Indecision Theory: Explaining Selective Abstention in Multiple Elections*

PAOLO GHIRARDATO
Division of the Humanities and Social Sciences
California Institute of Technology
Pasadena, CA 91125
e-mail: paolo@hss.caltech.edu

JONATHAN N. KATZ
Department of Political Science
University of Chicago
Chicago, IL 60637
e-mail: j-katz@uchicago.edu

October 15, 1999

Abstract

We address the so-called "roll-off" phenomenon: Selective abstention in multiple elections. We present a discuss a novel model of decision making by voters that explains this as a result of differences in quality and quantity of information that the voters have about each election. In doing so we use a spatial model that differs from the Euclidean one, and is more naturally applied to modelling differences in information.

1 Introduction

The rational choice (or positive political theory) literature since Downs (1957) has had difficulty explaining why an instrumentally rational individual will decide to vote if there are any costs to doing so — the so-called "Paradox of Voting." In general, the solution to this problem has been to make the expected (net) benefit of voting positive for at least a subset of the electorate. This feat has been accomplished by explicitly associating some direct utility gain (or citizen duty) to the act of voting (Riker & Ordeshook 1968).[1] What has been little

*We are grateful to Colin Camerer, Michel Le Breton and a seminar audience at the Department of Political Science at University of California, San Diego for helpful comments and discussion.

[1]This is also true of the game-theoretic models that account for voting, such as Ledyard (1984) or Palfrey & Rosenthal (1985), where in equilibrium in large elections the only citizens to vote are those with negative costs to voting.

recognized, however, is that these explanations fail to account for why a voter once in the ballot booth would ever choose to abstain from voting in any elections on the ballot. This selective abstention is often referred to as "roll-off" because voters are more likely to abstain in races for smaller offices, such as local judge or school board, that are typically listed lower on the ballot.[2] Roll-off is at odds with the explanation mentioned above, since it is in small elections that the individual's vote is most likely to make a difference in the outcome, hence making the expected benefit from voting highest. In this paper we present a formal, decision theoretic model that explains this "Paradox of Abstention."

So what explains this phenomenon? It cannot be the costs discussed above, since once in the ballot booth the voter has already paid all "fixed costs" to voting (or had negative costs to start with). Besides, such costs apply to all elections on the ballot equally, whereas to explain roll-off we need something that varies over the races. What differentiates elections is the amount and quality of information that the voter is likely to have about the alternatives in each of them. These differences in information generate what we could call *information costs*. The latter are the psychological costs that a voter faces in the event that he, not being completely informed about the candidates, "mistakenly" votes for the wrong candidate.[3] For example, think of a voter is *pro* death penalty but votes for a candidate for judge who has sworn never to impose the death penalty. The less informed the voter is, the greater the possibility for making such a mistake, so that higher informational costs are associated with the election. As voters acquire more information about an election, we think that they are more likely to vote on it, since these informational costs are lower.

Our model is an attempt to formalize the logic of this information costs argument for roll-off. The modelling exercise has two main components, each of which will be of some independent interest to those wishing to model political phenomena. First, we posit a preference structure of a voter under complete information, in order to be able to define a "mistake" in voting. Although we use a spatial framework, it is not the one of the standard Euclidean model (Enelow & Hinich 1984). We instead represent the policy space as a product of *binary* issues on which the voter and the candidates have well-defined positions — e.g., *pro* choice or *pro* life. As will be made clear below, this set-up is quite general and makes deriving comparative statics about levels of information easier. In particular, we do not have to require that there are only two — or finitely many — types of voters in the election (informed and uninformed), as is often done in voting models. Instead, a natural metric for how informed a voter is follows naturally from the set-up. Furthermore, we think that this binary structure provides a more plausible description of how voters actually view candidates.

The second feature of our modelling strategy is the introduction of a decision theoretic model that allows the possibility that the "ambiguity" associated with events — which informally depends on the quality and quantity of information at the disposal of the voter – affect behavior, something that is ruled out in the standard decision model employed in po-

[2]Strictly speaking roll-off is not the correct term, since it implies that the ballot order matters. Our model does not rely on ballot ordering nor does the empirical evidence suggest it matters much. See Cox & Munger (1990) for a discussion.

[3]Throughout, we use male pronouns for voters and female pronouns for candidates.

litical science. The model we use is an extension of Gilboa & Schmeidler (1988)'s "maximin expected utility" with multiple priors.[4]

These two components allow us to derive general conditions under which we expect an instrumentally rational voter to abstain. Abstention occurs when the voter is least informed, for example in elections for small offices in which the press is not interested in covering the campaign. It is important to note that since we present a decision theoretic model, our results on abstention specify a map from beliefs to behavior. They could therefore be extended to a game-theoretic model in which voters condition their action on other voters' (expected) actions (formally: strategies).[5] The added game structure would just impose restrictions on the voters' beliefs as a result of equilibrium reasoning, but itwould not change the predictions given the beliefs.To place our work in the context of the literature on voting, we should mention that while most of the formal theoretic literature on voting has sought to explain why anyone votes (see Aldrich (1993) for a recent review), there have been two recent papers that have attempted to model abstention. Matsusaka (1995) develops a decision theoretic model of abstention. The key assumption in his model is that more informed voters get a higher expected return for voting for the correct candidate then do their less informed counterparts. If we make voting costless within his framework, then all voters, informed and uniformed, will vote. Feddersen & Pesendorfer (1996*b*) assume no (fixed) costs to voting, as we do, but they generate abstention in a game-theoretic model by inducing a certain type of correlation in voters' beliefs about the election. We discuss their approach in more detail in section 6.

The paper proceeds as follow. We first describe more formally what we mean by ambiguity and how to model decision making when a voters are ambiguity adverse. The following three sections contain the theoretical development: In section 3 we present our model of voter behavior. Section 4 contains the main result describing how a voter decides to vote or abstain in a given election. The comparative statics are then derived in section 5. The last section discusses our results and how they might be extended and modified.

2 Information, Ambiguity Aversion and Multiple Priors

Information, or lack thereof, plays a crucial role in understanding voting behavior. Typically when a voter enters the polling booth on election day he is not completely informed about all of the policy stances of the candidates or the consequences of every proposition on the ballot. How does he decide what to do?

The standard approach in formal political models is to assume that the voter maximizes subjective expected utility (SEU), as described in Savage (1954)'s classical work. The decision maker chooses among actions, which have consequences that depend on which of several uncertain "states of the world" occurs. In the case of voting, for example, the set of actions

[4]See Ghirardato & Katz (1997) for the axiomatic underpinnings of our model. The model has some similarity to "minimax" regret model, used by Ferejohn & Fiorina (1974) to explain the opposite phenomenon of why people vote.

[5]An extension that we plan to pursue in future work.

available to the voter are: In each election, to vote for one of the candidates on the ballot or to abstain. The unknown states might be the policy positions of the candidates and how others are going to vote in the election. Although the exact state of the world is not known at the time of decision, the SEU decision maker forms beliefs on each state's relative likelihood, that are represented by a probability measure. These beliefs are subjective, so for instance it is legitimate for two different voters to have very different beliefs about the same candidates in an election.

More formally, we let $\mathcal{F}$ be the set of possible actions, Ω the state space, and $\mathcal{X}$ the set of possible consequences of actions (so that actions are functions mapping states into consequences). The voter's preferences over outcomes are characterized by a function, $u : \mathcal{X} \to \mathbf{R}$, called the *utility function.* A voter who maximizes SEU chooses the action $f \in \mathcal{F}$ which maximizes

$$U(f) \equiv \int_\Omega u(f(\omega))\, P(d\omega), \tag{1}$$

the expectation of the utility payoff with respect to a belief $P \in \Delta(\Omega)$.[6] The measure P is obtained by observing the decision maker's preferences among *bets* on events (subsets of Ω), so that it is correct to assert that P quantifies his confidence (willingness to bet) on each event happening.

As innocuous (and elegant) as SEU maximization seems, there is some reason to question the extent to which it actually describes choices under uncertainty. One central problem for our purposes is that it is very rare that the decision maker's preferences be well-specified enough to be represented by a probability measure P. In fact, this is less likely to happen the more "ambiguous" the structure of the uncertainty is: There is by now a wealth of empirical evidence (Camerer & Weber 1992) showing that the presence of such ambiguity has relevant consequences on decision maker's willingness to bet on events. The classical example of this phenomenon is the so-called *Ellsberg paradox* (Ellsberg 1961).[7]

2.1 The Ellsberg Paradox

Ellsberg (1961) presented a number of subjects with an urn containing 90 balls, of which 30 are red while the other 60 are black and yellow in unknown proportion. He then asked them to consider the following four bets on the urn:

1. \$100 if one red ball is extracted, \$0 otherwise;

2. \$100 if one black ball is extracted, \$0 otherwise;

[6] For a set X, $\Delta(X)$ denotes the set of all the finitely additive probability measures on the space $(X, 2^X)$, where 2^X is the power set of X. We could make weaker measurability assumptions but it would add small generality at the cost of additional notation.

[7] We should note that Ellsberg (1961) did these surveys under (in his words) "absolutely non-experimental conditions," so there could be a question as to whether the behavior to be described below is consistently observed in actual choices. So it seems: A large number of later experiments in both psychology and economics have given strong support to Ellsberg's findings (see Camerer & Weber 1992).

3. \$100 if either a red or a yellow ball is extracted, \$0 otherwise;

4. \$100 if either a black or a yellow ball is extracted, \$0 otherwise;

Typically subjects expressed the following preferences:

$$1 \succ 2, \qquad 4 \succ 3,$$

That is (assuming that each subject preferred \$ 100 to \$ 0), a typical subject was more confident on the extraction of a red ball than of a black ball, and on the extraction of a red or yellow ball than of a black or yellow ball. Yet this is not consistent with the assumption that the subject maximizes SEU, and that his confidence can be measured by some probability function P over states of nature. To see this, let $P(r)$, $P(b)$, and $P(y)$ be respectively the probabilities of the event that the ball extracted is red, black, or yellow. Then $1 \succ 2$ implies:

$$P(r) > P(b),$$

but $4 \succ 3$ implies (by the additivity of P):

$$\begin{aligned} P(r \cup y) &< P(b \cup y) \\ P(r) + P(y) &< P(b) + P(y) \\ P(r) &< P(b). \end{aligned}$$

As both inequalities cannot hold, we find a contradiction: SEU cannot rationalize these choices.

The reason for such choices by subjects in Ellsberg's surveys is obvious: People tend to prefer situations in which there is less *ambiguity*, that is situations in which the structure of uncertainty is more deeply known. In a sense it is as if the subject's willingness to bet on an event is a composition of a pure "likelihood" judgement, and a modifying factor, which accounts for the quality and quantity of the information that the subject has about the event. The requisite that the willingness to bet on an event be represented by a probability, a single number, is what makes the SEU model ill equipped to deal with situations like Ellsberg's urn, where there are significant differences in the ambiguity associated with different events. A SEU maximizer, by construction, does not mind ambiguity.

One might conjecture that ambiguity could be captured even in a SEU framework, by allowing the decision maker to be *uncertain* about his beliefs, i.e., to have possibly multiple conjectures (called first order beliefs), and then to have a belief as to which conjectures are more correct (a second order belief). This does not quite work, however, for such decision maker behaves *as if* he has a precise belief (the average, according to the second order belief, of his first order beliefs), and so once again does not care about ambiguity.

We should emphasize that the confidence in beliefs, or perception of ambiguity, that we just discussed is distinct from the spread or *risk* associated with a particular belief P. A decision maker could be very certain of his beliefs, but they could still imply a high degree of risk. An example of such a situation is betting on a particular number to come up on a roulette wheel. Clearly this is a risky bet, but assuming that the decision maker thinks the casino is honest he would be fairly confident of his judgment of the probability of winning. Bets 1 and 2 in Ellsberg's paradox are both *equally risky*, but they are (at least by a majority of people) associated with different levels of ambiguity.

2.2 Modelling Ambiguity

The task at hand then is to generate a plausible model of behavior when a voter is likely to have scarce information, and thus perceive ambiguity about the events which are relevant for his choice. That is, we need a decision theoretic model that can makes for the possibility that the voter cares about ambiguity. The formal approach we adopt here is based on Gilboa & Schmeidler (1988)'s model of "maximin expected utility" with multiple priors. The voter's beliefs, instead of being characterized by a single probability distribution, are given by a set of probability distributions $\mathcal{C}$. For every possible action, the voter calculates its expected utility for every probability in $\mathcal{C}$, and then chooses the action which obtains the largest *minimum* of these expectations. When $\mathcal{C}$ is a singleton, this is just a SEU decision maker. The larger the set $\mathcal{C}$, the more ambiguity there is in the decision problem — i.e., the scarcer he considers his information about the election.

It is easiest to see this property by considering a simple dichotomous event, say whether or not it will rain tomorrow. A SEU decision maker would summarize his (subjective) beliefs by the probability that it will rain tomorrow, say 0.5. A decision maker described by the model just outlined could instead summarize his beliefs by an interval, say by assessing that the probability of rain tomorrow is between 0.4 and 0.6. The larger the interval, the less confidence he displays in his judgement. In the limit his belief set could even be the whole interval $[0, 1]$, and then we would define him to be "completely ignorant" about the plausibility of that event. Intuitively, such a decision maker has no idea as what his beliefs should be, so that his set $\mathcal{C}$ is just the set of all possible probabilities on Ω.

While the preferences just described embody a heavy dose of pessimism (since the decision maker acts as if the worst belief is always the right one), it is this pessimism that generates the ambiguity aversion we need to explain behavior in situations like the Ellsberg paradox. Subjects do not like the second Ellsberg bet, for example, because there is some possibility that there are no black balls in the urn, or that there are only a few black balls. Similarly, they do not like the third bet because there might be few or no yellow balls in the urn either. The model described above can capture this, as it allows the decision maker to use different beliefs in evaluating different actions. We should note that this extreme form of pessimism is not necessary to generate ambiguity averse behavior, and we do not need the full strength of it for our results, however we postpone detailed discussion of possible relaxations until Section 6.

A last point on modeling ambiguity aversion: There is an alternative approach suggested by Schmeidler (1989) that uses a single non-additive probability to characterize beliefs instead of a set of probabilities. A non-additive probability, P, is between 0 and 1 and is "monotonic" — i.e., $P(E) \leq P(F)$ if E and F are events with $E \subseteq F$ — but not necessarily additive. That is, possibly $P(E \cup F) \neq P(E) + P(F) - P(E \cap F)$. The more non-additive the belief, the less confident the decision maker is about his beliefs. In many situations these two approaches yield the same behavioral predictions. However calculating expected utilities using non-additive probabilities requires using a non-standard notion of integral introduced by Choquet (1953–54). We chose to use the multiple priors model mostly because of its more immediate mathematical representation.

3 The Model

We imagine a voter in the ballot booth holding a blank ballot in his hands. Thus all "fixed costs" of voting are already sunk, and the only cost that the voter is facing is that of making up his mind on each item on the ballot (but see the discussion in Section 6). In general we are interested in describing his behavior if the ballot requires him vote on, say, M different elections. Which elections is he going to vote on, and which ones is he going to abstain on? To do so, we start by considering his behavior in a *single* election, given his knowledge on the issues and candidates at the moment in which he is looking at the blank ballot. We will then (in Section 5) make a comparison of his behavior in different elections, in which he has significantly different information. A limitation of this way of proceeding, worth pointing out from the outset, is that it implicitly assumes that the voter's choices in one election do not affect his behavior in another election. We think however that it provides a sufficiently realistic description of behavior in many circumstances.

So fix one election with two candidates A and B.[8] A voter is uncertain about the following facts: The policy position that either candidate would take if elected in office, and the outcome of the election in the absence of his vote. We start by delineating a voter's preferences under certainty: How he would rank the candidates if she knew exactly their policy position and could choose to put either of them in office.

3.1 The Voter's Preferences under Certainty

The policy space is modeled as a product of binary sets $\mathcal{Y} \equiv \prod_{i \in \mathbf{N}} Y_i$, where each $Y_i = \{0, 1\}$ and $\mathbf{N}$ is the set of the natural numbers. Each $i \in \mathbf{N}$ is a *policy issue* on which the candidate can have either a "yes" or a "no" position.[9] As will become clear from what follows, the discreteness of the space could be relaxed (at the cost of more complexity) without affecting the nature of our results.[10] We assume that the voter has an ideal point in policy space, which we take without loss of generality to be the 0 vector (so that $y_i = 1$ means that the candidate's position on issue i is different from the voter's). However he does not

[8]While we will stick to the standard case of two candidates in order to keep notation to a minimum (and to draw pictures on a two-dimensional page), nothing in the analysis to be presented depends on having two candidates. All results immediately generalize to the case of more than two candidates.

[9]We are not excluding that there are only finitely many issues, as will become clear presently when we discuss the voter's preference structure: In the notation to be introduced below, then we allow that possibility by letting $w_i = 0$ but for finitely many i's for every voter.

[10] For example we could assume $Y_i = [0, 1]$, which allows us to interpret y_i^j as the probability that candidate j disagrees with the voter on issue i.

necessarily care about, nor is he necessarily indifferent among, all the issues: His preferences are (additively) separable across issues as follows. There is a sequence $w = [w_1, w_2, \dots]$ of real numbers with the property that $w_i \geq 0$ and $\sum_{i \in \mathbf{N}} w_i = 1$. That is, w_i represents the subjective *weight* that the voter assigns to issue i in making his choice; for instance if $w_i = 1$ then he only cares about the candidates' position on issue i. If certain of the candidates' position, he will vote for the candidate $j \in \{A, B\}$ who is closer (according to the metric given by w) to his ideal point. That is, he will vote for, say, candidate A if

$$-\sum_{i \in \mathbf{N}} w_i y_i^A \geq -\sum_{i \in \mathbf{N}} w_i y_i^B, \tag{2}$$

where y^j is candidate j's *real* position. For $j \in \{A, B\}$, we let $\pi_j = \sum_{i \in \mathbf{N}} w_i y_i^j$, the "disutility" of having candidate j in office if her position is given by y^j.

It follows immediately from the structure of voter's preferences that $\pi_j \in [0, 1]$, so that we can always map the pair (π_A, π_B) in the square $[0, 1] \times [0, 1]$. Actually more is true: Suppose that the candidates' positions are given by the pair (y^A, y^B), and that the voter can correctly observe the first n coordinates of each vector. That is, he knows $y^j(n) \equiv [y_1^j, \dots, y_n^j]$ for both j's. Then it is clear to him that for every j, π_j must lie in the interval $I(j, n) = [l(j, n), r(j, n)]$, where

$$l(j,n) \equiv \sum_{i=1}^{n} w_i y_i^j \qquad \text{and} \qquad r(j,n) \equiv l(j,n) + \sum_{i=n+1}^{\infty} w_i.$$

So, when endowed with this information, he will know that the pair (π_A, π_B) lies in the square $S(n) \equiv I(A, n) \times I(B, n)$, the sides of which are both equal to the residual sum of weights $\sum_{n+1}^{\infty} w_i$. Figure 1 depicts all the relevant sets.

Clearly, as n increases, the square shrinks, eventually collapsing on the real values (π_A, π_B). Summing up, we have

Lemma 1 *For a voter whose preferences on $\mathcal{Y}$ are additively separable and given by the vector of weights w (where $w_i \geq 0$, $i = 1, \dots, n$, and $\sum_{i=1}^{\infty} w_i = 1$), the set $S(n) \subseteq [0, 1]$ decreases monotonically in n. That is, for $m < n$,*

$$S(n) \subseteq S(m).$$

The inclusion will be strict if $w_i > 0$ for some $i = m, m+1, \dots, n$. Moreover

$$\lim_{n \to \infty} S(n) = (\pi_A, \pi_B).$$

Two other implications of the preference structure are worth pointing out. The first is the trivial observation that in this model the event that two candidates are indifferent is far from being exceptional. Consider for simplicity the case of the voter discussed above for whom $w_i = 1$: Two politicians who have the same position on issue i will be indifferent to this

Figure 1: The set of possible pairs (π_A, π_B) and the square $S(n)$

voter. While this case is somewhat extreme, the fact that indifference is not rare conforms to our intuition.

The second implication is that quite generally the voter will not need to be perfectly informed about the candidates' positions in order to decide which one he likes best: Suppose that $\pi_A \neq \pi_B$, that is, the two candidates have different true policy positions, and the differences matter to the voter. Then there is (finite) n large enough so that either the interval $I(A, n)$ is all to the right of the interval $I(B, n)$ or *vice versa* (in the square $[0, 1]^2$ that happens respectively when $S(n)$ is properly below or above the diagonal, see Figure 1). In the former case, say, the voter knows that he definitely prefers B, whatever he will later know about her policy position on other issues.

The next step in the construction of the model is to outline the voter's decision problem: His possible choices, the state space describing the relevant uncertainty, and the possible outcomes.

3.2 The Voter's Decision Problem

The only nontrivial aspect of the exercise here is the description of the space of states of the world that the voter is facing for the election. One part of the uncertainty is clearly given

	ρ_1	ρ_2	ρ_3	ρ_4	ρ_5
a	A	A	A	T	B
b	A	T	B	B	B
ϕ	A	A	T	B	B

Table 1: Election Outcomes

by the results of the election in the absence of the voter's participation. That is, let V_j be the number of votes cast by all other voters in favor of candidate j (abstentions are of course allowed, i.e., we do not require that $V_A + V_B$ be equal to the number of eligible voters minus 1). Then the set of possible election results is given by $\mathcal{R} = \{\rho_1, \rho_2, \rho_3, \rho_4, \rho_5\}$, where ρ_1 is the event that $V_A > V_B + 1$, ρ_2 is $V_A = V_B + 1$, ρ_3 is $V_A = V_B$, ρ_4 is $V_A = V_B - 1$, and ρ_5 is $V_A < V_B - 1$. Thus the event that this particular voter is pivotal in the election is given by $(\rho_2 \cup \rho_3 \cup \rho_4)$. Table 1 summarizes the considerations made so far by plotting election outcomes (entries correspond to the winner and T stands for a tie) as a result of ρ and the of the voter's choice of voting for A (denoted by a), B (denoted by b), or abstaining (denoted by ϕ).

The entries of the matrix are however not the ultimate consequences of the decision problem, which are given by the voter's utility of having the winner j in office, $-\pi_j$, as described in the previous subsection. Here we introduce the second aspect of uncertainty for the voter: The policy position of the candidates. So in the absence of information on the candidates' policy positions, the voter considers the product $\mathcal{R} \times (\mathcal{Y} \times \mathcal{Y})$ to be the state space for his problem. As is tradition in the literature on voting, we assume that the voter acts *instrumentally*, that is, he eventually only cares about the election outcome. In the context of this model this translates into the following *state independence* assumption: The voter's preferences under certainty (i.e., the vector w) are *not* affected by the result ρ and by the real policy position pair (y^A, y^B). This assumption implies that we can without loss of generality take the voter's state space to be the set $\Omega \equiv \mathcal{R} \times [0, 1]^2$, so that a state is a triple $\omega = (\rho, \pi_A, \pi_B)$.

3.3 The Voter's Preferences under Uncertainty

We can now describe the voter's preferences in the case in which he knows only the first n coordinates of the candidates' positions (where n is possibly 0).

To capture our intuition that there is a cost to deciding to vote for either candidate, which abstaining does not entail, we assume that the voter looks at a problem in which the payoffs to every action $f \in \{a, b, \phi\}$ are renormalized as follows: For every state ω, the payoff to action f is given by $f(\omega) - \phi(\omega)$. This is what we call a model of choice with abstention *as a reference choice*. It corresponds to a form of focused "regret": A vote for candidate is *ex post* (that is, if the real state ω is known) considered a mistake if it yields a worse result than would be obtained by abstaining.[11] Notice that abstention is still potentially a mistake,

[11] It is different from the standard form of regret known in the literature (see Ferejohn & Fiorina 1974), as

if it turns out that by voting for an *ex post* better candidate the voter could have put her in office. However the renormalization by itself is not sufficient to obtain the result we need, as the following remark shows.

Remark 1 If the voter is a subjective expected utility (SEU) maximizer he forms a belief $P \in \Delta(\Omega)$, and chooses the action f which maximizes

$$U(f) \equiv \int_\Omega (f(\omega) - \phi(\omega))\, P(d\omega),$$

the expectation of the payoff $f(\cdot) - \phi(\cdot)$ with respect to P. In this case the renormalization of the payoffs does not affect his preferences, for it is immediate to use the linearity of the integral to show that for any two actions f and g, $U(f) \geq U(g)$ is equivalent to

$$\int_\Omega f(\omega)\, P(d\omega) \geq \int_\Omega g(\omega)\, P(d\omega).$$

We thus conclude that the existence of a reference choice does not affect the preferences of a voter who is a SEU maximizer. Later we will show (Proposition 1) that for this reason the SEU voter does not abstain in an election unless a certain type of correlation between election results and candidates' expected values obtains. $\diamond$

An additional feature is therefore required to obtain the type of behavior that conforms to our intuition: We need the voter to care proportionally more about losses (with respect to the yardstick set by abstention) than about gains.

To capture this we also assume that the voter is ambiguity averse, as formalized in the "multiple priors" model of Gilboa & Schmeidler (1988) that we discussed in Section 2.[12] Observe first that we can describe the voter as facing a *dynamic* choice situation, where the dynamic aspect is due to the possible different levels of information that might have. Regarding the process by which he gets information (to be discussed in greater detail below), we make here the following two key assumptions of *truthful and symmetric information*: The voter is given information about the position of *both* candidates on, say, the first n issues, and he believes this information to be truthful. We do not think that either assumption is crucial to our results,[13] but dispensing of them would certainly make the analysis more complicated. The first assumption implies that, for a fixed pair of policy positions (y^A, y^B), the information that the voter might obtain is given by the two n-dimensional vectors $y^j(n) \equiv [y^j_1, \ldots, y^j_n]$, for $j = A, B$, for some n. As a consequence, the complete description of the voter requires a specification of his preferences for every pair of n-dimensional truncations of (y^A, y^B).

the latter judges a mistake any action which is not *ex post* optimal, not only those that fare better than the reference.

[12]Gilboa & Schmeidler (1988) provides an *axiomatization* of the preferences discussed here, but without the reference choice property. It is simple to see how the latter property can be obtained by strengthening one of their axioms (see Ghirardato & Katz 1997).

[13]The first could be relaxed quite easily, while the second could be relaxed by letting policy coordinates be represented by $[0, 1]$, as discussed in footnote 10.

In line with the multiple priors model, we assume that his preferences at information n are represented as follows: There is a non-empty closed and convex set of probabilities $\mathcal{C}(n) \subseteq \Delta(\Omega)$ such that he chooses according to the mathematical functional U associating action f with the number

$$U(f) \equiv \min_{P \in \mathcal{C}(n)} \int_\Omega (f(\omega) - \phi(\omega))\, P(d\omega). \tag{3}$$

Clearly a SEU voter is one for whom $\mathcal{C}(n)$ is a singleton. We can now formalize the assumption of truthful information that we stated above, using what we call *consequentialist beliefs*: Each $P \in \mathcal{C}(n)$ has a support contained in $\mathcal{R} \times S(n)$. That is, all the measures in the belief set $\mathcal{C}(n)$ assign probability zero to the pairs of payoffs which are impossible given his information. There is a special case of these preferences which will provide us with an interesting benchmark:

Example 1 Suppose that the voter is completely ignorant, in the sense that his set of beliefs is $\mathcal{C}(n) = \Delta(\mathcal{R} \times S(n))$: Any probability satisfying the consequentialist beliefs assumption is in the voter's set of possible beliefs. In particular this implies that, for every state $\omega \in \mathcal{R} \times S(n)$, there is a $P \in \mathcal{C}(n)$ such that $P(\omega) = 1$. Thus, when applying Eq. (3) to an action f, we find

$$U(f) = \int_\Omega (f(\omega) - \phi(\omega))\, \hat{P}(d\omega),$$

where $\hat{P}$ is the probability which assigns weight 1 to a state ω which minimizes $(f(\omega) - \phi(\omega))$. In other words, the completely ignorant voter behaves in a "maximin" fashion. △

In general there is no reason to exclude the possibility that the voter's beliefs entail stochastic dependence of the result of the election ρ and the candidates' values (π_A, π_B). For instance in recent work Feddersen & Pesendorfer (1996*b*) (see also Feddersen & Pesendorfer 1996*a*) show that this will be the case in a game-theoretic model in which all the voters have SEU preferences and they are aligned in a certain way. In order to simplify the analysis, in this paper we rule this out by *imposing* a stochastic independence assumption. This obviously limits the generality of the model, but it helps putting the specific causes for the abstention we obtain here in sharper focus. In Ghirardato & Katz (1997) we discuss the general version of the model, and show that the intuition developed here carries on to the case where dependence is allowed. We shall therefore assume that the voter's beliefs satisfy the *stochastic independence* assumption: For every n the set $\mathcal{C}(n)$ is a "product" of a set of beliefs on election results, $\mathcal{D}(n) \subseteq \Delta(\mathcal{R})$, with a set of beliefs on candidates' values, $\mathcal{E}(n) \subseteq \Delta(S(n))$. Precisely:

$$\mathcal{C}(n) \equiv \text{Conv}(\{P \times Q : P \in \mathcal{D}(n),\, Q \in \mathcal{E}(n)\}),$$

where $\text{Conv}(X)$ is the convex hull of X, the smallest convex set containing X. This is intuitively a generalization of the usual property of stochastic independence to the case in which there is a set of probabilities, rather than a singleton. We show in Appendix A that in the context of this model, stochastic independence implies that the functional U can be written as

$$U(f) = \min_{P \in \mathcal{D}(n)} \int_{\mathcal{R}} F(\rho, \mathcal{E}(n))\, P(d\rho). \tag{4}$$

where

$$F(\rho, \mathcal{E}(n)) = \min_{Q \in \mathcal{E}(n)} \int_{S(n)} (f(\rho, \pi_A, \pi_B) - \phi(\rho, \pi_A, \pi_B))\, Q(d(\pi_A, \pi_B)). \tag{5}$$

This shows that, as it is natural to expect under independence, the set $\mathcal{E}(n)$ of beliefs on the product policy space $S(n)$ does *not* depend on ρ. It will make the analysis of the model simpler, as it will allow us to obtain conditions for abstention which do not depend on the specific election outcome ρ.

We conclude this Section by recapitulating the assumptions that we make about the voter (which, unless otherwise noted, we shall assume to hold throughout the rest of the paper). We assume that the voter faces no *fixed* (i.e., independent of the action and of his information) costs of voting, and that he has *symmetric information* about the candidates, i.e., he knows both candidates' position on the first n issues, with $n \geq 0$. The voter's preferences under certainty are determined by the vector of weights w, as represented by Eq. (2), and they are *state-independent*, in the sense that the vector w does not depend on the election outcome, and the candidates' real position (and also on the voter's choice). His preferences under uncertainty are given by the "maximin" functional with a set of priors $\mathcal{C}(n)$ where abstention is a *reference choice*, as represented in Eq. (3). Moreover his set of beliefs $\mathcal{C}(n)$ reflects *stochastic independence* of election results $\rho \in \mathcal{R}$ with the candidates' (values of) policy positions $(\pi_A, \pi_B) \in S(n)$.

4 To Cast or Not to Cast: Abstention in a Single Election

Having set up the model, we are now ready to ask the main question, which is under what conditions the voter will strictly prefer abstaining over voting for a candidate. As before, we are assuming that the voter is in the ballot booth, and we are fixing his information at the first n coordinates of the candidates' positions. Also, we assume that the tie-breaking rule for the election is a coin toss, so that the voter thinks that the T outcome corresponds to a $1/2$ probability of getting A and a $1/2$ probability of getting B.[14] As we will discuss in Section 6, our results would hold also if the tie-breaking rule were confirmation of the incumbent (assuming that there is one).

The first step is calculating the values of the function F representing the voter's expectation of action f under result ρ (given his beliefs on the values space $\mathcal{E}(n)$). Table 2 plots the values of $F(\cdot, \mathcal{E}(n))$ for every action f in the choice set $\{a, b, \phi\}$. In the table, we let

[14]We are assuming that the coin toss is perceived to be independent of the realization of the other relevant uncertainty. The fact that the voter has a single probability for the coin toss conforms with the model of Gilboa & Schmeidler (1988) (which is framed in an Anscombe & Aumann (1963) environment).

	ρ_1	ρ_2	ρ_3	ρ_4	ρ_5
a	0	0	$\min_{Q\in\mathcal{E}(n)} \frac{1}{2}\psi(Q)$	$\min_{Q\in\mathcal{E}(n)} \frac{1}{2}\psi(Q)$	0
b	0	$\min_{Q\in\mathcal{E}(n)} -\frac{1}{2}\psi(Q)$	$\min_{Q\in\mathcal{E}(n)} -\frac{1}{2}\psi(Q)$	0	0
ϕ	0	0	0	0	0

Table 2: The Graphs of Action Payoffs

$$\psi(Q) \equiv \int_{S(n)} (\pi_B - \pi_A)\, Q(d(\pi_A, \pi_B)).$$

The calculation of these values is explained as follows: Consider for instance action a in state $\omega = (\rho_3, \pi_A, \pi_B)$. Its payoff $a(\omega)$ is given by $-\pi_A$ (since A is elected). The payoff to abstention is instead

$$\phi(\omega) = (1/2)(-\pi_A) + (1/2)(-\pi_B),$$

since in such a case a coin toss decides who is elected. Subtracting we obtain

$$a(\omega) - \phi(\omega) = (1/2)(\pi_B - \pi_A),$$

which, when integrated with respect to a $Q \in \mathcal{E}(n)$, gives $(1/2)\psi(Q)$. The calculation of the other values is worked out similarly.

There is an immediate observation that we can make. Suppose that the set $\mathcal{E}(n)$ of voter's beliefs on the policy space is a singleton. That is, $\mathcal{E}(n) = \{Q\}$. Then it must be the case that if the non-zero values of, say, a are equal to $\alpha > 0$, the non-zero values of b are negative, being $-\alpha$. This immediately implies that

$$U(a) \geq 0 = U(\phi) \geq U(b),$$

since the three values correspond eventually to the integral of, respectively, a non-negative, zero, and non-positive function. In this case, while abstention could still be a weakly optimal choice, it will never be strictly preferred. In fact, it is easy to see under which conditions on the set $\mathcal{D}(n)$ choosing a is strictly better for the voter: Suppose that $\mathcal{D}(n)$ contains only measures P which assign positive probability to the event $(\rho_3 \cup \rho_4)$. Then clearly

$$U(a) = \alpha \left(\min_{P\in\mathcal{D}(n)} P(\rho_3 \cup \rho_4) \right) > 0.$$

Thus we have that if the voter's beliefs *on the policy space* are extremely precise, abstention is going only to be a knife-edge choice, which depends on fairly specific beliefs on the election result space. What is interesting is also that in proving this we have not really made use of some of the assumptions in our model. For instance, we made no use whatsoever of the assumptions of symmetric information and consequentialist beliefs (the support of the distribution Q is irrelevant). Summarizing we have

Proposition 1 *Consider a voter whose preferences are as described in Section 3, but whose beliefs are not necessarily consequentialist. Then if his beliefs on the policy space $\mathcal{E}(n)$ are given by a singleton Q, he will never strictly prefer abstaining over voting. Moreover, if his beliefs on the election result space $\mathcal{D}(n)$ are such that for every $P \in \mathcal{D}(n)$, $P(\rho_3) > 0$, then he will never choose to abstain if the expected difference in value of the candidates $\psi(Q)$ is different from zero.*

The following remark shows that the driving force behind this result is the stochastic independence assumption: As we observed above and discuss elsewhere (Ghirardato & Katz 1997), abstention can be obtained (as a unique optimum) also with a single prior, but only if a specific type of correlation between results in $\mathcal{R}$ and policy positions in $[0,1]^2$ obtains.

Remark 2 While we have to chosen to present Proposition 3 in the context of our model where abstention is a reference choice, even that assumption could be dispensed with. That is, it is possible to prove the following: *Consider a voter whose state-independent preferences under certainty are described by Eq. (2), and whose preferences under uncertainty are represented by the multiple prior model with stochastically independent beliefs. He will never strictly prefer abstention if his belief set on the policy space $\mathcal{E}(n)$ is a singleton.* In other words, under stochastic independence of the policy positions from election results, a voter with a single prior would not abstain in a nontrivial fashion. Notice that this result is true regardless of whether the voter is ambiguity averse on the election results space. So the result applies to a SEU voter, for whom $\mathcal{C}(n)$ is a singleton. $\diamond$

The previous considerations tell us that in order to obtain nontrivial abstention, that is,

$$U(\phi) > U(a) \qquad \text{and} \qquad U(\phi) > U(b), \tag{6}$$

we have to allow the set $\mathcal{E}(n)$ to contain more than one point. Let us however proceed in a slightly backward fashion, and start by assuming that we have conditions insuring that both a and b are non-positive functions on $\mathcal{R}$ (that is, all the entries on row a in Table 2 are zero or negative, and similarly for b). That is, assume that we found an $\mathcal{E}(n)$ such that

$$\min_{Q\in\mathcal{E}(n)} \psi(Q) < 0 \qquad \text{and} \qquad \min_{Q\in\mathcal{E}(n)} -\psi(Q) < 0. \tag{7}$$

Then we of course have that abstention is (weakly) optimal, in the sense that $U(\phi) \geq U(j)$ for $j \in \{a, b\}$. As for strict optimality, we have the following simple result:

Lemma 2 *Suppose that the voter is described by the model in Section 3, and that Eq. (7) holds. Then abstention is a strictly preferred action (i.e., Eq. (6) holds) if and only if the set $\mathcal{D}(n)$ contains a measure P_a such that $P_a(\rho_3 \cup \rho_4) > 0$ and a measure P_b such that $P_b(\rho_2 \cup \rho_3) > 0$.*

Clearly the condition in the Lemma is satisfied if the voter has some prior in $\mathcal{D}(n)$ assigning positive probability to a tie. This is true even if $\mathcal{D}(n)$ is a singleton, a fact that, when joined with Remark 2, shows that ambiguity aversion on $\mathcal{R}$ does not really play any significant role in explaining abstention.[15]

Now we come to the more interesting problem of showing when it is true that Eq. (7) holds. This turns out to have a nice graphical intuition. In fact observe that for every $Q \in \mathcal{E}(n)$, we can rewrite

$$\psi(Q) = \int_{S(n)} \pi_B \, Q(d(\pi_A, \pi_B)) - \int_{S(n)} \pi_A \, Q(d(\pi_A, \pi_B)) = \Pi_B^Q - \Pi_A^Q, \tag{8}$$

where we let $\Pi_j^Q \equiv \int \pi_j \, Q(d(\pi_A, \pi_B))$. Clearly every $Q \in \mathcal{E}(n)$ can be identified with the pair of expected values $(\Pi_A^Q, \Pi_B^Q) \in S(n)$. Let $\Pi(\mathcal{E}(n))$ be the set of all such pairs, that is,

$$\Pi(\mathcal{E}(n)) \equiv \{(x_A, x_B) \in [0,1]^2 : x_A = \Pi_A^Q \text{ and } x_B = \Pi_B^Q \text{ for some } Q \in \mathcal{E}(n)\}.$$

It is easy to see that the assumptions on $\mathcal{E}(n)$ imply that $\Pi(\mathcal{E}(n))$ is a closed convex subset of $S(n)$. Figure 2 depicts the $\Pi(\mathcal{E}(n))$ (the shaded polytope) induced by a set of priors $\mathcal{E}(n)$ generated by six measures. Using Eq. (8) we observe that the problem of minimizing $\psi(Q)$, for $Q \in \mathcal{E}(n)$, is just the problem of finding the point (x_A, x_B) in the set $\Pi(\mathcal{E}(n))$ which minimizes $x_B - x_A$, that is the point which touches the function in the linear family $x_B = x_A + k$, $k \in \mathbf{R}$, with the highest intercept k. In Fig. 2 this point is denoted by x. Symmetrically, minimizing $-\psi(Q)$ is tantamount to looking for the point in the set which touches the function in the same linear family with the smallest intercept, denoted by y in Fig. 2. Both the minimized values are negative (i.e., Eq. (7) holds) if the point x lies above and the point y below the diagonal of the $[0,1]^2$ square, for then clearly $x_B = x_A + k$ for a positive k, and $y_B = y_A + h$ for a negative h.

The situation depicted in Fig. 2 is representative of when Eq. (7) holds: This happens if and only if the set $\Pi(\mathcal{E}(n))$ is nontrivially separated in two parts by the diagonal, so that in $\Pi(\mathcal{E}(n))$ there is at least a point above, and at least a point below, the diagonal. We have thus proved the following

Lemma 3 *For a voter described by the model in Section 3, Eq. (7) will hold* if and only if *the set $\Pi(\mathcal{E}(n))$ contains (at least) a point above the diagonal $D \equiv \{(x_A, x_B) \in [0,1]^2 : x_A = x_B\}$ and (at least) a point below D. Formally: There are a point $x \in D$ and two measures $Q, Q' \in \mathcal{E}(n)$ and an $\alpha \in (0,1)$ such that, if we let $Q'' = \alpha Q + (1-\alpha)Q'$ and $\Pi^{Q''}$ be the point in $\Pi(\mathcal{E}(n))$ corresponding to Q'', $x = \Pi^{Q''}$.*

[15] In the context of this model, since there are *no* costs to voting. Obviously it would play a more significant role if there were costs to voting.

Figure 2: The set $\Pi(\mathcal{E}(n))$ and the points x and y

Clearly Lemma 3 implies a slightly stronger version of the first result of Proposition 1: If the set $\Pi(\mathcal{E}(n))$ is only a single point (as would be the case if $\mathcal{E}(n)$ were a singleton) then Lemma 3 trivially implies that Eq. (7) cannot be satisfied: One of the two minima must be non-negative. For instance, suppose that the set $\Pi(\mathcal{E}(n)) = \{y\}$. Then (the minimum of) $-\psi(Q) > 0$, while obviously $\psi(Q) < 0$.

Adding the two Lemmata together, we obtain necessary and sufficient conditions for the voter here described to abstain nontrivially in the election. Notice that regardless of whether the assumptions of Lemma 2 hold, if one of the two strict inequalities in Eq. (7) fails then clearly abstention cannot be strictly optimal. In fact then there is an action $f \in \{a, b\}$ which is non-negative, so that $U(f) \geq U(\phi)$ (see also Corollary 1 below).

Theorem 1 *Assume that the voter is described by the model of Section 3. Then he will strictly prefer abstaining over voting for either candidate (i.e., Eq. (6) holds)* if and only if *his belief sets* $\mathcal{D}(n)$ *and* $\mathcal{E}(n)$ *satisfy* both *the following conditions:*

(i) $\mathcal{D}(n)$ *contains a measure* P_a *such that* $P_a(\rho_3 \cup \rho_4) > 0$ *and a measure* P_b *such that* $P_b(\rho_2 \cup \rho_3) > 0$*;*

(ii) $\mathcal{E}(n)$ *contains two points* $Q, Q' \in \mathcal{E}(n)$ *for which there is* $x \in D$ *and* $\alpha \in (0, 1)$

such that, if $Q'' = \alpha Q + (1-\alpha)Q'$, $x = \Pi^{Q''}$.

It is immediate to characterize the cases in which abstaining is indifferent to voting for one (or both) candidates. Adding everything together provides necessary and sufficient conditions for abstention to be weakly optimal, i.e., there is no $f \in \{a, b\}$ such that $U(f) > U(\phi)$.

Corollary 1 *The voter weakly prefers abstention over voting for either candidate* if and only if either *the conditions of Theorem 1 hold,* or *both the minima in Eq. (7) are non-positive,* or finally *if one of the minima is positive and the set $\mathcal{D}(n)$ only contains measures which assign positive probability to the ρ on which the corresponding action is equal to zero.*[16] *(For instance if, say, we have $\min_{Q\in\mathcal{E}(n)} \psi(Q) > 0$, then we need that $P(\rho_3 \cup \rho_4) = 0$ for every $P \in \mathcal{D}(n)$.)*

The second condition in Corollary corresponds graphically to the case in which the the diagonal D is "tangent" to set $\Pi(\mathcal{E}(n))$: it weakly leaves it all to one side (this would happen for instance if the voter expected to be indifferent among the two candidates, so that $\Pi(\mathcal{E}(n)) \subseteq D$). The third corresponds to the case in which the diagonal leaves the set $\Pi(\mathcal{E}(n))$ properly to one side, but beliefs on $\mathcal{R}$ are such that the voter thinks it impossible that his vote will determine the election of the clearly better candidate.

Remark 3 Turning the conditions in Corollary 1 on their head tells us when the voter will strictly prefer to vote for one of the two candidates over abstaining. In particular suppose that the set $\Pi(\mathcal{E}(n))$ is properly to one side of the diagonal *and* every measure in $\mathcal{D}(n)$ assigns positive (possibly very small) probability to a tie (ρ_3).[17] Then there is a candidate such that voting for her is strictly better than abstaining: For instance suppose that $\Pi(\mathcal{E}(n))$ is all below the diagonal, then if the condition on $\mathcal{D}(n)$ holds voting for candidate B will be strictly preferred to abstaining, and to voting for A. ◇

Finally, it is interesting to observe when the "completely ignorant" voter described in Example 1 satisfies the conditions of Theorem 1, for that obviously provides an "upper bound" to the degree of abstention: If the completely ignorant voter does not abstain, then any voter with the same preferences (under certainty) and a smaller set $\Pi(\mathcal{E}(n))$ will not abstain.

Example 1 (continued) If the voter is completely ignorant (with strongly independent beliefs) then $\mathcal{D}(n) = \Delta(\mathcal{R})$ and $\mathcal{E}(n) = \Delta(S(n))$, so that in particular $\Pi(\mathcal{E}(n)) = S(n)$. Then condition *(i)* in Theorem 1 is clearly satisfied. Condition *(ii)* is satisfied if the square $S(n)$ is nontrivially intersected by the diagonal, as is for instance the case in Fig. 1. Formally, it is satisfied when $r(B, n) - l(A, n) > 0$ and $r(A, n) - l(B, n) > 0$. On the other hand, a voter who is completely ignorant on election results would never strictly prefer to vote for a candidate, say B, even if the square $S(n)$ were all below the diagonal (so that he is certain that candidate B is better than A). In fact then there is a measure in $\mathcal{D}(n)$ which assigns probability zero to the event ($\rho_2 \cup \rho_3$) on which the voter is pivotal and determines B's

[16] Notice that it is impossible that both minima in Eq. (7) be positive.

[17] More generally we could require: Every $P \in \mathcal{D}(n)$ is such that $P(\rho_2 \cup \rho_3) > 0$ and $P(\rho_3 \cup \rho_4) > 0$.

victory, and that measure will be used in evaluating the expectation of b, giving $U(b) = 0$. That is, "maximin" behavior on the matrix in Table 2 implies that the only action that *can* be strictly optimal is abstention. That is admittedly quite extreme, but as we mentioned before, we think of complete ignorance only as a limiting case. △

5 Clueless vs. Sherlock Holmes: Roll-off and the Comparative Statics of Information

Having characterized the conditions under which a voter will choose to abstain in a given election, we now move to analyzing how he will behave across different elections, when he is endowed with different information across them. In particular, we expect that he will have more information in *large* elections (elections with a large electorate, e.g., that for the President of the U.S.) than in *small* elections (e.g., school board elections), and so we are interested in seeing how his behavior will be affected by the size of the election.

Our analysis here is divided in two parts. The first discusses the comparative statics problem of how the behavior of the voter in *one* election is affected by his information on the two candidates. The second part uses the comparative statics result to compare the behavior across different elections.

5.1 Comparative Statics in a Single Election

We are interested in comparing the voter's behavior when he has m pieces of information on the candidates' policy position and when he has n pieces, with $m < n$. Let us start by making an assumption on the beliefs on $\mathcal{R}$ which will make the analysis less trivial and cleaner. This we dub the *relevance* assumption: For every n and every measure P in $\mathcal{D}(n)$, $P(\rho_2 \cup \rho_3) > 0$ and $P(\rho_3 \cup \rho_4) > 0$. It is for instance satisfied if (as we assumed in Remark 3) every $\mathcal{D}(n)$ contains only measures which assign a positive probability to a tie. More in general relevance requires that, regardless of his information (and hence of the size of the election), the voter believes that there is a positive (albeit possibly vanishing) probability that he is pivotal. Conditionally on being pivotal, the voter is also not certain that a *specific* candidate will win by one vote if he does not cast his vote. That is, there is a positive probability that the voter actually *determines* the winner of the election, rather than just causing a tie (this explains our choice of name for the assumption).

It is immediate to go back to the results of the previous section, and to observe why relevance makes our analysis sharper. First of all if relevance holds then condition *(ii)* of Theorem 1 is satisfied (by every P in this case). Second, (as we observed in footnote 17) if the set $\Pi(\mathcal{E}(n))$ is strictly to one side of the diagonal D, then under relevance the voter strictly prefers voting for one candidate over abstaining.

The comparative statics results follow immediately from Lemma 1 and the conditions for abstention. Suppose that given m pieces of information the square $S(m)$ has no intersection with the diagonal D. Then as observed above, the voter strictly prefers to vote for some candidate, say A. Imagine now that the voter learns more about the two candidates: He learns about the candidates' positions on $n-m$ additional policy issues. The set of possible candidates values is now given by the square $S(n) \subseteq S(m)$. By consequentialist beliefs, after "updating" his belief set he will have the possible values' set $\Pi(\mathcal{E}(n)) \subseteq S(n) \subseteq S(m)$, so that he will still prefer to vote for A. This will in particular be true for the voter who is completely ignorant on the policy space, as we discuss in the following

Example 1 (continued) Consider a voter whose beliefs on the election results space satisfy relevance (and hence cannot reflect complete ignorance), and whose beliefs on the policy space are completely ignorant, i.e., for every n, $\Pi(\mathcal{E}(n)) = S(n)$. The observation above immediately implies that the voter will strictly prefer abstaining for every n such that $S(n)$ has a nontrivial intersection with D. Suppose that the real policy positions of the two candidates, labeled $(\hat{y}_A, \hat{y}_B)$ are such that the associated values $(\hat{\pi}_A, \hat{\pi}_B)$ are distinct (so that they do not give a point on D), say $\hat{\pi}_A < \hat{\pi}_B$.[18] Then, by the observation after Lemma 1, there is a finite n such that after receiving n pieces of information $S(n) \cap D = \emptyset$, so that the voter switches to voting for A forever (i.e., for any additional piece of information). Thus he abstains with poor information, and votes with sufficient information.

A nice feature of completely ignorant beliefs is that, due to the observation above, the set $\Pi(\mathcal{E}(n))$ shrinks monotonically in n, which implies that whenever the voter decides to vote for one candidate, he will not change his mind after knowing more information. Also, abstention is naturally observed for low values of n, as long as the voter cares for more than very few issues. $\triangle$

For the voter who is not completely ignorant, things are unfortunately not so stark. When $S(n)$ has a nontrivial intersection with the diagonal (so that also $S(m)$ does), it is possible that more information will make the voter switch from, say, voting for A to abstaining. In fact, we have not excluded the possibility that the set $\Pi(\mathcal{E}(m))$ is all above the diagonal, and the set $\Pi(\mathcal{E}(n))$ has some point below it, which yields a vote for A with m pieces of information, and an abstention with n pieces of information. What *does* happen in general is that, assuming once again that $\hat{\pi}_A \neq \hat{\pi}_B$, there is a finite n such that after knowing the candidates' position on n issues the voter switches definitely to voting for the best one, and this regardless of how he dynamically modifies his beliefs. This is of course just a consequence of Lemma 1 and the consequentialist beliefs assumption.

Proposition 2 *Given a voter satisfying the assumptions of Section 3 and relevance, if the true values of the candidates are different, say $\hat{\pi}_A < \hat{\pi}_B$, then there is N such that when the voter knows $n \geq N$ pieces of information, he will vote for A, whatever his (consequentialist) beliefs.*

The problem is that not much more can be said about what he will do for values $n \leq N$, not unless we put some constraints on the dynamics of belief sets. Looking back at the example of the completely ignorant voter, one might think that an obvious sufficient condition for the voter to behave monotonically (i.e., he might start by abstaining, and then switch to voting

[18] Otherwise we will not be able to obtain strict preference for *any* action in the limit.

for one candidate forever) is that the belief sets shrink monotonically, $\Pi(\mathcal{E}(m)) \subseteq \Pi(\mathcal{E}(n))$. That is certainly going to yield the required result, but it is too strong a condition, as the following example shows.

Example 2 Suppose that the voter has a single prior $Q_n(\cdot) \equiv Q(\cdot|S(n))$ on policy space, and updates it according to Bayes's rule. Assume that his prior assigns positive probability to the actual pair of values $(\hat{\pi}_A, \hat{\pi}_B)$, so that Bayesian updating works for every information vector he might have. The condition that $\Pi(\mathcal{E}(m)) \subseteq \Pi(\mathcal{E}(n))$ then implies that for $j \in \{A, B\}$,

$$\int_{S(m)} \pi_j \, Q_m(d\pi) = \int_{S(n)} \pi_j \, Q_n(d\pi).$$

Thus the only way in which a Bayesian updater can satisfy the condition above is by having beliefs which always assign the same expectation to the values of the candidates. But it is immediate to see that, except in trivial cases, this will not be true: Bayesian updating will in general have the effect of moving the expectation of the candidates' values around. It also easy to construct an example in which the expectation bounces many times above and below the diagonal, moving the voter's preferences on candidates back and forth. △

A similar problem could arise if we had a voter whose beliefs are not necessarily represented by a single prior, but who updates his beliefs in the following "naive Bayesian" fashion: Suppose that with m pieces of information his beliefs are given by the set $\mathcal{E}(m)$, and that he learns $n - m$ additional coordinates of the candidates' policy position. Then he "updates" $\mathcal{E}(m)$ into $\mathcal{E}(n)$ as follows: For every $Q \in \mathcal{E}(m)$, he discards Q if it is incompatible with (i.e., assigned probability 0 to) the observation of the additional coordinates; He updates Q using Bayes's rule if it is compatible, and puts the resulting measure Q' in the set $\mathcal{E}(n)$. Then, once again, it can easily happen that the set $\Pi(\mathcal{E}(m+1))$ is larger than the set $\Pi(\mathcal{E}(m))$, violating monotonicity.[19] As for Example 2, it is easy to construct examples of dynamic preferences which bounce around between voting for the two candidates.

Remark 4 This is probably a good point to make a (fairly technical) comment about learning and Bayesian updating. A well-known result on Bayesian updating by Blackwell & Dubins (1962) shows that if the beliefs of a voter are given by a single prior Q and assign positive probability to the true values' pair $\hat{\pi} = (\hat{\pi}_A, \hat{\pi}_B)$, then his posterior Q_n will converge ("merge") to a probability degenerate on $\hat{\pi}$.[20] This in particular implies that the expectation according to Q_n will also get closer to the truth $\hat{\pi}$. So, while Lemma 1 insures that consequentialist beliefs

[19] The same would happen with an updating rule discussed by Gilboa & Schmeidler (1993), which is a refinement of the one discussed here. That rule discards all priors which do not give maximum likelihood to the observed evidence.

[20] In fact in the context of this model a much stronger result could be obtained. For instance, to obtain merging it is only necessary that truth is in the *support* of the voter's prior beliefs (and this also insures that the beliefs obtained by Bayesian updating are consequentialist). That is, every neighborhood of $\hat{\pi}$ is given positive probability by Q.

must converge to a small neighborhood of truth in finite time, the merging due to Bayesian updating *might* make the convergence to truth *faster*.[21]

A similar result could possibly be proved for the voter with multiple priors, if he, say, updates his beliefs according to the naive Bayesian rule outlined above. Assume that at least one prior in the voter's belief set $\mathcal{E}(n)$ assigns positive probability to truth. Then, we conjecture that as a result of the updating rule the *set* $\mathcal{E}(n)$ will merge to a measure degenerate on truth, in the sense that for every $\epsilon > 0$ there is an N such that for all $n \geq N$, the expected values with respect to Q, for every $Q \in \mathcal{E}(n)$, will be in a ball of radius ϵ around $\hat{\pi}$. This type of convergence however will likely take place at the same speed as the one due to Lemma 1, for if there are in the initial set priors which do not assign positive probability to truth (more in general, do not merge), these will eventually be ruled out due only to the collapse of the set $S(n)$. Faster convergence *might* again take place if all the priors in the initial set satisfy the condition for merging. Even then we have the problem that the speed of convergence will vary from prior to prior (and hence from set to set). Thus it seems that in general we can only say: The presence of Bayesian updating can make the shift towards a definite choice of one candidate take place a bit sooner than the monotonic reduction of $S(n)$, but not necessarily. $\diamond$

While the previous considerations show that the nice monotonicity of behavior that we observe in the case of the completely ignorant voter will not be observed in general, one might wonder whether there are conditions on his beliefs which will make the voter only oscillate between voting for one candidate and abstaining. Interestingly, this happens under a fairly general condition. Before seeing it, it is useful to look at a stronger assumption which is quite transparent. Suppose that $\hat{\pi} \notin D$, say $\hat{\pi}_A < \hat{\pi}_B$, and that for every n, $\hat{\pi} \in \Pi(\mathcal{E}(n))$, the true value is always in the set of values that the voter considers possible. Then by Theorem 1 and Remark 3, the only possibilities are that he either abstains or votes for A, the better candidate. This will depend on whether the set $\Pi(\mathcal{E}(n))$ contains also a point below the diagonal.

The condition above is however still a bit too strong. For instance it would not necessarily be satisfied even if the voter assigned a positive probability to truth. The following generalization might however be satisfied in a significant number of cases. We say that the voter has *sign-correct* beliefs if for every n his belief set $\mathcal{E}(n)$ contains at least a prior Q which has the same sign as truth. That is, if the true values are such that $\hat{\pi}_A < \hat{\pi}_B$ then

$$\int_{S(n)} \pi_A \, Q(d\pi) < \int_{S(n)} \pi_B \, Q(d\pi).$$

In other words, there is a prior in the voter's belief set which assigns sufficient weight to the event that $\pi_A < \pi_B$ (so that the expectations are also above the diagonal). Clearly this generalization works for the same reason as the proposal we discussed above, so we finally obtain

[21] Precisely, by faster convergence we mean the following: For every $\epsilon > 0$, if the set $S(n)$ is contained in the ball of radius ϵ around $\hat{\pi}$, then the expectation according to Q_n is also in the ball, and is possibly much closer than ϵ to $\hat{\pi}$.

Proposition 3 *Suppose that the voter satisfies the assumption of Section 3 and relevance, that $\hat{\pi} \notin D$ and that the voter has sign-correct beliefs. Then for every $n \geq 0$ he will either abstain or vote for the best candidate.*

5.2 "Large" vs. "Small" Elections

Having observed how the voter described in Section 3 will behave in a single election according to the amount of information he has regarding the candidates, we now move to comparing his behavior across different elections. Doing this properly requires a full model of how information gets transmitted from candidates to media, and then to voters. For the sake of brevity in this subsection we only develop a "reduced form" model of information acquisition, which we hope still captures some important features of the process by which voters receive information. The extremely simplified picture we have in mind is that of a voter who has, prior to relocating himself into the ballot booth, spent all his life in front of a TV, which every now and then provides him with news programs discussing the positions of candidates on different elections. As we pointed out before, we make the fairly innocuous assumption that information bits have the form of a comparison of the two candidates on their relative position regarding an issue. That is, they are of the form: "On the issue of whether permanent residents should be allowed to buy firearms, candidate A has restated his favorable position, while candidate B strongly opposes it." More substantially, we also assume that the voter receives stark information, and believes it.

Besides these assumptions, which we believe to a significant extent relaxable without altering the nature of our results, our caricatural description of the voter shows that in this subsection we implicitly view him as a passive subject in the information retrieval stage. Of course nothing in the model developed so far depends on the way information is gathered (except for the two assumptions we just mentioned), but in the analysis here we prefer just to assume that information is handed out by the media to the voter who collects it costlessly. Modelling active (and costly) search for information on part of the voter requires discussing in depth the issue of the value of information in this model, and introduces serious complications to the analysis. We thus prefer to leave it as an important task for future research.

The same is said for the part of the model dealing with the supply side of information. Leaving aside the question of the strategic revelation of information on part of the candidates (another crucial aspect to be included in future developments), it is clear that a complete model of the information about different elections must carefully describe the media's incentive to provide such information. We expect, though, the following result to hold in any model of information provision by media: *Except for very small electorates, the larger the electorate for a given election is, the more likely it is that a new piece of information regarding that election will be provided.*[22] The intuitive reason for this is that the larger the electorate, the more the media will be interested in broadcasting information about an election, for a

[22] The exception is due to the fact that for elections with a very small electorate, the monotonicity might be going in the opposite direction: The smaller the electorate, the higher the amount of information (everybody knows everybody) available about the election. As we shall see below, this is totally in line with the prediction of our model.

larger share of the audience will be (at least marginally) interested in knowing more about the candidates. The same could be said of elections which, for one reason or another, are more *salient* in the media's (and presumably public) opinion. For instance propositions on delicate issues, such as California's recent Prop. 187, restricting undocumented immigrants access to state services.

Suppose therefore that there are two elections on the ballot, which we denote 1 and 2. Each election i, $i = 1, 2$, has two candidates A_i and B_i competing. Imagine that the voter has m bits of information on election 1, and $n > m$ bits on election 2.[23] Since we are now in a truly multi-election set-up, we also have to specify the voter's preferences in this more complex environment. To do this, we take the easiest possible formulation and make the following assumptions. We assume that the voter's preferences under certainty for the pair of election outcomes are additively separable. That is, there are vectors w^1 and w^2, representing the voter's preferences among candidates in election 1 and 2 respectively, and the utility of the pair of candidates (j_1, j_2), $j_i \in \{A_i, B_i\}$, being elected is the *sum* of the two values obtained as in Section 3. Abstention is a reference choice in both elections, and the voter's (sets of) beliefs are assumed to be *stochastically independent* across the two elections and satisfy all the other assumptions of Section 3. This implies that the voter makes *separated* choices in the two elections: He chooses in election 1 the action which is optimal for that election alone, and then does the same for election 2. While additive separability of preferences is restrictive and material to the analysis here, stochastic independence is assumed only for expository purposes, and it could be relaxed without altering the nature of our results (which are based on the dynamics of the supports, and not on the specific form of the beliefs). Finally, as in the previous subsection we rule out some trivial cases by assuming that the voter's beliefs satisfy relevance for every election.

As before, we start by discussing the completely ignorant voter, who as usual provides the cleanest results and a useful benchmark case.

Example 1 (continued) Here the voter has beliefs on election 1 given by the set $\mathcal{E}^1(m) = S^1(m)$ and on election 2 given by $\mathcal{E}^2(n) = S^2(n)$. As we pointed out in the previous subsection, in this case the voter with n bits of information abstains in election i if the set $S^i(n)$ intersects the diagonal. That is, if we let $\hat{\pi}^i$ be he true values' pair for election $i = 1, 2$ we have

Proposition 4 *Suppose that $\hat{\pi}^1 \notin D$ and $\hat{\pi}^2 \notin D$. There is $m \geq 0$ small enough and n large enough such that the voter abstains in election 1 and votes in election 2.*

Clearly we have to allow $m = 0$ since it is possible that the voter only cares about the first issue on election 1, so that $w^1_1 = 1$. △

When the voter is not completely ignorant, we cannot conclude that he will abstain with little information, but we expect that to be the case if his beliefs are sufficiently spread out,

[23] Without any loss of generality, we label the issues on each election according to the order in which they are presented to the voter. That is, the candidates' position on issue 1 is the first bit of information given to the voter regarding the election, and so on.

so that condition *(ii)* of Theorem 1 holds. Also Proposition 2 implies that for n large enough the voter will strictly prefer one of the candidates (unless they are truly indifferent to him).

Proposition 5 *Suppose that the voter satisfies the assumptions mentioned above, and that* $\hat{\pi}^1 \notin D$ *and* $\hat{\pi}^2 \notin D$*. Then there is an* $m \geq 0$ *small enough and* n *large enough so that the voter might (if set* $\Pi(\mathcal{E}^1(m))$ *satisfies condition (ii) of Theorem 1) abstain in election 1 (he might vote for the correct candidate, if his beliefs are sign-correct), while he certainly votes in election 2.*

While it is as not as stark as one might like, we believe that this Proposition captures most of our intuition on the "paradoxes" of the patterns of voting. We expect a voter with little information about an election to be quite "uncertain" about the possible values (more generally: the possible *ranking*) of the candidates. That is, he has a fairly large $\Pi(\mathcal{E}(m))$ with a nontrivial intersection with the diagonal D, and thus strictly prefers abstaining over voting for the possibly "wrong" candidate. On the other hand, when well-informed about the positions of the candidates, the voter's beliefs are naturally more concentrated, and (especially if one candidate is significantly better than the other) he is likely to cast his vote for the "right" candidate.[24]

A couple of observations are now in order. Notice that the size of the m and the n in Proposition 5 depend quite crucially, via Lemma 1, on the form of the vectors w^i, $i = 1, 2$. If the voter cares only for few issues which are discussed early on, then he will switch to voting for the best candidate very soon.[25] Another factor which (for given vector w^i) affects the speed at which the voter switches definitely to voting is the distance of the two candidates in terms of values: The farther $\hat{\pi}^i$ is from D, the sooner the switch will take place. Thus for instance our model predicts lower abstention in elections in which there are few very salient policy issues (w^i is positive for a small number of issues for most voters), possibly heavily discussed by the media, like a number of California's Propositions in recent years. Analogously, we do not necessarily predict that voters will abstain in elections with a small electorate (say, the school board elections), especially when these elections are based on a small number of issues, and the election is so small that the voter might have good information (like "my neighbor knows him and says...") about the candidates' position on these issues.

6 A Discussion of the Other Candidates

Having seen the model and its implications for voting in multiple elections, it is now a good time to pause and analyze our choices and the possible alternatives. We start by

[24]It will not have escaped the reader's attention that in this model there is *by construction* "information aggregation" for high levels of information of the voters. Clearly the assumption of consequentialist beliefs is the driving force behind this. Whether elections aggregate information in other set-ups is however subject of current research (see Austen-Smith & Banks 1996, Ledyard 1989, McKelvey & Ordeshook 1985).

[25]Remember that the numbering of issues refers to the order in which they are discussed, not to their importance for the voter.

discussing some specific aspects of our model, and then move on to discussing some more radical departures which might also explain roll-off.

As for *preferences under certainty*, we do rely heavily on the separability across issues that is built in our structure. We do not believe that it will be generally satisfied, but neither we think that it is terribly restrictive. It is quite standard in the literature. Analogously, state independence is an assumption which is made in most rational choice models, but which *does* instead impose a strong restriction. It is crucial for our model, though: Forfeiting it would make the analysis of the voter's decision problem extremely difficult.

We have mentioned previously the possibility that $Y_i = [0, 1]$. Doing so would have the double advantage of allowing the possibility that: 1) The politician does not really have a well-defined position on every issue, her position y_i then represents the *probability* that she will have a negative (opposite to the voter's point of view) position whenever the issue is discussed; 2) The voter does not believe the candidate when she declares her position on a given issue, and again assigns probability y_i to the event that the candidate will have a negative position. The main reason why we chose not to include this in this model is that it would introduce a lot of complications in the updating procedure, which we totally avoid here. Complexity of the exposition aside, we do not believe that this extension would change our results significantly.

Coming to the structure of the voter's preferences under uncertainty, we have already mentioned that the stochastic independence assumption, while simplifying the analysis quite substantially, does not really modify the intuition behind our results. We relax it in a companion paper, Ghirardato & Katz (1997), and obtain similar results. In its absence however our argument that the voter with a single prior on values space (who is therefore not ambiguity averse) does not abstain is no longer true: There is a type of correlation between beliefs on $\mathcal{R}$ and $S(n)$ which can generate abstention even by a SEU maximizing voter. For that to be true, however, it is necessary that differences of one vote (that is, movements from ρ_2 to ρ_3 and from ρ_3 to ρ_4) significantly affect the voter's beliefs on the relative position of the candidates. This is for instance the case in the afore-mentioned papers by Feddersen and Pesendorfer (1996*b*, 1996*a*), where the right type of correlation results from an assumption that voters' preferences under certainty are aligned in a certain fashion, and that voters can be differentially informed. The problem with this justification of abstention is that, as it happens in most game-theoretic models, it relies heavily on the assumption that every voter knows the structure of the voting game. If a voter is not certain that all other voters have this particular type of preferences (and that they are differentially informed in a specific way) then it is not clear that he would place such importance in differences of one vote. After all, few people are as smart as game theorists make them, and if there are voters who are not as sophisticated as the model requires then a difference of one vote might not necessarily arise as a result of careful equilibrium reasoning. By making the clearly extreme assumption of stochastic independence, here we show that there is a *complementary* explanation of abstention which is more robust to changes in the voter's understanding of his environment.

Another important assumption for our model is that abstention is considered a reference choice. It is worth repeating that the bite of this assumption comes from its joining with the multiple prior model, as we saw in Remark 1. Our intuitive motivation for imposing it is the anecdotal observation that abstention is viewed by most people as an option which is "safe." It does not entail the same costs as voting for a candidate, which could lead to a situation of discomfort ("regret"), if one's vote determines the election and the candidate turns out to be the wrong person. A natural variant of the model is to assume full regret (*à la* Savage, see Savage (1951) and Ferejohn & Fiorina (1974)), in the sense that the voter considers a "mistake" any action which is not *ex post* optimal, rather than only those action which do worse than abstention. It is possible to see with some work that such a variant produces results analogous to the ones we obtain here, at the cost of less straightforward analysis. We chose the model presented here for the sake of simplicity and because we find a comparison with abstention more intuitive and plausible than a comparison with the optimal action. Strictly speaking, though, neither form of renormalization of payoffs is *needed* to obtain abstention. In fact abstention can be obtained in a pure ambiguity aversion model in which the voter does not renormalize the payoffs at all. Here abstention is a result of the pessimism in the valuation of the policy position of the two candidates and preference for "objective" randomization: If the voter is sufficiently ignorant then he favors abstaining over voting, basically because he prefers the "risky" randomization over A and B given by the coin toss over the purely "uncertain" prospects given by obtaining either candidate for sure. If the tie-breaking rule is not an objective randomization like a flip of an (unbiased) coin, say if the incumbent is confirmed in office, then we do not necessarily have abstention. The same happens if the voter's beliefs over the policy space are stochastically independent (as defined above) *across candidates*, a possibility which cannot be ruled out, for then the preference for randomization does not obtain. As we find neither of these results plausible and conforming to our intuition, we strictly prefer the model we present here to one where there is only ambiguity aversion.[26] Emphatically, we do not think that preference for objective randomization is what explains abstention, even in a world where voters do not have very good information on the candidates' positions.

Coming to the ambiguity aversion which is embodied in our "maximin" formulation, it is important to remark that such extreme attitude is not really required for obtaining abstention. Suppose for instance that the voter's preference functional is given by a convex combination (with weight $\alpha \in [0, 1]$) of the smallest possible integral with respect to measures in $\mathcal{C}(n)$, as in Eq. (3), and of the *largest* possible integral (substitute "max" for "min" in Eq. (3)). Then α can be taken as an index of the ambiguity aversion of the voter, ranging from extremely ambiguity averse ($\alpha = 1$) to extremely ambiguity loving ($\alpha = 0$).[27] It is immediate to see that our results hold in general for large values of α (while they clearly fail for α very small), and not only for $\alpha = 1$.

Moving away from our model, let us spend a moment discussing a different explanation

[26]As we remarked before, the results of this paper carry over with little modification to the case in which the incumbent staying is the tie-breaking criterion. Assuming that A is the incumbent, the only difference there is that $b(\rho_2) = 0$ and $a(\rho_3) = 0$, but the other entries of Table 2 remain identical. Hence the conditions on $\mathcal{E}(n)$ in Theorem 1 do not change, and the conditions on $\mathcal{D}(n)$ have to be strengthened slightly.

[27]A decision model with a similar preference representation is, e.g., in Ghirardato (1995). In the literature on decision theory, α is known as the *Hurwicz pessimism index*.

of roll-off, which is only apparently similar to ours. One could argue that the act of voting entails some "fixed" (independent of n in our model) costs that abstention does not entail. For instance a voter has to read the ballot and possibly turn many pages to find the actual place where the vote has to be inscribed. Suppose that a voter does not really care much for an election (which we might model by letting $w \equiv \sum_i w_i$ be variable and very small, rather than normalizing it to 1), and has no information on the candidates. Then his utility for voting can be much smaller than the fixed cost of voting, and he hence abstains. If instead, the voter cares for an election (and thus has high w) the utility might be higher than the cost, and so the voter chooses to vote. The problem with this explanation is that it runs into the traditional problem of obtaining *too much* abstention. If the utility of voting depends (as it probably would) on the probability that the voter is pivotal, and the voter is realistic in his calculation of the latter, it is almost impossible that he will vote on large elections, in which the probability of being pivotal is infinitesimal.[28] We frankly doubt that then we should ever see anybody voting for the President's election, even if the cost for casting that vote is very small and the voter possibly cares for its result.

We are therefore not totally convinced that such a model can adequately explain observed behavior. In our opinion the real fixed cost is that of placing the voter in the ballot booth, and we tend to think (in accordance with the received view in the literature) that such fixed cost is probably offset by some form of "citizen duty" value for going to the polls. Once there, we do not consider it likely that the voter will face significant costs, especially for elections in which he is well-informed. One could imagine enriching the model we presented by adding an (all-encompassing) "variable" cost factor to voting, which depends negatively on the amount of information that the voter has, and becomes zero with finitely many bits of information.[29] One would then obtain results similar to the ones we have here, with some more abstention in obscure elections due to the variable cost factor. The problem we have with such an extension is that we do not know how to model the variable cost, beyond the indirect way that we use here. We are certainly not keen on doing this just by assuming a given cost function. On the other hand we feel that our model captures a significant part of the "variable" costs to voting, so we leave it to future work to develop this aspect more fully.

Appendix

[28] This is for instance what happens in the game-theoretic models of voting of Ledyard (1984) and Palfrey & Rosenthal (1983), in which in large elections only voters with zero (or negative) fixed cost vote in equilibrium.

[29] That would probably require enriching our notion of information, to include also technical aspects of the election unrelated to the candidates' positions, like the standard format of the poll. For instance in Italy it is quite typical that media spend a considerable amount of time explaining what a vote in a *referendum* means when translated in practical terms. In fact the questions are formulated in an obscure bureaucratic language, so that a voter could easily find himself in the position of having decided what to vote for, but being unable to identify it on the poll!

A Separability

Here we show that given strong independence, the functional U can defined in Eq. (3) can be rewritten as in Eq. (4). This makes use of some results contained in Ghirardato (1997). There it is shown that, given sets of beliefs $\mathcal{D}$ and $\mathcal{E}$, respectively defined on two sets $\mathcal{R}$ and S,[30] if a "product" belief $\mathcal{C}$ on $\mathcal{R} \times S$ is the stochastically independent product of $\mathcal{D}$ and $\mathcal{E}$, if a function $F : \mathcal{R} \times S \to \mathbf{R}$ has *affinely related ρ-sections* then we have

$$\min_{R \in \mathcal{C}} \int_{\mathcal{R} \times S} F(\rho, \pi) \, R(d(\rho, \pi)) = \min_{P \in \mathcal{D}} \int_{\mathcal{R}} \min_{Q \in \mathcal{E}} \int_S F(\rho, \pi) \, Q(d\pi) \, P(d\rho), \tag{9}$$

which gives us Eq. (4) if, for $f \in \{a, b, \phi\}$, we substitute for F the function

$$\Gamma_f(\rho, \pi) \equiv f(\rho, \pi) - \phi(\rho, \pi).$$

So we are done if we show that for every $f \in \{a, b, \phi\}$, $\Gamma_f(\rho, \pi)$ has this property. First though, we need to give the formal definition:

Definition 1 *Given $F : \mathcal{R} \times S \to \mathbf{R}$, we say that F has* affinely related ρ-sections *if for all $\rho, \rho' \in \mathcal{R}$, $F(\rho, \cdot)$ and $F(\rho', \cdot)$ are affinely related functions from S into $\mathbf{R}$. That is, there exist $\alpha \geq 0$ and $\beta \in \mathbf{R}$ such that either $F(\rho, \pi) = \alpha F(\rho', \pi) + \beta$ for all $\pi \in S$, or $F(\rho', \pi) = \alpha F(\rho, \pi) + \beta$ for all $\pi \in S$, or both.*

As the name suggests, F has affinely related ρ-sections, if every pair of sections along the S axis are functions which are a positive affine transformation of each other (or at least one is a constant).

The fact that each Γ_f has affinely related sections follows immediately from the observation of Table 3, which plots Γ_f for every $f \in \{a, b, \phi\}$, at a given value $\pi \in S$. For instance, Γ_a has affinely related ρ-sections, for the sections are either identically 0, or they are equal to the function $1/2(\pi_B - \pi_B)$.

References

Aldrich, John H. 1993. "Rational Choice and Turnout." *American Journal of Political Science* 37(1):246–278.

Anscombe, Francis J. & Robert Aumann. 1963. "A Definition of Subjective Probability." *Annals of Mathematical Statistics* 34:199–205.

Austen-Smith, David & Jeffrey S. Banks. 1996. "Information Aggregation Aggregation, Rationality, and the Condorcet Jury Theorem." *American Political Science Review* 90:34–45.

[30] The notation is the one used in Section 3.3, with the exception that here we disregard the dependence on n.

	ρ_1	ρ_2	ρ_3	ρ_4	ρ_5
Γ_a	0	0	$\frac{1}{2}(\pi_B - \pi_A)$	$\frac{1}{2}(\pi_B - \pi_A)$	0
Γ_b	0	$\frac{1}{2}(\pi_A - \pi_B)$	$\frac{1}{2}(\pi_A - \pi_B)$	0	0
Γ_ϕ	0	0	0	0	0

Table 3: The Graphs of Γ_f for Fixed π

Blackwell, D. & L. Dubins. 1962. "Merging of Opinions with Increasing Information." *Annals of Mathematical Statistics* 38:882–886.

Camerer, Colin & Martin Weber. 1992. "Recent Developments in Modeling Preferences: Uncertainty and Ambiguity." *Journal of Risk and Uncertainty* 5:325–370.

Choquet, Gustave. 1953–54. "Theory of Capacities." *Annales de l'Institut Fourier* 5:131–295.

Cox, Gary W. & Michael C. Munger. 1990. "Putting Last Things Last: A Sequential Model of Turnout and Roll-off." mimeo, University of California, San Diego.

Downs, Anthony. 1957. *An Economic Theory of Democracy.* New York: Harper & Row.

Ellsberg, Daniel. 1961. "Risk, Ambiguity, and the Savage Axioms." *Quarterly Journal of Economics* 75:643–669.

Enelow, James M. & Mevin J. Hinich. 1984. *The Spatial Theory of Voting: An Introduction.* Cambridge: Cambridge University Press.

Feddersen, Timothy J. & Wolfgang Pesendorfer. 1996*a*. "Abstention and Common Values." mimeo, Northwestern University.

Feddersen, Timothy J. & Wolfgang Pesendorfer. 1996*b*. "The Swing Voter's Curse." *American Economic Review* 86(3):408–424.

Ferejohn, John & Morris Fiorina. 1974. "The Paradox of Not Voting: A Decision Theoretic Analysis." *American Political Science Review* 68:525–535.

Ghirardato, Paolo. 1995. Coping with Ignorance: Unforeseen Contingencies and Non-Additive Uncertainty. Social Science Working Paper 945 Caltech.

Ghirardato, Paolo. 1997. "A Note on Independence and Fubini's Theorem for Multiple Priors Functionals." mimeo, California Institute of Technology.

Ghirardato, Paolo & Jonathan N. Katz. 1997. "Free to Not Choose: Conditions for Abstention in Multiple Elections." mimeo, California Insitute of Technology.

Gilboa, Itzhak & David Schmeidler. 1988. "Maxmin Expected Utility with a Non-Unique Prior." *Journal of Mathematical Economics* 18:141–153.

Gilboa, Itzhak & David Schmeidler. 1993. "Updating Ambiguous Beliefs." *Journal of Economic Theory* 59:33–49.

Ledyard, John O. 1984. "The Pure Theory of Large Two-Candidate Elections." *Public Choice* 44(1):7–41.

Ledyard, John O. 1989. Informtion Aggregation in Two-Candidate Elections. In *Models of Strategic Choice in Politics*, ed. Peter C. Ordershook. Ann Arbor: University of Michigan Press pp. 7–30.

Matsusaka, John G. 1995. "Explaining Voter Tunout Patterns: An Information Theory." *Public Choice* 84:91–117.

McKelvey, Richard D. & Peter C. Ordeshook. 1985. "Sequential Elections with Limited Information." *American Journal of Political Science* 29:480–512.

Palfrey, Thomas R. & Howard Rosenthal. 1983. "A Strategic Calculus of Voting." *Public Choice* 41:7–53.

Palfrey, Thomas R. & Howard Rosenthal. 1985. "Voter Participation and Strategic Uncertainity." *American Political Science Review* 79(1):62–78.

Riker, William H. & Peter C. Ordeshook. 1968. "A Theory of the Calculus of Voting." *American Political Science Review* 62(1):25–42.

Savage, Leonard J. 1951. "The Theory of Statistical Decision." *Journal of the American Statistical Association* 46:55–67.

Savage, Leonard J. 1954. *The Foundations of Statistics.* New York and London: J.Wiley and Sons.

Schmeidler, David. 1989. "Subjective Probability and Expected Utility without Additivity." *Econometrica* 57:571–587.