Learning in Elections and Voter Turnout

Stefano Demichelis*and Amrita Dhillon†

January 2002

Abstract

Game theoretic models of voter turnout have recently fallen into disrepute because the crucial ingredient of the model, the probability of being pivotal for an individual voter, is infinitesimal in large elections. Moreover such models are plagued by the problem of multiple equilibria. We show that assuming voters to be boundedly rational instead of fully rational helps to ameliorate both these problems. Modelling the dynamics of voter learning in such a context leads to a unique equilibrium with a high probability, which increases with the number of voters. This enables the derivation of testable implications of the theory: increases in costs of voting affect turnout adversely but there may be persistence of turnout levels between elections even though costs and other parameters change. Increase in uncertainty increases

*CORE, 34 Voie du Roman Pays, 1348 Louvain la Neuve, Belgium, email: demichelis@core.ucl.ac.be

†Address for Correspondence: Amrita Dhillon, Department of Economics, University of Warwick, Coventry CV4 7AL, UK. Email:A.Dhillon@warwick.ac.uk. We would like to thank Alan Beggs, Thomas Palfrey, Matthew Jackson, V.Bhaskar, Myrna Wooders, Jonathan Cave, Michel Le Breton, Rabah Amir, Debraj Ray, and seminar participants at Ischia and the universities of Cambridge, Exeter, Essex, Warwick Mathematics, Montreal and Toulouse, Pennsylvania State University, ITAM Mexico, New York University, Florida International University, for comments.

turnout while increases in the size of the electorate decrease it, in line with intuition.
JEL Classification: C72, D72
Keywords: Voter Participation, Voter Learning, Asymptotically Stable Equilibrium, Markov chain, long run equilibria

1 Introduction

The question of why people vote in large elections has occupied political scientists and economists since a long time (see e.g. Downs (1957)). If people were rational and interested in voting in order to affect the outcome of the election (*instrumental voters*) then they would never vote. This is because voting is costly and the expected benefits small since the probability that one vote counts is infinitesimal. However, observed turnout rates are quite high. This is the "paradox of voting" (Downs(1957)), Ferejohn and Fiorina (1974)). The first game theoretic models of the participation game were those in Palfrey and Rosenthal (1983,1985) (henceforth PR). Recently the weight of scholarly opinion has been going against such "rational" models of turnout which depend on the probability of being pivotal to provide incentives to vote. There is increasing emphasis on voting as a consumption decision, i.e. people vote because they like to – it gives them a "warm glow". These are the "expressive" theories of voting. Unfortunately, these theories at present are not very developed[1] and do not offer testable predictions. In particular, expressive theories cannot explain stylised facts such as increasing abstention rates over time as observed in many countries or the dependence of the turnout on the perceived distance between the electoral platforms of the candidates, or on the transactions costs of voting. These stylised facts seem to suggest that instrumental motivations do play a role, even if such motivations are not the only relevant ones in the voting decision. Indeed, empirical studies (Mueller, 1989, Ch.18) support the instrumental voter paradigm, although they often contradict each other in the details.

The first game theoretic investigations of intrumental voting were developed by Ledyard (1981) and PR (1983,1985). The purpose of our paper is to show that a simple extension of the PR model still captures some of the essential features of real elections, at least at a qualitative level. In doing this, we can manage to solve, in a very natural way, the problem of mul-

[1]See however, Schuessler(2000) for a detailed analysis of expressive voting and a theory that does provide some predictions.

tiplicity of equilibria in the original model. We also clarify some common misconceptions about the levels of turnout predicted by PR , and show that these are not inconsistent with observed levels of turnout.

Since the question of which theory better explains the facts is crucial to decide between them, it is essential to be able to derive some comparitive statics, at least at a crude level.

Multiple equilibria become a serious problem when one wants to test the several different theories about voting, especially when different equilibria in the same game yield opposite comparitive statics! In addition, a correct estimation of the probability of being pivotal is vital, but, as Fischer (1999) remarks; "Since it could not be known in advance which equilibrium would be attained in any particular situation, it is implied that the probability of being decisive is not well defined in this framework. It would only make sense if probabilities could be assigned to the occurence of each of the possible equilibria ...".

This is the first contribution of our paper: giving a plausible and coherent selection criterion that will enable us to check whether the predictions of voter participation games are consistent with reality.

PR (1983,1985) examined the issue of voter turnout in two different settings, one with complete information and one with incomplete information. They found that even in the simplest model with complete information (two candidates, two types of voters, symmetric equilibria) there was a problem of multiple equilibria. In particular, even with an equal number of the two types of voters in the population and restricting attention to symmetric mixed strategy equilibria[2] they found two types of mixed strategy equilibria for plausible cost levels– one with low turnout and one with substantially high turnout. They say that this high turnout equilibrium "has the unappealing feature that there is another equilibrium with almost no one voting.

[2]There may be objections to studying mixed equilibria, but we believe that they are quite justified in this setting – as psychological experiments have shown that choices between alternatives that are perceived as being similar tend to be relatively random (Fudenberg and Levine (1998) Chapter 4), besides the usual purification arguments.

Apparently the only reason the upper one can be sustained is that the two electorates are of the same size so that for the probability of voting very close to 1, the probability of a tied election is very high. Again, the result rests on the fact that in equilibrium there is essentially no strategic uncertainty." (PR, 1985). Of course, there is no reason why this argument should apply only to the high turnout equilibrium[3]. They show that moving to the corresponding incomplete information game gets rid of this high turnout equilibrium. Indeed, in their model of the incomplete information game, they show that with some assumptions on the type of uncertainty allowed, the only symmetric mixed strategy equilibrium that survives is the low turnout one.

While we agree with PR (1985) that the low turnout equilibrium is intuitively more appealing, we argue that the reason they cite may not be the most compelling one. In fact, we show that removing some of the assumptions PR (1985) make on the type of voter uncertainty, the incomplete information model may still lead to multiple equilibria (at least for populations that are not "too" large). We show thatconsidering the dynamics of reaching a Nash equilibrium by boundedly rational agents and incorporating in the model an important feature of real elections, the information aggregation devices like pre-election polls that exist, and that can co-ordinate people's beliefs [4], would lead in a very simple and natural way to the low turnout equilibria in their example.

So why should an adaptive process be a good refinement criterion in this case? We think there are two types of reasons for believing this: methodological and empirical. We are studying an anonymous game with a large number of players who have little possibilities of communicating and who have little a priori information on what the others will do. It is not reasonable to assume that they will divine the Nash equilibria and that they will assume that everyone else will do. This applies 'a fortiori' to all the selection

[3]If we use the standard deviation as a measure of "strategic uncertainty", one can see that it is the same for the two equilibria.

[4]See Fey (1997)for a similar idea but in the context of costless voting and multiple candidate elections.

criteria that are used to refine Nash equilibria (it is easy to see that these have no cutting edge in our problem anyway). It seems more sensible to assume that during the period that precedes the election players will collect the information that is available to them (e.g. statistics from previous elections, opinion polls etc.), and then find what is their optimal behaviour via some type of Gedanken Experiment.

It is interesting to see that the outcomes of such processes not only are Nash equilibria, but satisfy the strongest known rationality criteria (although the converse is not true – see Demichelis and Ritzberger (2000)). On the empirical side there is good evidence (Cooper et al (1990), van Huyck et al (1990)) that players do not apply a rationalistic selection criterion when they select strategies to play. Thus, for instance weakly dominated strategies play a role in their choice, and adaptive considerations seem to play a strong role in guiding their behaviour.

We should remark here that the problem of multiplicity of Nash equilibria is not specific to the voting problem. The voting decision can be seen as a two sided version of the problem of private provision of a public good where the 'public' good is the fact of having one's candidate winning and the contribution is the cost of voting. Indeed both cases present similar problems (free-riding, instability of the equilibrium in which everybody contributes as the number of players increases etc.). There are obvious co-ordination problems in both these settings[5](Bagnoli and Lipman (1989)). The methodology of this paper is thus equally applicable to the more general "team games" (PR, 1983).

We show how the adaptive process we model leads to the selection of the "intuitive" equilibrium of PR, both for complete and incomplete information. We then move on to the comparitive statics analysis and qualitative predictions. This is done in Sections 3 and 4 for each of the two cases. The main results are: when costs increase or the size of the electorate increases, absention increases, in line with intuition; higher uncertainty about costs, in

[5]We thank Michel Le Breton for pointing this out to us.

contrast causes turnout to increase; social norms and habits are important in voting decisions: if turnout levels were historically high, they tend to stay high; for example, voter turnout in Italy is persistently high (close to 90% on average) compared to the United States (50% on average).

We also build a model of repeated elections that makes an interesting prediction: for small electorates, sudden downward jumps in turnout may be observed; while for large electorates turnout changes smoothly with cost and benefit parameters. This corresponds to a phenomenon that is commonly observed in Western democracies: that of steadily falling turnout rates over time (http://www.idea.int).

Finally, in Section 7 we discuss a common criticism of the PR model: that it predicts turnout levels that are much lower than empirically observed ones. On the one hand we show that at least as far as the orders of magnitude are concerned, the PR model with uncertainty has predictions that are not far from real elections. Moreover, we argue that several effects may cause turnout to be even larger than the figures we compute.
The framework of this paper is as follows: We discuss related literature in Section 2. We then introduce the model of voter learning in Section 3, then we consider the complete information game in Section 4 and the incomplete information game in Section 5. Section 6 and 7 have some discussion on how our model fares when faced with some stylised facts on elections and what we can say about expressive and instrumental voting. Finally Section 8 concludes.

2 Related Literature

There has been a lot of recent literature looking at the problem of how to explain empirically observed levels voter turnout (see Dhillon and Peralta (2002) for a recent survey). The PR (1983,1985) models are the basis for our paper. Ledyard (1981) preceded the PR models but is more general in the sense that it endogenizes party positions as well. However turnout levels are not characterised, although there are a few examples.

Bounded rationality models of voter turnout have been proposed by Conley, Toossi and Wooders (2001), and Sieg and Schultz (1995). They are different from our model in that they are pure evolutionary models where some citizens are pre-programmed to turn out with a high probability , and they consider the conditions under which the strategy of high turnout survives.

We assume that information about candidates is complete and that there are conflicts between citizens on which candidate is best (private values). If we relax the assumption of complete information and private values, the turnout decision acquires another dimension – abstention may become strategic in order to let informed voters make the correct decision. In the context of costless voting this problem is studied by Feddersen and Pesendorfer (1996,1999).

An excellent article on why turnout is related to the expected closeness of the election (and which does not rely on the probability of an individual voter being pivotal) is the model of Shachar and Nalebuff (1999). They model voters as passive followers who vote if they are contacted by a party. It is parties who make the strategic decision of how much effort to put into the campaign and who take account of how close the election is expected to be.

3 The Model

In this section we briefly describe the model due to PR (1983, 1985).There are two candidates (or two alternatives): 1 and 2. The voting rule is Simple Majority Rule: in case of tie either 1 is chosen or a coin toss takes place. There are N voters in the population. There are two groups of voters: T_1 (with N_1 voters who prefer 1) and T_2 (with N_2 voters who prefer 2) and all voters belong to one of these groups. Voting is costly and the cost of voting is the same for all voters. Voters have two pure strategies: vote (participate) or abstain – if they vote they always vote sincerely (i.e. for their best candidate).

We ignore the strategy of voting for the candidate that is less preferred as this strategy is strictly dominated by abstention, and cannot be a part of any Nash equilibrium.

Let R represent the expected net benefit from voting , p the probability of being pivotal, C the cost of voting, B the benefit from voting i.e. the difference between the benefits of i's more preferred alternative winning as opposed to the less preferred one, and D a fixed benefit from the act of voting (civic duty). Let n_r^i, $r = 1, 2$ denote the number of voters excluding voter i who turn out to vote for candidate r. Let α_1 denote the probability that $n_1^i = n_2^i$, α_2 the probability that $n_1^i = n_2^i - 1$, α_3 the probability that voter i is not pivotal and candidate 1 leads, while α_4 denotes the probability that voter i is not pivotal and candidate 2 leads. Note that $\sum_k (\alpha)_k = 1, k = 1, 2, 3, 4$. Thus the expected utility from voting for voter i is given by:

$$\alpha_1 B + \alpha_2 \frac{B}{2} + \alpha_3 B + \alpha_4 0 - C + D \tag{1}$$

The expected utility of voter i from abstaining is denoted by:

$$\alpha_1 \frac{B}{2} + \alpha_2 0 + \alpha_3 B + \alpha_4 0 \tag{2}$$

Therefore R is the difference between 1 and 2, and has the following expression:

$$R = (\alpha_1 + \alpha_2)\frac{B}{2} - C + D \tag{3}$$

Let $c = C - D$ represent the net cost of voting. We can normalise by dividing throughout by B so that only the ratio of costs to benefit matters. So, w.l.o.g set $B = 1$. Let $p = \alpha_1 + \alpha_2$. Thus a voter i will vote if and only if $p - 2c \geq 0$. We consider Nash equilibria of this game.

Let the probability for player i to choose to vote be q_i. Then $(q_1, q_2, ..., q_n)$ is a mixed strategy Nash equilibrium, if for all i voting and non-voting give the same expected payoff, given the mixed strategies of other players. As PR (1983) shows there are many Nash equilibria of this game, symmetric and asymmetric, pure strategy and mixed strategy. In our paper, we illustrate our arguments with the multiple equilibria that arise even when it is assumed that $N_1 = N_2$, and attention is restricted to symmetric equilibria. This case exhibits the same problems of multiplicity of equilibria as the general case, but allows us to make several technical simplifications in the solution that can be discussed at a very intuitive level.

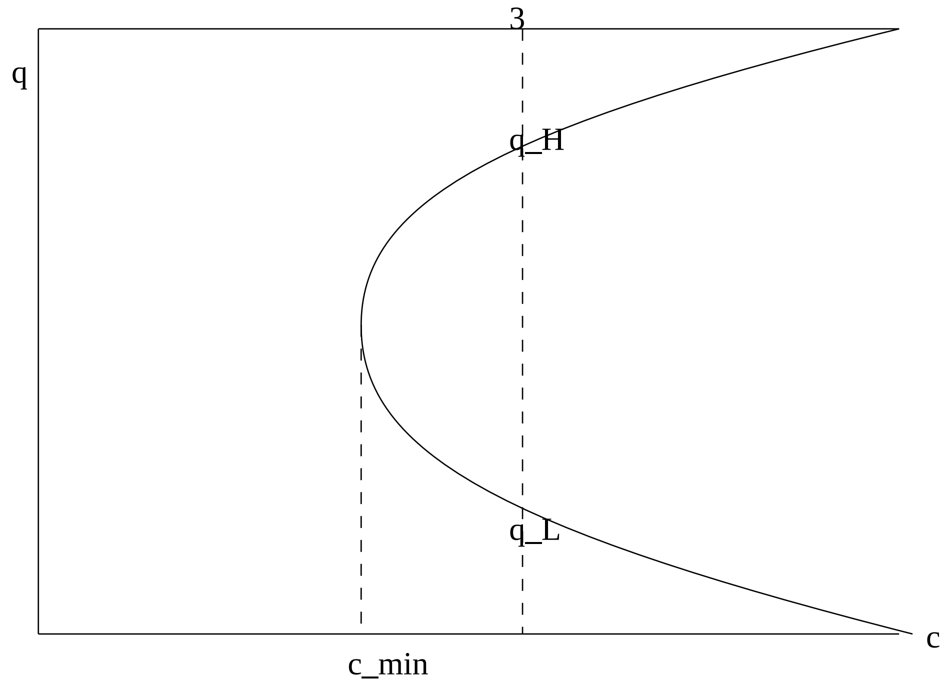
The symmetric mixed strategy equilibrium in this game with $N_1 = N_2 = M$ is denoted q and it satisfies the following equation:

$$\begin{aligned} 2c &= \sum_{k=0,\ldots,M-1} \frac{(M-1)!}{(M-1-k)!(k!)} \frac{(M)!}{(M-k)!(k!)} q^{2k}(1-q)^{2M-2k-1} \\ &+ \sum_{k=0,\ldots,M-1} \frac{(M-1)!}{(M-1-k)!(k!)} \frac{(M)!}{(M-1-k)!(k+1!)} q^{2k+1}(1-q)^{2M-2k-2} \end{aligned} \tag{4}$$

If $0 < c < 1/2$ then this equation has either no solution, 1 solution or two. Figure 1 shows (as in PR, 1985) the equilibria. Note that to plot the graph and discuss the equilibria we use the conjecture made by PR (1983), reproduced below. *Conjecture 1(PR, 1983) For each integer* $M > 1$, *there exists a minimum voting cost denoted by* c^M_{min} *such that: (1) The unique symmetric equilibria at* c^M_{min} *is* $q = 1/2$. *(2)* $c(q) = c(1-q) \geq c^M_{min}$ *for* $q \in (0,1)$. *(3) For every* $c \in (0, 1/2)$ *and every* $\epsilon > 0$, $\exists M_\epsilon$ *such that for all* $M \geq M_\epsilon$ *there exist exactly two equilibria,* $q_H(c), q_L(c)$ *such that* $0 < q_L < \epsilon$ *and* $1 - \epsilon < q_H(c) < 1$.

The proof is by numerical calculation for small N. For large N, the conjecture is supported by a heuristic argument based on approximating the binomial with the normal distribution[6]. To see how the function $f(q)$ changes as M increases see Figure 4 in PR (1983).

[6]This is available upon request from the authors.



The example shows that as the cost increases beyond c_{min} there are two types of mixed strategy equilibria: one where almost everyone votes (denoted q_H) and one with almost no one voting (denoted q_L). In addition there is a pure equilibrium in which everybody votes. Let us analyze them separately: in the pure strategy equilibrium everybody goes to vote believing that everybody else will go. Such a situation may arise when the number of voters is small but seems unlikely for electorates of large size. Note also that it relies heavily on complete information on the number of voters in each group[7]. The high turnout equilibrium q_H is counterintuitive, particularly if one considers comparative statics: it predicts an increase in turnout when the cost of voting increases. The low turnout equilibrium q_L, has good comparative statics properties and, apart from predicting "low" turnout, seems the one that should appear for large N. Note also that it Pareto dominates the other two. The outcome (50% win for each candidate) is the same in all three

[7] The assumption needed here is merely knowledge of the number of voters in each group, not equality: if $N_1 \leq N_2$ there is still a (mixed-pure) equilibrium in which everybody in group 1 votes based on similar beliefs.

equilibria, even though turnout is different (in any symmetric *mixed strategy* equilibrium with identical costs of voters, the probability that a voter is pivotal is the same for all voters).
PR, 1985 state that the high turnout equilibrium is not robust in that they arise because of the fact that there is almost no strategic uncertainty in the complete information model.
We claim that this is not the main problem with the high turnout equilibrium. In fact the problem of multiple equilibria remains in the incomplete information model: it will be shown below (see Section 4) that at least in small populations, unless the amount of uncertainty is quite large (probably too large to be observed in concrete applications), not only does one still get three equilibria near the original ones but, in certain cases, many more may appear. Standard refinements like trembling hand perfection do not have any bite.
We claim, instead, that the main problem with the equilibria other than the low turnout one is that they are inconsistent with reasonable models of voter learning. Indeed, the missing piece of the model, opinion polls, and the trial and error process that arise quite naturally in such a setting is a good way to select the "robust" low turnout equilibrium, and this is what our model does.
There is a very intuitive reason why the high turnout equilibrium is not a robust one, and this is true regardless of whether the game is of complete information or not. This is because of its stability properties with respect to learning dynamics. But first we will describe our model of voter learning.

4 Complete Information

In this section we will discuss learning in the PR model with complete information. Its primary purpose is to present the learning dynamics and show its use in selecting away "bad" equilibria. The artificial assumption of identical voting cost for everybody will be removed in section 4.

4.1 Dynamics: Single Elections

Learning can be modeled in several different ways (see Fudenberg and Levine (1998) for an overview of the literature). A good process of learning should satisfy the, possibly conflicting, requirements of assuming as little rationality as possible from the players and of being efficient enough to give acceptable outcomes (we want the voters to learn something before election day comes...). Typical models are the best reply dynamics (as in the Cournot adjustment model), and its improvement (discrete) fictitious play (Brown (1951)). These assume rather myopic players but have the defect that sometimes they forecast players behaving in a rather silly way, overshooting their beliefs and strategies so as to cycle for ever around an equilibrium without ever coordinating on it (Fudenberg and Kreps (1993)).

The problem can be circumvented, without having to assume a more sophisticated behaviour of the players, using continuous fictitious play, a model in which they gradually adapt their strategy instead of jumping suddenly from one pure strategy to another(Demichelis and Germano(2000), Fudenberg and Levine (1998) Chapter 4).

Let $f(q, M)$ denote the probability of being pivotal as a function of the equilibrium probability of voting of all voters (computed above in the RHS of expression (5) with the coin tossing tie breaking rule). Then, as before the best reply is to vote if

$$f(q, M) - 2c > 0 \tag{5}$$

and abstain in the opposite case with equality corresponding to indifference between the two strategies. We analyse this process for a fixed M.

A voter starts by observing a q, e.g. the share of the population who voted on the last election or the result given by a poll (we assume M is not too small so that the empirical observation of this quantities is a reasonable estimator of q) and checks whether the inequality 5 tells her to go to vote (or to abstain). She realizes that everybody will do the same and so expects the correct q to be higher (lower). Once she has adjusted q she checks again what is her best reply and so on.

This is consistent with PR's focus on symmetric equilibria. If this process converges to some point q^*, she will apply the corresponding mixed strategy[8].

The dynamics can be described by the differential equation[9] below, where the symbol M has been suppressed as the analysis is for a fixed M. The state variable is q, the symmetric mixed strategy for all players. The state space is thus the space of mixed strategies, $[0, 1]$. The function $K(q)$ is the vector field.

$$\frac{dq}{dt} = K(q) \tag{6}$$

and sign $K(q)$ =sign$(f(q) - 2c)$ The functional form of $K(q)$ depends on the particular speed of learning. Different individuals could have different K; as long as the sign requirement on the possibly different $K_i(q)$ is respected it will not affect the results. So, to simplify notation, we assume that everyone adjusts in the same way. We also assume that the initial point is the same for all voters. Again, the hypothesis of a common initial q is not very restrictive– all of our results go through with small differences in the initial q. The assumption that voters are able to compute $f(q, M)$ exactly may seem at variance with the bounded rationality hypothesis, but for our purposes we could as well assume that $f(q, M)$ is some approximation of the true function, as long as it has the same shape. For more on these points see Section 7.

We assume that the function $K(q)$ satisfies Lipschitz continuity and so there exists a unique solution for any initial q_0, for all $t \in R$. The solution of this equation for any initial condition, $q(q_0, t)$ is continuous[10]. Note that our requirements on the dynamics are very mild – any payoff consistent (and

[8]We may also think of voters using pure strategies and q describes the proportion of the population that votes – then our assumption is that population shares move continuously with changes in the expected benefit of voting. See Fudenberg and Levine (1998) Chapter 4 for more discussion on this point.

[9]If the adjustment steps are small enough, taking a discrete adjustment process would give essentially the same results

[10]See Weibull (1995) Appendix.

so any payoff monotonic or payoff positive (Weibull (1995), Demichelis and Ritzberger (2000)) dynamic satisfies it.

Now we examine the behaviour of the dynamics on the strategy space. The basic intuition is that the outcomes of the learning process that are likely to be seen in concrete cases are the limit points of $q(t)$ as t goes to infinity.

Now, we introduce some definitions adapted from Weibull (1995): Let $q(q_0, t)$ be the solution of the differential equation (6) with initial condition $q(0) = q_0$. Let $q(q_0, t) \in X = [0, 1], \forall t \in R$, i.e. the state variable q is a symmetric mixed strategy (the same for all players).

Definition 1: A state $q^* \in X$ is said to be Lyapunov stable if every neighborhood B of q^* contains a neighborhood B^0 of q such that $q(q_0, t) \in B$ for all $q_0 \in B^0 \cap X$ and $t \geq 0$.

Intuitively a state is Lyapunov stable, or just stable, if no small perturbation away from it induces a movement away from it.

Definition 2: A state q^* is asymptotically stable if it is Lyapunov stable and exists a neighborhood B^* such that the following holds for all $q_0 \in B^* \cap X$:

$$Limq_{t\to\infty}q(q_0, t) = q^* \tag{7}$$

A point that is not stable will be called unstable. While stability requires that there be no pull away from the state, asymptotic stability requires in addition that there be a local pull towards it as well.

Definition 3: Basin of attraction of state q^* : is the set of points $q_0 \in C$: $q(q_0, t)_{t\to\infty} \to q^*$.

Intuitively, the basin of attraction of q^* is the set of initial conjectures $q_0 \in C$ that, with learning, will lead to q^*.

We denote by equilibrium 1 in Figure 1 (also Figure 2 below) the low turnout mixed strategy equilibrium q_L, equilibrium 2 will be the high turnout mixed strategy equilibrium, q_H, and 3 will be the pure strategy full turnout equilibrium. Now we can state Proposition 1:

Proposition 1: For any learning dynamics of type (6) Equilibria 1 and 3 will be asymptotically stable, while equilibrium 2 will always be unstable.

We refer to the appendix for the (elementary) proof. Here is an informal discussion: If $c > 1$ or $c < c_{min}$ any trajectory trivially converges to the unique equilibrium which is the zero turnout or the full turnout equilibria respectively. When $c_{min} < c < 1$ the qualitative behaviour of the dynamics is as in Figure 2 below:

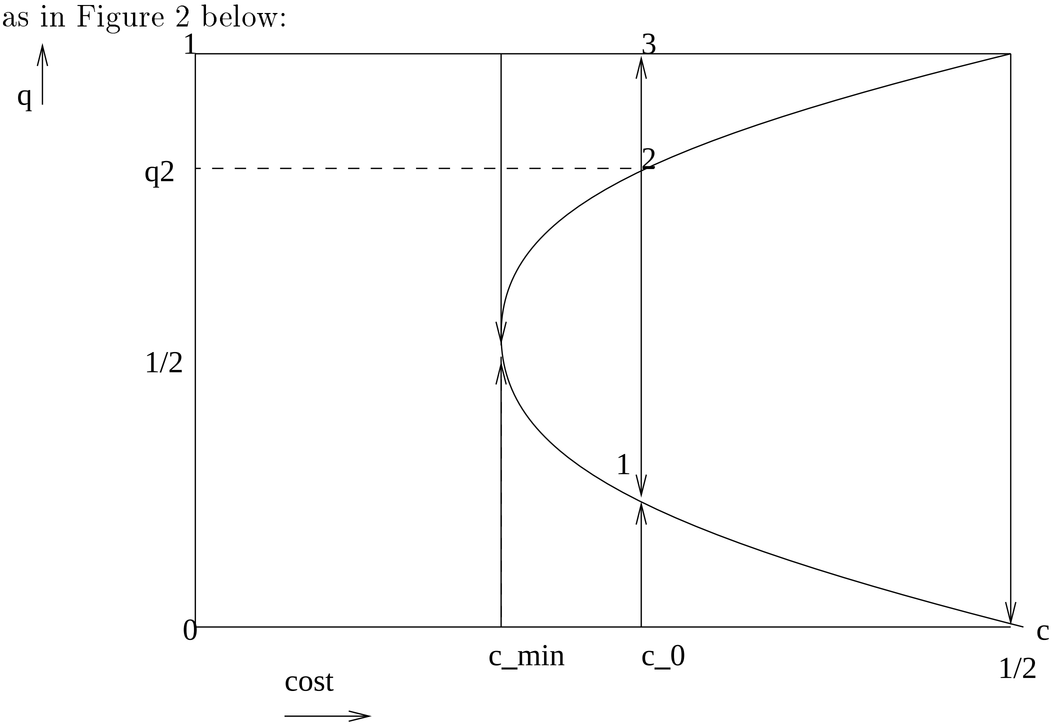


FIGURE 2

Any trajectory starting in the interval $[0, q_2)$ (its basin of attraction) converges to equilibrium 1, therefore it is stable ; equilibrium 2 is unstable, no trajectory leads to it; equilibrium 3 (the pure strategy equilibrium where everyone turns out) is stable with basin of attraction $(q_2, 1]$. Moreover, the basin of attraction of 1 is larger than that of 3 for any cost higher than c_{min}, and as c increases the basin of attraction for equilibrium 3 shrinks.

Considerations about the size of the basin of attraction allow us to improve the predictions of the model: for a stable equilibrium to have a large basin of attraction means having many initial conditions leading to it and therefore means that it has a high probability of being observed. Note also that we do not have to assume that players start the learning process at the same point, provided all the points are in the same basin of attraction.

So the prediction of the model in case of a single election is summarised in the following proposition.

Proposition 2: For any learning dynamics of type (6), any integer M and any cost level c, such that $c_{min}^{M} < c < 1/2$, equilibrium 2 will never be observed, equilibrium 1 will be observed with high probability and equilibrium 3 has a smaller chance to appear. If M is moderately large this probability goes to zero very fast.

The proof of this proposition is quite simple and is in the Appendix.

4.2 Repeated Elections

We mentioned earlier that the low turnout equilibrium has a large basin of attraction. This remark becomes crucial in the following extension of the model. Suppose that elections are repeated regularly: we can index elections with $i = 0, 1, 2, ...$ All other elements of the model remain the same as before. Now, at election i, voters begin the learning process at q_0^i. This depends on the turnout in the preceding election q_∞^{i-1} in the following (non-deterministic) way: q_0^i is a random variable uniformly distributed on the interval $[\underline{q}; \overline{q}]$ with $\underline{q} = max\{0; q_\infty^{i-1} - \delta\}$ and $\overline{q} = min\{1, q_\infty^{i-1} + \delta\}$; δ is a small positive number that gives a measure of the maximum possible mistake in ascertaining the turnout. Thus we get a random dynamical system, a Markov chain, since the system is time independent, whose states are the stable Nash equilibria, 1 and 3. The transition matrix for the Markov chain is denoted by: M and is as follows:

$$\begin{bmatrix} p_{11} & p_{13} \\ p_{31} & p_{33} \end{bmatrix}$$

where p_{ij} = the probability that the state moves from state i to state j, with $i, j \in \{1, 3\}$. Thus M^k is the k step transition matrix. The behaviour is given by proposition 3:

Proposition 3: For any $\delta > 0$, if the number of voters is larger than $N_0(\delta)$, the limit distribution of the outcomes q^i_∞ is concentrated on equilibrium 1, i.e. $lim_{k\to\infty}M^k =$

$$\begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$$

Proof in Appendix.

Note that the precise form of the probability distribution of the q^i_0 is, to a large extent, irrelevant to the result, provided the distribution is concentrated around q^{i-1}_∞. As before we give an informal discussion of the result: using a terminology borrowed from mechanics, equilibrium 3 is called "metastable". Intuitively the fact that the basin of attraction has positive measure but is very small means that if there are random disturbances, the equilibrium will be stable for a while but the probability of observing a jump to equilibrium 1 within time T goes to 1 when T goes to infinity. Equilibrium 1 is "stochastically stable" in the sense of Kandori, Mailath and Rob (1993). Proposition 3 can be generalised in the sense that for other distributions we get this result for any N. This argument carries over to our discussion of the incomplete information case in Section 5. The only difference is that there may be several metastable equilibria instead of one.

So, a consequence of metastability is that, if elections are repeated, equilibrium 3 tends to jump to equilibrium 1 after a long sequence if the cost is below 1 but higher than a certain value $1/2 > c_{min}$. If the cost is not in this range there is only one equilibrium[11].

[11] Equilibrium 3 tends eventually to jump to equilibrium 1 even if the number of voters N is small, the only change is in the average time it may take (e.g. if it is much more likely that people make small mistakes than larger ones, then for small N, if the initial state is 3, it remains 3, until a big enough mistake is made that puts voters in the basin of attraction of 1); such a behaviour is typical of risk dominated equilibria in the sense of Harsanyi and Selten, 1988. See also discussion in the introduction on public goods.

4.3 Comparitive Statics and Hysteresis

Note that, in all cases, equilibria with positive probability predict a non-decreasing q when c decreases, as intuition suggests. It is interesting to investigate further what happens when the cost, or the interest in the outcome, changes from one election to the other. In this case it is natural to ask that the fictitious play dynamics starts at the percentage of voting of the last election. As before, to simplify the discussion and isolate the different effects we will assume that there are no mistakes , i.e. that δ is zero. This allows us to see how q varies as a function of c. Suppose at some time we are given $c = c_0$ and we are in equilibrium 1. Then suppose that the introduction of electronic voting,for instance, causes c to decrease so that the predicted effect is a little increase of turnout (see Figure 2). But when c decreases below c_{min}, the mixed equilibria disappear and we suddenly fall in the basin of attraction of equilibrium 3, so q suddenly jumps up, i.e. at $c = c_{min}$ and below, equilibrium 3 is the only stable equilibrium. The equilibrium will tend to persist for a while even if cost increases again until we reach the point where $c = 1/2$ and the only equilibrium possible is the pure strategy one, where nobody votes; alternatively even if c stays below 1 but over 1/2, in the "very long run" we will see q jumping down, because of the phenomenon of metastability described in the previous section. This is phenomenon is known as hysteresis or "memory" of the system, and explains why the same values of parameters can cause the emergence of different equilibria, depending on the initial state.[12]

Again we refer to the Appendix for rigorous statements and proofs. Thus, we should observe phenomena of this type: in countries where there has been a large turnout in preceding elections one expects large turnout in the next election too even if cost has (moderately) increased or interest for the candidates has diminished. When a critical cost level is reached, or after a sequence of many elections, turnout will suddenly jump down and stay low even when cost decreases back to the original one. See section 4.1 for more

[12] We thank Jonathan Cave for pointing out this interesting feature of the model.

details in the case of incomplete information. Regarding the predictions for changes in population size: since the probability of being pivotal decreases as population size increases (the function $f(q)$ shifts towards the vertical axis so that for any q the probability of being pivotal is lower for a higher N), thus for any fixed level of cost we see lower turnout as population size increases.

5 Incomplete Information

We capture uncertainty by a model of incomplete information about costs, exactly as in the PR (1985) model[13]. Each voter i has a cost of voting c_i which is private information to him. Let the cumulative distribution of costs be denoted $F(c)$ and for simplicity we assume the distribution to be the same between the two groups. We look for the Bayesian Nash equilibria of this game (as in PR (1985)). Each voter then has a decision rule that specifies whether to vote or not as a function of his own cost c_i. It is easy to see that in any symmetric Bayesian equilibrium a voter votes if his cost is below a certain threshold level, c^*, so in this case a learning process will be a dynamic on the c^* to choose. Thus a (symmetric) Bayesian equilibrium is a cost level c^* such that $2c* = f(q(c^*))$, with the corresponding $q^* = F(c*)$. This corresponds to the equilibrium outcome in the game of complete information where all voters have cost c^* and vote with probability $q^* = F(c^*)$. Note that, although it is natural to assume that players choose the cost level c^*, this is equivalent, for symmetric equilibria, to choosing q^*. All dynamics in terms of one variable can be easily translated in dynamics in terms of the other. So let $C(q)$ represent the inverse of $q(c) = F(c)$. Then we need $2C(q^*) = f(q^*)$. In the graph below (Figure 3), this equilibrium is given by the point where the distribution function intersects the curve $f(q)/2$, which shows the probability of being pivotal (as in Figures 1 and 2 above).

[13] PR (1985) also have a section on incomplete information about preferences but this has similar results.

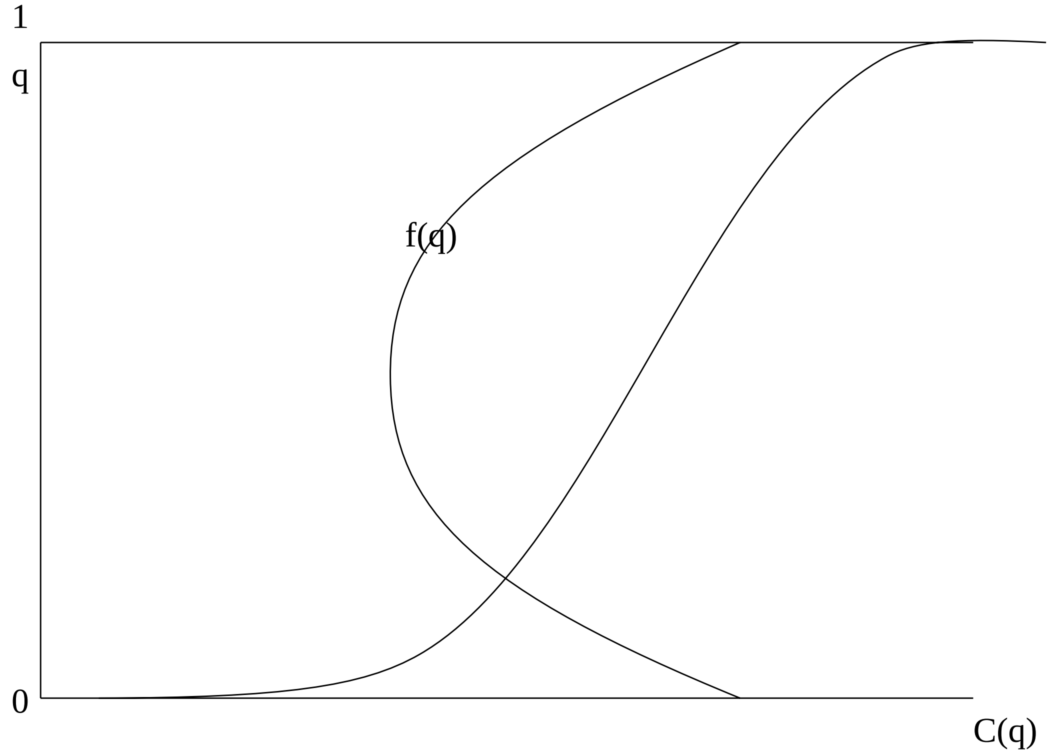


FIGURE 3

PR (1985) use assumptions under which there is only one intersection and the intersection converges to the point $q = 0$ as N becomes large. As is evident from the graph, however, everything depends on the shape of the distribution function $F(c)$ (or its inverse $C(q)$). The assumptions they use are: (1) $F(c)$ is continuous on $(-\infty, \infty)$, (2) $F(0) > 0$ and (3) $F(1) < 1$. The first assumption is rather natural and corresponds to assuming that the probability distribution of c has no atoms. Assumption 2 is quite realistic also, i.e. that there is a positive probability that cost will be negative (civic sense will prompt some people to vote regardless of their assessment about being pivotal.) However, the last assumption is stronger: it implies that there is a positive probability that a voter would not show up even if he were sure to be pivotal. It is also not innocuous and PR (1985) has an example where relaxing this assumption takes us back to the problem of multiple equilibria in the complete information case (see Figure 3). Moreover, there seems to be an implicit assumption that $F(c)$ is not too wiggly: Figure 4 shows a case where the curvature of $F(c)$ can change quite fast so that many additional equilibria are introduced. Note that such a multimodal probability distribution is not so pathological: it could model a population made of different groups each with different costs and with small variance within a group. Thus, incomplete information does not solve the problem of multiple equilibria, sometimes it even introduces more of them, some of which have high turnout and our intuition suggests that they should be "non-robust" in some sense. However we should mention that for sufficiently large N, the PR (1985) conclusion is still valid.

To address this we now consider learning as before: $\frac{dq}{dt} = K(q)$. We can isolate the stable equilibria as being the ones where $C(q)$ intersects $f(q)$ from below. These are the equilibria 1,3 and 5 in Figure 4.

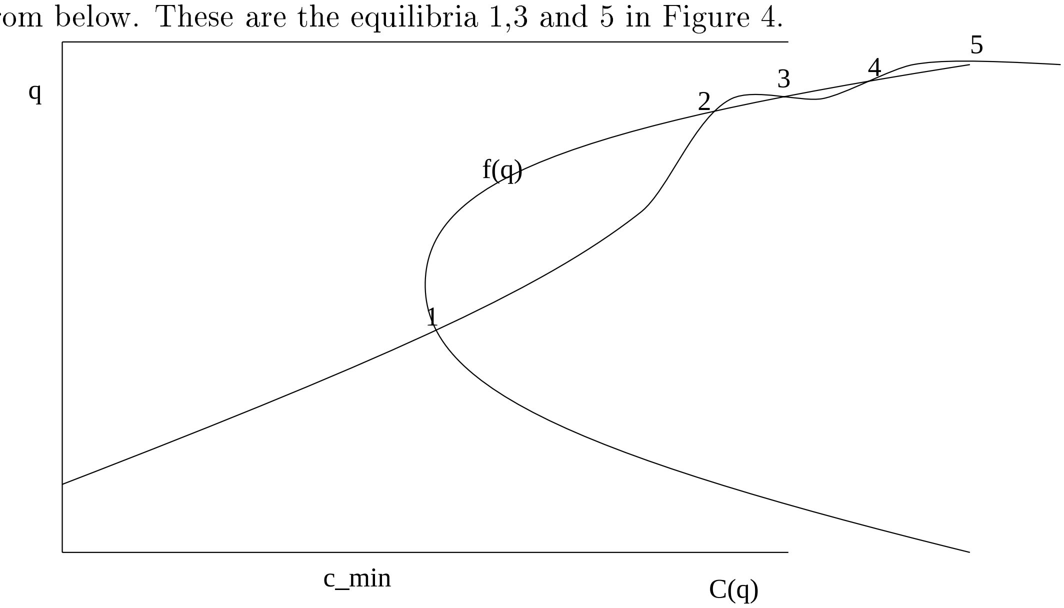


FIGURE 4

More formally we state this in the next Proposition:

Proposition 4: Let the graphs of $C(q)$ and $f(q)$ be in a generic position (this means that they intersect transversely i.e. $2C(q) - f(q) = 0 \Rightarrow K'(q) \neq 0$), *then the asymptotically stable points are those such that* $d/dq(2C(q)-f(q)) \geq 0$.

The proof is the same as for Proposition 1.

In the same way as in proposition 2, it is not hard to see that if N is large the only equilibrium with a large basin of attraction, containing at least the interval $[0, 1/2]$, is 1 – the low turnout equilibrium. All the others are either unstable, i.e. with zero probability, or metastable, i.e. with a small basin of attraction, whose size goes to zero when N goes to infinity. In the next section we do a similar analysis of the behaviour of equilibria when the parameters of the model are allowed to vary. We also illustrate with an example that the conditions under which there is a unique equilibrium with incomplete information may be very restrictive.

5.1 An Example

We now investigate how equilibria change when the distribution of c varies in a simple example. We shall assume that the c are uniformly distributed on the interval $[\overline{c}-s; \overline{c}+s]$, so that we have two parameters the average $\overline{c}$ and a measure of the dispersion s. Other distributions, such as the Gaussian, can be discussed in the same way and give qualitatively similar results (see the discussion at the end of the section). A uniform distribution corresponds to an F whose graph is shaped as in Figure 5.

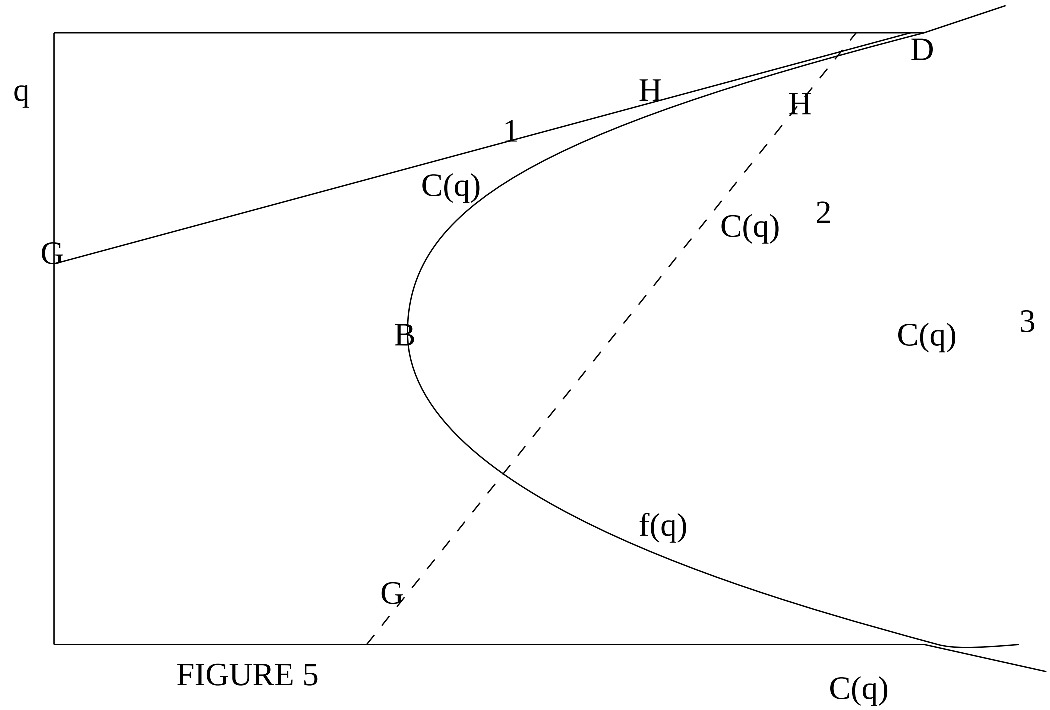


FIGURE 5

The intersections with the curve studied before give the equilibria. It should be geometrically intuitive, and it is easily proved, that if s is large enough so that the slope of the line GH is less than the slope of the arc BD at D, there is only one equilibrium, (see Figure 5). More precisely, we have the following proposition:

Proposition 5: If $s \geq (N-1)/2$, then there exists a unique equilibrium for any value of the average cost $\overline{c}$. In this range of s the unique equilibrium corresponds to nobody voting if $\overline{c} - s > 1/2$, everybody voting if $\overline{c} + s < 1/2$ and a percentage of voters that is a smoothly decreasing function of $\overline{c}$ in the cases in between.

The proof is obvious.

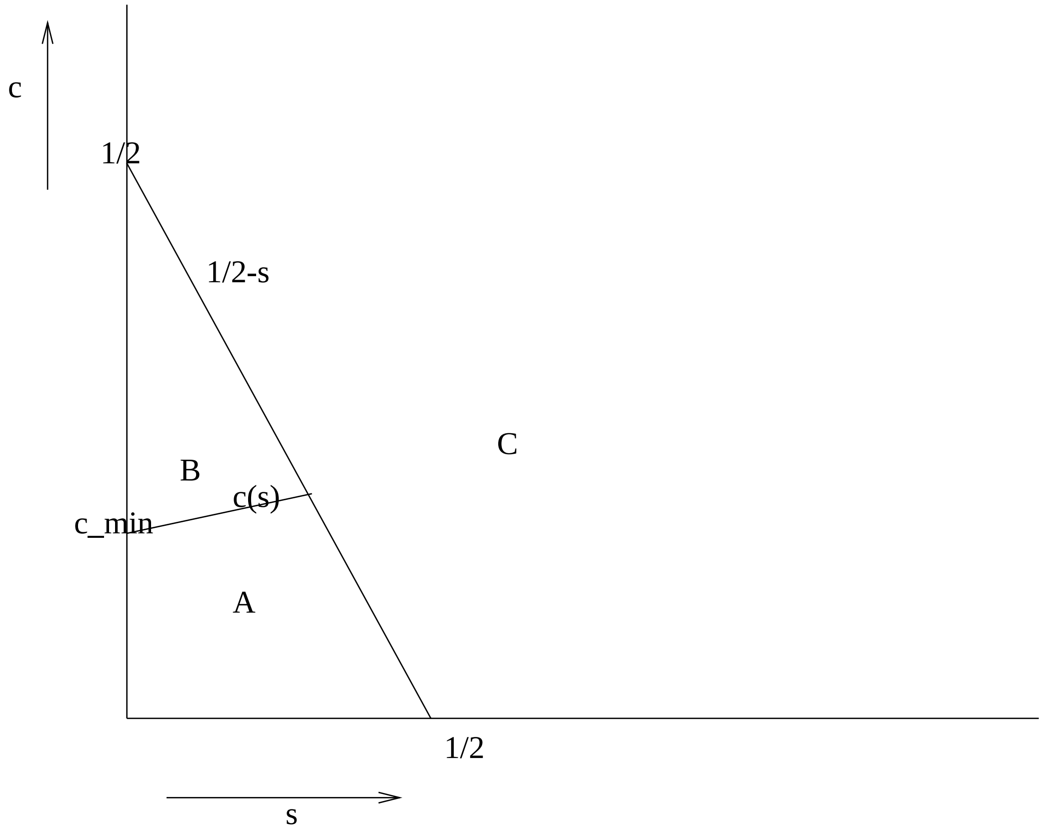
This corresponds to the case studied by PR (1985). Note that this is a very large value for s, for plausible values of N. The more realistic case of small (or rather not enormous), s , i.e. $s < (N-1)/2$, is more interesting and presents several analogies with the case of complete information, that corresponds to $s = 0$. For any s, let $c(s)$ be the value of $\overline{c}$ such that the line

GH is tangent to the curve BD. Note that $c(s)$ is equal to c_{min} for $s = 0$ and increases monotonically to $1/2$ when $s = (N-1)/2$. It is also clear from the picture that $\overline{c} < 1/2 - s$. We now have, as in the complete information case, three cases:

Proposition 6: If $\overline{c} < c(s)$ there is one equilibrium with $q = 1$ (everybody votes) (shown by area A in Figure 6). If $c(s) < \overline{c} < 1/2 - s$ there are three equilibria: $q = 1$ which is stable with a small basin of attraction (which decreases with c(s)), an unstable one with large q and a stable one with a low q and a large basin of attraction. This corresponds to area B in Figure 6. which makes the $q = 1$. If $\overline{c} + s > 1/2$, there is only one equilibrium with $q < 1$, and q smoothly decreasing with $\overline{c}$ (shown by area C in figure 6).

If the average cost, $\overline{c}$ is fixed, and s, the measure of uncertainty, increases, it is easy to see from Figure 5 that the stable equilibrium has increasing turnout. This gives a very interesting insight into the relation between uncertainty and turnout, which has not been investigated in previous literature. Indeed, PR (1985) implicitly suggest that adding uncertainty should lower the turnout.

The domain in the $\overline{c}, s$ plane corresponding to the three cases is shown in Figure 6.



Most of our discussion applies to other distributions as well provided they are regular enough. Note however that the tangency between the F curve and the q curve may not be unique in particular cases, such as costs concentrated around a finite number of values. This would make the jump in the hysteresis cycle split into the composition of several smaller ones.

Finally we remark that the analysis of repeated elections is the same as in Section 4, all results carry over.

6 What can be explained?

The honest answer to this question would be: not everything. However, the qualitative predictions of our model accord well with the facts and provide reasonable interpretations of them. This is rather surprising given that the

PR model is an oversimplified representation of real elections.

As mentioned before, one of our first comparitive static results is that turnout is decreasing with costs. Moreover, as we discussed in Section 5 we are able to predict a positive correlation between turnout and uncertainty, and this effect can be tested empirically.

Our model, and in particular hysteresis, can be used to see what happens when the cost c is changed by the introduction or removal of voting laws. For example: abstention is high in the U.S.A. and has been significantly lower in countries such as Belgium and Italy, even though there is no reason to expect big differences in cost of voting or interest in the elections. An explanation of this fact might be as follows: in the past Belgium and Italy had laws against abstention that made c quite low and possibly negative, so equilibria were high turnout equilibria. The abolition, or lack of enforcement, of such laws has moved the state to the segment of higher cost but with persistence of large q; in U.S.A. where there have never been such laws the more stable low turnout equilibria are observed.

A stylised fact about elections in Western democracies is that turnout rates have been falling steadily over time (http://www.idea.int). There are two possible explanations: one is the convergence of party platforms that leads to increasing apathy on the part of voters. An alternative explanation is that voters learn that there is another more efficient equilibrium with lower turnout. In our model the first explanation would correspond to an increase in the cost of voting (recall that this is a ratio of actual costs to benefits), while the second would correspond to a jump from equilibrium 3 to equilibrium 1 described in our section on repeated elections. It is possible to distinguish between the two explanations, as our comparitive statics predict a gradual decline in the first case but a sudden jump in the second case. Since the decline is indeed gradual with small oscillations, we may deduce that the first explanation is the correct one. The small oscillations would correspond to small changes in cost of voting and its distribution.

7 Instrumental or expessive voting?

As we remarked in the introduction, an important question in voting theory is what makes people vote. This question can actually be split in two: Is there a rational or rather consistent reason that makes them vote and if so, what is it?

More formally the question is whether the sign of the expression (3) triggers the decision to vote and, if it does, what are the relevant terms in it. The instrumental hypothesis claims that voters expect their vote to have an appreciable effect on the outcome of the election i.e. the term pB is relevant, while the expressive hypothesis claims that voters vote because they enjoy it, i.e. what matters is $D - C$. To defend the expressive theory it is usually claimed that $pB << C$ because p is negligible so that almost nobody would vote unless D were comparable and significantly larger than C.

Getting a reliable measure of p from statistics is problematic: basing themselves on official reports Mulligan and Hunter (2000) compute that one out of 100,000 votes matters, however they note that margin specific election procedures may have significantly lowered the observed p in comparison to the p voters use to make their choice. Decision theoretic models assume instrumental voters to compute p but then yield a p that is far too low to justify the assumption. Moreover p depends strongly on the method used. For instance, Beck (1975) has a model in which the probability r with which a candidate is voted is exogenous. Unless r is **exactly** equal to $1/2$ this gives a p that is exponentially decreasing with twice the size of the population: if we take $r = 0.51$ and a population of just million voters p becomes less than e^{-200} i.e. completely negligible. Good and Mayer (1975), have a more refined model in which r is a random variable extimated via an opinion poll on a small sample of the population, this gives a much larger p, roughly of the order of $1/N$, that is nevertheless too small to give an appreciable incentive to vote,since pB would be less than a cent.

We claim that, using the PR model, one can get at least orders of magnitude of the quantities pB and c that correspond to observed behaviour and support the instrumental voter hypothesis. We first begin by estimating B: in most electoral campaigns candidates promise advantages in terms of tax cuts, subsidies, cash tranfers from one group to another etc that correspond to an increase of income for their supporters amounting to several thousands euro (or dollars) per year. A legislation usually lasts 4 or 5 years. A conservative estimate (discounting some suspicion about electoral promises) gives more than 10,000 euro overall. In the PR model with uncertainty, observe that $p = F(c) \geq F(c_{min})$, so we begin by estimating c_{min}. In PR (1983) it is proved that the value of c_{min} in figure 1 is $2/\sqrt{\pi}(2N-1)$ and, with incomplete information, the model predicts that at least everybody whose cost is below this figure will go to vote. Estimating $2N$ is easy, voting populations are of the order of about 40 millions for large European countries and 200 millions for USA (http://www.IDEA). By substitution in the formula one obtains that c_{min} is about 1 euro, and maybe more for small countries. It is reasonable to believe that a non trivial part of the population (i.e. we can look at plausible values of $F(c_{min})$) will bother to cash a check of that amount e.g. on their way to shopping. So it seems that we don't have to postulate a significant D to make sense of appreciable turnout.

Our argument can be made stronger: it is not clear what should be the right N to take because people may get utility from their candidate winning just in their province or state or even their electoral college, i.e in much smaller units than the whole country. Another reason that may alter the size of N as the players perceive it is given by the tendency people have of thinking of the electorate as being composed of groups of individuals of the same size (e.g. women voters, ethnic minorities etc) whose electoral behaviour coincide, so that in this case one should think of N as the number of these types[14]. Moreover, thinking of groups rather than individuals would imply that the relevant variable is not the probability of a single voter being

[14]This type of reasoning (if I chose strategy X, it means that all people in my group or my co-player will probably do the same) is based on a wrong application of Bayes rule but seems to be common in human behaviour, it probably lies at the root of irrational behaviours observed in Newcomb's paradox or the prisoner's dilemma.

pivotal but that of the margin of votes being of the size of the group. In the latter case it would be the expected closeness of the election that mattered rather than the probability of being pivotal – a much easier interpretation.

It should also be remarked , as as often been done in the literature, that there is a tendency to overestimate very small probabilities. This implies that the functions $K_i(q)$ that voters use in their learning dynamics are based on $f_i(q, M)$ that are larger than the actual $f(q, M)$. This too contributes to make the perceived pB relevant.

Let us note, however that the weakness of the PR model is in the outcomes they predicts: a simple estimation of the binomial distribution in the case of type symmetric equilibria with a cost of voting of the order of some euro it gives a 50-50 outcome with a standard deviation of the order of less than $1/\sqrt{N}$, that means less than 0.1 percent! It can be proved that this holds also for unequal populations: i.e. for N1 dif from N2. In that case the only equilibria giving outcomes with an appreciable spread must be sustained by very low cost of voting. So, in our opinion, the model seems to give a consistent picture of the phenomenon of abstension while it fails to capture the behaviour of voters once they decide to go to the polls, in any case this is not what the model was devised for.

8 Conclusions

In this paper we showed how considerations based on learning dynamics and stability can select the intuitive equilibria in the PR (1985) model without having to resort to ad hoc arguments. In this way it is also possible to give the model some predictive value as we saw in the last section.

Clearly this stylised model misses many features of real elections: usually there are more parties, voters are not homogenous – they have different costs and benefits, and there exist many other ways to influence policy than voting, such as lobby groups and other forms of direct action. Moreover parameters such as the cost of voting or the measure of the interest in a candidate are not directly measurable with reasonable confidence, so that it seems hard to go beyond the rough estimates we made in the last section.

It should be obvious that in such cases asking for precise quantitative predictions is difficult – there are too many factors that can affect the predicted turnout that are hard to quantify exactly. So what can we hope to achieve from such simple models? The essence of the PR (1983,1985) models is that voters are strategic – the probability of being pivotal matters to them, as do costs and benefits of voting. Thus the power of the model comes from predictions about comparative statics and other qualitative predictions. Since our model gives some sharp qualitative predictions (jumps , hysteresis, long time drifting away from metastable equilibria) that are very robust with respect to the parameters involved, it makes the PR model more apt to be tested in this way.

One objection to this approach that we anticipate is that the model is not really suited to large populations because for large populations the probability of being pivotal is very small, and hence the game theoretic model is more suited to smaller groups e.g. committees. A partial answer has been provided above, anyway our model is quite applicable to voting in committees as well, except for the assumption that voters are myopic in their learning behaviour. We plan to extend this model allowing for more sophisticated learning on the part of voters.

The basic model of PR (1983) did not allow for uncertainty about candidates on the part of voters. Feddersen and Pesendorfer (1996) show that if there are common values among a group of informed and uninformed voters about the candidate (in a two candidate setting as in PR (1983)), then uninformed voters might prefer to abstain even when voting is costless, to ensure that informed voters decide the election. This can happen even with partisans in the population. It seems that many of these equilibria suffer from a similar problem in the sense that they require extreme co-ordination on the part of uninformed voters. It would be worthwhile to check if these equilibria satisfy our requirements of stability.

References

1. Bagnoli and Lipman, 1989, " Provision of Public Goods : Fully implementing the Core through Private Contributions", Review of Economic Studies, Vol 56, pp. 583-601.

2. Beck, N. "A Note on the Probability of a Tied Election", Public Choice 23, Fall 1975, pp. 75–80.

3. Brown, G.W., 1951, "Iterative solutions of games by fictitous play", in "Activity analysis of production and allocation", ed. T.C. Koopmans, New York: Wiley.

4. Conley, J., Toossi, A. and M.Wooders, (2001), "Evolution and Voting: how nature makes us public spirited", Warwick Economic Research Papers, No.601.

5. Cooper, R.W., D. DeJong, R. Forsythe, T.W.Ross,1990," Selection criteria in coordination games: some experimental results", American Economic Review ,March 1990, 80, pp.218-233.

6. Demichelis, S. and F.Germano, 2000, "On knots and dynamics in games", CORE Discussion Paper 00/10.

7. Demichelis, S. and Klaus Ritzberger, 2000, "From evolutionary to Strategic Stability", CORE Discussion Paper 00/59.

8. Dhillon, A. and S.Peralta, (2002) "Economic Theories of Voter Turnout", The Economic Journal, Vol.112 No.480, pp.F332–F352.

9. Downs, A., 1957, "An Economic Theory of Democracy", Harper, New York.

10. Feddersen, T. J. and W. Pesendorfer, 1996, "The swing voter's curse", The American Economic Review 86, pp. 408-424.

11. Feddersen, T. J. and W. Pesendorfer, 1999, “Abstention in Elections with asymmetric information and diverse preferences”, American Political Science Review, vol.93(2), June, pp.381–98.

12. Ferejohn, A. F. and M. P. Fiorina, 1974, ”The paradox of not voting: a decision theoretic analysis”, American Political Science Review 68, pp 525–536.

13. Fey, M. (1997), “Stability and Coordination in Duverger’s Law: A Formal Model of Preelection Polls and Strategic Voting”, American Political Science Review.

14. Fischer, A.J. 1999, “The Probability of being Decisive”, Public Choice 101, pp. 267-283.

15. Fudenberg, D. and D. Kreps, 1993, “Learning Mixed Equilibria”, Games and Economic Behaviour,5, pp.320-367 .

16. Fudenberg, D and David K.Levine, (1998) “The theory of learning in Games”, MIT Press.

17. Good, I.J. and Lawrence S.Mayer,1975 “Estimating the Efficacy of a Vote”, Behavioural Science, 20(1), January, pp.25-33.

18. Harsanyi, J. and R. Selten, 1988, “A General Theory of Equilibrium Selection in Games”, MIT Press.

19. Kandori, M., G.J.Mailath, and R.Rob (1993), Learning, Mutations and Long Run Equilibria in Games, Econometrica, Vol.61, No.1, pp.29–56.

20. Ledyard, J. O., 1981, ”The paradox of not voting and candidate competition: a general equilibrium ananysis”, in G. Horwich and J. Quirck (eds.), *Essays in contemporary fields of economics*, Purdue.

21. Mueller, Dennis C., 1989, “ Public Choice II”, CUP.

22. Mulligan, Casey B. and C.Hunter, 2000, "The Empirical Frequency of a Pivotal Vote", University of Chicago Discussion Paper 00-25.

23. Palfrey,T.R. and H. Rosenthal, (1983), A strategic calculus of voting, *Public Choice*, 41, pp.7–53.

24. Palfrey,T.R. and H. Rosenthal, (1985), Voter Participation and Strategic Uncertainty, *The American Political Science Review*, Vol.79, pp.62–78.

25. Shachar, R. and B. Nalebuff (1999), "Follow the Leader: Theory and Evidence on Political Participation", American Economic Review, Vol.89, No.3, pp.525–549.

26. Schuessler, A.A. (2000), *A Logic of Expressive Choice*, Princeton, Princeton University Press.

27. Sieg, G. and C.Schulz, (1995), "Evolutionary Dynamics in the voting game", Public Choice, vol.85, pp. 157–72.

28. van Huyck,R.B, R.C. Battalio, R.O.Beil, 1990, " Tacit Coordination games, Strategic uncertainty and Coordination failure", American Economic Review, March 1990, 80, pp. 234-248

29. Weibull, J.(1995), "Evolutionary Game Theory", MIT Press.

Appendix

Proof of Proposition 1: Let q_1 and q_2 represent the equilibria 1 and 2 respectively. We show that any trajectory beginning in the interval $[0, q_2]$ converges to q_1, thus $[0, q_2)$ is the required neighbourhood B^* for equilibrium 1, i.e. q_1, and any trajectory beginning in $(q_2, 1]$ converges to equilibrium 3 , i.e. to $q = 1$, hence the required neighbourhood for equilibrium 3 is $(q_2, 1]$.

Consider first a path starting in the interval $[0, q_2)$, i.e. $q(q_0, t) \in [0, q_2)$. By equation (6), $q(., t)$ is an increasing continuous function of t in this part of the domain, and remains so upto q_1. Hence $q(.)$ must converge to q_1. The other direction is the same.

Proof of Proposition 2: The basin of attraction for equilibrium $q_L(c)$ is given by $[0, q_H(c))$ and the basin of attraction for the equilibrium 3 for fixed c is given by $(q_H(c), 1]$. Thus if $c < 1/2$, $q_H(c) > 1/2$, hence the basin of attraction of equilibrium $q_L(c)$is bigger than $1/2$, while that of equilibrium 3 is smaller than $1/2$. As M increases moreover, by Conjecture 1 above, we have the basin of attraction of $q_L(c)$ tending to 1, and that of equilibrium 3 tending to 0 for any fixed cost level c.

Proof of Proposition 3: Let p_{ij} denote the probability of starting in state i and reaching state j. The transition matrix of the Markov Chain is given by:

$$\begin{bmatrix} p_{11} & p_{13} \\ p_{31} & p_{33} \end{bmatrix}$$

Since any point in $[0, q_H)$ converges to equilibrium 1 and any point in $(q_H, 1]$ converges to equilibrium 3, p_{11} is the probability that $q_0^i \in [0, q_H)$, the basin of attraction for equilibrium 1, if $q_\infty^{i-1} = 1$. p_{13} is the probability that $q_0^i \in (q_H, 1]$, the basin of attraction of equilibrium 3 if $q_\infty^{i-1} = 1$. p_{31} is the probability that $q_0^i \in [0, q_H)$, the basin of attraction for equilibrium 1, when $q_\infty^{i-1} = 3$ and p_{33} is the probability that $q_0^i \in (q_H, 1]$, the basin of attraction for equilibrium 3, when $q_\infty^{i-1} = 3$.

Note that when $N \to \infty$, $q_H \to 1$. Suppose that δ is small enough that $q_L + \delta \in [0, q_H)$, if $q_\infty^{i-1} = 1$, the probability that it stays in state 1 is 1 (recall

q_0^i is uniformly distributed in the interval $[\underline{q}, q_1+\delta]$ where $\underline{q} = min(0, q_1-\delta])$; thus $p_{11}=1$, and $p_{13}=0$. However if the system starts from state 3, it may stay in state 3, if $1-\delta < q_H$. However, by Conjecture 1 (PR, 1985), there is an $N_0(\delta)$ large enough (hence $q_H = q_0(\delta)$ large enough) that $1-\delta > q_0(\delta)$. Thus the probability $p_{31} = F(q_0(\delta) = \frac{q_0(\delta)-(1-\delta)}{\delta}$, which is strictly positive. Of course $p_{33} > 0$.

Suppose now that δ is large so that $p_{11} < 1$. Then, by the above reasoning $p_{31} > 0, p_{33} > 0$, and there is an $N_0(\delta)$ large enough that $p_{11} = 1, p_{13} = 0$. Let $p_{33} = \beta > 0$. Then M^k has the following form:

$$\begin{bmatrix} 1 & 0 \\ 1-\beta^k & \beta^k \end{bmatrix}$$

This clearly converges to

$$\begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$$

This conclusion is true as well with a more general distribution of q_0 as long as it is absolutely continuous with respect to the Lebesgue measure. In both cases it is easy to see that the invariant measure is concentration on state 1 in the first case or convergence to a measure on 1 in the general case.

To see this, let α_0^1 denote the probability that the system begins in state 1 (this is given by the basin of attraction of state 1), and α_0^3 is the corresponding probability that the system begins in state 3. Then, the probability that the state is in state 1 after k steps is given by $\alpha_0^1 p_{11}^k + \alpha_0^3 p_{31}^k$, and the probability that the state is in state 3 after k steps is given by $\alpha_0^1 p_{13}^k + \alpha_0^3 p_{33}^k$, Note that $\alpha_0^3 = 1 - \alpha_0^1 < 1/2$. Thus, $\frac{\alpha_0^1}{\alpha_0^3} > 1$. Consider the limit of the ratio $\frac{p_3^k}{p_1^k}$ as $k \to \infty$. If $p_{11} > p_{13}$ then this ratio tends to 0 as $k \to \infty$.