Economic Theories of Voter Turnout

Amrita Dhillon* and Susana Peralta†

November 14, 2001

Keywords: Voter Turnout, Instrumental, Expressive, Evolutionary, Learning, Pivot Probabilities.
JEL Classification: C72, D72, D82, D83

*University of Warwick. We would like to thank Toke Aidt and Iwan Barankay for comments.

†CORE, Université Catholique de Louvain. This paper was written with the financial support from the Sub-Programa Ciência e Tecnologia do 2^o Quadro Comunitário de Apoio, Ministério da Ciência e Tecnologia, Portugal. The second author would like to thank Jacques Thisse and Michel Le Breton for inspiring this work.

1 Introduction

The rationality of voting is the Achilles' heel of rational choice theory in political science. (Aldrich, 1997)

In modern democracies thousands of citizens are eligible to vote in each election. An individual vote cast in such an election will conceivably have a negligible effect on its outcome. This leads to 3 interesting paradoxes: (i) the voting paradox: why would the rational individual bother to spend time and resources to get well informed, go to the polls on the election day and maybe even wait in a queue? Why would a rational voter vote? (Downs, 1957) (ii) the paradox of indeterminacy: If they vote why do they care who they vote for? (Kirchgassner, 1992; Kirchgassner and Pommerehne, 1993), and finally (iii) the paradox of ignorance: To the extent that information is costly, the rational voter must be ignorant about relevant aspects of his decision making (Downs, 1957).

The answer to this question has serious implications. If the only alternative one is left with is to label the voter as an irrational individual, then how can democracy and its institutions be defended and studied on rational grounds? Moreover, what will be the normative properties of social equilibria stemming from voting institutions, if individuals have a basic incentive to abstain, and the ones who vote are not rational? This would imply a fundamental flaw in the voting system, if indeed the models capture reality!

The problem of explaining voter turnout provides fertile grounds to test the power of economic and political theory. It raises questions such as what is the right objective function for voters, what is the correct notion of rationality – are voters self interested or do they care about a larger community, are there informational problems that are crucial for the turnout decision?

Many authors have addressed this issue and tried to find a theory that can explain observed participation levels. This survey attempts to cover the theoretical literature on participation, up to the present time. We note here that we are mainly concerned with plurality rule elections – i.e first past the post systems as these are the ones for which the voting paradox was originally investigated[1].

We use the questions raised above to classify the theories we survey under the following broad headings:(i) the objective function: under this there are two main types : (A) Instrumental theories which assume that the main motivation of voters is to affect the outcome, (B) Expressive theories which believe that

[1]The problem of free riding, while pervasive in plurality rule elections, does exist even for proportional representation or for "voice" theories– one individual by himself cannot have more than a very minor impact on the share of votes that a candidate receives.

the act of voting gives utility to voters, and this utility can depend on various factors e.g. how many and which others are voting. (ii) Rationality assumptions: i.e. theories which weaken the full rationality assumptions on voters, this gives us category (C): Boundedly rational voter theories, (iii) group based explanations: clearly if voters voted as a group they would be more likely to affect the outcome. This co-ordination is hard to achieve but there are some evolutionary explanations as well as the use of parties as co-ordinating mechanisms: this gives category (D). Finally there are (iv) the information based theories (category (E)).

The paper is organised as follows: It will be convenient to first describe a general model to standardize notation. We do this in Section 2. Sections 3 through to 6 follow the four broad headings above. Finally Section 7 concludes.

2 The Model

The simplest model[2] has voters voting only over two candidates. One can think of two groups of voters T_1 and T_2, where T_i is the group that favours candidate i. The expected payoff from voting, R_i, is given by:

$$R_i = B_i P_i + D_i - C_i \tag{1}$$

where B_i=benefit expected to be derived from success of one's favorite candidate and is the difference in utility of voter i if his favourite candidate is elected and the utility if the opponent does, P_i=perceived likelihood that one's vote will make a difference, D_i = the expressive benefit that voter i gets from the act of voting and C_i=cost of voting of voter i. Let $c_i = C_i - D_i$, i.e. the net cost of voting. The basic calculus of voting is then denoted as:

$$R_i = B_i P_i - c_i \tag{2}$$

We can conveniently classify the various states that a voter faces into the following:

1. S1: Candidate 1 wins by more than one vote

2. S2: Candidate 1 wins by exactly one vote

3. S3: Tie between the candidates

[2]This is a combination of various models but borrows heavily from Palfrey and Rosenthal (1983, 1985).

4. S4: Candidate 1 loses by exactly one vote

5. S5: Candidate 1 loses by more than 1 vote

Let p_i represent the probability of state i where $i = 1, ..., 5$. Let N_1, N_2 and N denote the number of voters in groups T_1 and T_2 and the total number of voters respectively.

3 The Objective Function

3.1 A: Instrumental Voting

In these theories a voter is assumed to care mainly about the outcome of his vote usually identified as the policy platform of the winning candidate – in terms of the model above this is the term B_i.

The first few studies of voter turnout were decision theoretical models. In what follows we further classify intrumental models into (i)decision theoretic and (ii) game theoretic. Downs' (1957) was not a formal theory hence we do not classify this. In his 1957 book *An Economic Theory of Democracy*, Downs (1957) provides a theory of government behavior based on rationality, i.e. the actions of the government are assumed to arise from the rational pursuit of some goals. The specific goal proposed by Downs is to maximize political support. In this context, the act of voting has special importance:

> In order to plan its policies so as to gain votes, the government must discover some relationship between what it does and how citizens vote. In our model, the relationship is derived from the axiom that citizens act rationally in politics.

The rational citizen's decision to vote or not to vote is based on a number of factors. The first is the difference in the expected utility between voter i's utility if his favourite candidate is elected and his utility if the opponent wins (B_i). This should be discounted, since the rational individual must take into account that his vote alone won't decide the outcome of the election (by P_i). The magnitude of the individual's influence on the election depends upon the number of voters (P_i is a function of N) and also on the way they are expected to vote. In fact, the discounting factor will decrease with the number of voters and increase with expected closeness of the election. In very large electorates, as is the case in modern democracies, this probability of influencing the outcome of the election through an individual vote is likely to be very small. Hence the discounted party differential, which Downs (1957) calls the *vote value* (the term P_iB_i), will be very small for the rational voter. If we suppose that voting has no cost, then only strictly indifferent individuals should abstain. For all the others it is rational to vote, no matter how small the vote value is. But the act of voting is time consuming — spent on such activities as gathering information on the relative quality of the candidates, deciding who to vote for and whether or not to vote and going to the polls on the election day. Therefore, there is an opportunity cost of voting ($C_i > 0$). Given that the vote value is in most cases negligible, then even a very small cost will lead the rational individual to abstain. Two led Downs (1957) to question the validity of this conclusion. The first one is theoretical:

> Each citizen is thus trapped in a maze of conjectural variation. The importance of his own vote depends upon how important other people think their votes are, which in turn depends on how important he thinks his vote is. He can conclude either that (1) since so many others are going to vote, his ballot is not worth casting or (2) since most others reason this way, they will abstain and therefore he should vote. If everyone arrives at the first conclusion, no one votes; whereas if everyone arrives at the second conclusion, every citizen votes unless he is indifferent.

This seems like a precursor to the game theoretic models which show that pure strategy equilibria do not exist in such games, rather there are mixed strategy equilibria.

The second one was empirical– that the observed turnout levels are in fact quite high.

To solve this "problem", Downs (1957) included in the return from voting the long term value of democracy — supposing that all the citizens have a desire to see democracy working, voting prevents the collapse of the system caused by generalized abstention. The individual supports a short run cost to insure himself against a potential huge loss in the long term. This is a value that comes from the act of voting *per se*, independent of which candidate gets the individual's vote and who ultimately wins the election(D_i in the basic model (2)). This value is positive, increases with the benefits the individual expects from the democratic system and decreases with the expected number of voters — since the votes of others also avoid the collapse of democracy. Admitting this further component into the return from voting, one is able to accommodate the different possible strategies of individuals. If total return (the short term vote value and the value of voting *per se*) exceeds the cost of voting, the individual will vote. Formally, this means an individual who votes in a pure strategy equilibrium will be one who has a negative net cost (c_i) of voting.

3.1.1 Decision-theoretic models

The first attempt to formalise Downs (1957) was Tullock (1968). Concerned with the incentives for the individual to get well informed in order to take political decisions, Tullock's formalisation was as in the basic model (2) with the restriction that P_i is a decreasing function of N, and that the term B_i may differ for different voters, depending on whether they belong to pressure groups for example[3]. None of these factors is, nevertheless, able to overcome in a substantial way the negative value previously obtained. Now if we count explicitly the costs of obtaining information in order to decide accurately, then the decision to vote will depend upon this value — more information increases the accuracy of the judgement about which party to vote for and hence the term P_iB_i but is costly. The problem of the rational voter is then to find the amount of information that causes $R_i > 0$. It may be noted here that this conclusion in inconsistent with the assumption of an instrumental voter since the information has value only if the probability of being pivotal is significant[4].

Riker and Ordeshook (1968) further formalized the problem by giving an explicit closed form expression for the probability of the individual influencing the outcome and modelling this as a function of the number of voters. Previous papers had calculated the probability that an individual voter would be pivotal conditional on the fact that someone was pivotal: they implicitly assumed that the probability of being pivotal was one and that any voter had an equal chance of being pivotal. Thus (assuming that the voter is a supporter of candidate 1), either the probability that the individual casts the last necessary vote for candidate 1 is N^{-1}, or the probability that the individual on the winning side casts the last deciding vote $\left(\frac{N+1}{2}\right)^{-1}$ were used. Riker and Ordeshook (1968) instead made the following calculation: Let q $(= p_1+p_2+p_3)$ and $q\prime$ $(= p_1+p_2)$ the probability that candidate 1 wins for sure if the individual votes and if he abstains, respectively. Denoting the probability that candidate 1 gets exactly x votes as $Pr[1,x]$, we have: for N odd: $q = \sum_{x=\frac{N+1}{2}+1}^{N+1} Pr[1,x]$ and $q\prime = \sum_{x=\frac{N+1}{2}}^{N} Pr[1,x]$; and for N even: $q = \sum_{x=\frac{N}{2}+1}^{N+1} Pr[1,x]$ and $q' = \sum_{x=\frac{N}{2}+1}^{N} Pr[1,x]$. Imposing a weak axiom, namely that adding one more element to the set of voters doesn't

[3]For individuals belonging to pressure groups, the number of votes matters as well as influencing the outcome of the election.

[4]See Aidt (2000) for more details on the related paradox of ignorance.

change the preference order of the other voters, it can be proved that the probability that candidate 1 gets x votes out of N voters, $Pr_N[1,x]$ is exactly the same as the probability to get $x+1$ votes out of $N+1$ voters, if this new voter votes for candidate 1. They define the probability of being pivotal as $P = q - q' = p_3$, if N is odd, then $P = 0$; in fact,there is no way that the additional voter can make candidate 1 a winner, unless he was winning already[5]. If N is even, then $P = Pr_N[1, N/2]$, the probability that candidate 1 gets one half of the votes when there are N votes cast; in fact, if there was a tie before, the new voter will bring victory for candidate 1. Let $N_1 =$ the number of voters who vote for candidate 1. If the citizen believes that there are equal chances of N being odd or even, then we get: $P = \frac{1}{2} Pr_N[1, N/2] \simeq \frac{1}{2} Pr[1, N_1/N = 0.5]$. The likelihood of an individual voter influencing the outcome is now a function of the number of voters, N, as before, but also of the shape of $Pr_N[1,x]$. If it has a maximum at x_0, then P will be higher the closest is $N/2$ to x_0, and will decrease with the variance of the density around x_0. This calculation implicitly assumes that all voters turn out, otherwise some kind of population uncertainty would need to be incorporated.
They extended this calculation to show that even if voters care about vote shares of their preferred candidate, however the magnitude of the difference the voter makes by this marginal vote is so small that it would not explain turnout. Thus, this feature of the theory – free riding – carries over to proportional voting systems as well.
Strom (1975) proposed that voters had state dependent utilities – thus the term B was state dependent with voters deriving very high utilities from being pivotal. However, the discounted term $P_i.B_i$ is still too small to affect the conclusions.

A model proposed by Ferejohn and Fiorina (1974) models a rational voter as a regret minimaxer rather than an expected utility maximiser. Minimax regret is used as a decision rule under *uncertainty*, i.e. when the individual *does not estimate a probability distribution over the possible states of nature.*

For each pair of actions a_i and state of nature S_j, the individual will compute the regret r_{ij}. This is the difference between what he could have obtained under S_j if he knew it would happen with certainty and was able to choose the action that yielded maximum return and what he obtains by choosing a_i. Then for each action he computes the maximum regret possible, that is $\max_j r_{ij}$,over the

[5]For instance, if there are 51 voters, the number required for 1 to be a winner is 26; if an additional individual enters the universe and votes for 1, then we have a total of 52 voters, with 27 votes for candidate 1. From the perspective of this new voter, the probability of both events is the same.

different states and he chooses the action with the minimum of these maximum regrets.

Normalizing the utility from the preferred candidate to 1 and the utility from the opponent to 0, we have $B_i = 1, \forall i$. Doing this, we take the cost of voting to be measured in terms of *differential utility units* — if $c_i = 1/2$, it means that the cost of voting is one half of the utility difference between the favorite candidate and the opponent. Ties are decided by a coin toss. V_i stands for the action of voting for candidate i and A_ix for abstaining. Note that the intensities of preference that may differ between individuals is completely absent from the models we have considered – hence the normalisation of B_i. All subscripts are dropped from now on.

In this case, it is easily seen that with the regret table (Table 1), the maximum regret associated with actions V1, V2 and A are c, 1 and $1/2-c$, assuming that $c < 1/2$.

	S1	S2	S3	S4	S5
V1	c	c	0	0	c
V2	c	$\frac{1}{2}+c$	1	$\frac{1}{2}$	c
A	0	0	$\frac{1}{2}-c$	$\frac{1}{2}-c$	0

Table 1: Regret Matrix

The maximum regret of voting for one's favorite candidate is less than from abstaining (that is, V1 is the minimax regret action) whenever $c < \frac{1}{4}$. Clearly, the minimax regret rule yields higher turnout than the expected utility maximization. The reason for this result is that, although S3 and S4 are not likely to arise, the minimax regret voter considers the mere possibility and not the probability.

Grofman (1979) extended the minimax regret analysis by allowing the voter to use mixed strategies. Note that in Table 1, V2 is a weakly dominated strategy. In the reduced game, the author computes the optimal strategy using a rule in Luce and Raiffa (1957). The reduced regret matrix may be written:

	S1, S2, S5	S3, S4
V1	c	c
A	0	$\frac{1}{2}-c$

Table 2: Reduced Regret Matrix

The optimal mixed strategy is to mix between voting for 1 and abstaining with probability distribution $(1-2c, c)$ – this generalises the regret minimaxers behaviour above – he just weights his strategies according to the regret from doing any of them.

According to this result, a maximin regret voter votes with positive probability when $c < 1/2$, hence this result yields more expected turnout.

Although it does predict more turnout this model calls for extremely conservative behaviour of voters, for example a minimax agent should never cross the street.

The major drawback of the decision theoretical approaches is the way they treat the probability that a single vote influences the outcome – ignoring it (minimax regret) or taking it as exogenous (expected utility maximization). The probability that a vote is pivotal depends on the number of other voters: it is determined simultaneously with turnout. Downs (1957) clearly had in mind a more game theoretic situation: The first few papers that have taken this simultaneity into account use game theoretic models.

3.1.2 Game Theoretic Models

Ledyard (1981) was the first model in this strand of the literature. This model endogenised not only the probabilities, but also the candidates' platforms. The candidates are denoted $j = 1, 2$ and choose platforms θ_1, θ_2 on H, the issue space. There are $N + 1$ voters, indexed by i, with utility function $U^i(\theta)$. θ represents the vector of candidate platforms. The assumptions are standard:

- Candidate irrelevance: the individual prefers the candidate with highest $U^i(\theta_j)$;

- Expected utility: the individuals are expected utility maximizers;

- Utility parametrization: $U^i(\theta) = U(\theta, d_i)$ with $d_i \in D$ standing for the taste characteristics of the individual.

- No income effects: when candidate j wins, the individual receives utility $U(\theta_j, d_i) - c_i$ if he voted and $U(\theta_j, d_i)$ if he abstained.

The type of each individual is the vector $e^i = (d^i, c^i) \in E$. The cost of voting is bounded away from zero.

Using the states of the world S1–S5, defined in Section 2, we have the matrix of utilities in Table 3, where U_j is used as shorthand for $U(\theta_j)$.

	S1	S2	S3	S4	S5
V1	$U_1 - c$	$U_1 - c$	$U_1 - c$	$(U_1 + U_2)/2 - c$	$U_2 - c$
V2	$U_1 - c$	$(U_1 + U_2)/2 - c$	$U_2 - c$	$U_2 - c$	$U_2 - c$
A	U_1	U_1	$(U_1 + U_2)/2$	U_2	U_2

Table 3: Utilities

Defining $W(\theta, e^i) = \frac{U(\theta_1, d^i) - U(\theta_2, d^i)}{2c^i}$, we may deduce the optimal behavior of an expected utility maximizer:

Vote for 1 if $W(\theta, e^i) > \frac{1}{P_3+P_4}$

Vote for 2 if $W(\theta, e^i) < -\frac{1}{P_2+P_3}$

Abstain if $-\frac{1}{P_2+P_3} < W(\theta, e^i) < \frac{1}{P_3+P_4}$

First we discuss voters' equilibrium for a given θ. It is assumed that each individual is rational and believes the others are also rational, in the sense of deciding according to expected utility. Secondly, it is assumed that $e^i \sim iid\ \mu$ on E.

Every individual i can predict with certainty the decision of h, if he knows his type e^h; since e^h is unknown, i will compute the probability that h takes each of the possible actions, using the probability distribution of e^h.

Take p^i_k to be i's estimate of the probability of state k and let $j = 0$ represent abstaining. Given θ_1, θ_2, μ and p^i, voter i believes that an arbitrary voter $h \neq i$ votes for $j = 1, 2, 0$ with probability q_j:

$$q_1 = 1 - G(\frac{1}{\alpha}, \theta) \tag{3}$$

$$q_2 = G(-\frac{1}{\beta}, \theta) \tag{4}$$

$$q_0 = G(\frac{1}{\alpha}, \theta) - G(-\frac{1}{\beta}, \theta) \tag{5}$$

where $\alpha = P^i_3 + P^i_4$, $\beta = P^i_2 + P^i_3$ and $G(r, \theta) = \mu\{e^i \in E \mid W(\theta, e^i) \leq r\}$
The proof is trivial, noting that G is in fact the cumulative distribution function of W, implied by the probability distribution of the characteristics of the individuals. For instance, all the individuals with $W < \frac{1}{\alpha} = \frac{1}{P_3+P_4}$ will vote for candidate 1. This is the typical description of a Bayesian Nash equilibrium, where rules are of the form: vote for 1 if your net welfare from candidate 1 is above a certain threshold level.

Both α, the probability that a 1 supporter casts a decisive vote, by breaking or creating a tie and β, the correspondent probability for a B supporter, depend on how the electorate is behaving, that is, the values of q_1, q_2 and q_0. α equals all the possible combinations of N voters (the electorate excluding the individual) that allows the voter to create a tie between 1 and 2 or to lead 1 to victory by one vote. This is given by the following expression using the binomial distribution: $\alpha = f(q_1, q_2)$ and $\beta = f(q_1, q_2)$ with $f(x,y) = \sum_{k=0}^{N/2} \binom{N}{k}\binom{N-k}{k} x^k y^k (1-x-y)^{N-2k} + \sum_{k=0}^{(N-1)/2} \binom{N}{k}\binom{N-k}{k+1} x^k y^{k+1} (1-x-y)^{N-2k-1}$

A *symmetric voters' equilibrium* $(q_1^*, q_2^*, \alpha^*, \beta^*)$ is a symmetric Bayesian Nash equilibrium of this game. Existence of equilibria is proved under some conditions on continuity of the distribution function $G(r,\theta)$ in r for all $r(-\infty,\infty)$, given θ_1, θ_2, μ.

The interesting issue, however, is to predict the levels of turnout in equilibrium. He describes a full abstention equilibrium – this occurs when $G(1,\theta) - G(-1,\theta) = 1$, i.e. there is no voter who has $W(\cdot) > 1$ or $W < -1$, so that $q_1 = q_2 = 0$. Normally however there would be positive turnout in equilibrium. With G continuous in r and $G(1,\theta) - G(-1,\theta) < 1$ an equilibrium exists with positive turnout. If the distribution of tastes in the electorate is such that there are some costs sufficiently low as compared to the utilities derived from elected platforms, that is, some $W < -1$ and/or some $W > 1$, then some agents will vote and some will abstain.

Finally an electoral equilibrium is defined, assuming that parties are interested in maximising expected plurality. Given that a voters equilibrium exists, this induces a zero sum game between the two parties. A Nash equilibrium of this game in pure strategies exists under some technical conditions. Under some plausible assumptions the two parties platforms converge (the median voter theorem) and this involves zero turnout in equilibrium.

While turnout levels are not characterized for the general case, some examples are provided. We end the description of Ledyard's model with a quote about the predictions on turnout (from Ledyard (1984)):

> A major issue raised by many who see this model for the first time is the lack of turnout in equilibrium. While I see this as good (the deadweight cost of voting costs is avoided), many see this as a prediction of the model clearly contradicted by the facts. It must be remembered that, because of the many possible frictions, actual elections will rarely match this theory. Among other things, most elections are held to decide several contests simultaneously and political activists, ignored in my model, operate to interfere with the natural

forces.

Ledyard (1984) basically uses the same model to get further results on maximum turnout and uniqueness of equilibria.
Palfrey and Rosenthal (1983) study a simplified version of the Ledyard model with complete information. They later (1985) extended this to the incomplete information case. We now review both these important contributions. They model the voting decision as a *participation game*. The basic idea behind this type of game is the existence of *teams* which share the same preferences and the same payoffs. The strategy is a binary decision by each player to participate by voting or not.

The payoff of a member of one team is non-decreasing with the number of participants in this team and non-increasing with those of the other. This model can be applied easily to the electoral process: if one is a supporter of one party, his payoff is higher the more people vote for that party and the less vote for the other; the outcome of the elections affects the supporters of the same party in the same way.

With such a game, both the competitive aspect between supporters of opposite parties and the incentive to free-ride on the possible votes of the supporters of the same party can be captured. The former induces voting, the latter induces abstention. To illustrate the two aspects, Palfrey and Rosenthal (PR) present the two following examples of 2X2 games – one where only the free riding aspect is present (Prisoner's Dillemma) and the other where only the competitive aspect is present (Chicken).

Let's suppose, as usual, that a fair coin decides a tie, $c < 1/2$ and the utility is 1 if one's team is elected, and 0 if the other team wins.

If we consider a game without free-riding, that is, with one member in each team, we have a prisoners' dilemma: the only Nash equilibrium is both individuals voting. Each team wins with 1/2 probability, hence abstaining Pareto dominates the voting equilibrium, since players face the same expected payoff of 1/2 without incurring the cost of c[6].

This simple game shows that the competitive aspect of the election induces voting.

[6] There might be an objection to using a coin toss in the case where no one votes, but this is just an artificial way to capture the competition involved rather than a statement of how democracy works.

	Vote	Abstain
Vote	$1/2-c, 1/2-c$	$1-c, 0$
Abstain	$0, 1-c$	$1/2, 1/2$

The pure public good or free riding problem is captured by the chicken game, with both players being on the same team. The game has two pure strategy equilibria, with one player voting and the other abstaining (free-riding on the other's vote) and a mixed strategy equilibrium, i.e. vote with probability $1-2c$. This shows that multiple equilibria are possible in election games and that asymmetric equilibria may also arise, some of which are quite reasonable.

	Vote	Abstain
Vote	$1-c, 1-c$	$1-c, 1$
Abstain	$1, 1-c$	$1/2, 1/2$

The PR model is a generalisation of these two aspects. Payoffs are given by the basic calculus of voting discussed in Section 1(equation (2)). All subscripts vanish in the game of complete information: all voters are assumed to be indentical, thus we can use the normalised version as in the discussion of Ferejohn and Fiorina (1974), dividing all variables by B, with $B=1$. Voters will vote if the expected utility from voting is higher than the cost of voting – the main difference from Riker and Ordeshook (1968) is that the probability of being pivotal is endogenously determined, as in Ledyard (1981). Unlike Ledyard, however this model has exogenously given party platforms and is a game of complete information. Free riding occurs within the group as voters would not want to vote if they are not pivotal, while the competitive aspect arises because they want to ensure that their preferred candidate wins. The authors study three types of equilibria, according to the types of strategies used by players: pure, pure-mixed and totally mixed strategy equilibria (TMSE). The pure-mixed equilibria 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. There are, in general, multiple equilibria.

Since we are interested in predictions for large electorates, we discuss the equilibria that survive in large electorates. These can be classed into two categories: symmetric (members from both teams play the same mixed strategy), TMSE with low turnout, symmetric TMSE with high turnout (these arise only when $M=N$) and pure-mixed equilibria in which, in the pure strategy team, either no one votes (and in the other the probability of voting approaches zero) or there is a number of voters equal to the size of the minority team [7]

The TMSE have the property that as electorates become large, the probability of voting becomes smaller, but this is not true of the pure-mixed equilibria.

There is an apparently counterintuitive fact for the pure-mixed equilibria: turnout is increasing with the cost of voting[8].

This model shows that it is possible to have a high level of turnout in equilibrium, namely, twice the size of the minority team, which is equivalent to full turnout if $N_1 = N_2$, even with large electorates and strictly instrumental behavior on behalf of voters.

However, the kind of information that the Palfrey and Rosenthal model requires to be common knowledge is very demanding in large electorates. If the number of agents increases indefinitely, it is not reasonable to assume that everyone will know how many of them belong to each team. On the other hand, the equilibria that lead to substantial turnout with large electorates in the participation game both rely on a certain degree of certainty on the outcome of the election.

This led the authors to question the role of full information in their model and to generalize it by allowing for incomplete information (PR, 1985). They introduce incomplete information in the model through the cost of voting. The cost of voting is assumed different across individuals (c_i) and is private information. Pure strategies in this model are functions from types to strategies. They compute symmetric Bayesian equilibria for this game of incomplete information and show that under some assumptions on the shape of the distribution function of costs, there is essentially a unique low turnout equilibrium for large electorates, which corresponds moreover to the low turnout equilibrium in the complete information game with $N_1 = N_2$[9].

[7]Say the minority team is the one playing pure strategies. Then it votes as a whole and the other team votes with a probability close to the ratio of the size of minority to its own size. If, on the contrary, it is the majority team who plays pure strategies, then only a number of players equal to the size of the minority vote and in the minority side they vote with a probability close to 1. The minority votes as a whole and the majority exactly offsets the minority.

[8]Given a strategy from the other team, if voting cost increases, abstaining becomes a dominant strategy for randomising players; to get back indifference, they must increase the probability of voting, hence increasing turnout.

[9]The Bayesian strategies are of the form: vote if cost is below a certain threshold level, c^*, which is the same for all voters in a symmetric equilibrium if we assume that both populations are of the same size and both have the same prior distribution of costs. Hence we can find the low turnout equilibrium corresponding to the cost level c^*.

As already pointed out a big problem with the complete information model of PR was that of multiple equilibria – some of which are very fragile in the sense that they require voters to have very precise beliefs about what others are doing. PR(1985) was an answer to this problem since moving to a game with incomplete information did select the low turnout equilibria, but only when the degree of uncertainty is sufficiently high[10]. Demichelis and Dhillon (2001) suggest another way to capture the robustness of the low turnout equilibria. This is discussed later in Section 4. A rather plausible case of multi modal distribution functions is discussed – which could arise for example if there are groups in the population, the members of which have very similar costs but which are different between groups – may give rise to multiple equilibria, some of them with high turnout even when there is incomplete information. However only the low turnout equilibria survive in the long run, in their model of fictitious play learning (see Section 4), with or without incomplete information.

The major concern with the paradox of voting arises in big elections and moreover when there is uncertainty about the population size. PR (1985) assumed that the total population was known, although the uncertainty about turnout is captured by any kind of incomplete information that affects the turnout. Population uncertainty is yet another source of incompleteness in information. Myerson (1998A, 1998B, 2000), models *population uncertainty* as a *Poisson game*. Working under the hypothesis that the number of players is a Poisson random variable has a number of advantages, in particular to analyse voting games – which have the special feature that only the numbers of each type turning out to vote (in the two candidate setting) matter. This model of population uncertainty however stipulates a symmetry between any two players of the same type on the beliefs about other players which is a restriction as compared to a general Bayesian game. The main contribution for the purposes of this survey is to show that in this setting it is technically much easier with a Poisson game to show that the low turnout equilibrium is unique in the examples of PR (1985) and Ledyard (1981). Poisson games, moreover have the attractive feature that type-splitting does not change the equilibria, thus justifying somewhat the restriction imposed on strategy choice for people in the same type. The main features of the model and the important technical results are presented in the Appendix.

Costly and instrumental voting therefore seems to be inconsistent with significant turnout.

[10]See the section on incomplete information in Demichelis and Dhillon (2001) for a full discussion of this point.

3.2 B: Expressive Motivations

One of the first models to bring in expressive features was that of Fiorina (1976). Although most earlier models did allow for some fixed and exogenously given expressive motivations, this was the first to explicitly consider how expressive utilities varied. This model took account of party loyalties by the citizens. By introducing psychological factors — namely, a benefit from voting according to one's party identification and a cost from voting against it — the expressive component of the return from voting was posited to vary across strategies, which was not the case with the D term (citizen duty) in Instrumental theories.

All citizens are assumed to be members of some (single) party. Thus voters may prefer to vote for their own party even when they would get a negative benefit out of it (because of their policy position). They categorise such voters as partisan, cross pressured, while those who do not have this conflict are called partisan consistent voters.

Let a be the benefit from a vote that affirms one's party identification and d as the cost from voting for the opposing party. D, as before in the model (2) is the expressive benefit of voting as opposed to not voting. One can then compute the payoffs of every type of voter from voting for one of the two candidates or not voting. One can further distinguish between voters who are loyalist, i.e. voters such that if they vote, they vote for their own party and voters who are disloyalist. Loyalist (disloyalist) citizens vote if the payoff from voting for their own (other) party is higher than the cost. The expected utility is calculated as before taking the 5 states of nature as described in Section 2. The main contribution of this theory is that the comparitive statics for different types of voters are different, as summarised in Table 4.

	$D \nearrow$	$\lvert B\rvert \nearrow$	closeness $\nearrow$	$a \nearrow$	$d \nearrow$
Consistent	increases	increases	increases	increases	no effect
Cross-pressured, loyalist	increases	decreases	decreases	increases	no effect
Cross-pressured disloyalist	increases	increases	increases	no effect	decreases

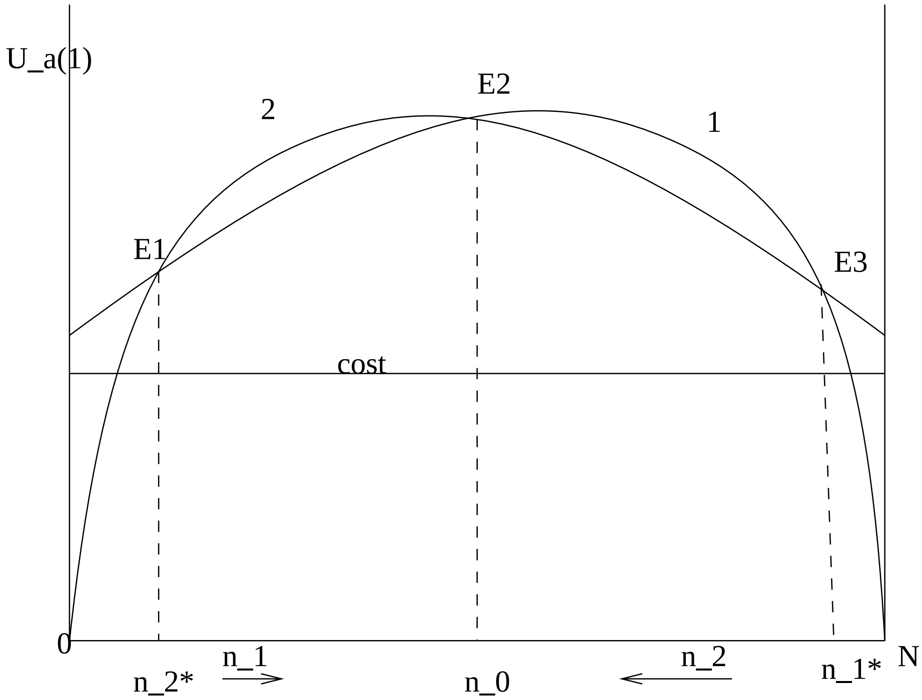
Table 4: Variations in Turnout

If one takes into account that the estimates of the probability of being pivotal are very small in large electorates, the conclusion is, as before, that only the expressive factors count.

The most detailed analysis of expressive voting is in a recent book by a political scientist – Schuessler (2000). Schuessler believes that expressive motivations in voting are very important in explaining not only the observed turnout levels but also empirically observed features of elections like bandwagon effects and the polarisation of voters due to negative campaigning. He believes that "voting in large scale elections...represent instances in which individuals express and reaffirm, to others and to themselves, who they are." An expressive voter is one who votes not because of his influence on the outcome but because by doing so he attaches to an outcome. He does not rule out instrumental motivations however, rather he believes that the two aspects are used strategically to win elections by politicians. He argues that politicians recognise that people want to attach to a candidate that they know (and everyone knows) is attractive to "many" others but not too many because that would lower the intrinsic worth of the attachment. Besides the "how many" dimension to election campaigns therefore there is also a "who" dimension – voters want to belong to the "right" club. Since politicians want as many voters to attach to them as possible, they must practice the politics of ambiguity – i.e. try to be as many things to as many people. One can note in passing that this theory has much in common with an incomplete information based theory. If voters do not know for sure the type of a particular candidate they may take the number of voters who support him as a signal of his worth – yet if everyone supports him, the signal is non revealing in the sense that a candidate clearly cannot satisfy all (or too many) voters. Indeed, a possible micro foundation for the theory might involve such informational issues.

We describe the model briefly (details are in Chapter 6 of the book): we assume as in the basic model (2) that there are two candidates, 1 and 2. Let n_j denote the turnout for candidate j. Then the ith individual's utility for candidate 1 is determined by an expressive and an instrumental component: $u_i^1(n_1, q_i) = f_1(n_1) + q_i$, where $q_i = P_i B_i$. The expressive component of an individuals utility is captured in $f_1(n_1)$ and the instrumental one in $P_i B_i$, which we can take to be the same as for previous purely instrumental models, in its dependence on the probability of being pivotal. Using the same logic as the decision theory models discussed above, i.e that the probability of being pivotal goes to zero as the electorate size increases, he reaches the conclusion that turnout will only occur if the expressive part of the utility for 1 can overcome the cost of voting, which is a "threshold". Note the problem with assuming the probability of being pivotal to be exogenous, as the game theoretic literature pointed out.

In order to obtain a theory of aggregate participation, the benchmark model assumes that while people have heterogenous instrumental preferences, they have identical expressive preferences which are candidate specific. Hence the average utility for candidate 1, u_a^1 is given by: $u_a^1 = f_1(n_1) + q_a^1$. Using empirical evidence, of bandwagon effects, he postulates a concave utility function $f_a^1(n_1)$, i.e. the average voter gets a higher utility from voting for 1 the higher the number of others voting for him, until a point n_1* is reached. He postulates moreover that this point is to the right of the 50% mark. The utility curve for candidate 2 is the mirror image of the utility curve for candidate 1. In the figure below this defines some "tipping" points for participation in 1(2) – E1, E2, E3 and some equilibria – the two stable equilibria are E1 (where candidate 2 wins) and E3 (where candidate 1 wins). The cost is low enough that any of these two can be an equilibrium with associated turnout given by n_2*, and n_1* respectively.



Turnout is determined by the threshold level of utility and which of the stable equilibria are reached: voters will turn out only if their expressive utility from some candidate is higher than the cost of voting. In the general model, the assumption of expressive homogeneity is removed – instead there are assumed to be different types of voters each of whom would have different tipping points and a distribution over tipping points. Equilibria are points where expectations are realised.

Unlike most other models, apart from Ledyard (1981), this model makes party positions and party campaigning endogenous. It is used to show how the producers of mass participation will manipulate perceived support to maximise their chances of winning. Lowering the costs of voting or shifting up the expressive returns from voting for a group of voters may not always yield an increase in turnout because of "expressive crowding out" or the congestion costs of having too many people in the club. The heterogeneity of voter preferences on the other hand may explain "ambiguity" i.e the practice of seeming to be everything to all voters. Unlike Ledyard (1981) however, the (general) equilibrium notion is not defined e.g. it is assumed that one party (candidate) is passive, rather than analysing a Nash equilibrium of the game where parties are players, given voter behaviour.

With large numbers of voters, in order to get rid of the problem of free riding in any collective action problem, it seems that one has to assume that the act of voting gives some utility – the difference in Schuessler's expressive theory of voting is that the utility from voting depends on the number and identity of other participants– turnout may increase due to bandwagon effects in elections. The crucial element of the theory is still the fact that the act of voting gives utility: other theories of turnout all seem agreed on this.

4 Rationality Assumptions

4.1 C: Bounded Rationality Models

The inconsistency of rational choice with empirical turnout leads us to question whether voters are really rational in the sense assumed by the theory. Are they really able to do complicated calculations of pivot probabilities before they decide whether to vote or not? In this section we discuss theories which replace the assumption of full rationality of voters. Sieg and Schulz (1995) replace the full rationality of a voter assumed in a game theoretic model, by a model of adaptive learning. No voter knows optimal strategies but learns voting strategies which are of high economic fitness. The social position of the individual voter is the ratio of his payoff to the average payoff for other voters. Learning is through trial and error: a voter changes his strategy by accident. One trial shows whether the status quo or the new strategy is better. A voter is able to learn a new strategy if this increases his social position. The results on turnout are ambiguous and depend on relative shares of the two populations and on costs. The equilibrium concept used is a version of Symmetric Evolutionary Equilibrium, which they call Evolutionary Voting Equilibrium (EVE).

An attractive result is that small groups are in a better position to solve the free rider problem. It is not clear, however, why this specific dynamic is used : that voters are interested in improving their *relative* position does not seem particularly appropriate for a voter turnout story.

One of the problems with using the Nash equilibrium concept in the complete information game of Palfrey and Rosenthal (1983) was that of the multiplicity of equilibria. We showed this in Section 3 above with an example for the complete information game when the size of the two populations is the same. Demichelis and Dhillon (2001) have considered instead (symmetric) equilibria from a learning model (fictitious play[11]) that are much more intuitive: turnout decreases with cost in such equilibria. In this model, voters will increase the probability of voting if the payoff from doing so given the prior beliefs on the proportion of the voting population is bigger than the costs. Thus turnout depends on perceived turnout through opinion polls or the last election, and on costs. The conclusion is negative in the sense that it is the low turnout Nash equilibria that are "robust", in the sense that as the size of the electorate increases, the low turnout equilibria are aymptotically stable states of the system and moreover the likelihood that the society is in a low turnout equilibrium gets higher the bigger the size of the electorate. They also make the point that no model can hope to predict the actual levels of turnout observed as many features of actual elections are not taken account of – as also pointed out by Ledyard (1984) and Feddersen and Sandroni (2001)– and because we cannot know the parameter values. What is important is to predict the "right" comparative statics, or predict other qualitative phenomena. The Demichelis and Dhillon (2001) model does have such qualitative predictions (e.g. hysteresis in turnout as costs change) that are apt to be a better test of the theory than actual turnout levels. If the theory is correct– i.e. that individuals vote depending on costs and perceived probabilities of being pivotal, then the comparitive statics predicted by the model should be empirically correct.

Conley, Toossi and Wooders (2001) consider the problem of voting from an evolutionary standpoint. Their model is an attempt to microfound the main insight from previous game theoretic literature that individuals who vote are those who get a "warm glow"(expressive motivations) from voting. Their ques-

[11]The crucial element of the model which is in the spirit of continuous fictitious play is the dependence of actions on empirical frequency of past actions. The players are naive and myopic, an assumption that seems in keeping with large elections. Voters perceive the same starting level of turnout and adjust in exactly the same way. They then update their priors in the same way and adjust again. This process continues till it converges and they play the corresponding mixed strategy.

tion is: where does the warm glow come from? They show that more public spirited agents have an evolutionary advantage, under some parameter values. In their model with two types of voters, this relies on the assumption that voters who have a high propensity to vote also have the same preferences and similarly for voters who have a low propensity to vote. The attractive feature of this model is that it goes back a step to examine what determines voters preferences which themselves determine the level of turnout. If expressive motivations are the only reason why people vote then this is indeed the right question to ask. Technical details of the model appear in the Appendix.

The results differ for the models of Sieg and Schulz and Demichelis and Dhillon (learning models) above in part because of the different dynamics assumed. Conley, Toossi and Wooders (2001) models voters with inherently different propensities to vote, rather than voters who differ only in their preferences for the policy. A starting position in their paper is the share of the population who are high (propensity to vote) type, rather than some observed turnout level (Demichelis and Dhillon). Moreover, the two types of voters behave asymmetrically in Conley et al (2001).

However, none of the bounded rationality models predict unconditional high turnout.

5 D: Group Based Theories

Close to Schuessler's model, is a new model proposed by Feddersen and Sandroni (2001). They develop a model of voting in which voters are motivated by ethical considerations. Following Harsanyi (1980), they view voters as rule utilitarians, who think of social welfare rather than their individual benefits from voting. Individuals turn out if the "ethical" benefit they get from voting is higher than the cost. Like Schuessler their model would predict zero turnout if there was unanimity among voters on which candidate was best for society, although the mechanism by which this happens differs. Some philosophical conflict is needed to generate turnout in their model. The main feature of their model seems to be the coordination between members of a group of voters – thus if there are two candidates, the game between voters can be represented as a two player game between the groups, where each group decides on its level of turnout. However the level of turnout not deterministic but determined randomly by the fraction of voters who have higher ethical returns from voting than costs. They obtain the "right" comparitive statics, which is one of the main objectives of the paper – thus voter turnout decreases as cost increases, or if the share of votes to one candidate becomes very large or if the payoff to acting ethically is decreased.

In contrast to many of the models of instrumental voting (with the exception of Ledyard (1981, 1984), Schuessler endogenises party positions. Hence equilibrium turnout depends on parties' strategic choices as well.

Shachar and Nalebuff (1999) take this point of view further – in their model voters are completely passive (followers) and will vote for the party they are ideologically close to, if they are contacted by the party (leaders). Leaders choose to make effort based on some (imperfect) information about vote shares. Thus it is parties that determine the level of turnout and who wins. It is parties that have expectations about the closeness of the election and then decide to influence participation. We should thus see higher turnout when the election is expected to be close. While this model captures an important aspect of elections, it is difficult to believe that this is the *only* feature in the decision to vote.

6 E: Information based theories

If we eliminate costs of voting, then voting is a (weakly) dominant strategy. The opposite side of the coin is the question – why (if) there is any abstention in this case? Motivated by the fact that when voting is costless (e.g. voting on several issues at one time) people still abstain, Feddersen and Pesendorfer (1996) show that voters who are uninformed about candidates may strategically delegate their decision to informed voters.

According to Feddersen and Pesendorfer (1996), *a swing voter is an agent whose vote determines the outcome of an election.* Even if this decisive agent has a strictly preferred candidate, abstaining may be a best response for him. This happens because his beliefs about the state of the world are changed conditional on the information that his vote is decisive.

A simple example (example 1 for our purposes) illustrates this point (Feddersen and Pesendorfer, 1999). There are two states of the world, 1 (more likely) and 2. Three voters are supposed to decide between two candidates and all of them prefer candidate 1 in state 1 and candidate 2 in state 2 (common values). All voters know exactly one of them is perfectly informed about the true state of nature. This voter will vote for the right candidate. An uninformed voter is better off abstaining than voting for candidate 1, the one that maximizes his expected utility. The only way his vote might matter is if there is a vote for candidate 2 and 0 for candidate 1. But this only happens if the true state of the world is 2 and the informed voter cast his vote for 2. The uninformed voter has an incentive to abstain, delegating the decision to the informed ones. This example of course does not have any population uncertainty, nor uncertainty about the number of informed voters in the population, but it carries over to the case where voters have uncertainty about the numbers of informed voters. It is clear that the example works because there is an implicit assumption that all voters of the same type (all uninformed voters are identical) necessarily believe that all of them will act in the same way. Otherwise there could be asymmetric equilibria where abstention is not the best response. The Feddersen and Pesendorfer (1996) model generalizes this basic idea to allow for different preferences in the electorate, namely, there will be three types of individuals – partisans who favour one of the two candidates regardless of the state of nature, and independents, and the same two states of nature.

Another crucial assumption in this paper is that of common values among independent voters who are informed and uninformed. Consider the same example as above but assume now that uninformed voters have exactly the opposite preferences: they prefer candidate 1 in state 2 and candidate 2 in state 1 (example 2 for our purposes). Then, conditional on being pivotal they can be sure (in a symmetric equilibrium) that they are voting for the right candidate!

Feddersen and Pesendorfer (1999) generalises the model to the case of a continuum of types and states of the world– agents are differentially informed and perfect common values is not assumed. The probability of an informed voter (i.e. one who does not get a completely uninformative signal) is assumed to be strictly positive. They also assume that higher signals are more likely in higher states (strict monotone likelihood ratio property). A Poisson game with population uncertainty is used to model this, as in Myerson (1998B). They show that there will always be a symmetric equilibium in this game, and if information is not perfect, it supports abstention. Necessary and sufficient conditions for abstention in this setting are provided as are comparitive statics results which show that increased levels of education are consistent with lower participation (among uneducated voters). However, abstention decreases with the closeness of the election. Moreover the election works well as an information aggregation device, i.e. even though some voters are uninformed, the candidate chosen in large elections is the one preferred by the majority in that state of the world. Example 2 is ruled out in their general model – the crucial assumption ruling it out is Assumption 1: that all voters prefer the higher candidate in the higher state.

7 Concluding Remarks

In this paper we survey some of the most important contributions to the theoretical literature explaining (or not) voter turnout. We classified these theories in some broad headings in Section 1. Among instrumental theories, as best put in the Palfrey-Rosenthal (1983) paper, there are two driving forces in a rational voters calculation when deciding whether to turn out or not – free riding and competition. The free riding incentive dominates in the older decision theoretic models of Tullock (1968) and Riker and Ordeshook (1968), since the probability of affecting the outcome is assumed to be inversely related to the number of voters. In the game theoretic models (Palfrey- Rosenthal, 1983, 1985), where voters are assumed to be fully rational, there are some equilibria that have high turnout even for large populations. However, it is only the low turnout equilibria that are robust to small changes in beliefs as shown by PR (1985), Myerson (1998B), and Demichelis and Dhillon(2001). Evolutionary models (Conley et al, 2001 and Sieg and Schulz, 1995) do not really fare any better, and cannot predict high levels of turnout unambiguously i.e without some initial condition being satisfied. Shachar and Nalebuff (1999) suggest that voters are too small to make strategic decisions and it is really the parties' efforts that cause high or low turnout in any particular election. Thus parties help voters to coordinate votes. Feddersen and Pesendorfer (1996) do not believe that cost is such an important factor in the voting decision – they show that voters have information based reasons to vote or abstain. All models seem agreed that some expressive factor is causing the relatively high levels of turnout. In the third section we focused on some of the important expressive theories of voter turnout. The more sophisticated theories postulate that expressive benefits are related to other voters participation and can generate bandwagon effects among other useful comparitive static results.

References

[1] Aidt, T., 2000,"Economic voting and Information", Electoral Studies, 19, pp. 349–62.

[2] Aldrich, J. A., 1997, "When is it rational to vote?", in Dennis Mueller (ed.), *Perspectives on Public Choice, a Handbook*, Cambridge University Press, Cambridge, 373–390.

[3] Conley, J., A. Toossi and M.Wooders, 2001, "Evolution and Voting: How Nature makes us public spirited", Warwick Economic Research Papers No.601.

[4] Downs, A., 1957, "An economic theory of democracy", New York, Harper.

[5] Demichelis, S. and A. Dhillon, 2001, "Learning in Elections and Voter Turnout Equilibria", Warwick Economic Research Papers No.608.

[6] Feddersen, T. J. and W. Pesendorfer, 1996, "The swing voter's curse", The American Economic Review 86, 408–424.

[7] Feddersen, T. J. and W. Pesendorfer, 1999, "Abstention in elections with asymmetric information and diverse preferences", American Political Science Review 93(2), June 1999, 381–398.

[8] Feddersen, T. J. and A.Sandroni, 2001, "A Theory of Ethics and Participation in Elections", http://www.stanford.edu/group/SITE/sandroni.pdf.

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

[10] Fiorina, M. P., 1976, "The voting decision: instrumental and expressive aspects", Journal of Politics 38, 390–413.

[11] Fudenberg, D. and David K.Levine, 1998, "The theory of learning in games", MIT Press.

[12] Grofman, B, 1979, "Abstention in two-candidate and three-candidate elections when voters use mixed strategies", Public Choice 34, 187–200.

[13] Harsanyi, 1980,"Rule Utilitarianism, Rights, Obligations and the theory of rational behaviour", Theory and Decision, Vol.12, pp. 115–33.

[14] Kirchgassner, G., 1992, “Towards a theory of low cost decisions”, European Journal of Political Economy 8(2), 305–320.

[15] Kirchgassner, G., and W.W.Pommerehne, 1993, “Low cost decisions as a challenge to public choice”, Public Choice 77(1), 107–117.

[16] 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 University Press, 54–80.

[17] Ledyard, J. O., 1984, ”The pure theory of large two-candidate elections”, Public Choice 44, 7–41.

[18] Luce, R. D. and H. Raiffa, 1957, ”Games and Decisions”, New York, Wiley.

[19] Myerson, R., 1998A, ”Extended Poisson games and the Condorcet jury theorem”, Games and Economic Behaviour 25(1998), 111-131.

[20] Myerson, R., 1998B, ”Population uncertainty and Poisson games”, International Journal of Game Theory 27, 375–392.

[21] Myerson, R., 1999, ”Large Poisson games”, Journal of Economic Theory 94 (2000), 7–45.

[22] Palfrey, T. and Rosenthal, H., 1983, ” A strategic calculus of voting”, Public Choice 41, 7–53.

[23] Palfrey, T. and Rosenthal, H., 1985, ”Voter participation and strategic uncertainty”, American Political Science Review 79, 62–78.

[24] Riker, W.A. and P.C. Ordeshook, 1968, ”A theory of the calculus of voting”, American Political Science Review 62, 25–42.

[25] Schuessler, Alexander A., 2000, “A Logic of Expressive Choice”, Princeton University Press.

[26] Shachar, R. and B. Nalebuff, 1999, “Follow the Leader: Theory and Evidence on Political Participation”, American Economic Review 89(3), 525–549.

[27] Sieg, G. and C. Schulz, 1995, ”Evolutionary dynamics in the voting game”, Public Choice 85, 157–172.

[28] Strom, G. S., 1975, "On the apparent paradox of participation: a new proposal", American Political Science Review 69, 908–913.

[29] Tullock, G., 1968, "Toward a mathematics of politics", University of Michigan Press.

8 Appendix

8.1 Demichelis and Dhillon (2001)

Since this paper offers an equilibrium selection of the basic PR model, we first demonstrate the problem that arises with (multiplicity of) 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. The symmetric mixed strategy equilibrium is denoted q which satisfies the following equation:

$$\begin{aligned} 2c &= \sum_{k=0,\ldots,N-1} C_{N-1,k} C_{N,k} q^{2k} (1-q)^{2N-2k-1} \\ &+ \sum_{k=0,\ldots,N-1} C_{N-1,k} C_{N,k+1} q^{2k+1} (1-q)^{2N-2k-2} \end{aligned} \tag{6}$$

If $0 < c < 1/2$ then this equation has either no solution, 1 solution or two. The graph below shows this:

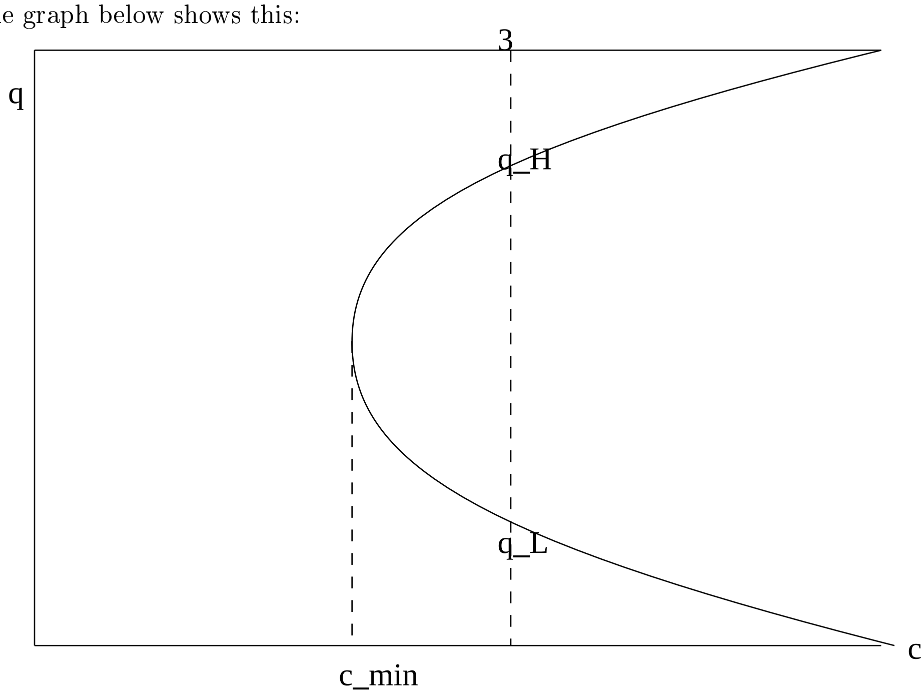


FIGURE 1

The example shows that as the cost increases beyond c_{min} there are two types of mixed strategy equilibria: one where almost everyone votes and one with almost no one voting. 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.

Assume now that voters 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.
The type of learning by adjustments is the standard one assumed in several contexts. The details can be specified in several different ways, for example see Fudenberg and Levine (1998). In general they can all be described by the differential equation[12]:

$$\frac{dq}{dt} = K(q, c) \tag{7}$$

and sign $K(q,c)$ =sign$(f(q) - 2c)$ The functional form of $K(q)$ depends on the particular model of learning. However the main result holds for all of them that satisfy the assumptions on $K(q)$. The main result of course is that the low turnout equilibria are the only ones that are asymptotically stable and with a basin of attraction that increases as the size of the population increases.

This basic idea is applied to the incomplete information case as well, and selects the low turnout equilibria.

8.2 Myerson (1998b, 1999): Population Uncertainty

We present the main features of Myerson's model and some of his (mainly technical) results below, as these are very useful for voting games:

The total number of players is assumed to be a Poisson random variable with mean n: any integer number of players, k may arise with probability

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

$$\frac{e^{-n}n^k}{k!}$$

T is the set of possible types indexed by t. Each player is independently assigned one type according to the probability distribution $r = (r(t))_{t\in T}$.

Consider a vector y such that each of its elements represents the number of players of each type, that is, y is a *type profile.* Since the types are allocated independently, the probability of the random vector y, $Q(y)$, is:

$$Q(y) = \prod_{t\in T} Q(y(t))$$

where each $Q(y(t)) = \frac{e^{-nr(t)}(nr(t))^{y(t)}}{(y(t))!}$, i.e the probability that a randomly selected player is of type t. The utility of each player depends on his type, his action, and the number of players who choose each possible action in the set of actions C (that is, the *action profile*):

$$U : Z(C)\ X\ C\ X\ T \longmapsto R$$

where $Z(C)$ denotes the set of action profiles.

The Poisson game is fully described by (T, n, r, C, U).

The *strategy function is* $\sigma : T \longmapsto \Delta(C)$ that prescribes each type a certain strategy. That is, we force every player of the same type to follow the same strategy, since there is no individual characteristic of each agent that allows them to be distinguished.

The *decomposition property* of a Poisson random variable[13] guarantees that:

1. $\widetilde{Y}(t)$, the random number of players of type t is distributed according to a Poisson distribution with parameter $nr(t)$.

2. if players choose actions independently according to the strategy function $\sigma(c|t)$, then the number of players of type t who play action c is itself a Poisson with expected value $nr(t)\sigma(c|t)$.

[13]Roughly, it states that in a Poisson population in which each member is independently assigned some characteristic according to a known probability distribution, the number of elements with each characteristic s_0 is itself a Poisson with mean equal to the mean of the population times the probability of s_0.

To deduce the distribution of $\widetilde{X}(c)$ the total number of players who choose action c, we apply the *aggregation property* of the Poisson distribution — the sum of independent Poisson random variables is itself a Poisson with expected value equal to the sum of the means. $\widetilde{X}(c)$ is a Poisson with expected value $\sum_{t \in T} nr(t)\sigma(c|t)$.

A game is said to have the *independent-actions property* if and only if, for every strategy function σ the random variables $\widetilde{X}(c)$ for all $c \in C$ are independent random variables. Let $Z(T) =$ set of possible type profiles in the game. Thus, for any vector y in $Z(T), Q(y)$ denotes the probability that for every $t \in T$, $y(t)$ is the number of players of type t in the game. The following theorems are proved in Myerson (1998b):

Theorem 1 *Suppose that the game (T, Q, C, U) satisfies the independent-actions property. Then for every strategy function σ, each $\widetilde{X}(c)$ must be a Poisson random variable with mean $\sum_{t \in T} \sigma(c|t) \sum_{y \in Z(T)} Q(y)y(t)$ and the total nuber of players in the game must be a Poisson random variable.*

Another important property of a Poisson game is *environmental equivalence.* It relates to the perception each player formulates about his environment — the number of players in the game, excluding himself — as compared to the outside viewer assessment of the probability distribution over the number of players.

The fact that the player excludes himself from the environment makes it seem smaller, whereas the fact that he finds himself recruited as a player consists in a private signal of a bigger game. In Poisson games, this two factors exactly cancel each other, such that the player considers the same probability distribution over the types of the other players as the game theorist over the whole game.

Theorem 2 *A game with population uncertainty satisfies environmental equivalence if and only if it is a Poisson Game.*

The third theorem presented by Myerson concerns existence.

Theorem 3 *Any game with population uncertainty (T, Q, C, U) (where T and C are finite and U is bounded) must have at least one equilibrium.*

Myerson (1998A, theorem 1) provides an existence result for extended Poisson games, in which global uncertainty is introduced.

When we defined the strategy function, it was said that all the players of the same type must choose the same equilibirum behavior. It would be problematic if the splitting of types into sub-types with the same payoff would change the equilibria of the game.

However, this is not the case.

Theorem 4 *If the Poisson game Γ' is derived from the Poisson game Γ by splitting a type into two utility-equivalent types that have the same total probability, then the set of marginal distributions on the action set C that are generated by equilibria is the same in Γ and Γ'.*

Myerson (2000) derives two useful formulas to compute pivot probabilities in large Poisson games. These are then used to derive some theorems in Ledyard (1984), imposing population uncertainty, in a much simpler and cleaner way.

8.3 Feddersen and Pesendorfer (1996)

The set of states is $Z = \{0, 1\}$, the set of candidates is $X = \{0, 1\}$ and the type space is $T = \{0, 1, i\}$. Type-0 and type-1 agents are partisans: they prefer candidate 0(1) whatever the state of nature. Type i agents are independents (*with common values*): they like candidate 0(1) in the state of the world 0(1). The utility of an independent agent is:

$$U(x, z) = \begin{cases} -1 \text{ if } x \neq z \\ 0 \text{ if } x = z \end{cases}$$

Nature chooses the state 0 with probability α and the state 1 with probability $1-\alpha$. Without loss of generality, it is assumed that $\alpha \leq 1/2$, state 1 is at least as likely to happen as state 0. This probability distribution is common knowledge. Nature selects a player with probability $(1 - p_\phi)$ from a population of $N + 1$: the number of voters follows a binomial distribution with parameters $N + 1$ and $(1 - p_\phi)$. Then the agent is of type j with probability $\frac{p_j}{1-p_\phi}$: the number of type j voters follows a binomial distribution $(N + 1, p_j)$. The vector of probabilities $p = (p_i, p_0, p_1, p_\phi)$ is common knowledge.

Every player will get a message from the set $M = \{0, \alpha, 1\}$. The message represents the probability of state 0 occurring. Agents who receive message $m = \alpha$ do not update the common prior and are therefore called the uninformed agents. The probability that an agent is informed is q.

Each agent may take an action $s \in \{\phi, 0, 1\}$, where ϕ represents abstaining and 0(1) voting for candidate 0(1).

Each agent chooses a strategy depending on his type and private signal.

A pure strategy is a map $s : T\ X\ M \longmapsto \{\phi, 0, 1\}$.

A mixed strategy is a map $\tau : T\ X\ M \longmapsto [0, 1]^3$, where τ_s is the probability of taking action s.

The set of equilibria is restricted to symmetric Nash equilibria, that is, agents of the same type that receive the same message will follow the same strategy.

Since the number of players ranges from 0 to $N+1$, there is a strictly positive probability that any agent is pivotal. With costless voting, type-0, type-1 and informed independents have a strictly dominant strategy to vote, since there is a positive probability that one will influence the outcome and elect his favorite candidate.

We now analyze the behavior of uninformed independent agents (UIA's).

Note that if the state of nature does not correspond to the candidate, the informed independents will never vote for that candidate: only UIA's and partisans may do it. Let $\sigma_{z,x}(\tau)$ denote the probability a random draw of nature results in a vote for candidate x if the state is z. This is:

$$\sigma_{z,x} = \begin{cases} p_x + p_i(1-q)\tau_x & \text{if} \quad z \neq x \\ p_x + p_i(1-q)\tau_x + p_i q & \text{if} \quad z = x \end{cases} \tag{8}$$

¿From (8), it is clear that the probability of a mistaken vote is always smaller than the probability of a correct vote. The only way a random draw by nature may result in abstention is if no agent is drawn or if an UIA is drawn and he abstains — type-0, type-1 and informed type-i always vote.

This probability is independent of the state of nature:

$$\sigma_{0,\phi}(\tau) = \sigma_{1,\phi}(\tau) = \sigma_\phi(\tau) = p_i(1-q)\tau_\phi + p_\phi \tag{9}$$

An UIA will be pivotal in the situation of a tie or when a candidate is ahead of the other by exactly one vote. The probability of a tie in state of nature z and given UIA's are playing strategy τ is:

$$\pi_t(z,\tau) = \sum_{j=0}^{N/2} \frac{N!}{j!j!(N-2j)!} \sigma_\phi(\tau)^{N-2j} (\sigma_{z,o}(\tau)\sigma_{z,1}(\tau))^j$$

Candidate x receives exactly one vote less than candidate y with probability:

$$\pi_x(z,\tau) = \sum_{j=0}^{N/2-1} \frac{N!}{(j+1)!j!(N-2j-1)!} \sigma_\phi(\tau)^{N-2j-1} \sigma_{z,y}(\tau)(\sigma_{z,x}(\tau)\sigma_{z,y}(\tau))^j \tag{10}$$

In order to determine the best response for UIA's we need to define the differences in expected utility of each action.

	Vote for 1	Vote for 2	Difference
State 1	$\pi_t 0 + \pi_1 \frac{1}{2}(-1) + \pi_1 \frac{1}{2} 0$	$\pi_t \frac{1}{2} 0 + \pi_t \frac{1}{2}(-1) + \pi_1(-1)$	$(\pi_1 + \pi_t)\frac{1}{2}$
State 0	$\pi_t(-1) + \pi_1 \frac{1}{2}(-1) + \pi_1 \frac{1}{2} 0$	$\pi_t \frac{1}{2} 0 + \pi_t \frac{1}{2}(-1) + \pi_1 0$	$-(\pi_1 + \pi_t)\frac{1}{2}$

Table 5: Expected Utilities

Let's take, for instance, voting for candidate 1 and abstaining, in Table 5. With analogous arguments, we get the following relations:

$$Eu(1,\tau) - Eu(\phi,\tau) = \frac{1}{2}[(1-\alpha)(\pi_t(1,\tau) + \pi_1(1,\tau)) - \alpha(\pi_t(0,\tau) + \pi_1(0,\tau))] \tag{11}$$

$$Eu(0,\tau) - Eu(\phi,\tau) = \frac{1}{2}[\alpha(\pi_t(0,\tau) + \pi_0(0,\tau)) - (1-\alpha)(\pi_t(1,\tau) + \pi_0(1,\tau))] \tag{12}$$

$$Eu(1,\tau) - Eu(0,\tau) = (1-\alpha)[\pi_t(1,\tau) + \frac{1}{2}(\pi_1(1,\tau) + \pi_0(1,\tau))] - \alpha[\pi_t(0,\tau) + \frac{1}{2}(\pi_1(0,\tau) + \pi_0(0,\tau))] \tag{13}$$

When the voter will be indifferent between the two candidates (for a fixed τ) it implies that no matter what the state of nature is, if he votes for candidate 1 or candidate 2 the conditional probability that he makes a mistake given that he is pivotal is higher than the probability that he votes correctly given that he is pivotal. This is the intuition underlying the first proposition.

Theorem 5 *Let $p_\phi > 0$, $q > 0$, $N \geq 2$ and N even. For any symmetric strategy profile τ in which no agent plays a strictly dominated strategy, $Eu(1,\tau) = Eu(0,\tau)$ implies $Eu(1,\tau) < Eu(\phi,\tau)$.*

Moreover since this is true for any symmetric strategy profile τ, it implies that the only equilibrium (when an uninformed voter is indifferent between the two candidates) is a pure strategy one where all uninformed voters abstain. Again, this does not imply that there are no other asymmetric equilibria where the swing voters curse does not hold.

In large elections, this result is generalised to show that whenever the proportion of UIA's is large enough to overcome the effect of partisans, some abstention will be predicted in (symmetric) equilibria. The intuition is as before: if the probability that a random draw leads to a mistaken vote in state 1 is bigger than in state 0, the probability that the state is 1 goes to 1 as the size of the electorate increases: i.e. the probability that the true state is 1 conditional on the UIA being pivotal goes to 1, so he should vote for 1 and vice versa if the probability of a mistaken vote is larger for state 0. Because of the strategic abstention by uninformed voters, this implies that for sufficiently large populations, the election fully aggregates information with a very high probability (Proposition 4 in their paper).

8.4 Conley Toossi and Wooders (2001)

The model has a continuum of agents who are of two types designated "high"(H) and "low"(L). H voters have the same preferences and have a higher propensity to vote while L voters have a low propensity to vote (and have identical preferences). The share of type j in the population is denoted S_j. In each period agents vote on a randomly generated public proposal, which produces a cost or benefit, denoted B_j to each type of agent that is uniformly distributed on the interval $[-1, 1]$. The preferences of the two types of agents are positively or negatively correlated: thus $B_L \in \{B_L^p, B_L^n\}$, where B_L^p denotes positively correlated preferences, and B_L^n denotes negatively correlated preferences. The degree of correlation is measured by a parameter α. The propensity to vote is denoted $V_j \in (0, 1)$. The relative public spiritedness of the two types is the ratio, $\frac{V_H}{V_L} \geq 1$. Voters turnout is assumed to be given by the function $TO_j = S_j V_j B_j$, which can be positive or negative (it measures the number of yes votes if positive and no votes if negative). The cost of casting a vote is the same for everyone, and denoted C. The expected cost is given by $C_j = V_j |B_j| C$. Expected Payoff can then be calculated for each type and each period.

The evolution of shares over time is modelled using replicator dynamics. They find that depending on parameter values, there are three possible outcomes, one of which has $S_H = 1$ as the only stable steady state. Theorem 2 however shows that likelihood of the evolutionary success of the high types depends on parameters like costs (negatively related), relative public spiritedness (positively related), the degree of positive correlation in the preferences of the two types (negatively related because of the free rider problem). The crucial assumption they need for the high type to do well evolutionarily, is that all high types have the same preference over public policy and similarly for the low types.