ELECTORAL COMPETITION WITH LOCAL EXTERNALITIES

ANTONELLA IANNI

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 opinion choices, as well as in that of the aggregate process. We then analyze the implication that these findings have in terms of space-time allocation of fundings in an electoral campaign, where two candidates run in a winner-take-all election.

JEL: **C72, C92, D83.**

1. Introduction

This paper analyzes a simple stochastic dynamic process of public opinion formation, by studying the way in which this evolves over space and over time. 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, at the time when the election come. Voters may be *informed* or *uninformed*. If informed, a voter is policy motivated and casts a vote depending on the electoral platform of each candidate. If uninformed, a voter is insufficiently informed about policies (and/or uninterested) and casts a vote depending on the configurations of voters in his or her neighbourhood. The fact that neighbourhoods are overlapping introduces an element of heterogeneity across voters. Candidates are policy motivated. The policy space is unidimensional, but may also incorporate a spatial element that takes into account voters' address.

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

---

Date: This version: May 2012 .

Address for Correspondence: Economics Division, School of Social Sciences, University of Southampton, Southampton SO17 1BJ, UK. .

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 model. The collective processes of opinion formation 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 behaviour of uninformed voters is endogeneized in terms of a simple statistic of the opinions adopted in a voter's neighbourhood. The study of the process of public opinion relies on the characterization of the long-run properties of the process, as well as those of its dynamics. Section 3 analyses the ensued political equilibrium. Specifically, we consider the implications of our model of public opinion on the strategic incentives driving candidates' electoral camoaign. Finally Section 4 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, T. Two candidates, L and R, run and the winner will be decided through simple majority voting

There is a countable set of voters, V. Each $v \in V$ is provided with a utility function decreasing in the distance between the implemented policy $\pi \in [0, 1]$, and her bliss point, $\mathbf{v} \in [0, 1]$:

$$u(\pi, \mathbf{v}) = - \mid \pi - \mathbf{v} \mid \tag{1}$$

Bliss points, $\mathbf{v}$ are randomly distributed with cdf $F(\mathbf{v})$ and density $f(\mathbf{v})$. Each voter $v \in V$ is also provided with a location, $v(x)$, i.e. an *address* on a 1-dimensional lattice Z^1. Addresses are fixed and independent of preferences; bliss points are fixed and independent of the the spatial location, x.

The two candidates, L and R, propose a policy, denoted by $l \in [0, 1]$ and $r \in [0, 1]$. After the elections take place, the policy proposed by the winning candidate is implemented. Parties are policy motivated and have Euclidean preferences, with bliss points $\mathbf{0}$ for party L and $\mathbf{1}$ for party R. Suppose parties propose policies (l, r) with $l < r$, respectively.

When informed, a voter with bliss point $\mathbf{v}$ chooses candidate R iff $\mathbf{v} \geq 0.5(l+r)$., which occurs with probability $1 - F(\mathbf{v} \leq 0.5(l + r))$.

When uninformed, a voter located at x observes the opinion of a randomly chosen nearest neighbours, i.e. any voters at locations $y :\mid y - x \mid= 1$, and tends to imitates it. If $p(x, v)$ denotes the fraction of x's neighbours favouring party R, then a voter who is uninformed voter will favour party R with probability $g_\sigma(p(x, v))$.

The preference specification used to model choices of an informed voter is entirely standard. The behavioural specification used to describe choices of an uninformed voter is designed to capture the amount of correlation that seems to be particularly pertinent in processes of opinion formation and will be elaborated further later.

2.1. **Public Opinion Formation.** As anticipated, we are going to model a dynamic process of public opinion formation, in which we take the voting specifications as primitives.

The general notation of the model we study has individual $x \in V$ choosing ballot $v(x) \in \{0, 1\}$, where $v(x) = 1$ (vs. $v(x) = 0$) denotes voter x supporting candidate R (vs. L). 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.

The dynamic aspect of the model stems from the fact that an uninformed voter's probability of observing an opinion in favour of candidate R depends on x's address and varies endogenously over time. Our general model assumes that the probability that voter x's ballot favours candidate R, given that parties propose policies (l, r) is:

$$\Pr[v_t(x) = 1 \mid (l, r)] = (1 - \alpha)(1 - F(\frac{l + r}{2})) + \alpha g_\sigma p(x, v_t) \tag{2}$$

where $g_\sigma(p(x, v_t))$ is defined as follows:

$$g_\sigma(p(x, v_t)) = \frac{1}{1 + \exp[-4\sigma(2p(x, v_t) - 1)]} \tag{3}$$

and

$$p(x, v_t) = \frac{1}{2} \sum_{y:|y-x|=1} v_t(y) \tag{4}$$

Note that, for any value of $\sigma < \infty$, $g(0.5) = 0.5$, $0 < g(0) < g(1) < 1$) and $g(0) + g(1) = 1$. We think of this specification as a noisy form of imitation, in the sense that, for any $\sigma < \infty$ an uninformed voter is likely to imitate his or her neighbours, but could do otherwise with positive, albeit small, probability. Clearly, as $\sigma \to \infty$, $g_\sigma(p(x, v_t)) \to p(x, v_t)$, reflecting the fact that behaviour is entirely driven by the imitation of a randomly chosen neighbour.

Also, our assumptions ensure that the probability that more than one voter revises opinion at the same time is negligible, since within a small time interval dt:

$$\Pr[v_{t+dt}(x) = 1 \mid (l, r)] = \Pr[v_t(x) = 1 \mid (l, r)]dt + o(t)$$

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 V. For any $t \geq 0$, let $\alpha_t \in [0,1]$, and $g_\sigma(p(x,v_t))$ defined in (3) . At each random exponential time t, with mean one, voter x chooses ballot R at rate:*

$$\Pr[v_t(x) = 1 \mid (l,r),\sigma] = (1-\alpha_t)(1-F(\frac{l+r}{2})) + \alpha_t g_\sigma(p(x,v_t))$$

In general, we shall denote by $v_t \in \{0,1\}^V$ public opinion at time t, we shall assume that V is countable and consists of a unidimensional lattice (i.e., a line), and we are interested in characterizing its evolution over time and over space.[1]

$\alpha = 0$ **Voters are always Informed.** Clearly, in the special case where $\alpha_t = 0$ for all t, the model looses any spatial dimension and collapses to the standard framework, with $m : F(m) = 0.5$ being the median voter. Each voter supports party R with probability $p(l,r) \equiv 1 - F(l+r)/2$ and the asymptotic behaviour of the process of public opinion formation is described by the product measure of mutually singular countable Markov chains.

$\alpha = 1$ **Voters are always Uninformed.** Interesting questions arises in the case when $\alpha_t = 1$ for all t, as 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), although the behaviour of a single uninformed voter no longer is (due to the local nature of the interaction).

In words, the Theorem that follows asserts that in the absence of any informed voter, the process admits an invariant measure that can be fully characterised. Specifically, this invariant measure posits higher limit probability to configurations of public opinion where opinions among voters tend to be spatially homogenous. Moreover, the process is ergodic, meaning that this invariant measure is unique and independently of any initial condition, the process of public opinion will converge to it over time.

Theorem 1. *Consider the process of Public Opinion as in Definition (1). and suppose $\alpha_t = 1$ for all $t \geq 0$.*

The following measure is the unique invariant measure for the process:

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

[1] 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 V and configuration $v^{\mathbf{1}}$ where $v(x) = 1$ for all x in V) 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.

where K is such that $\sum_v \mu^\sigma(v) = 1$ and $\lim_{t\to\infty} \mu_t^{\sigma,\mu_0} \equiv \mu_\infty^\sigma = \mu^\sigma$ for any initial distribution μ_0. Furthermore, for all σ, $v^i = \{v^{\mathbf{0}}, v^{\mathbf{1}}\}$ and $v \neq v^i$:

$$\frac{\mu_\infty^\sigma(v^{\mathbf{0}})}{\mu_\infty^\sigma(v^{\mathbf{1}})} = 1 \qquad \textit{and} \qquad \lim_{\sigma\to\infty} \frac{\mu_\infty^\sigma(v)}{\mu_\infty^\sigma(v^i)} = 0$$

Proof. See Appendix. □

The above Theorem provides a characterization of the limit behaviour of the dynamics of public opinion when $\alpha_t = 1$ at any point in time t. Given that, under $g_\sigma(p_t(x, v))$, each uninformed voter can support either candidate with strictly positive probability, 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 and the process does not show any path dependence. In addition, the process is *monotonic* or *attractive*, in the sense that voters tend to agree in opinion with their nearest neighbours. As the limit distribution (5) 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, notice that the sum of which in the square brackets of (5) 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 $v^{\mathbf{1}}$.

It is clear that this reasoning only relies on a comparative statics exercise over the limit distributions of a sequence of identical processes, that differ only in the parameter σ, obtained by taking the limit for $\sigma \to \infty$ of the limit distribution obtained for $t \to \infty$. To gain a better understanding of the dynamics, what we do next is to reverse the order of the limits, by first looking at the limit for $\sigma \to \infty$ and second at what happens along the dynamics of this process (i.e. asymptotically for $t \to \infty$).

Theorem 2. *Consider the process of Public Opinion as in Definition (1). Take $\sigma = \infty$ and suppose $\alpha_t = 1$ for all $t > 0$.*

For $z \in [0, 1]$, let $\mu \equiv \mu_z$ be the product measure with density z, i.e. $\mu_z\{v(x) = 1\} = z$ for all $x \in V$. Suppose that the process is started with μ_z at time 0. Then the process is path dependent and:

$$\mathfrak{I}_e = \{\mu^{\mathbf{0}}, \mu^{\mathbf{1}}\} \ \textit{and} \ \lim_{t\to\infty} \mu_t^{\mu_z} = (1 - z)\mu^{\mathbf{0}} + z\mu^{\mathbf{1}}$$

Furthermore, convergence obtains at rate $\sqrt{t}$

Proof. See Appendix. □

The above result shows that if voters are uninformed and formulate an opinion simply by imitation one of their neighbours (which, with our parameters, corresponds to the limit case, for $\sigma \to \infty$), then the process is clearly path-dependent, since both configurations where all voters adopt the same opinion are invariant for the process. Moreover, and perhaps surprisingly, these are the *only* two invariant measures for the process, and we are able to characterize their basins of attraction in terms of the initial condition: if z is the probability with which each voter favours candidate R at time 0, then z also identifies the set of initial conditions that lead the process of public opinion to asymptotically show uniform agreement on candidate R.

One thing worth noticing is that, along the dynamics, the process shows *consensus*, in that if we look at any possible couple of voters, x and y in V, the probability that they choose different opinions approaches zero asymptotically:

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

Clearly, for any $z \in (0,1)$, 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.

Since our process is defined in the two dimensions of time and space, we can relate these two dimensions in a space-time analysis, by looking at the emerging *clustering* process. With the term “cluster” we mean a connected group of individuals holding the same opinion, and measure it by 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 can express the length of a cluster as a function of t. The result we obtain state that consensus grows slowly, at rate $\sqrt{t}$. This means that, for t large, any cluster of voters supporting the same candidate tends to be *almost stationary*, in the sense that the rate at which it changes is slower than the rate at which time changes. In other words, although the process we analyse does not admit any stationary distribution where both opinion co-exist indefinitely, any such configuration can indeed be observed along the dynamics, and when viewed locally, remains almost stationary. Notice that the study of this *clustering* process, and the associated space-time analysis can only be done relying on the local specification of the model.

3. Political Equilibrium and the allocation of funding in an electoral campaign

Parties are policy motivated and aim at maximizing expected utility. Given policies (l, r), and bliss points (x_L, x_R), let $\Pr[v(x) = R \mid (l, r)]$ denote the probability that voter x casts a vote in favour of R. The expected support for candidate R is then $\sum_{v\in V} \Pr[v(x) = R \mid (l, r)] \equiv R(l, r)$ and the expected support for candidate L is $\sum_{v\in V} \Pr[v(x) = L \mid (l, r)] \equiv L(l, r)$ [2]. A Political Equilibrium is defined as follows:

[2]Since our model allows for a countable population of agents, an equivalent formulation would rely on *share of the votes* to be intended as the limit of its natural restriction to $[-N, N]$ as

Definition 2 (Political Equilibrium). *A Political Equilibrium is a pair of policy choices,* (l^*, r^*) *such that* $l^* \in \arg\max_l u_L(l.r)$ *and* $r^* \in \arg\max_r u_R(l.r)$*, where* $u_L(l.r)$ *and* $u_R(l,r)$ *are given by:*

$$u_L(l,r) = L(l,r)(-\mid l - x_L \mid) + R(l,r)(-\mid l - x_L \mid)$$
$$u_R(l,r) = L(l,r)(-\mid l - x_R \mid) + R(l,r)(-\mid l - x_R \mid)$$

Voters are Informed at time 0, but Uninformed at time $t > 0$.

$\alpha_0 = 0, \quad \alpha_t = 1 \quad \forall t > 0$. Since the process is ergodic, no matter where the process starts, the probability with which each configuration could be observed asymptotically is given by the limit distribution in (5). 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 supports party R with probability $p(l,r) \equiv 1 - F(l+r)/2$, as any informed voter would choose to do, upon candidates announcing policies (l,r). This formally corresponds to the case where $\mu_0 \equiv \mu_p$, that is a product measure with density p (such that $\mu_p\{v(x) = 1\} = p$ for all $x \in V$. Hence, if voters' behaviour were policy motivated at time $t = 0$, but became driven by neighbours' preferences from then on (for any $t > 0$), then limit behaviour of the public opinion process will be entirely unaffected by the initial condition. In an analogous manner, if policies were to be announced at a finite point in time $0 < T < \infty$ and taken into account (with weight $0 < \alpha \leq 1$) by all informed voters at any time $t > T$, the asymptotic properties of the process of public opinion would not be affected. 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 policy considerations in each of the voters' assessments, the reversibility properties upon which the full characterization relies fail to hold.

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 distribution of votes matters in the long run, as well as in the short run. In particular, simply by looking at the limit distribution for the ergodic process generated by the dynamics of the *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,

$N \to \infty$ whenever this limit exists. As it will become clear, this will not play a key role in our results.

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:

$$\begin{array}{lllllll} v_A: & .... & v(x-2)=1 & v(x-1)=0 & v(x)=1 & v(x+1)=0 & .... \\ v_B: & .... & v(x-2)=1 & v(x-1)=1 & v(x)=0 & v(x+1)=0 & .... \end{array}$$

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

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

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 $v(x-1) \neq v(x)$:

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

Suppose, for simplicity, that the process is started deterministically at configuration v. In this case the duality equation (10) (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 (9), 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 **??**, 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 **??**, 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 **??**, Part 3.), but it also increases the rate at which support grows in the population (Remark **??**) 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[3]), 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

[3]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.

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.

4. 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.

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 (??) by substituting $\sigma = \frac{1}{4}\log(2^{2(m+1)} - 1)$:

$$\Pr{}^{nc}[1 \mid m, p(x,v)] \equiv \Pr{}^{nc}[1 \mid \sigma, p(x,v)] = \tag{6}$$

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

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 (??) 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

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

$$\begin{aligned}\frac{\mu_\infty^{(m)}(v_{x=1})}{\mu_\infty^{(m)}(v_{x=0})} &= \exp[4\sigma^{(m)}(p)d(2p-1)]\\ &= \frac{1}{1+\exp[-2\sigma^{(m)}(p)d\sum_{\{y:\|y-x\|=1\}}(2v(y)-1)]}\cdot\\ &\quad\cdot[\frac{1}{1+\exp[2\sigma^{(m)}(p)d\sum_{\{y:\|y-x\|=1\}}(2v(y)-1)]}]^{-1}\\ &= \frac{\Pr^{nc}[1\mid p,m,\,v(x)=0]}{\Pr^{nc}[0\mid p,m,\,v(x)=1]}\equiv\frac{\Pr^{nc}[1\mid p,m]}{1-\Pr^{nc}[1\mid p,m]}\end{aligned}$$

where $\Pr^{nc}[1 \mid p,m]$ is as in equation (**??**). 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 (**??**), 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 (**??**).

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:

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

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 ??

1. Recall that the processes $v^{nc}(m)$ and $v^{c}(m)$ are ultimately defined by the transition probabilities of which in (??) and (??) 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:

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

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:

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

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 (??). 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:

$$c^c(x,v) = v(x) + p(x,v)(1 - 2v(x)) \tag{8}$$

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 (8) 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 (8) 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:

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

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 $\mid A \mid= 2$, $g_t(A) \to_{t\to\infty} 1$, 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 $\mid A \mid= 2$ for some $t^* < \infty$.

Let $A = \{x \in S : v(x) = 1 \quad \text{for all } x \in A\}$ 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)):

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

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 $\mid A_t \mid$ random walks, started at A, are still alive at time t.

By using this duality relation, we now show that, given a product measure μ^θ, $\lim_{t\to\infty} \mu_t^\theta = (1-\theta)\mu^{\mathbf{0}} + \theta\mu^{\mathbf{1}}$.

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 $\mid A_t \mid= 1$ (which is finite with probability one by our assumption that $g_{t^*}(A) = 1$ when $\mid A \mid= 2$) the duality equation (10) 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 $\lim_{t\to\infty} \mu_t^\theta \to (1-\theta)\mu^{\mathbf{0}} + \theta\mu^{\mathbf{1}}$ 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:

$$c^c(x,v,\alpha,\pi,m) = v(x) + (1-2v(x))[\alpha \Pr{}^c[1 \mid p(x,v),m] + (1-\alpha)\pi] \tag{11}$$

for $v(x) = \{0,1\}$ where $\Pr^c[1 \mid p(x,v),m]$ is given in equation (**??**). 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 $v^{\mathbf{1}} = \{v \in \{0,1\}^{Z^1} : v(x) = 1\}$ is absorbing, since, form the flip rates of which in (11), 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[4].

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

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

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

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

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 ($v_N^{\mathbf{1}} \equiv \{v(x) = 1 \text{ for all } x \in S_N\}$) 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 ??

[4]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 $\alpha \in (0, 1)$.

a. We have already shown that, under the assumptions of Theorem **??**, 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 **??**, Part 3.) that, starting from any time T distribution, the system converges to $v^{\mathbf{1}}$. 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 $v^{\mathbf{0}}$ (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 $\gamma < 1$.