LEARNING IN ELECTIONS AND VOTER TURNOUT EQUILIBRIA

Stefano DeMichelis

And

Amrita Dhillon

No 608

WARWICK ECONOMIC RESEARCH PAPERS

DEPARTMENT OF ECONOMICS

Learning in Elections and Voter Turnout Equilibria

Stefano DeMichelis*and Amrita Dhillon†

September 2001: PRELIMINARY AND INCOMPLETE.

Abstract

Both complete and incomplete game Theoretic Models of Voter Turnout (Palfrey and Rosenthal, 1983,1985) have the problem of multiple equilibria, some of which seem unreasonable. How can the counter intuitive high turnout equilibria be explained? Palfrey and Rosenthal (1985) suggest that the main reason is that strategic uncertainty is too low in a complete information model. We show that this is not the main problem with these equilibria– incomplete information may exacerbate the problem of multiple equilibria. We propose a very intuitive criterion based on voter learning to distinguish reasonable equilibria. This paper makes precise the sense in which the high turnout equilibria in the Palfrey-Rosenthal model are not robust. We show how the model can be used to qualitatively explain several phenomena observed in reality.
JEL Classification: C72, D72
Keywords: Voter Participation, Voter Learning, Stable Equilibrium, Asymptotically Stable Equilibrium, Markov chain, long run equilibria

*CORE, 34 Voie du Roman Pays, 1348 Louvain la Neuve, Belgium, email: Stefano@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

1 Introduction

Downs (1957) pointed out a basic paradox in voting theory: the fact that large numbers of people vote is contradictory to the idea of a rational voter who takes the costs of voting into account. Rational voters must realise that the probability of their being pivotal is very small in a large electorate (hence the benefits of instrumental voting are very small) so that they would be better off abstaining. Is this borne out by a game theoretic model with rational voters? Palfrey and Rosenthal (1983,1985) examined this issue in two models, 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 in the simplest case with an equal number of the two types of voters in the population and restricting attention to symmetric mixed strategy equilibria they found two types of mixed strategy equilibria for costs in the interval $(0, 1/2)$ – one with low turnout (as expected) but one with substantially high turnout. They claimed that this high turnout equilibrium was very unsatisfactory because it "has the unappealing feature that there is another equilibrium with almost no one voting. Apparently the only reason the upper one can be sustained is that the two electorates are of the same size so that for q 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." (Palfrey and Rosenthal, 1985) Of course, we do not see any reason why this argument should apply only to the high turnout equilibrium.[1].

They claim 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 Palfrey and Rosenthal (1985) that the low turnout equilibrium is the more appealing one, we argue that the reason they cite may not be the most compelling one. There is a sense in which the high turnout equilibria are not robust – they require precise beliefs about what other voters are doing in equilibrium. This paper attempts to make this claim more precise. We show that removing some of the assumptions on voter uncertainty of the Palfrey

[1] When analysing the general game where the size of the two electorates is different they still get some "quasi-symmetric mixed-pure strategy equilibria" which have high turnout even as the size of the electorate increases. But all symmetric totally mixed strategy equilibria have the property that as the size of the electorate increases the turnout decreases.

Rosenthal (1985) incomplete information model leads to multiple equilibria even in the incomplete information case. We show that introducing voter learning about equilibrium (i.e. considering the dynamics of reaching a Nash equilibrium by boundedly rational agents) in a very simple and natural way would lead to the low turnout equilibria in their example. It would rule out the high turnout mixed strategy equilibria. The argument for this example is close to the risk dominance arguments of Harsanyi and Selten (1988).
We introduce the model of voter learning in Section 2, then we consider the complete information game in Section 3 and the incomplete information game in Section 4. Section 5 concludes.

2 The Model

We use the model due to Palfrey and Rosenthal (1983, 1985) (henceforth PR). There is a total of N voters in the population, and there are 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 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 $= c$. Voters have full information. Voters thus decide whether to vote (participate) or not – if they vote they always vote sincerely (i.e. for their best candidate).
Let R represent the expected net benefit from voting , p the probability of being pivotal, C the (same across voters) 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).
Then we have: $R = pB - C + D$. Let $c = C - D$, the fixed net cost of voting. We can normalise so that only the ratio of cost per unit benefit matters: $\frac{R}{B} = p - \frac{c}{B}$.W.l.o.g let $B = 1$.
We consider Nash equilibria of this game.
Let the probability for player i to choose to vote be q_i. Then $(q_i^*, q_2^*, ..., q_n^*)$ 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: let n_1^i and n_2^i denote the number of votes received in equilibrium by alternative 1 and 2 if player i is not included. So, we have that $0 < q_i^* < 1$ iff Prob(i is pivotal)$(1 - c_i)+$ $(1-$Prob(i is pivotal)$)(0 - c_i) = 0$. Also i is pivotal whenever $|n_1^i - n_2^i| \leq 1$. This gives the following equation if the tie-breaking rule is a coin

toss: $c =$ [Prob $(n_1^i = n_2^i - 1)$+Prob$(n_1^i = n_2^i)$](1/2). As Palfrey and Rosenthal (1983) show, there are many Nash equilibria to this game: they can be divided into two categories: the first where all voter's probabilities of voting are strictly between 0 and 1 is called a Totally Mixed Strategy Equilibrium (TMSE). The equilibria in this category have the property that as electorates become large, the probability of voting becomes smaller. The other category has all voters in one group using a mixed strategy while in the other group, voters are divided into two subgroups, one in which voters definitely abstain and the other in which voters definitely vote. This does not have the property that turnout becomes smaller as the size of the electorate becomes larger. The problem with this equilibrium, according to PR, is that it is very fragile and requires very precise beliefs about the number of votes each alternative will receive. If there is too much uncertainty in equilibrium then the probability that the election is close approaches zero as the electorate size increases.

2.1 An Example with $N_1 = N_2$:

Let us demonstrate first the problem that arises with Nash equilibria in a simple example where costs are equal for all voters and $N_1 = N_2$. Assume that we have the coin toss tie breaking rule. Let $N_1 = N_2$. The symmetric mixed strategy equilibrium is denoted q which satisfies the following equation:

$$\begin{aligned} 2c &= \sum_{k=0,\ldots,N_1-1} C_{N_1-1,k} C_{N_1,k} q^{2k} (1-q)^{2N_1-2k-1} \\ &+ \sum_{k=0,\ldots,N_1-1} C_{N_1-1,k} C_{N_1,k+1} q^{2k+1} (1-q)^{2N_1-2k-2} \end{aligned} \quad (1)$$

If $0 < c < 1/2$ then this equation has either no solution, 1 solution or two. We can plot a graph (as in PR, 1985) to see the equilibria and how they change as N becomes large:

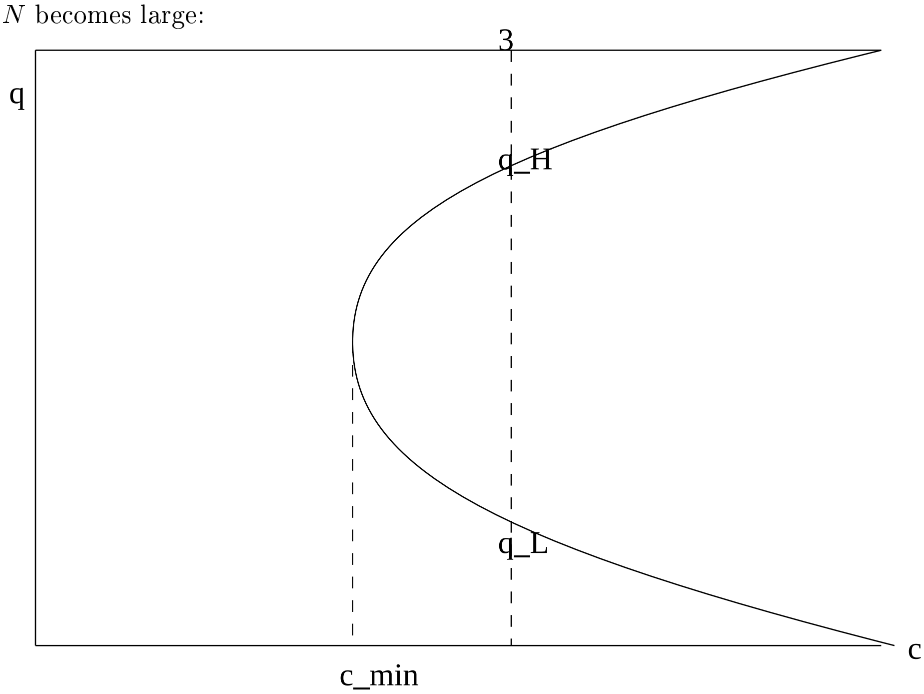


FIGURE 1

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! Thus the complete information model can generate a symmetric mixed strategy high turnout equilibrium. This equilibrium looks implausible for several reasons: among other things it would predict that abstention decreases when the cost of voting increases,which is rather counterintuitive behaviour.

As we mention above (see Introduction) PR, 1985, claim 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 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. Palfrey and Rosenthal obtain uniqueness in the incomplete information game *only when uncertainty on the cost is high enough.*

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 such as fictitious play. We suggest, in this paper an alternative way to select the voting equilibria that gives the "good" prediction even in the case of an equal number of voters in the two electorate. 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.

3 Dynamics: Single Elections

3.1 Complete Information

Consider the PR example above. Let $N_1 = N_2 = N$ and as before let q denote the probability that a voter participates in the election and c the cost of voting. Let $f(q)$ denote the probability of being pivotal (computed above in (2) with the coin tossing tie breaking rule). Then, as before the best reply is to vote if $f(q) - 2c > 0$ and abstain in the opposite case with equality corresponding to indifference between the two strategies. We assume that people learn the equilibrium through a process of "continuous fictitious play". A voter starts by conjecturing a q, e.g. the share of the population who voted on the last election or the result given by a poll (we assume N is not too small so that the empirical observation of this quantities is a reasonable estimator of q) and checks whether the inequality before 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. We assume that everyone begins with the same q and everyone adjusts in the same way. This is consistent with Palfrey-Rosenthals' focus on symmetric equilibria. If this process converges to some point q^*, she will apply the corresponding mixed strategy.

With a large population the influence of each voter is very small so that players would play myopically. The type of learning by adjustments is the standard one assumed in several contexts such as e.g. neural networks. The details can be specified in several different ways, for example see Fudenberg and Levine (98). In general they can all be described by the differential equation[2]:

$$\frac{dq}{dt} = K(q, c) \tag{2}$$

and sign $K(q, c)$ =sign$(f(q) - 2c)$ This is the *Monotonicity property* assumed in Kandori-Mailath-Rob (93)(henceforth KMR).

The functional form of $K(q)$ depends on the particular model of learning. However our result holds for all of them that satisfy the assumptions on $K(q)$. We assume that the function $\frac{dq}{dt}(.)$ satisfies continuity at any initial point t_0, q_0, hence 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. Now, we examine the the behaviour of the dynamics on the strategy space. The basic intuition is that points which are the limit of $q(t)$ as t goes to infinity should be the outcomes of the learning process that are likely to be seen in concrete cases. To be more precise we will introduce some definitions taken mostly from Weibull (1995): To begin with let $q(q_0, t)$ be the solution of the differential equation (2) with initial condition $q(0) = q_0$, we : 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 neighbourhood B of q^* contains a neighbourhood 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

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

exists a neighbourhood B^* *such that the following holds for all* $q_0 \in B^* \cap X$:

$$lim_{t\to\infty} q(q_0, t) = q^* \tag{3}$$

A point that is not aymptotically 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^*.

Recall that equilibrium 1 in Figure 1 (also Figure 2 below) is the low turnout mixed strategy equilibrium,q_L^*, 2 is the high turnout mixed strategy equilibrium, q_H^*, and 3 is the pure strategy full turnout equilibrium. Now we can state Proposition 1:

Proposition 1: For any learning dynamics of type (2) 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 equilbria respectively. When $c_{min} < c < 1$ the qualitative behaviour of the dynamic is as in Figure 2 below:

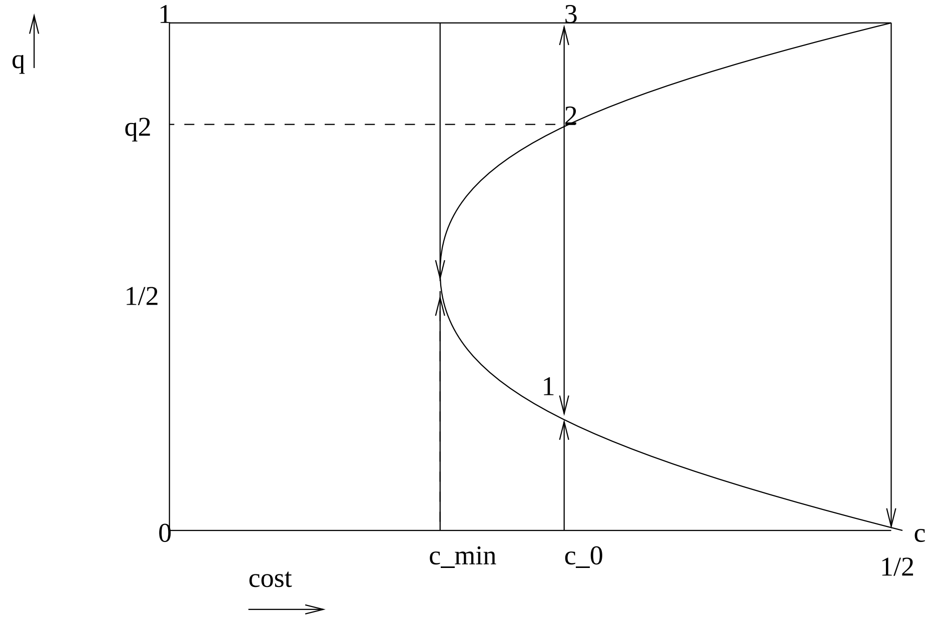


FIGURE 2

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 so has a high probability of being observed.

So the prediction of the model in case of a single election is that equilibrium 2 will never be observed, equilibrium 1 will be observed with high probability and equilibrium 3 has a smaller chance to appear (if N is moderately large this probability goes to zero very fast).

The case where $N_1 = N_2$ is easy to analyse since it involves one-dimensional dynamics. We could extend the analysis to look at both the asymmetric equilibria of the game with $N_1 = N_2$ and also to consider equilibria of the general game with $N_1 \neq N_2$. We have the following examples, in this regard. The learning model is now generalised to: $\frac{dq_i}{dt} = K_i(q_1, q_2, c)$ for $i = 1, 2$, where sign $K_i(q_1, q_2, c)$ = sign $(f_i(q_1, q_2) - 2c)$. $K(\cdot)$ is a function with continuous partial derivatives. This is a non-linear two dimensional system, but we can use the Linearisation theorem. q_i can denote the (different) strategies of two players who are identical orthe symmetric strategies of two players who belong to different groups (when $N_1 \neq N_2$). PR point out that the asymmetric strategy equilibria when $N_1 = N_2$ disappear in large electorates. In our example, we see that the dynamics forsmall N show cycles, i.e they are stable but not asymptotically stable.

Lets look at asymmetric equilibria of the game with $N_1 = N_2 = 2$. We take an example from PR (83). $N_1 = N_2 = 3$ and cost $= \frac{20}{81} < c_{min}$. The equilbria: $q_1, q_2 = (2/3, 1/3)$ and $q_1, q_2 = (2/3, 1/3)$.

The equations for these are:

$$2c = f_1(q_1, q_2) = (1-q_1)^2(1-q_2)^3 + 3(1-q_1)^2 q_2 (1-q_2)^2 +$$

$$6q_1q_2(1-q_1)(1-q_2)^2 + 6q_1(1-q_1)q_2^2(1-q_2) + 3q_1^2q_2^2(1-q_2) + q_1^2q_2^3$$

(for the first group – for the second group just permute q_1 and q_2.)

This gives the the Jacobian with the following terms:

$J_{11} = -2(1-q_1)(1-q_2)^3 - 6q_1q_2(1-q_2)^2 + 6(1-q_1)q_2^2(1-q_2) + 2q_1q_2^3$, $J_{12} = -6q_2(1-q_2)(1-q_1)^2 + 6q_1(1-q_1)(1-q_2)^2 - 6q_1(1-q_1)q_2^2 + 6q_1^2q_2(1-q_2) + 3q_1^2q_2^2$, $J_{22} = -J_{11}$ and $J_{21} = 6q_1^2q_2^2 - J_{12}$.

With the first set of solutions, we get $J_{11} = \frac{-51}{81}$, $J_{12} = \frac{84}{81}$ and $J_{21} = \frac{-60}{81}$ So the eigenvalues are not real numbers and have zero real parts.

Next we consider an example with $N_1 \neq N_2$. We consider a TMSE in this casewhich turns out to be unstable: $N_1 = 2, N_2 = 3$. Let q_1, q_2 denote a symmetric TMSE. The mixed strategy equilibrium involves: $f_1(q_1, q_2) = (1-q_1)(1-q_2)^3 + 3(1-q_1)q_2(1-q_2)^2 + 3q_1q_2(1-q_2)^2 + 3q_1q_2^(1-q_2) = 2c$, and $f_2(q_1, q_2) = (1-q_1)^2(1-q_2)^2 + 2q_1(1-q_1)(1-q_2)^2 + 4q_1q_2(1-q_1)(1-q_2) + 2q_1^2q_2(1-q_2) + q_1^2q_2^2 = 2c$ where $f_1(q_1, q_2)$ represents the probability of being pivotal for type 1 players and $f_2(q_1, q_2)$ for type two players. The Jacobian matrix J has the following elements: $J_{11} = (1-q_2)[3q_2^2 - (1-q_2)^2]$, $J_{12} = 3(2q_2^2 - 2q_1q_2^2 + q_1 - 2q_2)$, $J_{21} = 4q_2 - 2q_1 - 4q_2^2 + 4q_1q_2^2$, $J_{22} = -2 + 2q_2 + 4q_1 + 4q_1^2q_2$.

The asymptotic stability properties at any solution therefore depend on whether the eigenvalues are both negative or not at this solution. Let $c = 1/4$. Real solutions for q_1 and q_2 are: $(0.4, 0.6)$ and $(0.83, 0.16)$. The eigenvalues λ_1, λ_2 are given by

$$\frac{-(a_{11} + a_{12}) + -\sqrt{(a_{11} + a_{22})^2 - 4a_{22}a_{11} - a_{12}a_{21}}}{2}.$$

On calculating these eigenvalues for the two solutions above: the first solution has $\lambda_i = \frac{0.71+-\sqrt{0.5-4(1.3)}}{2}$(no real roots) and the second:$\lambda_i = \frac{1.55+-\sqrt{76.4}}{2}$., which has one positive solution-hence there is a positive eigenvalue. Hence the totally mixed equilibrium is unstable, at least for small values of N_1, N_2. The expression above for $F(q_1, q_2)$ is too complicated to permit us to evaluate the properties of it as N_1, N_2 become large, but we can use a Poisson approximation introduced by Myerson (98,99)in the more general case with population uncertainty.

3.2 Incomplete Information

We capture uncertainty by a model of incomplete information about costs. 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^*. 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, althought 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 equilbrium 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).

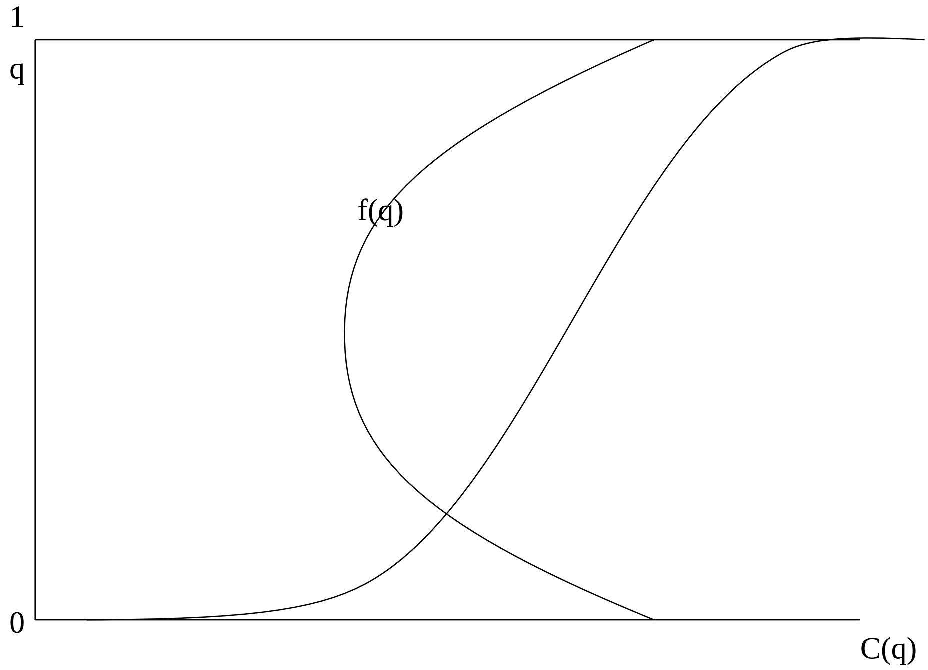


FIGURE 3

Palfrey and Rosenthal (1985) show that under some assumptions 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 make 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 Palfrey and Rosenthal have 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 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.
To achieve this we now consider learning as before: $\frac{dq}{dt} = K(q)$, with $\text{sign}(K(q)) = \text{sign}(f(q) - 2c(q))$. 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.

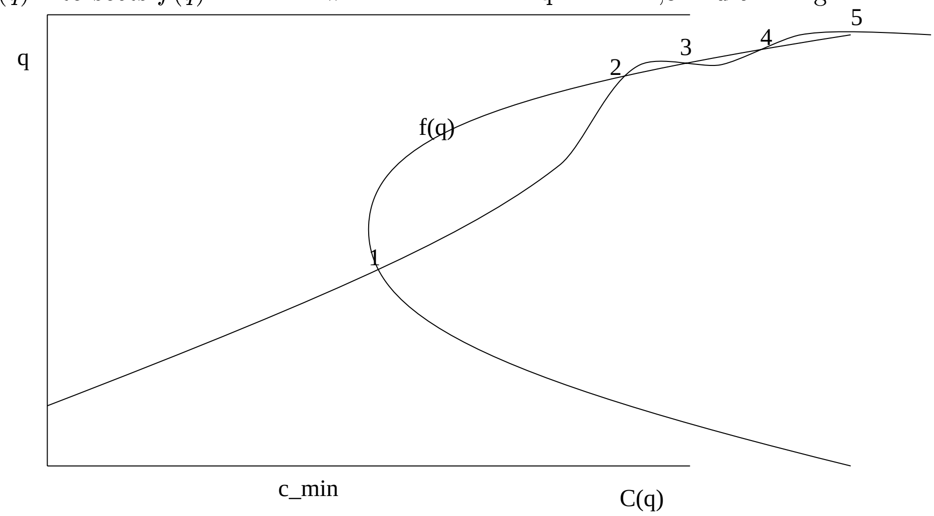


FIGURE 4

More formally we state this in the next Proposition:

Proposition 2: Let the graphs of $C(q)$ and $f(q)$ be in a generic position (this means that they intersect transversely i.e. $2c(q) - f(q) = 0 \Rightarrow d/dq(2c(q) - f(q)) \neq 0$), 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.

4 Comparitive Statics

We now investigate how equilibria change when the distribution of c varies; For simplicity 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

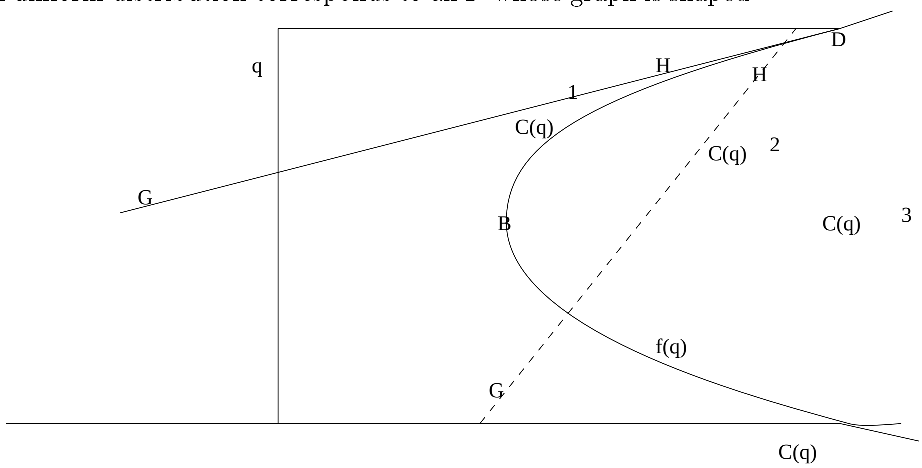
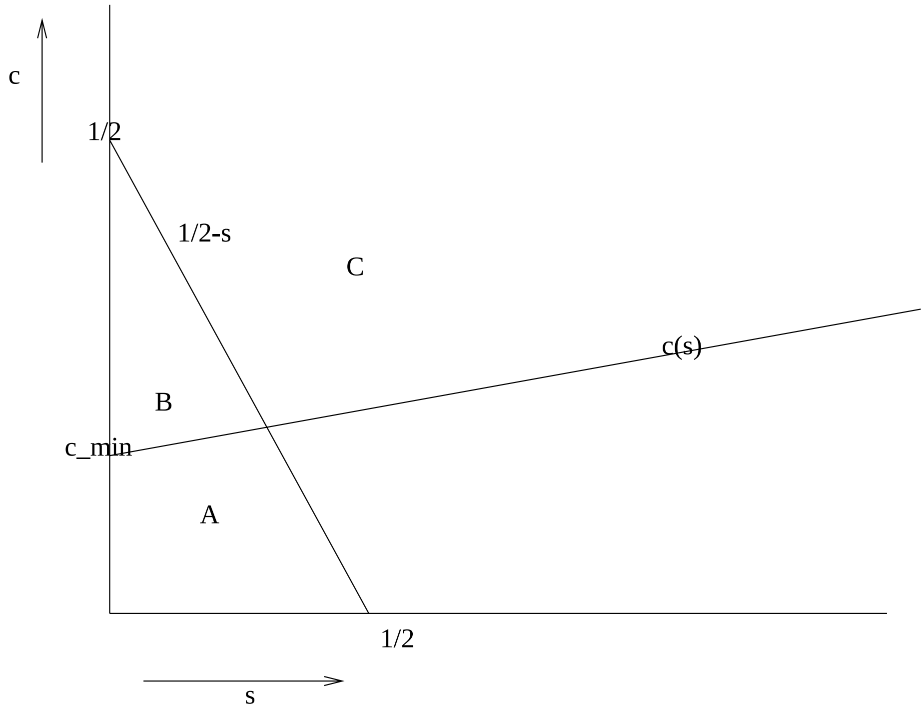


FIGURE 5

as in 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). This slope can be computed explicitly ($\frac{N-1}{2}$). This tells us that the condition for such a behaviour is that $s \geq (N-1)/2$. Note that this is a very large value for s, for plausible values of N. 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, in good agreement with intuition. This corresponds to the case studied by Palfrey and Rosenthal. 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: 1) For $\overline{c} < c(s)$ there is one equilibrium with $q = 1$ (everybody votes) (shown by area A in Figure 6) 2) For $c(s) < \overline{c} < 1 - s$ there are three equilibria: the previous one, that is stable, an unstable one with large q and a stable one with a low q, with a large basin of attraction, (shown by area B in Figure 6) which makes the $q = 1$ equilibrium stable at first and then makes it disappear when 3)$\overline{c} + s > 1/2$ and there is only one low turnout equilibrium (shown by area C in figure 6) . 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.

5 Repeated Elections

5.1 Complete Information

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 $i = 0, 1, 2, 3, ...$, all other elements ofthe model being 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 $\underline{q} = max\{0; q - \delta\}$ and $\overline{q} = min\{1, q + \delta\}$ with δ a small positive number that gives a measure of the possible mistakes in ascertaining the turnout. in this way we get a random dynamical system, actually a Markov chain since the system is time independent, whose states are the stable Nash equilibria, 1 and 3. The behaviour is given by proposition 2:
it Proposition 2: If the number of voters is larger than $N_0(\delta)$, the limit distribution of the outcomes q_∞^i is concentrated on equilibrium 1.
Proof in Appendix.

Note that the precise form of the probability distribution of the q_0^i is, to a large extent, irrelevant to the result. 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 after a sufficient long time we should observe a jump out of it towards equilibrium 1 (think of a golf ball in a bowl in a shaky train coach).
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 value $1/2 > c_{min}$. If the cost is not in this range there is only one equilibrium .

Note that, in all cases, equilibria with positive probability predict a nondecreasing 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 fictitous 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 mistake , 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 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. 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.[3] 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. Our model, and in particular hysteresis, can be used to see what happens when the cost c is changed by 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 significant 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, 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 persistence of large q; in U.S.A. where there have never been such laws the more stable low turnout equilibria are observed. Metastability gives an explanation of another phenomenon often observed: in old democracies (countries with a long history of voting, with approximately the same c) like the USA abstention is often high. This may not stem from the closeness of the electoral platforms (or high cost of voting) or indifference of the voters about the issues, which is a common (and a little tautological) explanation. To see what happens in our model, assume that c is between $1/2$ and 1.If this is the case even if one starts with high turnout equilibria, random fluctuations will make them eventually jump down to low turnout equilibria.

In the incomplete information case, the low turnout equilibrium has a large

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

basin of attraction that increases with $\overline{c}$ making first the $q = 1$ equilibrium metastable and then making it disappear when 3) $\overline{c} + s > 1$ and there is only one, low turnout equilibrium. The domain in the $\overline{c}, s$ plane corresponding to the three cases are shown in Figure 6.

One would see the same phenomena of hysteresis as in the complete information case when one moves back and forth on a line crossing the regions as AB does.

6 Conclusions

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

Usually it is very hard to test models of this type because parameters such as the cost of voting or the measure of the interest in a candidate are not directly measurable with reasonable confidence. It is not even clear what should be the right N to take: 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 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. It should be obvious that in such cases asking for quantitative predictions is meaningless; on the other side, 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 Palfrey and Rosenthal model more apt to be tested in this way. One could even get deduce from a sequence of electoral behaviours something about the shape of the distribution of the $c's$, for instance a case in which a gradual change of c would introduce several severe jumps in the turnout should point to the existence of a multimodal distribution, as described at the end of the last section, while a unique jump would be evidence of a more homogeneous electorate as far as costs are concerned.

References

1. Fudenberg, D and David K.Levine, (1998) "The theory of learning in Games", MIT Press.

2. 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.

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

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

5. 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, B^* 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 (2), $q(., t)$ 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: Since any point is $[0, q_2]$ converges to equilibrium 1 and any point in $[q_2, 1]$ converges to equilibrium 3, the transition matrix of the Markov Chain is given by:

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

where p_{11} is the probability that $q_0^i \in [0, q_2)$, the basin of attraction for equilibrium 1, if $q_\infty^{i-1} = q_1$, p_{31} is the probability that $q_0^i \in (q_2, 1]$, the basin of attraction of equilibrium 3 if $q_\infty^{i-1} = q_1$, p_{13} is the probability that $q_0^i \in [0, q_2)$, the basin of attraction for equilibrium 1, when $q_\infty^{i-1} = 1$ and p_{33} is the probability that $q_0^i \in (q_2, 1]$, the basin of attraction for equilibrium 3, when $q_\infty^{i-1} = 1$.

Note that when $N \to \infty$, $q_2 \to 1$. So given the distribution we have: $p_{1i} \to 1, p_{3i} \to 0, \quad i = 1, 3$. 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 converges to a measure on 1 in the general case.