R. ROB (1993), "Learning, Mutations, and Long Run Equilibria in Games," *Econometrica,* **61**, 29-56.

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.

KURAN, T. (1995), *Private Truths, Public Lies,* Harward University Press.

LEE, I.H. AND A. VALENTINYI (2000), "Noisy Contagion Without Mutations," *The Review of Economic Studies,* **67**, 47-56.

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.

MCKELVEY R.D., AND T.R. PALFREY (1995), "Quantal Response Equilibria for Normal Form Games," *Games and Economic Behavior,* **10**, 6-38.

MEEHL, P.E. (1977), "The Selfish Voter Paradox and the Thrown-Away Vote Argument," *The American Political Science Review,* **71**, 11-30.

MORRIS, S. (2000), "Contagion," *The Review of Economic Studies,* **67**, 57-68.

MORRIS, S. (1997b), "Interaction Games: A Unified Analysis of Incomplete Information, Local Interaction and Random Matching," , mimeo, University of Pennsylvania.

PIKETTY, T. (2000), "Voting as Communicating," *The Review of Economic Studies,* **67**, 169-191.

RIKER, W.H. AND ORDERSHOOK P.C. (1968), "A Theory of the Calculus of Voting," *American Political Science Review,* **62**, 25-42.

SCHUESSLER A. A. (1997), "Expressive Voting," , mimeo, New York University.

SIMON, H.A. (1954), "Bandwagon and Underdog Effects and the Possibility pf Equilibrium Predictions," *Public Opinion Quarterly,,* **18**, 245-253.

YOUNG, H. P. (1993), "The Evolution of Conventions," *Econometrica,* **61**, 57-84.

ZUCKERMAN, A.S., N.A. VALENTINO AND E.W. ZUCKERMAN (1994), "A structural Theory of Vote Choice: Social and Political Networks and Electoral Flows in Britain and in the United States," *The Journal of Politics,* **56**, 1008-1033.