DISCUSSION PAPER SERIES

No. 3126

ON THE (NON) PARADOX OF (NOT) VOTING

Micael Castanheira

PUBLIC POLICY

Centre for Economic Policy Research

www.cepr.org

Available online at: **www.cepr.org/pubs/dps/DP3126.asp**

ISSN 0265-8003

ON THE (NON) PARADOX OF (NOT) VOTING

Micael Castanheira, ECARES, Université Libre de Bruxelles and CEPR

Discussion Paper No. 3126
January 2002

Centre for Economic Policy Research
90–98 Goswell Rd, London EC1V 7RR, UK
Tel: (44 20) 7878 2900, Fax: (44 20) 7878 2999
Email: cepr@cepr.org, Website: www.cepr.org

CEPR Discussion Paper No. 3126

January 2002

ABSTRACT

On the (Non) Paradox of (Not) Voting*

Why do people vote? This question received a lot of attention for more than thirty years, and yet remains unanswered. In this Paper, we take stock of existing empirical regularities and argue that we can use them to improve the model of instrumental voting. Once this is done, we show that purely rational/instrumental factors actually explain a large fraction of turnout variations. To perform our analysis, we use Myerson's (1997, 2000) advances on Poisson Games and generalize the Riker and Ordeshook (1968) seminal model of instrumental voting. Applying our results to US data, we show how our model can explain several stylized facts, like the secular fall in turnout rates in the US.

JEL Classification: C72, D72 and D81
Keywords: paradox of voting, poisson gamex and rational voter hypothesis

Micael Castanheira
ECARES
CP114
50 Avenue Roosevelt
1050 Bruxelles
BELGIUM
Tel: (32 2) 650 4467/3
Fax: (32 2) 650 4474
Email: mcasta@ulb.ac.be

For further Discussion Papers by this author see:
www.cepr.org/pubs/new~dps/dplist.asp?authorid=126467

* A previous version of the Paper circulated under the title ‘Should I Stay or Should I Go: Poisson Games and the Paradox of Voting’. I wish to express my gratitude to Juan Carrillo who taught me how to deal with Fishes, Gerard Roland for pushing this project so much and Laura Bottazi, Isabelle Brocas, Francesco Corrielli, Mathias Dewatripont, Francesco Giavazzi, Eliana La Ferrara, Giovanni Peri, Michele Polo, Anouk Riviere, Georges Siotis, Guido Tabellini, Françoise Thys and Philippe Weil for their helpful comments and encouragements. The comments of the audiences at Université Catholique de Louvain and at IGIER were also very helpful. All remaining errors will disappear at a Poisson rate determined by the average sense of duty of the readers (if any).

Submitted 27 November 2001

NON-TECHNICAL SUMMARY

Why do people vote? Because they feel they have to? Or because they hope to affect the outcome of the election? Arguably, each single ballot only has a negligible weight on the outcome of the election when turnout is large. This in turn suggests that only psychological factors could explain the decision to turn out.

Such an argument has been repeatedly stressed in the literature, to turn down the so-called 'instrumental' approach of voting. In contrast to such conclusions, this paper argues that instrumental motivations (meaning that voters use their ballot as an instrument to affect the outcome of the election) remain a strong determinant to the voting decision. To reconcile the instrumental model of voting with empirical regularities, however, it proves necessary to test each of the implicit assumptions that are generally made. Using existing empirical results, we show that (at least) three of those assumptions are clearly not admissible, and adapt the model accordingly.

First, an oft-used simplifying assumption is that all voters have a common (positive) cost of voting. This clearly limits potential turnout and biases the model towards large abstention rates. In opposition to this assumption, the empirical literature shows that voting costs widely differ across voters. Hence, we model the cost of voting as a random variable and show how the predictive power of the model is affected.

Second, the preferences of the electorate are often modeled as being common knowledge. We show that such an assumption is incompatible with the population's betting behaviour prior to the elections. Accordingly, we generalize the model to incorporate aggregate uncertainty and show that the relationship between (expected) vote shares and turnout can be better explained in this way.

Finally, the outcome of the election is generally modeled as an all-or-nothing composition of the platforms proposed by the parties. That is, if a party announces a platform and wins the elections, it exactly implements what it had announced. Again, it appears that this assumption is far-fetched and that post-electoral bargaining tends to moderate parties' positions when the margin of victory is small. Incorporating this last observation in the model, we show that the power of one vote is larger than usually thought, and that turnout rates should accordingly increase.

Having derived those analytical results, we calibrate the model on US presidential election data and show that, once properly generalized, the model can actually explain why turnout rates fell as they did since 1960. Hence, we conclude that instrumental motivations may actually explain a substantial

fraction of voting behaviour, and thus usefully complement 'psychological' approaches.

1 Introduction

"To measure the political preferences of legislators by their votes at year 1 and, then, to use those very votes to explain their behavior at year 1 is to argue that legislators vote the way they do because they vote the way they do."

(Epstein and Mershon, 1996 pp261-262)

Any recipe for a representative democracy requires several ingredients: elections as a blender, a few candidates for the topping, and voters (lots of them) for the dough. Then, electoral rules determine how voters' ballots translate in a given representation and thus how candidates compete, how budget will be spent, and so on. Electoral rules thus determine the value of voting, which should in turn influence turnout rates.

This paper studies precisely this last problem: why do people vote? Explaining turnout is extremely difficult, because each single vote has a negligible weight on the outcome of the election and, if this weight is so low, why not always abstain? Put differently, the mere fact that turnout can be large (105 million voters turned out at the U.S. presidential elections, in 2000) sounds as a paradox. Wouldn't voting be one of the symptoms of people's irrationality?

The literature trying to reconcile (bounded) rationality with large turnouts is extremely vast, and still provides little agreement, except on one point: turnout does not seem to be motivated by instrumental arguments. "Instrumental voting" means that a voter goes to the polls in order to affect the outcome of the election. That is, voters use elections as an instrument to influence policy. Riker and Ordeshook (1968) were the first ones to identify the exact weight of a vote. As they show, a vote is only effective if it is *pivotal*, i.e. if it changes the outcome of the election. Therefore, the instrumental value of a ballot is proportional to the probability that it is pivotal. In a large population, this probability is extremely small and cannot match any positive cost of voting. Consider for instance the 1996 presidential elections in the U.S.: 96 million voters turned out; Democrats won with 49.2% and Republicans lost with 40.7% of the votes. In that example, the probability of a tie (a single vote can only be pivotal if there is a tie) was around $10^{-169814}$. Could a rational voter ever turn out on that premise?

Since this seems unlikely, research has explored other rationales for the turnout/abstention decision. As alternatives to the rational-instrumental voter hypothesis, one could imagine that voters are actually not rational (e.g. Ferejohn and Fiorina (1974) assume that, since voters cannot evaluate probabilities, they behave as if this probability was equal to one), or at least not instrumental. Under the latter assumption, variations in turnout are caused by some mechanism that directly affects the **cost** of voting. For instance, Aldrich (1993) argues that parties voluntarily reduce the cost of voting if the election is very important to them; Harbaugh (1996) presents a model where voters turn out because of peer pressure; or Feddersen and Pesendorfer (1996) argue that voters may instead rationally abstain when they are ill-informed about the relative merits of the candidates.[1]

In other words, there is already a wealth of approaches to explain the voting paradox, which is why we do not want to offer "yet another" theory of voting. Instead, the aim of the present paper is to deepen the analysis of the former hypothesis, which states that voters are purely instrumentally motivated. We argue that this approach may have been discarded too easily and could actually better complement the above theories than what is generally thought. For instance, a lot of weight has been put on the psychological motivations for turning out, but couldn't these motivations be somewhat driven by more rational motivations? The view offered by this paper is that an appropriate generalization of *a)* the cost of voting, *b)* the information set of voters and *c)* the institutional set-up, restore the credibility of the rational-instrumental approach.

In this paper, we use Myerson's (1997, 2000) advances on *Poisson Games* as a tool that allows us to deepen the analysis of the voting paradox. Note that *per se*, Poisson Games do not provide any solution to the problem. For instance, Myerson (1997) shows in a numerical example that his model cannot explain why more than 64 voters turn out. However, Poisson Games make the analysis much more tractable, and hence allows us to easily generalize several (hidden but objectionable) assumptions that are usually made in models of instrumental voting.

[1]For a more complete overview of the literature, see for instance Mueller (1989) or Aldrich (1993, 1997).

To this end, section 2 presents the standard model of voting and highlights three important assumptions that bear most of the responsibility for its low predictive power. After highlighting these assumptions and presenting the results that they generate, we relax them one after the other. In section 3, we use existing empirical evidence to model the cost of voting as well as the information set of the voters, in a more sensible way. In section 4, we exploit Stigler's (1972) and Razin's (2000) results to study the case in which the "mandate" given to the politician matters in determining the implemented policy.

Then, in section 5, we apply those results to concrete numerical examples and show that, when we apply our theoretical results to U.S. data, most of the fall in turnout since 1960 can actually be explained by our model. Finally, section 6 concludes. Some of the computations are relegated to the appendix, which also presents a short survey of some of the empirical regularities highlighted in the literature.

2 The model

We consider a game in which voters must decide whether to turn out or to abstain in the upcoming election, and we shall be looking for Bayesian Nash Equilibria of this game. Our reference model is that of Riker and Ordeshook (1968) – RO for short. In that model, the utility of the voters depends both on the platform that is implemented and on their decision to turn out or to abstain. Assuming that the policy space is unidimensional, voter i's preferred platform can be represented by some real number θ_i. Similarly, the implemented Policy is defined as another real number P. We define the utility of voter i as:

$$U(\theta_i, \mathsf{P}, v_i) = -(\theta_i - \mathsf{P})^2 + c(v_i), \tag{1}$$

in which the first member on the right-hand side of (1) denotes the utility derived from the implemented policy P and $c(v_i)$ represents the cost of voting.

Assume that there are two candidates, with policy platforms A and B. Denoting abstention by $\varnothing$, voters can choose one in three actions: vote for A (action A), vote for

B (action B) or abstain (action $\varnothing$):

$$v_i \in \{A, B, \varnothing\}.$$

Introducing the cost of turning out is made by posing:

$$c(A) = c(B) = -c_i,$$
$$c(\varnothing) = 0.$$

Thus, if $c_i > 0$, the voter has a positive cost of voting and if $c_i < 0$, the voter instead has a positive cost of abstaining (c_i is thus the gross cost of voting *net of* any sense of duty).[2]

From those definitions, one can immediately derive the value of a ballot. By turning out, there is some probability that the voter is pivotal, and thus affects the outcome of the election. Hence, the value of the ballot depends on this *pivot probability*:

$$\begin{aligned} \mathsf{E}\left[U(\theta_i, \mathsf{P}, A) - U(\theta_i, \mathsf{P}, \varnothing)\right] &= \mathsf{E}_{\mathsf{P}}\left[-(\theta_i - \mathsf{P})^2 | v_i = A\right] - \mathsf{E}_{\mathsf{P}}\left[-(\theta_i - \mathsf{P})^2 | v_i = \varnothing\right] - c_i \\ &= \Pr\left[piv_{AB}\right]\left(-(\theta_i - A)^2 + (\theta_i - B)^2\right) - c_i, \end{aligned}$$

in which piv_{AB} denotes the event that voter i's ballot is pivotal in favor of A (see RO for more detail). That is, had she abstained, platform B would have been implemented, whereas her additional vote changes the outcome, and platform A is implemented instead of B. Similarly, the value of a vote for B is given by:

$$\mathsf{E}\left[U(\theta_i, \mathsf{P}, B) - U(\theta_i, \mathsf{P}, \varnothing)\right] = \Pr\left[piv_{BA}\right]\left((\theta_i - A)^2 - (\theta_i - B)^2\right) - c_i,$$

in which piv_{BA} denotes the event that voter i's ballot is pivotal in favor of B.

To shorten subsequent notations, we define $W(\theta_i, v_i)$ as the value of action $v_i = A, B$ gross of the cost of voting:

$$W(\theta_i, v_i) = \Pr\left[piv_{v_i p}\right]\left(-(\theta_i - v_i)^2 + (\theta_i - p)^2\right), \text{ for } v_i, p \in \{A, B\}, \; v_i \neq p \qquad (2)$$

[2]The literature generally presents the utility of a voter under the following shape: $P \times B + D - C$, in which P represents the pivot probability, B the benefit of changing the outcome of the election, D the sense of duty, which gives some utility to turning out (or cost to abstaining) and C is the cost of going to the polls. In our modelling, $c_i = C_i - D_i$, because costs and benefits can be idiosyncratic.

Using (2), the voter abstains as soon as:

$$\max_{v_i \in \{A,B\}} W(\theta_i, v_i) \leq c_i, \tag{3}$$

which gives us a simple rule to identify which voters turn out or abstain.

Now, we turn our attention to the aggregate distribution of preferences in the population. We distinguish between left-wing voters with preferences $\theta_i \leq 0$ and right-wing voters with preferences $\theta_i \geq 0$. Setting $A < 0 < B$ and $A = -B$, all voters to the left of 0 prefer platform A to platform B, and conversely for right-wing voters. We also define $\gamma_A \in (0,1)$ as the fraction of the population that is left-wing and $\gamma_B = 1 - \gamma_A$ as the fraction of right-wing voters. Then, we assume that preferences are uniformly distributed over $[-\alpha, 0]$ and over $[0, \alpha]$. Put differently:

$$\begin{aligned} \forall \theta_i &\in [-\alpha, 0),\ f(\theta_i) = \frac{\gamma_A}{\alpha}, \\ \forall \theta_i &\in [0, \alpha],\ f(\theta_i) = \frac{\gamma_B}{\alpha}, \\ \forall \theta_i &\notin [-\alpha, \alpha],\ f(\theta_i) = 0. \end{aligned} \tag{4}$$

This implies that $\int_{-\alpha}^{\alpha} f(\theta_i)\, d\theta_i = 1$ for any $\gamma_A > 0$. Moreover, combining (1), (2) and the assumed positions of the parties yields:

$$\begin{cases} W(\theta_i, A) = 4 \Pr[piv_{AB}]\ \theta_i\, A \\ W(\theta_i, B) = 4 \Pr[piv_{BA}]\ \theta_i\, B. \end{cases} \tag{5}$$

The size of the population, $\tilde{n}$, is also random and follows a Poisson distribution of argument λ: $\tilde{n} \sim \mathcal{P}(\lambda)$, where λ is the *expected* size of the population (see Myerson, 1997, 2000). Once the number of voters has been drawn, each voter is assigned a type θ_i by *i.i.d.* draws out of the piece-wise uniform distribution (4).

2.1 Base Model: Key Assumptions

We have now all ingredients ready to analyze voters' behavior. It only remains to characterize the institutional set-up and identify some key assumptions that are made in the literature. As we shall see below, under those assumptions, the recipe fails to explain turnout. In order to stress the weight of these assumptions on the results of the model let us spell them out clearly here. Denote by n_p the number of votes obtained by party p. The three working assumptions in the literature are:

Key Assumption 1: the election rule is "first-past-the-post", and the implemented platform is A if $n_A > n_B$ and B if $n_A < n_B$.

Note that some tie-breaking rule must be devised. In Great-Britain, the Queen (or the King) decides who wins when there is a tie. The more standard modelling assumption is that each candidate be elected with probability one half. To simplify computations, we assume here that A wins for sure when $n_A = n_B$ (our results extend to the other rules but expressions would become heavier). Using the Poisson distribution and this tie-breaking rule, we can use Myerson's results to derive the pivot probability that one single vote is pivotal if all voters turn out: for λ sufficiently large,[3]

$$\Pr\left(piv_{BA}|\gamma_A\right) = \frac{e^{-(\sqrt{\gamma_A}-\sqrt{\gamma_B})^2.\lambda}}{2\sqrt{\pi\,\lambda}\,\sqrt[4]{\gamma_A{\cdot}\gamma_B}} \tag{6}$$

$$\Pr\left(piv_{AB}|\gamma_A\right) = \sqrt{\frac{\gamma_B}{\gamma_A}}\Pr\left(piv_{BA}\right) \tag{7}$$

Key Assumption 2: the distribution of preferences is common knowledge.

In the specification of our model, this implies that the share of the population who prefers A, γ_A, is constant and known.

Key Assumption 3: the cost of voting is positive and identical for all voters:

$$c_i = c > 0,\ \forall i.$$

Note that, though this third assumption is standard in the theoretical literature, it is completely at odds with empirical measures of the cost of voting, which has been proved to vary substantially across voters, population classes, education levels, etc...[4] Again, we are not trying to make assumptions realistic for the time being; each of those

[3] See appendix 1 for a derivation.

[4] Such empirical regularities are reviewed in appendix 4.

assumptions will be relaxed in the next sections. However, it is by comparing the results obtained with and without those assumptions that we will identify their role and hence how we can partly solve the paradox of voting.

Now, under those assumptions, the expected turnout is easily identified: amongst voters, those with the most intense preferences are the most extreme types: $-\alpha$ and α. The value of their ballot is respectively:

$$\begin{cases} W(-\alpha, A) = 4 \Pr[piv_{AB}] \ \alpha \ |A|, \\ W(\alpha, B) = 4 \Pr[piv_{BA}] \ \alpha \, B. \end{cases}$$

Since they turn out if and only if (3) is violated, it is straightforward to see that:

Lemma 1 *Under Assumptions 1-3, a necessary condition to observe positive turnout is*

$$\Pr[piv_{AB}] > \frac{c}{4\alpha \ |A|} \ \text{and} \ \Pr[piv_{BA}] > \frac{c}{4\alpha \, B}$$

Proof. Immediate. ■

Since those pivot probabilities, by (6) and (7), are decreasing in the number of voters who turn out, this also imposes an upper bound on expected turnout. Our first proposition derives the (Bayesian Nash) Equilibrium of this voting game:

Proposition 1 *Under Assumptions 1-3,*

1) *only sufficiently extremist voters turn out;*

2) *the share of votes received by the winning party is smaller than its actual support in the population;*

3) *for very large populations, equilibrium turnout $\bar{n}$ is independent of the actual size of the population;*

4) *turnout can be large iff the share of votes going to each party is expected to be 50%.*

Proof. See Appendix 2. ■

This proposition relies on a very simple intuition: since the cost of voting is positive and identical for every voter, the only parameter that differentiates voters is their ideology. As preferences are concave by assumption, extremists are more involved than moderate voters, and thus more likely to turn out. Still, for all voter types, the value of a ballot is decreasing in turnout size, which imposes an upper limit to the expected number of voters who turn out. Put differently, under Assumptions 1-3, the model displays very limited predictive power, which casts doubt on its validity.

Table 1 illustrates those results with the help of numerical simulations. We assume in these simulations that the (expected) size of the population goes to infinity, so that population size plays no role in determining absolute turnout. We compare predicted turnout for two sets of parameters: $4B/c = 1000$ or 2000 (that is, the stake conditional on being pivotal is 1000 or 2000 times larger than the cost of voting for a voter with $|\theta_i| = 1$), and we define $\bar{n}_p$ as the equilibrium expected number of votes for party p:

Table 1: Highest possible turnouts in the base case.

$\bar{n}_A/\bar{n}$	α	Highest possible expected turnout (in # votes) if $4B/c = 1000$	if $4B/c = 2000$
50%	5	3,978,872	15,915,158
	10	15,915,158	63,610,304
52%	5	4,271	5,035
	10	5,035	5,811
55%	5	844	968
	10	968	1,094

As this Table shows, if the shares of votes differ from 50%, none of the important variables in the model can explain turnout: the size of the population here is infinite, but about one thousand people vote if the shares are 55 vs. 45%. Doubling voters' valuation of the election (increasing α) or doubling the importance of changing the outcome of the election (B/c) only marginally affects turnout. Despite such results, and instead of discarding instrumental motivations, we shall discuss the assumptions we highlighted, relax them, and show how the limitations of this model can be overcome.

3 A Generalized Model of Costly Voting

The simplest way to generate higher turnout would be to assume that voters "like" voting. That is, we could find some reason why c_i (the cost of turning out for voter i) is non-positive. This would generate a large turnout (every voter with a non-positive cost of voting always finds it optimal to turn out), but lead to no *explanation* for turnout, nor provide any rationale for turnout variations. This is precisely the criticism by Epstein and Mershon in the opening quote of this paper: such an explanation would be tautological and the actual cost of voting could only be estimated ex-post. To avoid this problem, we shall relax the three Key Assumptions we have spelled out and maintain voters' preferences constant across elections. The goal of this method is precisely to derive turnout as a function of the characteristics of the election, not as a function of voters' preferences.

3.1 Generalizing the cost of voting

Key Assumption 3 states that the cost of voting is positive and identical for all voters. Here, we maintain the assumption that costs are non-negative, but we allow these costs to be different across voters:

Assumption 3': all voters have positive (net) voting costs: $c_i = C_i - D_i \geq 0$, $\forall i$. **Costs are idiosyncratic to each individual, and uniformly distributed between 0 and** C:

$$\tilde{c}_i \sim \mathcal{U}\left[0, C\right] \tag{8}$$

where $\mathcal{U}$ represents the uniform distribution and $C > 0$ is a parameter that represents the highest cost in the population, that is, $c_i \in \mathcal{C} = [0, C]$. To contrast this assumption with the initial one, Key Assumption 3 imposes $\mathcal{C} = [c, c]$, with $c > 0$. Now, under Assumption 3', type is two-dimensional: voters are defined both by their cost of voting c_i and their ideology θ_i. Hence, the Bayesian Nash equilibrium is characterized by a cost-cut-off function where voters with a cost lower than a given threshold turn out, while those with a cost higher than this threshold abstain. This threshold is endogenous in equilibrium and depends on θ_i.

Lemma 2 solves for the shape of this threshold function, and Figure 1 illustrates graphically the difference between the two equilibria. Note that we are not yet endogenizing overall turnout, because the other two assumptions are also crucial to this end.

Lemma 2 *Under assumptions 1, 2 and 3', for any given expected turnout, the probability that a voter does not abstain is proportional to her degree of extremism,* $|\theta_i|$.

Proof. By (3) and (5), a voter turns out if:

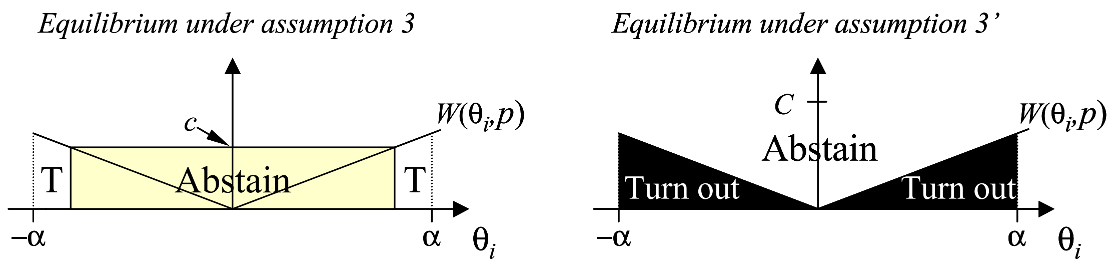
$$\begin{aligned} c_i &< W(\theta_i, B) = 4 \Pr(piv_{BA}) \; B\,\theta_i, \; \forall \theta_i \geq 0 \\ c_i &< W(\theta_i, A) = 4 \Pr(piv_{AB}) \; |A\,\theta_i|, \; \forall \theta_i \leq 0 \end{aligned}$$

For any given θ_i, given the distribution of costs, the probability that this happens is given by:

$$\begin{aligned} \Pr(c_i \leq W(\theta_i, B) \,|\theta_i \geq 0) &= \frac{4 \Pr(piv_{BA}) \; B}{C} \theta_i \\ \Pr(c_i \leq W(\theta_i, A) \,|\theta_i \leq 0) &= \frac{4 \Pr(piv_{AB}) \; |A|}{C} \; |\theta_i| \end{aligned}$$

■

Figure 1: Comparison of the equilibria under assumptions 3 and 3' (T denotes the voters who turn out in the former equilibrium)



Note that, in Lemma 2, turnout probabilities are linear in θ_i only because we assumed quadratic preferences and a uniform distribution of costs. Without these assumptions, the distribution of turnout would become more complex, but aggregate turnout would remain unaffected; only the distribution of turnout rates across types would be altered. Therefore, we only take lemma 2 as a convenient result, knowing that it is fragile and depends on the specification of the model. The important result in lemma 2 is that all types are represented in the election, though unevenly. Moreover, if turnout increases, $\Pr(piv_{BA})$ decreases, and representation becomes more even.

The reader might consider our alternative specification as being as "particular" as that of a fixed-and-constant cost of voting. Yet, the uniformity of the distribution is only chosen for convenience; it does not affect the essence of our results. The more relevant part of the assumption is that costs can be nil. However, this assumption is actually based on consistent empirical measures, which show that a substantial share of voters (about 40-45%) have a *negative* cost of voting (see e.g. Brody and Page, 1973). Since we are only interested in instrumentally motivated voters, we simply truncated the distribution to focus on voters with $c_i \geq 0$.

3.2 Compound Poisson Games

The second assumption we generalize is the one of known and fixed support for the two parties:

Assumption 2': the share of voters having left-wing and right-wing preferences is random and is unknown to the voters prior to the election. I.e. voters only have priors on the *distribution* of γ_A and γ_B:

$$\begin{aligned} \tilde{\gamma}_A &\sim \mathcal{L}(\bar{\gamma}_A) \\ \tilde{\gamma}_B &= 1 - \tilde{\gamma}_A \end{aligned}$$

That is, the share of voters supporting A is random, with an average $\bar{\gamma}_A$. We do not need to specify the distribution of $\tilde{\gamma}_A$, and leave it to be distributed under some "law" $\mathcal{L}$.

Why do we assume random distributions? Why is the assumption of fixed margins necessarily wrong? Strategic positioning by parties somewhat justifies that expected shares should be around 50%. But, if the distribution of preferences were fixed, opinion polls should leave no doubt about the identity of the winner of the election. Instead, uncertainty remains most of the time. For instance, two weeks before the 1997 election in the U.K., bookmakers were taking bets at 4 to 1 for John Major in his fight against Tony Blair, while the latter was expected to have a lead of 10%. With a Poisson distribution with fixed expected voting shares, we can easily compute the number of votes such that the leader is elected with a probability of 99.9%:

Table 2.

Total number of votes that ensures victory with a probability of 99.9% when the expected share is:				
50.1%	50.5%	51%	52%	55%
$\bar{n} = 2,427,000$	$\bar{n} = 97,000$	$\bar{n} = 24,000$	$\bar{n} = 6,000$	$\bar{n} = 975$

That is, with 975 votes, the odds against J. Major should have been at 1000 to 1, and much higher with a larger turnout. We readily see that, for a "standard" turnout, i.e. a turnout of a few million votes, the assumption of known margins is incompatible with the slightest uncertainty on the identity of the winner, unless expected shares are exactly 50%. A Poisson distribution with a random argument is called a "Compound Poisson distribution" in statistics, upon which we name this game.

Already now, even before relaxing the first of the three Key Assumptions, we can already derive the behavior of average turnout under Assumptions 1, 2' and 3'. It substantially differs from the results of Proposition 1:

Proposition 2 *Under Assumptions 1, 2' and 3', and if* $\Pr(\tilde{\gamma}_A \simeq \tilde{\gamma}_B) > 0$, *absolute turnout is proportional to* $\lambda^{2/3}$ *and the turnout* rate *decreases in the cubic root of population size:* $\bar{n}/\lambda \propto 1/\sqrt[3]{\lambda}$.

Proof. From Myerson (2000)'s *Magnitude Theorem*, we know that the distribution of states of the world *conditional on being pivotal* degenerates to a point mass around the set of states of the world that make pivotability most likely. That is, if the voter wants to assess the probability that she will be pivotal in the election, she must only consider the states of the world where pivotability is most likely. Hence, by (6) and (7), she only has to consider values of $\tilde{\gamma}_A$ that are close to 50%. The main determinant to turnout is thus the probability of the event $\tilde{\gamma}_A \simeq 1/2$:

$$\begin{aligned} W(\theta_i, A) &= 4\ \Pr(piv_{AB})\ |A\,\theta_i| \\ &\simeq 4\ \Pr[\tilde{\gamma}_A \simeq 1/2] \cdot \frac{1}{\sqrt{2\pi\,\bar{n}}} \cdot |A\,\theta_i|\,,\ \forall\theta_i \leq 0; \end{aligned}$$

$$\begin{aligned} W(\theta_i, B) &= 4\ \Pr(piv_{BA})\ B\,\theta_i \\ &\simeq 4\ \Pr[\tilde{\gamma}_A \simeq 1/2] \cdot \frac{1}{\sqrt{2\pi\,\bar{n}}} \cdot B\,\theta_i,\ \forall\theta_i \geq 0, \end{aligned}$$

where $\bar{n}$ is the equilibrium number of voters who turn out, which remains to be determined. By the definition of the Bayesian Nash Equilibrium of the game and by Lemma 2, voters who turn out are the ones who have a sufficiently low cost of voting. Assuming the expected size of the population (λ) to be sufficiently large, the equilibrium is defined as the value of $\bar{n}$ that solves:

$$2\lambda \int_0^{\alpha} \frac{4\ \Pr[\tilde{\gamma}_A \simeq 1/2] \cdot B\,\theta_i}{C\sqrt{2\pi\bar{n}}} d\theta_i = \frac{4\lambda \Pr[\tilde{\gamma}_A \simeq 1/2]}{\sqrt{2\pi\bar{n}}\,C}\,B\,\alpha^2 = \bar{n}.$$

Solving for $\bar{n}$ yields:

$$\bar{n} = \left(\frac{8\lambda^2}{\pi\,C^2}\ \Pr[\tilde{\gamma}_A \simeq \tilde{\gamma}_B]^2\ B^2\,\alpha^4\right)^{1/3},$$

and the participation *rate* is thus given by:

$$\frac{\bar{n}}{\lambda} = 2\left(\frac{\Pr[\tilde{\gamma}_A \simeq \tilde{\gamma}_B]^2\ B^2\,\alpha^4}{\pi\,C^2\,\lambda}\right)^{1/3}$$

Consequently, for continuous distributions of $\tilde{\gamma}_A$, forecasted turnout will be rather low (compared to observations), but increasing in the probability of close margins, that is in $\Pr(|\tilde{\gamma}_A - \tilde{\gamma}_B| < \varepsilon)$, with $\varepsilon \underset{\bar{n}\to\infty}{\longrightarrow} 0$. ■

As we can see, relaxing Key Assumptions 2 and 3 already highlights two separate elements. First, generalizing the cost of voting allows to explain why turnout should be monotonically increasing in population size, instead of being constant, like in Proposition 1. Second, introducing uncertainty on the preferences of the electorate shows that total turnout can increase rather quickly in population size. This is, in a sense, the positive lesson of the exercise. However, there is also a negative lesson; namely, that turnout is only large when voters believe that, with a high enough probability, parties have (almost) equal support in the electorate: $\Pr[\tilde{\gamma}_A \simeq 1/2]$ must be sufficiently large. It thus appears that, even though relaxing Assumptions 2 and 3 does improve our results, we still need an extra step to reconcile the model with empirical regularities. Hence, we now turn to the first assumption, which will prove much less innocuous than it seems at first glance.

4 If Mandate Matters

Under the first Key Assumption, the utility of voters is given by:

$$\begin{aligned} U(\mathsf{P}, \theta_i | n_A < n_B) &= U(B, \theta_i) \\ U(\mathsf{P}, \theta_i | n_A \geq n_B) &= U(A, \theta_i). \end{aligned}$$

As one can read from these expressions, the utility of any given voter can thus only take two values, with a discrete jump in $n_A = n_B$. However, if ex-post policymaking involves a bargaining of some sort once the election is over, or if the elected candidate/party pays attention to the share of votes he obtained, the implemented platform may not be an "all-or-nothing" composition of the proposed platforms. For instance, Ronny Razin (2000) demonstrates how this vote share actually communicates information to the candidates, who consequently have an incentive to moderate their policy when their margin of victory shrinks.[5]

When this happens, a lot of "activity" happens around the change in majority, i.e. policy changes can become substantial even when $n_A \neq n_B$. In other words, the implemented platform will be a smoother function of the vote shares received by the parties than under Key Assumption 1. Accordingly, we take the implemented policy to be a convex combination of the share of votes received by the two parties: define n_B as the realized number of votes for party B and n as the realized turnout. We pose:

Assumption 1': the implemented platform, P, is a continuous function of

[5] Other types of equilibria can also arise. See Razin (2000) for more detail.

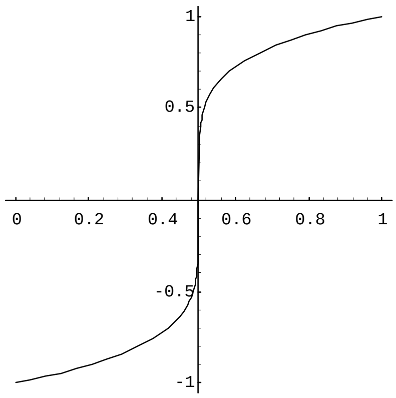
the share of votes going to the two parties, such that $A \leq \mathsf{P}\left(\frac{n_B}{n}\right) \leq B$**, and** $\mathsf{P}'(.) > 0$.

More precisely,

$$\mathsf{P}\left(\frac{n_B}{n}\right) = \left[1 - \sigma\left(\frac{n_B}{n}\right)\right] \cdot A + \sigma\left(\frac{n_B}{n}\right) \cdot B \tag{9}$$

where σ is a "power sharing function", in the spirit of Stigler (1972) or Grillo and Polo (1991), with $\sigma \in [0,1]$, continuous, differentiable and where σ' is positive and reaches its maximum in 1/2. An example is illustrated in figure 2.

Figure 2: Implemented platform if $A = -1$ and $B = 1$. Horizontal axis: share of B.



Denote the realized vote shares of the two parties by $s_p \equiv n_p/n$, $p = A, B$. For n expressed votes, a voter will change the share of votes going to party B by

$$\begin{aligned} \frac{n_B}{n+1} - \frac{n_B}{n} &= -\frac{s_B}{n+1}, \text{ if her vote goes to } A \\ \frac{n_B+1}{n+1} - \frac{n_B}{n} &= \frac{s_A}{n+1}, \text{ if her vote goes to } B. \end{aligned}$$

Therefore, for any given n and n_B, the *expected effect of a vote* on the implemented platform is:

$$-\mathsf{P}'(s_B) \times \tfrac{s_B}{n+1}, \quad \text{if the vote goes to } A \tag{10}$$

$$\mathsf{P}'(s_B) \times \tfrac{s_A}{n+1}, \quad \text{if the vote goes to } B. \tag{11}$$

Now, we need to derive the expected effect of a vote in case n_B and n are random. Quite clearly, $s_B = n_B/n$ is a priori not independent of $n+1$, which prevents us from easily deriving the values of (10) and (11) in a Compound Poisson Game. However, the following lemma demonstrates that the different elements in (10) and (11) are asymptotically independent one from another:

Lemma 3 *If turnout, n, is distributed according to a Poisson distribution of argument λ and $\tilde{n}_A \sim \mathcal{P}(\gamma_A.\lambda)$, $\tilde{n}_B \sim \mathcal{P}(\gamma_B.\lambda)$, with $0 < \gamma_A = 1 - \gamma_B < 1$, then*
1) If σ is linear, there is regressive *independence between the number of votes cast and the expected effect of a vote;*
2) For any twice continuously differentiable function σ, there is asymptotic stochastic *independence between the expected number of votes $\bar{n}$, and the marginal effect of a vote $\mathsf{P}'(s_B) \times s_p$.*

Proof. For any random X and Y, we know that

$$\mathsf{E}[XY] = \int_{-\infty}^{+\infty} Y\ f(Y)\,\mathsf{E}(X|Y)\,dY,$$

where $f(Y)$ is the density of Y. Applying this expectation operator to (10), let $X = a \cdot n_B/n$ and $Y = 1/(n+k)$. Holding Y constant (that is, holding n constant), the distribution of n_B is a binomial $B(\gamma_B, n)$. In that case, $\mathsf{E}[n_B/n|Y] = \gamma_B$ independently of n. Therefore, $\mathsf{E}[X|Y]$ is independent of Y if X is a linear function of n_B/n. In statistics, this is saying that the share of B is *regressively* independent of the number of votes. However, it is also clear that n_B/n is not *stochastically* independent of the number of votes. For instance, if $n = 2$, the share of B can be only 0, 0.5 or 1. Instead, the realization $n_B/n = 0.5$ is impossible to obtain when n is odd. However, by the *De Moivre-Laplace limit theorem* (see Mood (1974), pp120-121), we know that for n large and $\forall \varepsilon > 0$

$$\Pr(\gamma_B - \varepsilon \leq n_B/n \leq \gamma_B + \varepsilon) \simeq 2\,\Phi\left(\frac{\varepsilon\sqrt{n}}{\sqrt{\gamma_B(1-\gamma_B)}}\right) - 1 \underset{n\to\infty}{\longrightarrow} 1$$

where Φ is the C.D.F. of the centered-reduced normal distribution. In that case, for all twice-continuously differentiable functions σ,

$$\int_{-\infty}^{+\infty} \phi(n_B/n|n)\,\sigma(n_B/n)\ d(n_B/n) \underset{n\to\infty}{\longrightarrow} \sigma(\gamma_B),$$

where $\phi(.)$ is the P.D.F. of the centered-reduced normal distribution. ■

This lemma has a very simple meaning: if the power sharing function is continuous, in contrast to the results of RO, a vote can always affect the implemented platform at the margin, and the voter can consider two effects separately. On the one hand, the more voters turn out, the smaller the effect of her own ballot on the implemented platform. On the other hand, she can evaluate whether she prefers A or B only by looking at the distribution of preferences in the electorate. The former and the latter are stochastically independent one from another.

Of course, if the function σ had discontinuities, one would also have to consider the probability to be on a point of discontinuity of the function, as in previous sections. In this case, turnout would increase both in the expected effect of a vote *and* in the probability that the two parties are tied. Another point highlighted by this Lemma is that, at the margin, the effect of a vote only depends on the shape of $\mathsf{P}(s_B)$. So to say, the effect of a vote becomes deterministic and not probabilistic, which means that the importance of a vote is much higher than under previous specifications:

Proposition 3 *Under Assumptions 1'-3', total turnout is proportional to the square root of expected population size and increases in the expected value of* $|\mathsf{P}'(.)|$.

Proof. See Appendix 3. ■

What do we learn from this proposition? First, we see that instrumental voting is compatible with large turnouts when the mandate matters, even if the probability to be pivotal is minimal in the commonly accepted sense. Second, if the power sharing function is similar to that in figure 2, turnout falls smoothly in the share of the winner (and not abruptly as in section 2). Third, this proposition shows that the importance of the election is directly linked to the shape of the power sharing function, for which we do not have accurate empirical estimates.

It is also interesting to note that, under the assumption of concave utility functions, there is an *underdog* effect on the day of the elections. That is, if a given share of the population, say $\gamma_A > .5$, prefers party A to party B, election results will tend to give A a fraction of the votes that is lower than γ_A. To see this, assume for a moment that γ_A is fixed and certain. By (10) and (11), the expected effect of a vote for A is smaller than the expected effect of a vote for B. Therefore, *ceteris paribus,* the value of a vote for A is smaller than that of a vote for B. Moreover, this cumulates with a lower interest in the election for the majority, as the average distance between the A-voters and the platform is smaller. Hence, voters on the side of the loser will have higher turnout rates than voters supporting the expected winner of the election.

5 Numerical simulations

Generalizing the three key assumptions we underlined at the start of this paper has proven a tedious work and led to rather abstract results. To shed some light on the usefulness of these generalizations, we apply the results of Proposition 3 to two empirical regularities, and we show that our model has much higher predictive power than the model in section 2. The first empirical fact we explore is that of the downward trend in participation observed in the U.S. since World War II. Second, it has been regularly verified that turnout is higher in richer classes of the population. We show below that also this fact can be explained by our model, for some set of utility functions.

5.1 The downward trend in turnout partly explained

The fall in turnout in the American elections since World War II has been largely documented, not very much explained. Putnam (1995) details with scrutiny the fall in participation across generations, time, etc... for all kinds of social activities. Clearly, our model cannot explain all these but, in the case of elections, gives surprisingly good results.

To confront the model with data,[6] we must further generalize the model to take

[6] The data are taken from David Leip's Atlas of U.S. presidential elections, which are available on the Worldwide Web at the address: `http://uselectionatlas.org`

one of its limitations into account: several empirical studies have demonstrated that a substantial part of the population votes for non-instrumental reasons. We thus need to extend the distribution of costs to account for a negative minimal cost of voting. According to Brody and Page (1973), about 43% of the population has negative costs. Hence, we assume that costs are uniformly distributed between -\$2$\cdot\frac{.43}{.57}$ and \$2, which means that the highest possible cost of voting is \$2, and 43% of the population has a negative cost of voting.

There remains to explain why the other voters turn out or abstain. That is, on the theoretical side, we need to compute the turnout rate of instrumental voters, taking account of the 43% of the population that turns out anyway. From our model,[7] we compute the "theoretical instrumental turnout rate", $T(\cdot)$:

$$T(\lambda_t) = \frac{-\gamma/2 + \sqrt{\gamma^2 + 4(1-\gamma)\ K/\lambda_t}}{1-\gamma}, \tag{12}$$

where K is the marginal effect of a vote and γ is the share of the population with negative costs (43% in the simulations). Note that, in (12), $T(\lambda_t)$ represents the fraction of instrumental voters (i.e. of voters with $c_i \geq 0$) who is predicted to turn when the total number of registered voters ($\#Registered_t$) in the population is λ_t.

On the side of the data, we perform a similar exercise and compute the share of instrumental voters who actually turned out in each election:[8]

$$T_t^{Data} = \frac{Actual\ Turnout_t - 0.43\ \#Registered_t}{\#Registered_t}$$

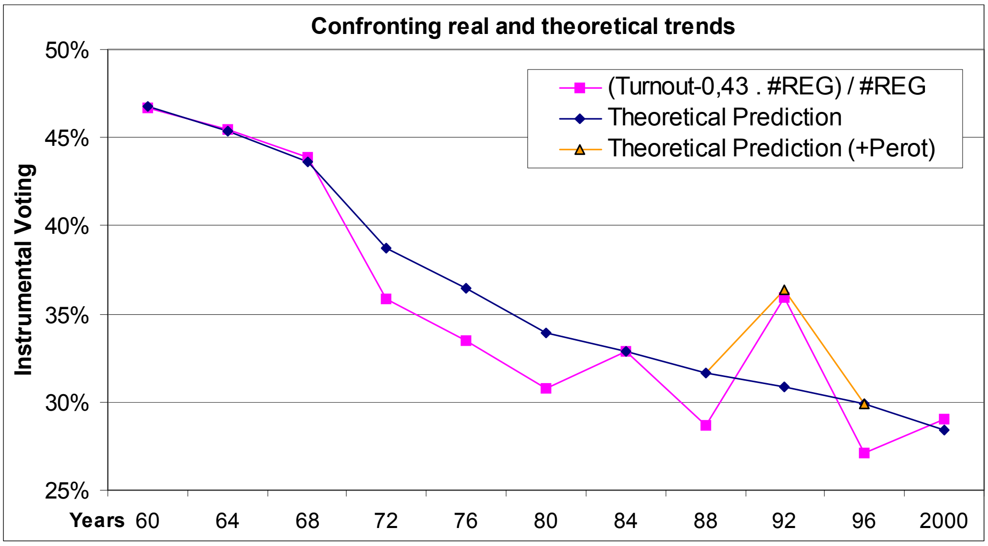
Now, how do we perform numerical simulations? To make the simulations figures comparable with actual data, we set K so as to make the theoretical and actual values of T_{1960} coincide ($T_{1960}^{Data} = 46.7\%$). This value of K is then maintained constant, and we only let population size vary across time. The results of this simple exercise are displayed in figure 3, and provide a surprisingly good fit.

In this figure, the curve with the "squares" displays the actual evolution of turnout.

[7] Theoretical developments are available from the author upon request.

[8] Note that our transformation of the turnout rate is only meant to reveal the fraction of *instrumental* voters who turn out. One can directly recover total turnout by computing $Actual\,Turnout_t = \left(T_t^{Data} + 0.43\right) \#Registered_t$ and $Predicted\,Turnout_t = (T(\lambda_t) + 0.43)\ \lambda_t$

Figure 3: Explaining the continuous fall in turnout rates in U.S. Presidential elections.



In turn, the curve with the "diamonds" displays the predictions of the model, and we see that the general downward trend is quite well captured. Even the acceleration in the fall of turnout occurring in 1972 is partly present, simply by taking the newcomers of the baby boom into account. Putnam (1995) documents that these "baby-boomers" typically turn out less often than previous generations, which probably explains why the actual fall in turnout since 1972 is not entirely explained by the model. The biggest error of prediction happens in 1992, when turnout is underestimated by 5 points. However, this is also the year when Ross Perot received 19% of the votes, who should thus not be counted in the number of votes received by the main two parties. If we only take $81\% = 100\% - 19\%$ of the registered population in 1992 into account, we get the third curve (the curve with "triangles"), which perfectly fits turnout that year.

This simulation thus shows that, according to Proposition 3, the sole increase in the size of the population can explain the bulk of the fall in turnout rates since 1960.

5.2 Public Goods as a Rationale for the Class Bias

Another unexplained stylized fact is that richer voters vote more often than poorer voters (see Appendix 4 for more detail). We address this point here.

We want to show that, for a wide class of utility functions, richer people should turn out with a higher probability than poorer ones, even for purely instrumental reasons. Let us assume a utility function $U(x_i, G, \theta_i)$, where x_i is the consumption of private goods, G is the consumption of public goods and θ_i is a preference parameter:

$$\begin{aligned}
\frac{\partial U(x_i, G, \theta_i)}{\partial x_i} &> 0 \\
\frac{\partial U(x_i, G, \theta_i)}{\partial G} &\propto \theta_i \text{ that is, } \theta_i \text{ is a parameter of taste concerning } G \\
\frac{\partial^2 U(x_i, G, \theta_i)}{\partial x_i \, \partial G} &> 0 \text{ if } \theta_i > 0 \\
&< 0 \text{ if } \theta_i < 0
\end{aligned}$$

The key feature of this utility function is that it displays a cross derivative which is of the same sign as the marginal utility for public goods. In that case, since x_i increases in income, so does the *intensity* of preferences with respect to public goods (i.e. the absolute value of the marginal utility of public good provision). Note for instance that a "standard" Cobb-Douglas utility function satisfies such properties: $U(x_i, G, \theta_i) = x_i^{\beta} G^{\theta_i(1-\beta)}$.

Now, if two candidates propose two different levels of public good provision, G_A and G_B, we have:

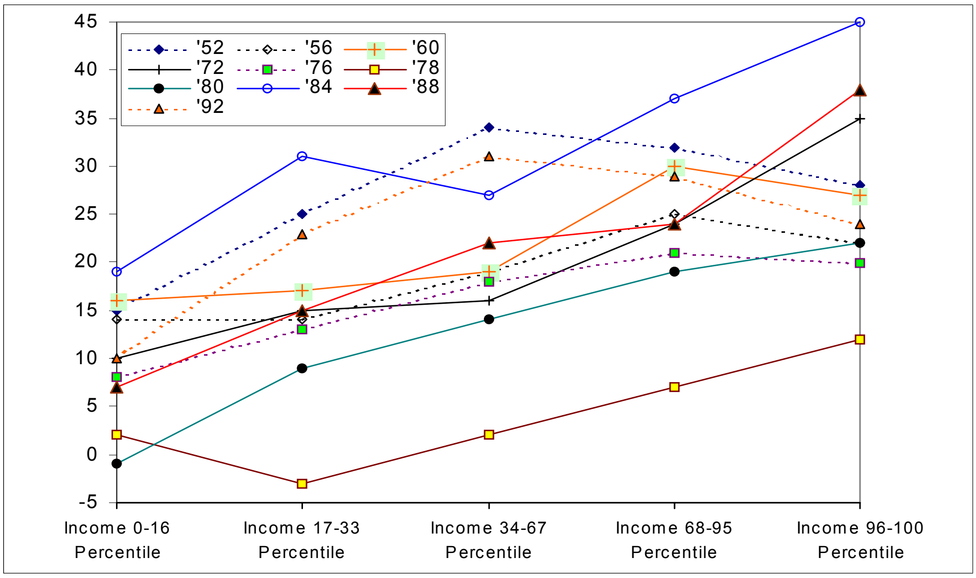
$$W(\theta_i, x_i; p) \propto \left| \frac{\partial U(x_i, G, \theta_i)}{\partial G} \frac{G_A - G_B}{1 + \bar{n}} \right|,$$

which is increasing in income. As we showed in Proposition 3 that voters' participation increases in their valuation of affecting the implemented policy, we see immediately that higher-income voters will participate more in the election:

Corollary 1 *If the vote is about the provision of public goods, and if private and public goods are complementary, there will be a class bias, such that turnout rates are higher in the wealthier classes of the population.*

Complementarity between private and public goods simply implies that voters who *a priori* value public goods, when they become wealthier, value them even more. Instead, voters who *a priori* dislike public goods, when they become wealthier, feel more and more opposed to their provision. Clearly, this Corollary is not claiming that utility functions necessarily have this shape. It means that empirical work should look at the potential consequences of different political platforms on the utility of each voter to predict turnout rates. The key would be to check whether promises made by candidates on some targeted points of the political debate tend to increase the turnout of those targeted voters. Strömberg (2001), for instance, confirms such a point.

Figure 4: Turnout rates for instrumental voters, by income percentile (Source: NES. Each curve represents a different election. Data were not available for all years).



Clearly, our purpose is not to perform a complete empirical analysis of such questions. Nonetheless, data seem again to confirm our point: using data from the National Elections Survey of the U.S., we compute instrumental turnout rates by class as the difference between the share of people who claim they turned out and that of people who think one "should vote even if one is not interested in the election" (the latter being a proxy for non-instrumental voting).[9] Results are displayed in figure 4 and hint again

[9]We computed the difference between the share of people who say they voted and those who answer

at the fact that, not only overall turnout, but also *instrumental* turnout is larger in the upper percentiles of the income distribution, like Corollary 1 suggests.

6 Conclusions

This paper assesses the real limits of the model of instrumental voting. We showed that, in opposition to usual perceptions, assuming a (even narrowly defined) rational behavior of voters can reproduce several stylized facts. Our conclusions are that, even though "standard" models fail to explain empirical regularities, the instrumental voting hypothesis is not at the roots of the problem. Instead, we argue that, on the contrary, it remains one of the main driving forces behind the turnout decision. More precisely, we show that several unjustifiable assumptions are generally made, and that they alone are responsible for the failures of the model to reproduce those empirical regularities.

More precisely, we showed that:

1. The cost of voting cannot be modelled as being the same for each voter;
2. The shares of votes going to each party cannot be modelled as fixed;
3. The effect of vote shares on the actually implemented policy must be taken into account.

We showed that, once those three elements are considered, the theoretical predictions of our extended model are consistent with the downward trend in turnout rates since the 1960s, and with the tendency for richer voters to turn out more often than poorer

"Yes", "Don't know" or "no opinion" to the question "do you think you should vote if you don't care about the outcome?" This leaves us with a measure of the people who should abstain if $A = B$ but do vote in equilibrium. Note that the steepness of the curves is much more dramatic for the undisplayed part of the sample, as one could expect. This shows again that both instrumental factors (illustrated in the figure) and non-instrumental factors (not displayed) move in the same direction and complement each other.

voters. Moreover, we motivate why our assumptions are more realistic than the ones usually made.

In other words, the main contribution of this paper is theoretical and shows that instrumental voting altogether may have tended to be rejected, not because it is *highly suspect in nature*, as Fiorina (1997, p403) puts it, but rather because of the excessive simplifications used in those models. Altogether, and even though we doubt that instrumental motivations alone can explain all variations in turnout rates, we infer from our results that they might lie behind several other, more psychological, motivations.

We hope that the theoretical developments of this paper also have an empirical reach. Our conclusions may seem to contradict some empirical findings, and the reader may thus reject them on these grounds. Such a shortcut is however incorrect, since most of the existing empirical evidence cannot be used to test these results. For instance, we claim that, if the implemented policy is a continuous function of the "*mandate*" given to the winner, closeness does not matter in the same way as in the standard Riker and Ordeshook (1968) model. One could thus infer that, in order to test this prediction, we should compare turnout rates during presidential and mid-term elections. Clearly, policy changes should be more "continuous" after mid-terms than after presidential elections. Hence, following our results, mid-term elections should generate *higher* turnout rates than presidential elections, which appears to contradict our predictions. Still, another argument would be that potential policy changes are also more limited after mid-term elections, which should decrease turnout. As this example shows, though our model is consistent with some empirical findings, additional empirical research is needed to test it, because existing empirical evidence is somewhat orthogonal to the questions raised by the paper (in the case of our example, an accurate assessment of platform changes after mid-terms and after presidential elections is needed to test the model). Another point raised by the empirical literature is whether closeness increases or decreases turnout. If we believe in the power-sharing function we proposed in section 4, the large expected victory of a candidate should lower turnout. However, there is also another effect, that goes in the opposite direction: if large electoral support reflects larger expected gains from electing the winner, turnout can be increased. Again, without additional empirical work, we cannot tell in which direction turnout should go.

Summing up, we hope that our theoretical contribution also points to more micro-founded ways to pursue future empirical research: if we assess which parts of the population are more targeted by parties, what are the potential gains for different classes of the population, how uncertain the outcome of the election is, and so on, we would have better tools to estimate the expected value of a vote and, from there, use our results to derive potential turnout rates.

References

Aldrich, John H., 1976. Some Problems in Testing Two Rational Models of Participations. American Journal of Political Science, Vol. 20, 713-734.

Aldrich, John H., 1993. Rational Vote and Turnout. American Journal of Political Science, Vol. 37, No 1, Feb. 246-278

Aldrich, John H., 1997. When is it Rational to Vote? In: *Perspectives on Public Choice, a Handbook.* Dennis Mueller Editor. Cambridge, Cambridge University Press, pp373-390

Ashenfelter, O. and S. Kelley, 1975. Determinants of Participation in Presidential Elections. Journal of Law and Economics, December, Vol. 18. 695-733

Bartels, Larry M., 1996. Uninformed Votes: Information Effects in Presidential Elections. American Journal of Political Science, Vol. 40, No 1, February.194-230

Barzel, Y. and E. Silberberg, 1973. Is the act of voting rational? Public Choice, Fall, 16. 51-58.

Brody, R. A. and B. I. Page, 1973. Indifference, Alienation and Rational Decisions. Public Choice, Summer, Vol. 15, 1-17.

Cavanagh, Thomas, 1982. Changes in American Voter Turnout, 1964-1976. Political Science Quarterly, Vol. 96. 53-65

Downs, Anthony, 1957. An Economic Theory of Democracy. Harper and Row

Epstein, Lee and Carol Mershon, 1996. Measuring Political Preferences. American Journal of Political Science, Vol. 40, No 1, February. 261-294

Feddersen, Timothy and Wolfgang Pesendorfer, 1996. The Swing Voter Curse. American Economic Review, 86(3). 408-424

Ferejohn, John A. and Morris. P. Fiorina, 1974. The Paradox of Not Voting: A Decision Theoretic Analysis. American Political Science Review, June, Vol. 68. 525-536

Fiorina, Morris P., 1981. Retrospective Voting in American National Elections. New Haven: Yale University Press

Foster, Carroll B., 1984 The Performance of Rational Voter Models in Recent Presidential Elections. American Political Science Review, September, Vol. 78. 678-690.

Harbaugh, W.T., 1996. If People vote because they like to, then why do so many of them lie? Public Choice 89. 63-76.

Leighley, Jan E. and Jonathan Nagler, 1992. Socioeconomic Class Bias in Turnout, 1964-1988: The Voters Remain the Same. American Political Science Review, Vol. 86. 725-736.

Leip, David, 2001. Atlas of U.S. Presidential Elections, `http://uselectionatlas.org`

Mood, Alexander McFarlane, 1974. Introduction to the Theory of Statistics. McGraw-Hill, 3rd edition.

Mueller, Dennis C., 1989. Public Choice II, Cambridge University Press

Myerson, Roger B., 1997. Population uncertainty and Poisson Games. Center for Mathematical Studies in Economics and Management Science Discussion Paper 1102R, Northwestern, Revised from 1994

Myerson, Roger B., 2000. Large Poisson Games. Journal of Economic Theory, Vol. 94, No. 1, Sept. 7-45

National Election Studies, Center for Political Studies, University of Michigan. Electronic resources from the NES World Wide Web site (`www.umich.edu/~nes`). Ann Arbor, MI: University of Michigan, Center for Political Studies [producer and distributor], 1995-2000.

Putnam, Robert D. 1995 Tuning In, Tuning Out: The Strange Disappearance of Social Capital in America. Political Science and Politics, Vol. 28 No4. 664-683

Quid? 1994 Robert Laffont Editions

Razin, Ronny, 2000. Swing Voters and Mandate Makers: Voting When Margin of Victory Matters. Mimeo New York University

Riker, William H. and Peter C. Ordeshook, 1968. A Theory of the Calculus of Voting. American Political Science Review, No 62, March. 25-42.

Settle, R. F. and B. A. Abrams, 1976. The Determinants of Voter Participation: A More General Model. Public Choice, Fall, 27. 81-89.

Shields, Todd G. and Robert K. Goidel, 1997. Participation Rates, Socioeconomic Class Biases, and Congressional Elections: A Crossvalidation. American Journal of Political Science, Vol. 41, No 2, April. 683-691.

Silberman, J.I. and G. C. Durden, 1975. The Rational Behavior Theory of Voter Participation. Public Choice, Fall, Vol. 23. 101-108.

Smithies, A., 1941. Optimum Location in Spatial Competition, Journal of Political Economy, June, No 49. 423-439

Stigler, George J., 1972. Economic Competition and Political Competition. Public Choice, 13, Fall. 91-106.

Strömberg, David, 2001. Radio's Impact on Public Spending, IIES mimeo

Teixeira, Ruy A., 1987. Why Americans Don't Vote: Turnout Decline in the United States, 1960-1984. New York: Greenwood

Wolfinger, Raymond E. and Stephen J. Rosenstone, 1980. Who Votes? New Haven: Yale University Press.

Technical Appendix

Appendix 1: General Properties of Poisson Games

Poisson Distribution

A variable $\tilde{n}$ is distributed according to a Poisson distribution of parameter λ if:

$$\Pr\left(\tilde{n}=k\right)=\frac{e^{-\lambda}.\lambda^{k}}{k!},\quad (k\in\mathcal{N}). \tag{13}$$

It follows that, for $\tilde{n}_A\sim\mathcal{P}(\lambda_A)$ and $\tilde{n}_B\sim\mathcal{P}(\lambda_B)$:

$$\begin{aligned}\Pr(\tilde{n}_A=\tilde{n}_B+c) &= \sum_{k=0}^{\infty}\Pr(\tilde{n}_A=k+c)\ \Pr(\tilde{n}_B=k)\\ &= \sum_{k=0}^{\infty}\frac{e^{-\lambda_A-\lambda_B}.\lambda_A^{k+c}.\lambda_B^{k}}{(k+c)!\ k!} \qquad (14)\end{aligned}$$

Using Bessel Functions in Pivot Probabilities

The definition of a modified Bessel function I of degree c is given by:

$$I_c(2z)=\sum_{k=0}^{\infty}\frac{z^{2k+c}}{k!\ (k+c)!},\ \text{with}\ \lim_{z\longrightarrow\infty} I_c(z)=\frac{e^{z}}{\sqrt{2\pi z}} \tag{15}$$

Rewriting (14) thus obtains:

$$\begin{aligned}\Pr(\tilde{n}_A=\tilde{n}_B+c) &= e^{-\lambda_A-\lambda_B}\frac{\left(\sqrt{\lambda_A}\right)^{c}}{\left(\sqrt{\lambda_B}\right)^{c}}\times\sum_{k=0}^{\infty}\frac{\left(\sqrt{\lambda_A}\right)^{2k+c}.\left(\sqrt{\lambda_B}\right)^{2k+c}}{(k+c)!\ k!}\\ &= e^{-\lambda_A-\lambda_B}\left(\frac{\lambda_A}{\lambda_B}\right)^{\frac{c}{2}}\times I_c\left(2\sqrt{\lambda_A.\lambda_B}\right) \qquad (16)\\ &\underset{\lambda_A,\lambda_B\to\infty}{\longrightarrow} e^{-(\sqrt{\lambda_A}-\sqrt{\lambda_B})^{2}}\left(\frac{\lambda_A}{\lambda_B}\right)^{\frac{c}{2}}\left(2\sqrt{\pi}\sqrt[4]{\lambda_A.\lambda_B}\right)^{-1} \qquad (17)\end{aligned}$$

So, if $\lambda_i=\gamma_i\cdot\lambda$, one obtains:

$$\Pr(\tilde{n}_A=\tilde{n}_B+c)\simeq e^{-(\sqrt{\gamma_A}-\sqrt{\gamma_B})^{2}\lambda}\left(\frac{\gamma_A}{\gamma_B}\right)^{\frac{c}{2}}\left(2\sqrt{\pi.\lambda}\sqrt[4]{\gamma_A.\gamma_B}\right)^{-1}. \tag{18}$$

That is, when $\gamma_A=\gamma_B$,

$$\Pr(\tilde{n}_A=\tilde{n}_B+c)\simeq\left(2\sqrt{\pi.\lambda}\ \sqrt[4]{\gamma_A.\gamma_B}\right)^{-1}, \tag{19}$$

which decreases at speed $\sqrt{\lambda}$. By contrast, when the γ_i's differ, the probability decreases exponentially towards 0.

Appendix 2: Proof of Proposition 1

For a given pivot probability π and a given cost of voting $c > 0$, a voter with type θ_i participates in the election iff $2\pi |(A - B)\ \theta_i| \geq c$, that is:

$$|\theta_i| \geq \frac{c}{4\pi\ B}, \tag{20}$$

in which π depends on the expected number of "active" voters. Hence, for any pivot probability π, there is a pair of cut-offs θ_A and θ_B such that:

$$\begin{array}{lll} \forall\theta_i < \theta_A, & |\theta_i| > \frac{c}{4\pi\ B} & \Rightarrow v_i^* = A \\ \forall\theta_i \in [\theta_A, \theta_B] & |\theta_i| \leq \frac{c}{4\pi\ B} & \Rightarrow v_i^* = \varnothing \\ \forall\theta_i > \theta_B & |\theta_i| \geq \frac{c}{4\pi\ B} & \Rightarrow v_i^* = B. \end{array} \tag{21}$$

This proves the first part of the proposition.

The Bayesian Nash Equilibrium of the game is thus determined by the equilibrium value of these two cut offs, such that:
i) the equilibrium turnout rate T must be equal to the ratio between the expected number of extremist voters and the expected total number of voters in the population:

$$T \equiv \frac{\bar{n}}{\lambda} = \left(\gamma_A \frac{\alpha + \theta_A}{\alpha} + \gamma_B \frac{\alpha - \theta_B}{\alpha}\right)$$

ii) the equilibrium share of votes going to A and B, denoted respectively by s_A and s_B must be consistent with (21):

$$\begin{aligned} s_A &= \frac{\gamma_A (\alpha + \theta_A)}{\gamma_A (\alpha + \theta_A) + \gamma_B (\alpha - \theta_B)} \\ s_B &= \frac{\gamma_B (\alpha - \theta_B)}{\gamma_A (\alpha + \theta_A) + \gamma_B (\alpha - \theta_B)} \end{aligned}$$

iii) voters must compute the pivot probabilities that result from those vote shares and turnout rate, and these probabilities must be compatible with (21): using (18),

$$\Pr(piv_{AB}) = e^{-(\sqrt{s_A} - \sqrt{s_B})^2 T\lambda} \sqrt{\frac{s_B}{s_A}} \left(2\sqrt{\pi.T\lambda}\sqrt[4]{s_A.s_B}\right)^{-1} \tag{22}$$

$$\Pr(piv_{BA}) = \sqrt{\frac{s_A}{s_B}} \Pr(piv_{AB}) \tag{23}$$

and

$$\theta_A = \frac{-c}{4\ \Pr(piv_{AB})\ B} \tag{24}$$

$$\theta_B = \frac{c}{4\ \Pr(piv_{BA})\ B}. \tag{25}$$

Since $|\theta_i| \leq \alpha$ by assumption, positive turnout requires:

$$\begin{aligned} \alpha &\geq \frac{c}{4\ B}, \text{ and} \\ \Pr(piv_{AB}), \Pr(piv_{BA}) &\leq \frac{c}{4\,\alpha\, B}. \end{aligned}$$

Since pivot probabilities decrease in turnout, this imposes an upper bound on $\bar{n}$, which proves the second part of the proposition.

Next, using (23) and (25) shows that the majority of the population gets underrepresented in the results of the election. We show this by contradiction: assume $\theta_A = -\theta_B$, which implies that $\Pr(piv_{BA}) = \sqrt{\frac{\gamma_A}{\gamma_B}} \Pr(piv_{AB})$, and hence

$$\gamma_A \gtrless \gamma_B \Rightarrow \Pr(piv_{BA}) \gtrless \Pr(piv_{AB}) \Rightarrow \theta_B \lesseqgtr -\theta_A.$$

Put differently, s_A/s_B will always be closer to 1 than γ_A/γ_B, the true preferences of the population.

Finally, (23) and (25) directly show the fourth part of the proposition. ■

Appendix 3: Proof of Proposition 3

Exploiting the results of Lemma 3, the expected utility of a voter who abstains is given by:

$$\begin{aligned} \mathsf{E}U(\theta_i, \mathsf{P}|v_i = \varnothing) &= \mathsf{E}_{s_B}\left[-(\theta_i - \mathsf{P}(s_B))^2\right] \\ &= -\theta_i^2 + \int_0^1 f_{s_B}(s_B) \cdot \mathsf{P}(s_B) \cdot (2\theta_i - \mathsf{P}(s_B)) \cdot ds_B \qquad (26) \end{aligned}$$

where s_B is the share of votes cast on B and $f_{s_B}(s_B)$ is the density of s_B. By (11) and Lemma 3, for $\bar{n}$ sufficiently large, the value of an additional vote for B can be computed using the derivative of (26) with respect to n_B:

$$\begin{aligned} W(\theta_i, B) &\simeq \mathsf{E}_{\bar{n}} \frac{2}{1+\bar{n}} \int_0^1 f_{s_B}(s_B) \cdot \mathsf{P}'(s_B) \cdot (1 - s_B) \cdot (\theta_i - \mathsf{P}(s_B)) \cdot ds_B \qquad (27) \\ &\simeq \frac{2}{\bar{n}} \Omega^B(\theta_i) \qquad (28) \end{aligned}$$

where $\Omega^B(\theta_i)$ is equal to the integral in (27).

Following the same reasoning, a vote for A is worth:

$$\begin{aligned} W(\theta_i, A) &\simeq \mathsf{E}_{\bar{n}} \frac{2}{1+\bar{n}} \int_0^1 f_{s_B}(s_B) \cdot \left[-\mathsf{P}'(s_B)\right] s_B \, (\theta_i - \mathsf{P}(s_B)) \cdot ds_B \\ &\simeq \frac{2}{\bar{n}} \Omega^A(\theta_i). \qquad (29) \end{aligned}$$

The Nash response of voter i will thus be:

$$\begin{aligned} &\text{Play } A \text{ if } W(\theta_i, A) > \max[W(\theta_i, B), c_i], \\ &\text{Play } B \text{ if } W(\theta_i, B) > \max[W(\theta_i, A), c_i], \\ &\text{Abstain if } \max[W(\theta_i, A), W(\theta_i, B)] \leq c_i. \end{aligned}$$

Now that we know the Nash responses of each voter, we can solve for the Bayesian Nash Equilibrium of the game. From (28) and (29), the value of turning out is inversely proportional to $\bar{n}$, the expected turnout. The Bayesian Nash Equilibrium of the game is thus obtained when $\bar{n}$ is consistent with the Nash responses of all voters:

$$\lambda \cdot \int_{-\alpha}^{\alpha} \Pr\left[c_i \leq \frac{2}{\bar{n}} \max\left\{\Omega^A(\theta_i), \Omega^B(\theta_i)\right\}\right] d\theta_i = \bar{n}.$$

Substituting for the distribution of costs and solving for $\bar{n}$ yields:

$$\frac{\bar{n}}{\lambda} = \left(\frac{\int_{-\alpha}^{\alpha} \max\left\{\Omega^A(\theta_i), \Omega^B(\theta_i)\right\} d\theta_i}{\alpha C \lambda}\right)^{1/2}$$

$$\bar{n} = \left(\frac{\int_{-\alpha}^{\alpha} \max\left\{\Omega^A(\theta_i), \Omega^B(\theta_i)\right\} d\theta_i}{\alpha C}\lambda\right)^{1/2}$$

■

Appendix 4: stylized facts

Appendix 4.1: Abstention in the U.S.

As mentioned already, there are very detailed surveys and a huge amount of research on abstention in the U.S. The goal of this appendix is not to summarize these findings exhaustively but to stress its most salient aspects. More detail can be found in Mueller (1989) and Aldrich (1993, 1997).

From a survey study on presidential elections, Riker and Ordeshook (1968) conclude that, in the formulation $P\,B + D - C$, where P is the pivot probability, B is the benefit of changing the outcome of the election, D is some fixed benefit of voting and C the gross cost of voting, all variables influence the probability that a voter turns out at the election. The following table summarizes their findings:

Table 3: Riker and Ordeshook's main conclusions.

		% turnout		% turnout
Close margin:	yes:	78%	no:	72%
Benefit (B):	high:	82%	low:	66%
Sense of duty (D):	high:	87%	low:	51%

Ashenfelter and Kelley (1975) give more mixed evidence regarding the rational voter hypothesis: closeness does not seem to affect turnout (correct sign, but insignificant). Still, benefit and costs matter strongly (a \$6-voting tax was abolished in 1972. The probability for a given individual to turn out was 42% lower with the tax) as well as the sense of duty (people who feel "obliged to" turn out with a probability increased by 30%). There is however controversy about some variables used to proxy B. If those variables were instead relating to D, the rational voter hypothesis would not be supported anymore.

Concerning the variable P, Foster (1984) used data from presidential elections to reestimate –in cross-section– some previous time-series analyses. His conclusion is that the closeness of the election and the size of the electorate (which decreases P) are in general insignificant or of the wrong sign.

To summarize: "*Many applications, especially those that use aggregate data find that the P term is a significant predictor (e.g. Barzel and Silberberg 1973; Settle and Abrams 1976; Silberman and Durden 1975). Other tests using survey data (e.g., Ferejohn and Fiorina 1975; Foster 1984) have found it to be unrelated to the vote*" (Aldrich, 1993). But, in cross-section, any variation in perceived closeness must come from personal errors in answering P-like questions, since all different answers regard the same event, while time-series studies compare true differences in closeness (Aldrich 1976, 1993).

Finally, there is a clear upward trend in abstention rates.

Another set of studies has underlined the **class bias** of turnout: richer and more educated voters tend to participate more than others. This observation is made in presidential as well as in legislative elections. There is less clear evidence about the way this class bias evolves over time: for some, it is stable. For others, it is falling, even if there is more evidence about the latter (e.g. see Wolfinger and Rosenstone 1980, Cavanagh 1982, Teixeira 1987, Leighley and Nagler 1992, Shields and Goidel 1997).

This leads us to the following stylized facts:

Stylized Fact 1: Sense of duty motivates participation, costs of voting reduces participation.

Stylized Fact 2: There is a negative trend in turnout since the end of World War II.

Stylized Fact 3: Participation increases in education and income.

But we also want to keep in mind the uncertainty about the effect of closeness on turnout:

Point of debate: Does closeness of elections increase turnout?

The third stylized fact is considered as counterintuitive. Education and income "pick up the wrong sign" in regressions. Indeed, more educated people should be aware that pivot probabilities are negligible. Richer people should also have higher opportunity cost of time, leading to higher costs of voting and consequently to lower turnout rates. Mueller (1989) argues this can be explained by the fact that successful people also learn to comply to social rules, which may explain their behavior. In the text, we instead argue that the PB part alone may motivate this behavior. In other words, under the appropriate conditions, the sign of these regressors are not "wrong".

Appendix 4.2: Abstention in France

We did not find any empirical study of turnout in France. The data on turnout rates since 1962 are given in Figure 5.

This documents several points: first, turnout in France also seems to display a downward trend, though of a lesser extent than in the U.S. Second, the importance of the election affects turnout: European elections are of small effect on the daily life of citizens and display the highest rate of abstention. Presidential elections are decisive for the general orientation of policy and generate the highest turnout (a t-test reveals that equality is rejected with a probability of error of $p = 1.7\%$ in the first round. In second rounds, there is a clear outlier in 1969, when De Gaulle pulled out of the election. Including this outlier gives a p-value of 9%. Removing the outlier increases the significance level to a p-value of 0.3%. Pooling the two rounds displays a probability of 0.3% including the outlier). This is another hint that the importance of the election, which determines B, influences turnout. The same happens in the U.S. (mid-term elections are typically affected by lower turnout rates).

Stylized Fact 4: Turnout varies with the type of election.

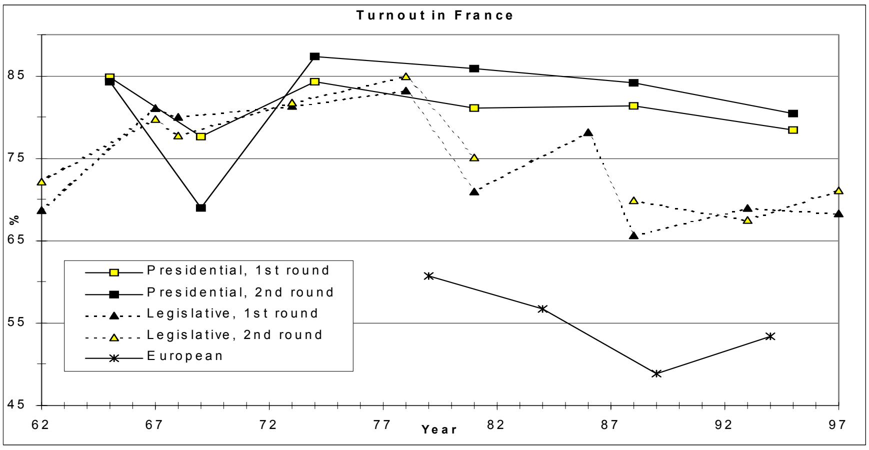


Figure 5: Turnout in France (Source: Quid 1994 and Le Monde, various issues)