Ethical Voters and Costly Information Acquisition*

Timothy Feddersen and Alvaro Sandroni
MEDS Department, Kellogg School of Management
2001 Sheridan Road, Evanston IL 60208
tfed@kellogg.northwestern.edu; sandroni@kellogg.northwestern.edu

August 5, 2005

Abstract

Game theoretic and statistical models have emphasized the desirable information aggregation properties of large elections. However, such models do not explain why voters choose to acquire costly information. In this paper we use an ethical voter model to endogenize the decision to acquire information. We show that a significant fraction of the electorate will acquire costly information. However, even with negligible costs of acquiring information a fraction of the electorate will remain uninformed. Moreover, we show that as the quality of information increases information aggregation properties of election improve, but the fraction of informed voters may decrease. This result stands in contrast to previous models where the information aggregation properties of elections are insensitive to changes in the fraction informed. In addition, changes in the quality or cost of information affect the relative likelihood that each candidate wins the election.

*Sandroni gratefully acknowledges the financial support from the National Science Foundation Grant SES 0109650. All errors are ours.

1. Introduction

What impact, if any, does an electorate composed of apparently ill-informed citizens have on the quality of the collective decisions made via elections? The formal theory literature suggests that, with high probability, election outcomes would not change if everyone were perfectly informed.[1] On the other hand, empirical work by Bartels (1996) suggests that election outcomes would change if all voters were informed.[2]

Existing models (empirical and theoretical) typically assume exogenously determined information levels. The problem with this assumption is that information levels are ultimately the result of a choice by voters. The same factors that are relevant in choosing whom to vote for or whether to vote at all may play an important role in the decision to acquire information.

To understand the significance of information levels on election outcomes it is necessary to develop a theory that endogenizes the choice to become informed. This cannot be easily done by extending game-theoretic models of elections because, in large elections, the likelihood a vote is pivotal is negligible.[3] Introducing even small costs to acquire information would eliminate almost all incentives to become informed.[4] On the other hand if there are no information costs it is hard to justify uninformed voters.

In this paper we adapt the ethical voter model of Feddersen and Sandroni (2004a) to develop a model of large elections in which agents decide first whether to acquire costly information and, second, whether and for whom to vote. Ethical agents are motivated to vote and to acquire information out of a sense of civic duty. Such agents behave strategically even though pivot probabilities play no role in our analysis. In fact, we assume a continuum of agents so a single vote is never pivotal.

With no costs to acquire information our model replicates the game-theoretic results of Feddersen and Pesendorfer (1996): i.e., a fraction of less informed agents will strategically abstain even when voting is costless. However, in their model

[1] See, for example Austen-Smith and Banks (1996); Feddersen and Pesendorfer (1996, 1999), Grofman and Feld (1988), Miller (1996), and Young (1998).

[2] See also Berelson, Lazerfield and McPhee (1954), Page and Shapiro (1992), and Popkin (1991) for evidence related to the information possessed by voters and the implications of information on outcomes.

[3] See Feddersen (2004) and Feddersen and Sandroni (2004a) for a discussion of the problem of strategic pivotal voter models.

[4] See Martinelli (2002) and Persico (2004).

the fraction of informed voters is exogenously fixed and election outcomes are not sensitive to changes in the fraction of informed voters. In our model, we endogenize information acquisition and show that a positive fraction of the electorate will choose not to acquire information–even if the cost of information is minimal. We show that some level of political ignorance is not only unavoidable, but also socially desirable. Moreover, we show that changes in the quality and cost of information impact both the fraction informed and election outcomes.

In order to explain information acquisition several parameters are important: the expected difference in partisan support and uncertainty about this difference; the absolute level of partisanship within the population; the costs of acquire information and the quality of information. We develop comparative statics as a function of these parameters for levels of information acquisition, the probability the election outcome fully aggregates information and the probability each candidate wins the election. In particular, we show that elections may not perfectly aggregate information. As the quality of information increases information aggregation properties of election improve, but the fraction of informed voters may decrease. Moreover, such changes may benefit one candidate at the expense of the other.

In the next section, we give a detailed (but still informal) description of our model and results.

2. Informal Description of Model and Results

In our model society must choose between two alternatives by majority voting. Some voters have state dependent preferences (*independents*) while others always prefer one or the other of two candidates (*partisans*). There are two states–one in which all independents prefer the first candidate and a second state where all prefer the other candidate. Agents choose whether to be *informed* or *uninformed.* Informed agents receive a private signal correlated with the state.

We assume that *ethical* agents endogenously determine the behavior that they *should* take by evaluating the merits of different behavioral *rules* for their preference type. The rule that defines ethical action for an agent is the rule that produces the best outcome for the agent's preference type taking as given the behavior of other types of agents. We say that a behavioral rule profile is *consistent* if it defines rules that each type of agent decides they must follow, given a proper anticipation of the actual behavior of other types of agents.

To illustrate our model consider the following example. Suppose that there are two candidates 0 and 1 and that 10 percent of the voters in the population are partisans for 0; 30 percent are partisans for 1; and 60 percent are independents. Now assume that 1/2 of the independents are perfectly informed about the state while the other half are uninformed. Since voting is costless partisans will vote for their preferred candidate while informed independents will vote for the candidate they know to be best. In a consistent profile 2/3 of the uninformed independents will vote for candidate 0 and the rest will abstain. The votes of the uninformed independents compensate for the partisan advantage of candidate 1 and ensures that the votes of the informed independents will decide the election. This rule is optimal for independents because their preferred candidate wins with probability one in each state.[5]

An election outcome satisfies *full information equivalence* if the candidate preferred by the majority for the given state wins. Full information equivalence holds whenever the outcome of the election would not be altered if the state were common knowledge. In our baseline model, as in the example above, election outcomes satisfy full information equivalence with probability one. This is the same result obtained in Feddersen and Pesendorfer (1996).

Now suppose that ethical independents must choose not only how to vote but whether to become informed at some small cost. Consider a rule that directs 1/3 of the independent agents not to become informed but to vote for candidate 0; a very small fraction of the remaining independents are to become informed and the rest abstain. In this case the costs associated with information acquisition are minimal while, nevertheless, the candidate preferred by the independent agents for the realized state wins with probability one. This is the case because the vote by the uninformed independents exactly balances out the partisan advantage to candidate 1 and allows the very small fraction of informed independents to be decisive. When information is costly this rule is better than any rule that requires a larger fraction of the independents to become informed.

In the example above even arbitrarily small costs to acquire information result in almost no information being acquired. In our baseline model, full information equivalence holds even when a very small fraction of the independents are informed. As a consequence there is no incentive for ethical agents to gather information in the first place. In order to motivate substantial information acquisition we need an additional source of uncertainty.

[5] The reader might wonder if there are not multiple consistent profiles in this example. Our formal model is slightly more complicated than the example and this profile is the unique consistent one.

We now show that when partisan support is uncertain then costly information acquisition will occur. It is worth noting that the introduction of uncertainty will not lead to significant acquisition of costly information in fully rational models.

Suppose in the example above that between 0 and 20 percent of the voters in the population are partisans for 0; between 20 and 40 percent are partisans for 1; and 60 percent are independents. Thus, there is always a partisan advantage for candidate 1 but the magnitude of that advantage is uncertain. Given this uncertainty, it is impossible for uninformed independents to balance out the partisans with probability one. Full information equivalence is still possible, but only if a significant fraction of the independents become informed.

Consider the following rule. Let 1/3 of the independents become informed. That leaves 40% of the population as uninformed independents. Let 50% of the uninformed independents vote for candidate 0 while the rest abstain. Without the informed independents, neither candidate can ever be ahead by more than 20% of total votes. Since 20% of the electorate are informed independents this rule guarantees that the candidate favored by the independents always wins the election. In this example, the uncertainty over the partisan advantage requires at least 1/3 of the independents to become informed in order to guarantee full information equivalence with probability one.

However, even when full information equivalence is possible it may not occur. Agents evaluate rules not only on the basis of probability of electing their preferred candidate but they also take into account the total cost of acquiring information. More precisely, in a consistent profile the ethical rule for independents maximizes the probability their preferred candidate wins the election minus the total costs of acquiring information. Hence, if there is uncertainty over the difference in partisan support then full information equivalence requires sufficiently low costs of information. This follows because full information equivalence can only be achieved with a significant fraction of independents informed.

On the other hand, even when costs to acquire information are arbitrarily low it is *not* the case that *all* independents will choose to become informed. Indeed, whenever one candidate enjoys an advantage in partisan support the best rule for independents requires some of them to vote for the candidate with smaller partisan base (candidate 0) and reduce the difference in partisan support. Clearly, an agent who is going to vote for candidate 0 regardless of the signal should not acquire costly information. It follows that some level of political ignorance is both unavoidable and desirable.

When signals are perfectly correlated with states full information equivalence occurs with probability one for sufficiently low information costs. However, when information is not perfect then full information equivalence may not always be achievable. In the example above we required at least 20% of the electorate to be perfectly informed independents. Now suppose that signals coincide with the true state with probability only slightly greater than 50%. Then, even if all independent agents were to become informed the advantage in votes to their preferred candidate would be less than 20% of the electorate.

In our example, when the two states are equally likely it follows that, if the electorate were fully informed, half the time candidate 0 would win and half the time candidate 1 would be the winner. We say that the election is *biased* if one candidate is more likely to win the election than the other. In our model, as in the example, there are more partisans for candidate 1 than for candidate 0. As a consequence, candidate 1 will always be at least as likely to win the election as candidate 0. Hence we measure bias as the probability candidate 1 wins the election (minus 1/2).[6]

We show that when information acquisition is endogenous and the quality of the signals available to each voter is very poor then the probability of full information equivalence is low but the election is nevertheless unbiased. As the quality of the signals increase the probability of full information equivalence increases without introducing bias. However, when the quality of the signals achieves a certain threshold then a further increase results in both an increase in the probability of full information equivalence and bias. At some point the bias of the election is maximized. A further increase in the quality of the signals leads to an increase in the probability of full information equivalence and a decrease in bias.

It follows that bias is maximized when the quality of information is in an intermediary range and also that the relationship between bias and quality of information is not monotonic. An increase in the quality of information may benefit either candidate (and partisan group) depending upon parameter values.

The relationship between the quality of information and the fraction informed is also non-monotonic. The fraction of independents who become informed first increases and then decreases as the quality of information increases. To see why this must be the case consider the level of information that is sufficient to ensure victory of the candidate with larger partisan support (candidate 1) in the state that independents prefer candidate 1 (assuming that uninformed agents vote to

[6]In our general model there are some parameter values such that even if the electorate were fully informed candidate 0 would win less than half of the elections. As we will explain later, such cases complicate the relationship between bias and full information equivalence.

compensate for the difference in partisan support). If the level of information crosses this threshold then additional information no longer has any marginal benefit for independents.[7] In this case, even when information is almost costless, further increases in the quality of information lead to a lower fraction choosing to become informed.

Now consider the case in which signals are of very poor quality. Then, information has almost no value and only a very small fraction will become informed. As the quality of the signals increase information becomes more valuable and the acquired level of acquired information increases. This continues up to the point in which the acquired level of information reaches the aforementioned threshold.

The paper proceeds as follows. Section 3 contains the baseline model (when information levels are exogenously given) and the results in this framework. Section 4 endogenizes costly information acquisition and contains our main results. Section 5 provides a brief conclusion.

3. Baseline Model: Exogenous Level of Information

There are two equally likely states, 0 and 1, and two candidates, candidate 0 and candidate 1. There is a continuum of voters (or agents) each of whom is one of three types. Type 0 and type 1 voters are *partisans* for candidates 0 and 1 respectively and prefer their candidate to win the election independent of the state. Type i voters are *independents* who prefer candidate $w \in \{0,1\}$ in state w. Let $T \equiv \{0,1,i\}$ be the set of types. If candidate $w \in \{0,1\}$ is elected then partisans of type w receive a payoff of 1; partisans type $w' \neq w$ receive a payoff of zero; independents receive a payoff of 1 when the state is w and zero when the state is not w. The fraction of partisans in the electorate is $\beta \in (0,1)$. A fraction $k \in (0,0.5)$ of partisans support candidate 0. So, there are more partisans for candidate 1 than for candidate 0.

Every voter receives a message $m \in M \equiv \{0,1,\phi\}$. Voters who receive a message 0 or 1 are *informed* and all others are *uninformed.* The fraction of informed voters in the population is given by $q \in (0,1)$. Among the informed voters, the fraction who observe message $m \in \{0,1\}$ in state m is $\rho \in (.5,1]$. Thus, when ρ is close to 0.5 the message is a very noisy signal of the true state. When ρ is close to 1 the message almost perfectly conveys the true state.

[7] This is an alternative way of understanding why a positive fraction of independents will always remain uninformed.

Every voter chooses an action $s \in S \equiv \{a, 0, 1\}$, where a indicates abstention while 0 or 1 indicates a vote for candidate 0 or 1 respectively. The winner of the election is the candidate that receives the majority of votes cast.

We assume that voters derive strictly positive utility from acting *as they determine they should*. An agent votes *if* he understands that he should vote and abstains *if* he understands he should abstain.[8] Each agent independently determines whether or not he or she should vote (and whom to vote for) by conducting the following moral calculation. Taking the behavior of partisans as given, independent agents consider the outcomes produced by behavioral rules that specify an action for each independent agent as a function of their information. Each independent agent determines they should act according to a rule that produces the best outcome. A similar logic applies to partisans. Note that this does not mean that either all independent agents determine they should vote or that all should abstain. A rule may direct some agents to vote and others to abstain. Below we formalize these ideas.

Let $\Delta(S)$ be the set of probability measures over the action space S. A behavioral rule for type $t \in T$ voters is a function $e_t : M \longrightarrow \Delta(S)$ that maps messages into probabilities of taking each possible action. We assume that voters do not consider rules that require agents of the same type and with the same information to vote for both candidates with strictly positive probability. This restriction is justified in the appendix. All elements of $\Delta(S)$ can be described by a probability of voting and the candidate receiving the votes. By convention, an element of $\Delta(S)$ is a number $x \in [-1, 1]$. If x is positive then x is probability of voting for candidate 0. If x is negative then $-x$ is the probability of voting for 1.

A rule e_t for type $t \in T$ denotes the behavior type t voters *should* take. The following notation is useful: let $e = (e_0, e_1, e_i)$ be a rule profile, let $e_{-t} = (e_{t'})_{t' \neq t}$ be the rules for voters who are not type t (i.e., $e_{-0} = (e_1, e_i)$), and let E be the set of all rules.

Voters of type $t \in T$ rank behavioral rules e_t according to the outcome produced when *all* voters of type t follow rule e_t (taking the behavior of voters of other types as given). So, partisans for candidate $w \in \{0, 1\}$ rank rule e_w above e'_w if the probability that candidate w wins the election is higher under e_w than under e'_w. Analogously, independents rank rule e_i above e'_i if their expected payoff is higher under e_i than under e'_i. So, both partisans and independents' ranking of

[8] In this model there are no costs to vote. We believe that introducing such costs would unnecessarily complicate the analysis. See Feddersen and Sandroni (2004a) for an ethical voter model with costs to vote.

rules is ultimately implied by their preferences.

Let $R_t(e_t|\ e_{-t})$ be the expected payoff for agents type t if voters follow the rule profile $e = (e_t,\ e_{-t})$. Let $p(e,\omega)$ be the probability that candidate 0 wins the election when the state is $\omega \in \{0,1\}$ and all voters behave according to the rule profile e. Then,

$$R_0(e_0|\ e_{-0}) \equiv 0.5p(e,0) + 0.5p(e,1);$$
$$R_1(e_1|\ e_{-1}) \equiv 0.5(1 - p(e,0)) + 0.5(1 - p(e,1));$$
$$R_i(e_i|\ e_{-i}) \equiv 0.5p(e,0) + 0.5(1 - p(e,1)).$$

We say that a rule profile e^* is *consistent* when it specifies the maximally ranked rule for each voter type given the behavior of other types.

Definition 1 (Consistency Requirement). *The profile e^* is consistent if for any type $t \in T$*

$$R_t(e_t^*|\ e_{-t}^*) \geqslant R_t(e_t|\ e_{-t}^*) \text{ for all rules } e_t \in E.$$

If the rule profile is *not* consistent then all voters of at least one type will conclude that an alternative rule would better determine how they should behave. We now analyze the properties of consistent rule profiles.

3.1. Comments on the term "ethical."

The reason for applying the term "ethical" to some agents in our model is not that their preferences over outcomes necessarily satisfy some normative criteria. Rather, agents are described as ethical for two reasons. First, ethical agents evaluate alternative behavioral rules in a Kantian manner by comparing the outcomes that would occur if everyone who shares their preferences were to act according to the same rule. Second, they receive a positive payoff for acting according to a behavioral rule they determine is best using this type of reasoning. For an extensive discussion of our terminology and related conceptual issues see Feddersen (2004), and Feddersen and Sandroni (2005a and b).

3.2. Baseline Results

Recall that there are more partisans for 1 than for 0. The *difference in partisan support* (i.e., the fraction of the population that is partisan for candidate 1 minus the fraction of partisans for candidate 0) is $\beta(1-2k)$. Assume that there are more uninformed independent voters than the difference in partisan support. That is, $(1-\beta)(1-q) \geqslant \beta(1-2k)$ or, equivalently, the fraction of partisans in the electorate is sufficiently small so that $\beta \leqslant \beta(q,k) \equiv \frac{(1-q)}{(1-q)+(1-2k)}$.

Partisans for each candidate maximize the chance their preferred candidate wins by voting for their candidate. If each uninformed independent agent votes for candidate 0 with probability $\frac{\beta(1-2k)}{(1-\beta)(1-q)}$ and abstains with probability $1-\frac{\beta(1-2k)}{(1-\beta)(1-q)}$ then the difference in partisan support is compensated for by the votes of uninformed independents. It follows that the election outcome will be decided by the votes of the informed independents. If informed independents vote according to their message then candidate $w \in \{0,1\}$ is elected in state w with certainty. An ideal outcome for independent voters.

Formally, let $\hat{e} = (\hat{e}_0,\ \hat{e}_1,\ \hat{e}_i)$ be the rule profile where $\hat{e}_0 \equiv 1$, $\hat{e}_1 \equiv -1$, $\hat{e}_i(1) = -1$, $\hat{e}_i(0) = 1$ and

$$\hat{e}_i(\phi) = \frac{\beta(1-2k)}{(1-\beta)(1-q)}. \tag{*}$$

That is, the rule profile $\hat{e}$ is such that partisans vote for their candidate, informed independents vote according to their message, a fraction given by $(*)$ of uninformed independents vote for the candidate least supported by partisans and other uninformed independents abstain. Proposition 1 below is straightforward.

Proposition 1. *Assume that $\beta \leqslant \beta(q,k)$. The rule profile $\hat{e}$ is consistent.*[9]

By proposition 1 abstention occurs even though there are no voting costs. Uninformed independents vote against the candidate most supported by partisans, but only up to a point that compensates for the difference in partisans' votes. As in the Feddersen and Pesendorfer (1996), independents vote in such a way as to effectively delegate the electoral decision to informed independents. This maximizes the chances that candidate $w \in \{0,1\}$ is elected in state w.

[9] As mentioned in the introduction the reader might be concerned about the existence of other consistent rule profiles. However, all relevant multiplicity can be eliminated in a slightly more complex model. Assume that the fraction of informed agents who observe message $w \in \{0,1\}$ in state w is a random variable $\tilde{\rho}$ with full support on $(0.5, 1)$. No voter observes the realization of $\tilde{\rho}$ before the votes are casts. To ensure the victory of candidate $w \in \{0,1\}$ in state w the votes of the uninformed independent voters must exactly offset the partisans' vote. Hence, equation $(*)$ must be satisfied in a consistent rule profile.

Proposition 1 holds if independents' payoffs for having candidate 1 elected in state 1 are higher than the payoff for having candidate 0 elected in state 0. Then, uninformed independents ex-ante prefer candidate 1. However, in a consistent profile no uninformed independent agents vote for 1 and a fraction of them strategically vote for 0. This demonstrates that strategic behavior may be a consequence of moral calculation.

Consistent profiles exist even if the assumption in Proposition 1 does not hold. Consider a variation on the example in the introduction. Ten percent of the voters in the population are partisans for 0; 30 percent are partisans for 1; and 60 percent are independents. Now assume that 5/6 of the independents are perfectly informed about the state while the rest are uninformed. Clearly, there are insufficient uninformed independents to compensate for the difference in partisans and the assumption of Proposition 1 doesn't hold.

However, there is a consistent profile. Suppose that all uninformed independents vote for candidate 0. As before, partisans vote for their preferred candidate. Now, let informed independents who observe the signal for candidate 0 vote for candidate 0 but those who observe the signal for 1 will randomize between abstention with probability 2/5 and vote for 1 with probability 3/5. To see that this is a consistent profile note that if $\rho = .5$ i.e., the signal is completely noisy then each candidate will receive votes from 45% of the population while 10% of the population abstains. Now suppose that $\rho = .5 + \varepsilon$ where ε is positive. In state 0 the vote total for candidate 0 will be $.45 + .5\varepsilon$ while the votes for 1 is $.45 - .3\varepsilon$. So, candidate 0 wins. A similar calculation shows that candidate 1 wins in state 1. This is a consistent profile because full information equivalence is satisfied.

This example poses a problem for a theory of abstention due to lack of information because there are consistent profiles in which the only agents who might abstain are informed.[10] This problem disappears when the choice to become informed is endogenized as we show in section 4. We show that it is not optimal to pay a cost for information and then not use it by abstaining or voting against it. Unless otherwise specified, in the remaining part of section 3 we restrict ourselves to parameters where this problem does not arise.

[10] A related, but not identical, problem with the theory of abstention for informational reasons have been indentified in Feddersen and Pesendorfer (1996).

3.3. Comment on compensating for partisan support

Some readers may find the result that uninformed agents vote to compensate for one candidate's partisan advantage implausible. The main assumption that underlies the compensation result is that uninformed independents are sufficiently informed to know the partisan balance without knowing which candidate they prefer. There is a large empirical literature (see for example Bartels (1996); Campbell, Converse, Miller and Stokes (1960) among others) suggesting that many voters possess minimal information. If uninformed independents have no information about the balance of partisans within the electorate then our model would predict that uninformed independents abstain.

Abstention by uninformed agents seems plausible because uninformed votes add noise to election outcomes increasing the likelihood that a less preferred candidate wins the election. Thus, the idea that uninformed agents act so as to let informed agents determine the election outcome is not, in itself, implausible. Now consider the case that uninformed voters know the partisan balance. Then, if uninformed agents strategically balance out the partisan advantage then they increase the likelihood that their preferred candidate wins the election.

If the fractions of informed and uninformed agents are exogenously given then there is no substantive difference between the voting behavior of ethical agents and agents in standard models such as Feddersen and Pesendorfer (1996). To see this consider the following example: assume that there is a committee with three voters and two candidates. Voter 1 is a partisan for 1. Voters 2 and 3 are independents. Voter 2 is perfectly informed. Voter 3 is uninformed. So, voters 1 and 2 have weakly dominant strategies. Voter 1 votes for 1 and voter 2 votes for the candidate that is best for the independents. Given the play by voters 1 and 3, it is optimal for voter 3 to vote for 0 and balance out the partisan vote. Voter 3 is only pivotal when voter 2 has voted for candidate 0. In that case, candidate 0 is best for independents. Hence, uninformed independents vote so as to compensate for the partisan advantage of one candidate in both rational and the ethical models.

Finally, there is extensive empirical evidence that, in multi-candidate elections, voters strategically compensate for predicted behavior of others by voting for candidates that are not their first choice (see, for example, Cox (1997)). It is an open empirical question whether the particular form of strategic behavior involved in balancing out partisan advantage occurs in reality.

3.4. Turnout and Margin of Victory

In this sub-section we demonstrate that our baseline model produces the same comparative statics on turnout and margin of victory as in Feddersen and Pesendorfer (1996). However, our analysis is much simpler. Turnout (T) in the election is the fraction of partisans plus the fraction of informed independents, $\beta + (1-\beta)q$ (because they all vote), plus the fraction of uninformed independents who vote, $\beta(1-2k)$. Hence,

$$T = \beta + (1-\beta)q + \beta(1-2k).$$

The votes of the partisans and uninformed independents cancel out. Therefore, in each state, the margin of victory is the fraction of independents whose message is the same as the state minus the fraction of independents whose message is different from the state, divided by the turnout. That is,

$$MV = \frac{(1-\beta)q(2\rho-1)}{\beta + (1-\beta)q + \beta(1-2k)}.$$

It follows from the expressions above that turnout is increasing in the fraction of partisans (β), but the margin of victory is decreasing in β. Hence turnout and margin of victory are inversely related through changes in β.

Turnout increases as β increases because uninformed voters should vote to offset the difference in partisan support. So, the fraction of uniformed voters who vote must increase as β increases. All other agents vote. On the other hand, as β increases the fraction of informed independents in the electorate must decrease and, therefore, the margin of victory also decreases because the votes of informed independents ultimately decide the election. It is worth noting that turnout and margin of victory are not necessarily inversely related in this model because they are both increasing functions of the probability of being informed (q).

We refer the reader to Blais (2000) and Nalebuff and Shachar (1999) for a discussion of the empirical relationship between margin of victory and turnout. We also refer the reader to Feddersen and Pesendorfer (1996 and 1999) for the relationship between the empirical literature and the comparative statics results of the baseline model.

3.5. Uncertainty over the Difference in Partisan Support

As anticipated in the introduction, in order to motivate information acquisition we need to introduce an additional source of uncertainty. Assume, for simplicity, that the fraction of partisans who support candidate 0, $\tilde{k}$, is uniformly distributed over $(\bar{k}-v, \bar{k}+v)$ where $\bar{k}+v < 0.5$.[11] Let $\hat{e}$ be the rule profile in section 2.1, but replacing k with $\bar{k}$. If $\beta \leqslant \beta(q, \bar{k})$ then $\hat{e}$ is a consistent profile. In this case there are sufficient uninformed independents to nullify the expected rather than the actual partisan bias. The complete definition of a consistent profile for all parameter values is more complex and can be found in Lemma A in the Appendix.

The result below will be useful in the next section. For *all* parameter values, the expected payoff of independent agents (evaluated under a consistent rule profile) is

$$0.5 + 0.5\frac{(1-\beta)(2\rho-1)}{2\beta v}q \tag{**}$$

if $q \leqslant L\left(\beta, \rho, \bar{k}, v\right)$ and reaches a maximum at $q = L\left(\beta, \rho, \bar{k}, v\right)$ where,

$$L\left(\beta, \rho, \bar{k}, v\right) \equiv \begin{cases} \frac{2\beta v}{(1-\beta)(2\rho-1)} & \text{if } \beta \leqslant \beta_1; \\ \frac{1+2\beta(\bar{k}+v-1)}{(1-\beta)2\rho} & \text{if } \beta_1 \leqslant \beta \leqslant \beta_2; \\ 0 & \text{if } \beta_2 \leqslant \beta. \end{cases}$$

with

$$\beta_1 \equiv \frac{(2\rho-1)}{2(v+\left(1-\bar{k}\right)(2\rho-1))}; \text{ and } \beta_2 \equiv \frac{0.5}{1-(\bar{k}+v)}.$$

Equation $(**)$ holds because candidate $w \in \{0,1\}$ wins or loses the election in state w depending on the informed independents support for w $((1-\beta)(2\rho-1)q)$ and the difference in partisan support bias (for 1) from its mean (with a maximum value of $2\beta v$).

The payoff for the independent types is weakly increasing in the fraction that are informed (q). This must be the case because the informed independents can always mimic the behavior of the uninformed. However, independents' welfare does not increase with the fraction of informed independents (q) above a threshold $L\left(\beta, \rho, \bar{k}, v\right)$. When q is greater than this threshold we show that there are three cases to consider. If the fraction of partisans in the electorate is small (i.e.,

[11] The assumption that $\bar{k}+v < 0.5$ simplifies our proofs because it implies that there are always more partisans for 1 then for 0. The main results remain qualitatively unchanged if we relax this condition. The uniform distribution permits explicit formulas, but the main results generalize to a density function of $\tilde{k}$, $f_{\tilde{k}}$, that is symmeric around the mean, i.e., $f_{\tilde{k}}(\bar{k}-\eta) = f_{\tilde{k}}(\bar{k}+\eta)$ and peeks at $\bar{k}$, i.e., $f_{\tilde{k}}(\bar{k}) \geqslant f_{\tilde{k}}(\eta)$ for $\eta \geqslant 0$.

$\beta \leqslant \beta_1$) then w is elected in state w with certainty. If the fraction of partisans in the electorate is large (i.e., $\beta \geqslant \beta_2$) then the candidate with the most support among the partisans (1) is always elected. Hence in these two cases the payoffs of independents cannot increase when q is above the threshold $L(\beta, \rho, \bar{k}, v)$.

Now consider the case in which β is in the intermediary range $\beta_1 \leqslant \beta \leqslant \beta_2$. Assume that uninformed independents all vote for candidate 0. Suppose further that all informed independents vote their signal. Because the fraction of partisans in the population is large and q is above the threshold $L(\beta, \rho, \bar{k}, v)$ there are not enough uninformed independents to counterbalance the difference in partisan support for all realizations of k.

In state 1 the difference in partisan support combined with the greater support for candidate 1 by informed independents guarantees that candidate 1 will be elected for all q above the threshold. In state 0, however, either candidate might win. It follows that such a rule creates an excess of votes for candidate 1 and a shortage of votes for candidate 0. An optimal rule compensates by transferring votes by informed independents who observe signal 1 into either abstention or voting for candidate 0. It follows that an increase in q will only result in an increase in such vote transfers without having any impact on the probability candidate 0 is elected in state 0.

4. Costly Information Acquisition

We now endogenize the decision to become informed. We assume that prior to voting each agent must decide whether or not to purchase a signal with the properties described in the previous section. Let $A \equiv \{u, i\}$ be the set of actions where u denotes the choice not to purchase a signal and i denotes the choice to purchase a signal. Each agent's cost to purchase information is given by $C > 0$ times an independent value drawn from a uniform distribution over the interval $(0, 1)$.[12] Each agent knows her own realized cost to acquire information, but not the realization of other agents' costs.

A rule is a pair (r, e), where $r = (r_0, r_1, r_i)$, r_t is a function mapping a realized cost $c \in (0, C)$ into an action $a \in A$ and e is defined as in section 2. We assume that all agents receive a payoff $D > C$ for acting as they determine they should. Given that D is greater than the highest possible cost, agents will act as they determine they should.[13] Agents rank rules as in section 2, but subtracting the

[12] The main results can be extended to non-uniform distributions.

[13] See Feddersen and Sandroni (2004b) for a model where this assumption is relaxed.

induced cost of information acquisition.

In a consistent rule profile, partisans will choose to be uninformed because information is costly and does not influence their behavior. Independents face a trade-off. When more independents acquire information the chances that $w \in \{0,1\}$ is elected in state w may increase, but at a higher cost.

The most economical way to have a fraction $q \in [0,1]$ of independent agents informed is to have all those for whom the cost to acquire information is below Cq informed and those for whom the costs to acquire information is above Cq uninformed. Hence, the expected cost of having a fraction $q \in [0,1]$ of independent agents informed is given by

$$(1-\beta)\int_0^{Cq} x\frac{1}{C}\partial x = C(1-\beta)\frac{q^2}{2}.$$

In a consistent rule profile, the fraction of independents that acquire information ($\hat{q}$) is chosen fully anticipating that once this fraction is determined, behavior is given by $\hat{e}$ (as defined in section 2). By equation (**), $\hat{q}$ must maximize

$$0.5 + 0.5\frac{(1-\beta)(2\rho-1)}{2\beta v}q - C(1-\beta)\frac{q^2}{2}, \qquad (***)$$

subject to $q \leqslant L\left(\beta, \rho, \bar{k}, v\right)$. Hence,

$$\hat{q} = \min\left\{L\left(\beta, \rho, \bar{k}, v\right), G\left(\beta, \rho, C, v\right)\right\},$$

where $G(\beta, \rho, C, v) \equiv \frac{(2\rho-1)}{4\beta C v}$ maximizes $(***)$ unconstrained.

So, the profile $(\hat{r}, \hat{e})$ where $\hat{e}$ is defined in section 2, $\hat{r}_0 \equiv u$, $\hat{r}_1 \equiv u$ (i.e., partisans do not get informed), and independents get informed if and only if their cost to gather information is below $C\hat{q}$, i.e.,

$$\begin{array}{ll} \hat{r}_i = i & \text{if } c \leqslant C\hat{q} \\ \hat{r}_i = u & \text{if } c > C\hat{q} \end{array}$$

is consistent.

4.1. Main Results

The main properties of the consistent rule profile $(\hat{r}, \hat{e})$ are as follows:

Fact 1. *Informed independents vote their signal. Uninformed independents are the only agents who might abstain.*

Recall from the example in section 3 that when the fraction informed (q) is exogenous there are some parameter ranges in which some informed agents abstain. It is even possible that informed agents will vote for candidate 0 when they observe signal 1. This cannot happen when information is costly and endogenously obtained.

If information is costly then agents will prefer rules in which information is not collected above the point at which it ceases to be useful. Consistent rules require that only information that will be used should be collected. Hence, once the decision to acquire information is endogenized the model delivers a clear prediction relating abstention and information: less informed voters will sometimes abstain while informed voters participate.

Corollary to Fact 1. *Turnout and the fraction of informed agents are positively related.*

One of the central arguments for the Feddersen and Pesendorfer (1996) model was that it provided an informational explanation for why better informed agents would be more likely to vote. As mentioned above, when the fraction of informed independents is exogenously given this result holds for some, but not all parameters. In contrast, when information acquisition is endogenized informed agents always vote and the uninformed may vote or abstain. It follows that the informational explanation for turnout is strengthened.

Fact 2A. *If uncertainty over the difference in partisan support (v) goes to zero the fraction of informed independents ($\hat{q}$) also goes to zero.*

Fact 2B. *Assume that the costs of acquiring information are sufficiently high (i.e., $C > \dot{C}$, where $\dot{C}$ is a threshold level). As the uncertainty over the difference in partisan support (v) increases, the fraction of informed independents ($\hat{q}$) initially increases, reaches a maximum and then decreases.*[14]

With no uncertainty over the difference in partisan support then (as demonstrated in section 3) full information equivalence holds with probability one. Either it is possible to elect candidate $w \in \{0, 1\}$ in state w with *few* informed independents or the candidate most favored by partisans (1) will be elected. Hence, even if the costs to acquire information are very low, the fraction of informed independents will be small when the uncertainty over the difference in partisan support is small. Fact 2A shows why a second source of uncertainty (beyond the uncertainty over the state) is important for explaining information acquisition.

[14] If the costs of acquiring information are sufficiently low (i.e., $C \leqslant \dot{C}$) then the fraction of informed agents monotonically increases with the uncertainty over the difference in partisan support.

As the uncertainty increases, it is necessary to have a larger fraction of informed independents to ensure the election of candidate $w \in \{0, 1\}$ in state w. This increases the incentives to gather information and leads to a significant fraction of the electorate gathering information. On the other hand, if the uncertainty over the difference in partisan support and the costs to gather information is sufficiently high then independents will not choose a behavioral rule that implies the victory of candidate $w \in \{0, 1\}$ in state w with certainty (or such a rule is not feasible). Then, an increase in uncertainty reduces the marginal benefits of information. This reduces the incentives to gather information.

4.1.1. Information Aggregation

The formal theory literature has focused on one primary metric of information aggregation: *full information equivalence.* As shown above, elections do not always satisfy this requirement with probability one. A second, weaker, requirement is *informed electorate equivalence.* It holds when the winning candidate would be the same if every agent had observed a private signal. The two definitions are the same when information is perfect (i.e., $\rho = 1$) but may differ otherwise.[15]

Fact 3. *For sufficiently low cost of becoming informed (i.e., below a certain threshold), the fraction of informed independents does not change with costs and is less than one. In addition the election outcome satisfies the informed electorate equivalence condition (but not necessarily full information equivalence) with probability one.*

Full information equivalence is not necessarily always satisfied simply because (as illustrated in the introduction) rules for independents that assure victory of candidate $w \in \{0, 1\}$ in state w may not exist. Nevertheless, informed electorate equivalence must hold if information is virtually free because it must be acquired up to the point that it ceases to be useful. Moreover, when the fraction of informed independents goes above a threshold then information no longer has any marginal benefit. To see this assume that all independent agents acquire a signal and vote accordingly. Then, in state 1 candidate 1 wins the election with an excess of votes, whereas in state 0, candidate 0 may win but with a smaller excess of votes or may lose. So, consider an alternative rule that directs a small fraction of agents to not acquire information and always vote for 0. This new rule still ensures that

[15] In general, if full information equivalence holds then informed electorate equivalence must also hold. The converse, however, is not true. In some cases, informed electorate equivalence is satisfied and full information equivalence does not occur with probability one.

candidate 1 wins in state 1 and does not reduce the chances that 0 wins in state 0. However, the new rule imposes lower information costs. Thus, in a consistent profile, it cannot be the case that all independents will acquire information.

The substantive implication is that some level of political ignorance is not only unavoidable but desirable. There is no reason that everyone in society should be informed when there is no possible way for independents to benefit in terms of an increased likelihood of a better outcome.

The following facts ($4A - 4C$) relate changes in the quality of information to information acquisition, information aggregation and election bias (recall that the election bias is defined as the probability that candidate 1 wins the election minus 0.5). All these facts hold for the case in which the costs of acquiring information are sufficiently low (i.e., below a threshold $\bar{C}$). The results when the cost of acquiring information are above the threshold $\bar{C}$ are described in footnotes. We also assume that the fraction of partisans in the electorate is sufficiently small so that $\beta < \beta_2$. Otherwise, candidate 1 always wins the election and no independent acquires information ($\hat{q} = 0$).

Fact 4.A *As the quality of the signals (ρ) increases, the fraction of independents that acquire information ($\hat{q}$) initially increases, reaches a maximum and then decreases.*[16]

The intuition behind Fact $4A$ is as follows. For the moment suppose that the fraction informed (q) is exogenously fixed. Increasing q increases the chances that candidate $w \in \{0, 1\}$ wins the election in state w. However, (as demonstrated in Fact 3 and in section 3.3) there is a threshold L such that increasing q above L no longer produces any marginal benefits for independents. The threshold L is defined as the fraction of informed independents that suffice to elect candidate 1 in state 1 with certainty (assuming that uninformed independents vote for 0 to the point that minimizes any difference in expected partisan support). As the quality of signals (ρ) increases fewer informed independents are needed to ensure the victory of candidate 1 in state 1. It follows that the threshold L must be decreasing in ρ. Also note that independents will never choose a rule that directs them to acquire information above the threshold L because to do so would mean to incur additional costs with no additional benefits. Hence, the fraction of independents that acquire information ($\hat{q}$) is either strictly below L or is equal to L. In the later case we say that $\hat{q}$ is binding.

Now assume that the quality of information is very poor, i.e., ρ is close to

[16] If costs of aquiring information are above the threshold $\bar{C}$ then the fraction of independents that acquire information increases with the quality of the signals.

1/2. Then, information has a very low value and full information equivalence is impossible. In this case almost all independents will be uninformed. As the quality of information increases, the marginal benefits of information increases and the marginal cost stays the same. Therefore, the optimal fraction of informed voters increases. On the other hand, an increase in ρ also leads to a smaller threshold L. So, if the quality of signals becomes sufficiently large then $\hat{q}$ becomes binding. At this point, the fraction of informed independents is maximized. A further increase in the quality of the signals will produce a smaller binding solution $\hat{q}$ (because the threshold L is decreasing in ρ).

By definition, the fraction of informed independents does not change with the quality of the signals in our baseline model. Once the decision to acquire information is endogenized these two variables become related in a non-monotonic way. This non-monotonic relationship stands in contrast to the negative, monotonic relationship between the costs of acquiring information (C) and the fraction of informed independents ($\hat{q}$).[17]

Fact 4.B *As the quality of the signals (ρ) increases, the election bias is initially weakly increasing in ρ until a threshold $\bar{\rho}$ is reached. After this threshold, the election bias is weakly decreasing in ρ.*[18]

The proviso "weakly" in Fact $4B$ stands because it is possible that a marginal change in ρ will not alter the chances that candidate 1 wins the election and, therefore, leaves the election bias unchanged. If the fraction of partisans in the electorate is very small (i.e., below a threshold) then, in a consistent profile, there will be enough uninformed independents to compensate for the partisan advantage. It follows that the election will be unbiased for all values of $\rho \in (0.5, 1)$. On the other hand, if the fraction of partisans in the electorate is very high (but still below β_2) then the election will always be biased and this bias is either strictly increasing or strictly decreasing in the quality of the signals. What follows is an informal description of the relationship between the election bias and the quality of the signals when the fraction of partisans in the electorate is at an intermediary value.

Assume that ρ is close to 1/2. Then, information has very low value and almost all independents will be uninformed. The uninformed will vote for candidate 0 at a level to balance out the expected partisan advantage to 1. It follows that the probability of electing each candidate is equal. Now, as we increase the quality

[17] A proof for this statement can be found in the proof of Fact 3 in the Appendix.

[18] If costs of aquiring information are above the threshold $\bar{C}$ then the election bias increases with the quality of the signals.

of the signals the incentive to acquire information increases and the fraction of uninformed independents decreases. This continues up to the point at which uninformed independents cannot balance out the expected partisan advantage to 1. As ρ increases further more independents become informed. The effect is to increase the chances that candidate 1 wins the election (because uninformed independents had been voting for candidate 0). This continues up to the point at which candidate 1 is guaranteed to win in state 1. After this point, when ρ is increased even further, independents do better by reducing the fraction informed and giving additional votes to candidate 0. Candidate 1 continues to win with probability 1 in state 1, but candidate 0 has a higher chance of winning the election in state 0. The overall effect is to increase the chances that candidate 0 wins the election. This continues up to the point that candidate $w \in \{0, 1\}$ wins the election in state w with certainty. After this point, both candidates have an equal chance of winning the election.

The substantive point in Fact $4B$ is that the impact of improving the quality of information is not neutral from a partisan perspective. The expected payoff to partisans is affected by the quality of the information. This suggests that our model can be used as a foundation for explaining the partisan provision of information in campaigns. An exploration of the strategic incentives of partisan information provision is beyond the scope of the current paper. However, Fact $4B$ shows that there are conditions under which either candidate might benefit from an improvement in the quality of information.

Fact 4.C *The probability the election outcome satisfies full information equivalence is weakly increasing in* ρ.[19]

Fact $4C$ is an intuitive result. However, it is worth noting that the provision "weakly" here is needed only because there is a threshold on the quality of the signals above which the election satisfies full information equivalence with probability one. It follows that increasing the quality of the signals does not have the same effect as decreasing the cost of acquiring information. As shown in Fact 3, reducing the costs of acquiring information does not necessarily ensure full information equivalence. On the other hand, making the quality of information sufficiently high does.

Assume that full information equivalence holds. Then, in the case of an exogenously given information structure, the welfare of independents does not change with the quality of information. In the case of an endogenously determined fraction of informed independents agents, an increase in the quality of information

[19] This result holds for all costs to acquire information.

increases the welfare of independent agents because fewer informed agents are required to achieve the same outcome. So, the social costs of acquiring the necessary information is smaller. Figure 1 below illustrates facts $4A - 4C$.

Insert Figure 1 about here

Figure 1 shows that it is theoretically possible to have full information equivalence holding even when the fraction informed is very small. Conversely, full information equivalence may not hold when a large fraction of voters are informed. In fact Figure 1 shows the probability of full information equivalence may decrease as the fraction informed increases. On the other hand, Figure 1 shows that the aggregation properties of elections unambiguously improves with the quality of information (ρ).

4.2. Turnout and Margin of Victory

In section 3.2, we showed comparative static results that relate turnout and margin of victory with the parameters of the our baseline model. In this section, we show that some of these comparative statics results may change when the decision to acquire information is endogenized. We start with a preliminary result.

Fact 5. *As the fraction of partisans in the electorate* (β) *increases, the fraction of informed independents* ($\hat{q}$) *initially increases, reaches a maximum and then decreases.*

The intuition behind Fact 5 is as follows: If the level of partisanship is low then independents can elect candidate $w \in \{0, 1\}$ in state w with a small fraction of informed independents. So, the incentives to gather information are low. If the level of partisanship is very high the incentive to gather information is also low. This follows because independents have little influence on the outcome of the election. The incentive to acquire information is maximized when the fraction of partisans in the electorate is at a middle level.

Fact 6. *Turnout may decrease when the fraction of partisans in the electorate* (β) *increases.*

If the fraction of informed independents is fixed then, as demonstrated in section 3.2, overall turnout is also either constant or increases with the fraction of partisans in the electorate. On the other hand, if the fraction of informed independents is determined endogenously then (as mentioned in fact 5) that fraction may decrease with the fraction of partisans in the electorate. Since uninformed independents have lower turnout overall turnout may decrease. This shows that the comparative statics results derived for an exogenously given information structure do not necessarily hold once the decision to gather information is endogenized.

Fact 7.A *If the fraction of partisans in the electorate is small (i.e., below a threshold) then a marginal increase in the quality of the information (ρ) decreases turnout and increases the expected margin of victory.*

Consider our baseline model in which the fraction of informed independents is fixed. In this model, as demonstrated in section 3.2, overall turnout does not depend on the quality of information. However, if the decision to acquire information is endogenous then turnout changes with the quality of information.

The intuition behind Fact 7.A is as follows: If the fraction of partisans in the electorate is sufficiently small then full information equivalence holds. The fraction of informed independents is decreasing in the quality of information because a smaller fraction of informed voters are required to elect candidate $w \in \{0, 1\}$ in state w. Therefore, a higher fraction of independents remain uninformed and these are the only voters who might abstain. Hence, overall turnout decreases.[20] The margin of victory increases because the election is decided by informed independents. The difference in votes in favor of the winning candidate increases as the quality of information increases.

Fact 7.B *If the fraction of partisans in the electorate is small then a marginal increase in the uncertainty over the difference in partisan support (v) increases turnout and the expected margin of victory.*

Fact 7.B shows that turnout and margin of victory may be positively related due to changes on the uncertainty over the difference in partisan support. This stands in contrast with the negative relationship between turnout and margin of victory obtained in Fact 7.A. It is also worth noting that, in our baseline model, neither turnout nor the expected margin of victory depend on the uncertainty over the difference in partisan support

The intuition behind Fact 7.B is as follows: When the fraction of partisans in the electorate is small, a marginal increase in the level of uncertainty leads to an increase in the fraction of informed independents because more informed agents are needed to ensure that candidate $w \in \{0, 1\}$ wins in state w. This leads to an increase in expected turnout because there are fewer uninformed independents. It also leads to a greater margin of victory because the difference in independent votes in favor of the winning candidate is greater.

Fact 7.C *If the fraction of partisans in the electorate is small (i.e., below*

[20] For other parameter values, overall turnout could increase with the quality of information because (as mentioned in fact 4.A), the fraction of informed independents may increse with ρ.

a threshold) then a marginal increase in the expected balance of partisans ($\bar{k}$) decreases turnout.

When the balance of partisans is minimized, at $\bar{k} = v$, the uninformed independents must vote at a high rate in order to compensate for the large partisan advantage for candidate 1. When expected balance of partisans is maximized, at $\bar{k} = .5 - v$, the proportion of partisans is closest to equal, uninformed independents vote at lower rate to compensate for the smaller partisan advantage for candidate 1. In addition, the fraction of informed agents does not change with $\bar{k}$ when the overall fraction of partisans (β) is low.

4.3. Relationship of Results to the Empirical Literature

Fact 1 is one the central findings in our paper. It relates the fraction informed and turnout. Empirically testing this relationship is difficult because the decision to acquire information is related to the decision to vote. Lassen (2005) takes account of this problem and empirically shows a positive correlation between information and turnout.

Bartels (1996) demonstrates empirically that voting behavior of the informed is significantly different from the voting behavior of the informed. This observation is consistent with our model. Bartels also claims that election outcomes might be different if all uninformed voters were to vote as if they were informed. This counterfactual exercise assumes that changing the fraction informed won't change the voting behavior of informed and uninformed alike. In our models (as in Feddersen and Pesendorfer (1996)) changing the fraction informed may change the voting behavior of both informed and uninformed. Hence, our model raises question about the validity of this counterfactual exercise.

Perhaps the most counterintuitive prediction of our model (as well as the Feddersen and Pesendorfer (1996) model) is that uninformed independents either abstain or vote against the candidate with strongest partisan support. We are unaware of any empirical test of this prediction.

4.4. Introducing Costs to Vote

In Feddersen and Sandroni (2005) we developed an ethical model of costly voting. Combining costs to vote with costs to acquire information is beyond the scope of the current paper. It is possible that, if combined, the two models will interact in non-trivial ways. For example, in the model with costly voting decreasing the partisan imbalance leads to higher turnout, whereas in this model Fact 7C shows that decreasing the partisan imbalance may decrease turnout.

A basic problem in combining the two models is establishing the existence of consistent profiles. Subject to establishing existence, we believe that the balancing result, i.e., that uninformed voters either abstain or vote for the candidate supported by partisan minority, will continue to hold.

5. Conclusion

The contribution of this paper is to produce a logically consistent model in which strategic voters acquire costly information. Voter behavior in our model is not driven by events with negligible probability (e.g. a pivotal vote), but nevertheless voters behave strategically. We show that some voters will acquire costly information. However, even when costs to acquire information become arbitrarily low a fraction of the electorate will choose to remain uninformed. As the quality of information increases information aggregation properties of election improve, but the fraction of informed voters may decrease. Moreover, such changes may benefit one candidate at the expense of the other. This suggests that our model may be used as a foundation for explaining political campaigns.

6. Appendix

In section 2, we require agents of the same type to not vote for different candidates with positive probability. Without this assumption an independent voter would give equal rank to a rule that requires all independents to abstain and another rule that requires all to vote for both candidates with the same probability. This is so because with a continuum of voters these two rules lead to identical outcomes. This indifference may produce multiple consistent rules. However, this indifference would be broken (in favor of the rule with smaller turnout) if we had arbitrarily small voting costs or if we had a suitably redefined model with a finite, perhaps large, number of voters.[21] We make this restriction on rules instead of these alternative modelling choices to avoid unnecessary complications.

[21] In a model with finitely many voters, if each voter votes for both candidates with equal probability then there is a chance, perhaps small, that the difference in votes will not be zero.

Lemma A.0 Assume that $\beta \leqslant \beta(q,\bar{k})$. Under the rule profile $\hat{e}$, in state $w \in \{0,1\}$ candidate w is elected with probability greater than 0.5.

Proof: The votes of uninformed independents for candidate 0 are $x \equiv (1-\beta)(1-q)e_i(\phi)$. In state 0 the difference in informed independents votes favoring candidate 0 is $z_0 \equiv (1-\beta)q\left(\rho e_i(0)+(1-\rho)\,e_i(1)\right)$. In state 1 the difference in informed independents votes favoring candidate 0 is $z_1 \equiv (1-\beta)q\left((1-\rho)\,e_i(0)+\rho e_i(1)\right)$. In state 0 candidate 0 is elected when $\beta\tilde{k}+x+z_0 > \beta(1-\tilde{k})$ which is equivalent to $\beta(1-2\tilde{k}) < x+z_0$. In state 1 candidate 1 is elected when $\beta\tilde{k}+x+z_1 < \beta(1-\tilde{k})$ which is equivalent to $\beta(1-2\tilde{k}) > x+z_1$. Hence, the expected payoff for independents is

$$0.5P\left(\beta(1-2\tilde{k}) < x+z_0\right)+0.5P\left(\beta(1-2\tilde{k}) > x+z_1\right). \qquad (\Upsilon_1)$$

Without loss of generality, we can assume that $z_1 \leqslant z_0$ because if $z_1 > z_0$ the payoff in (Υ_1) increases when the values of $e_i(0)$ and $e_i(1)$ are exchanged. Therefore, (Υ_1) is equivalent to

$$0.5+0.5P\left(x+z_1 < \beta(1-2\tilde{k}) < x+z_0\right). \qquad (\Upsilon_2)$$

The random variable $\beta(1-2\tilde{k})$ is uniformly distributed around mean $\beta(1-2\bar{k})$, with upper bound $\beta^u \equiv \beta(1-2\left(\bar{k}-v\right))$ and lower bound $\beta^l \equiv \beta(1-2\left(\bar{k}+v\right))$. The interval $(x+z_1; x+z_0)$ can be written as $(m-l; m+l)$ where

$$m \equiv x+0.5(1-\beta)q\left(e_i(0)+e_i(1)\right) \text{ and } l \equiv 0.5(1-\beta)q\left(2\rho-1\right)\left(e_i(0)-e_i(1)\right).$$

If $\beta \leqslant \beta(q,\bar{k})$ or equivalently, $q \leqslant \frac{(1-\beta)-\beta\left(1-2\bar{k}\right)}{(1-\beta)}$ then, $\hat{e}_i(\phi) = \frac{\beta(1-2\bar{k})}{(1-\beta)(1-q)}$; $\hat{e}_i(0)=1$; and $\hat{e}_i(1)=-1$ puts m at the center of the distribution of $\beta(1-2\tilde{k})$ and maximizes l. This maximizes (Υ_2). Moreover, evaluated at $\hat{e}$, z_0 is positive and z_1 is negative. So,

$$P\left(\beta(1-2\tilde{k})-\beta(1-2\bar{k}) < z_0\right) = P\left(\beta(1-2\tilde{k})-\beta(1-2\bar{k}) > z_1\right) > 0.5.$$

$\boxdot$

Lemma A The profile $\hat{e}$ such that $\hat{e}_i(0)=1$, $\hat{e}_i(\phi)=\min\left\{\frac{\beta(1-2\bar{k})}{(1-\beta)(1-q)},1\right\}$ and

$$\hat{e}_i(1) = \begin{array}{ll} -1 & \text{if } q \leqslant \frac{1+2\beta(\bar{k}+v-1)}{(1-\beta)2\rho}; \\ \frac{\beta(1-2\left(\bar{k}+v\right))-(1-\beta)(1-q)-(1-\rho)(1-\beta)q}{(1-\beta)q\rho} & \text{if } q \geqslant \frac{1+2\beta(\bar{k}+v-1)}{(1-\beta)2\rho}. \end{array}$$

is consistent.

Proof: All notation used in the proof of Lemma $A.0$ is maintained. Without loss of generality, we can assume that $\beta \leqslant \frac{0.5}{1-(\bar{k}+v)}$. Otherwise then $m+l$ is smaller than β^l even at $e_i(1) = e_i(0) = e_i(\phi) = 1$. Hence, (Υ_2) is always 0.5 and $\hat{e}$ can be arbitrarily defined.

Note that

$$\frac{(1-\beta)-\beta\left(1-2\bar{k}\right)}{(1-\beta)} \leqslant \frac{1+2\beta(\bar{k}+v-1)}{(1-\beta)\,2\rho}.$$

If $q \leqslant \frac{(1-\beta)-\beta\left(1-2\bar{k}\right)}{(1-\beta)}$ then the proof is immediate from Lemma $A.0$. For the remaining of this proof we assume otherwise. So, $(1-\beta)(1-q) \leqslant (1-2\bar{k})\beta$. Then, (Υ_2) is maximized at $\hat{e}_i(\phi) = 1$. Assume by contradiction that this is not true and let $\ddot{e}_i(0)$, $\ddot{e}_i(1)$ and $\ddot{e}_i(\phi) < 1$ be the maximum of (Υ_2). Without loss of generality we can assume that $0.5\left(\ddot{e}_i(0)+\ddot{e}_i(1)\right) \leqslant 0$. Otherwise, we can decrease both $e_i(0)$ and $e_i(1)$ and increase $e_i(\phi)$ leaving m and l unchanged until $e_i(\phi)$ reaches 1 or $0.5(e_i(0)+e_i(1))$ reaches zero. If $\ddot{e}_i(0)$, $\ddot{e}_i(1)$ is fixed and $e_i(\phi)$ increases to 1, l remains fixed and m increases to a point m^* which is smaller than $(1-\beta)(1-q) \leqslant (1-2\bar{k})$. If m^*+l is less than β^u then (Υ_2) does not decrease as $e_i(\phi)$ increases. If m^*+l is greater than β^u then $m^*-l \leqslant \beta^l$ and, at $e_i(\phi) = 1$, (Υ_2) is 1.

(Υ_2) is maximized at $\hat{e}_i(0) = 1$. If not then let $\ddot{e}_i(0) < 1$, $\ddot{e}_i(1)$ and $\hat{e}_i(\phi) = 1$ be the maximum of (Υ_2). Without loss of generality we can assume that $\ddot{e}_i(1)$ $= -1$. Otherwise, we can increase $e_i(0)$ and decrease $e_i(1)$ leaving m unchanged and increasing l until $e_i(0)$ reaches 1 or $e_i(1)$ reaches -1. As $e_i(0)$ increases to 1, m and l increase to m^* and l^*, where m^* is equal to $(1-\beta)(1-q) \leqslant (1-2\bar{k})$. Hence, as $e_i(0)$ increases to 1 either (Υ_2) does not decrease or it reaches 1.

l decreases with $e_i(1)$, but $m-l$ and $m+l$ increases with $e_i(1)$. So, if $m-l$ is greater than β^l then (Υ_2) does not decrease as $e_i(1)$ decreases. If $m-l$ is smaller than β^l then (Υ_2) does not increase as $e_i(1)$ decreases. Assume $e_i(\phi) = e_i(0) = 1$. If $q \geqslant \frac{1+\beta\left(-1+2\left(\bar{k}+v\right)\right)}{((1-\beta)2\rho)}$ then $\hat{e}_i(1) \in (-1,1)$ solves $m-l = \beta^l$ and maximizes (Υ_2). Otherwise, $m-l$ is greater than β^l even at $e_i(1) = -1$ and, hence, (Υ_2) is maximized at $\hat{e}_i(1) = -1.\boxdot$

Proof of Equation $(**)$: Assume that $\beta \leqslant \beta_1$. Then, $\frac{2\beta v}{(1-\beta)(2\rho-1)} \leqslant \frac{(1-\beta)-\beta\left(1-2\bar{k}\right)}{(1-\beta)}$. If $q \leqslant \frac{2\beta v}{(1-\beta)(2\rho-1)}$ then $q \leqslant \frac{(1-\beta)-\beta\left(1-2\bar{k}\right)}{(1-\beta)}$. So, $\hat{e}_i(1) = -1$, $\hat{e}_i(\phi) = \frac{\beta(1-2\bar{k})}{(1-\beta)(1-q)}$ and

$\hat{e}_i(0) = 1$. It follows that evaluated at $\hat{e}$, $m = \beta(1-2\bar{k})$ and l is smaller than $2\beta v$. Thus, $(m-l, m+l) \subseteq \left(\beta^l, \beta^u\right)$. Hence, if $q \leqslant \frac{2\beta v}{(1-\beta)(2\rho-1)}$ then (Υ_2) is l divided $2\beta v$, and if $q = \frac{2\beta v}{(1-\beta)(2\rho-1)}$ then (Υ_2) is 1.

Assume that $\beta_1 \leqslant \beta \leqslant \beta_2$. If $q \leqslant \frac{1+2\beta(\bar{k}+v-1)}{(1-\beta)2\rho}$ then $\hat{e}_i(1) = -1$, $\hat{e}_i(\phi) = \min\left\{\frac{\beta(1-2\bar{k})}{(1-\beta)(1-q)}, 1\right\}$, $\hat{e}_i(1) = 1$. It is also the case that $q \leqslant \frac{2\beta v}{(1-\beta)(2\rho-1)}$ and, evaluated at $\hat{e}$, l is smaller than $2\beta v$ and $m-l$ is greater than β^l. So, $(m-l, m+l) \subseteq \left(\beta^l, \beta^u\right)$. Therefore, (Υ_2) is l divided $2\beta v$. On the other hand, if $q > \frac{1+2\beta(\bar{k}+v-1)}{(1-\beta)2\rho}$ then $\hat{e}_i(\phi) = \hat{e}_i(0) = 1$ and $\hat{e}_i(1) > -1$ is given by the formula in Lemma A. Hence, evaluated at $\hat{e}$, $m - l = \beta^l$ and

$$l = 0.5\,(2\rho-1)\left(\frac{\left(1-2\beta+2\left(\bar{k}+v\right)\right)}{(1-\rho)}\right)$$

which does not depend on q.⊡

6.1. Proof of Facts in section 4.

Proof of Fact 1. Note that $\hat{e}_i(0) = 1$, $\hat{e}_i(1) \neq -1$ only if $q > L\left(\beta, \rho, \bar{k}, v\right)$. However, $\hat{q} \leqslant L\left(\beta, \rho, \bar{k}, v\right)$.

Proof of Facts 2A and 2B. L increases with v and G decreases with v. Also, L goes to zero as v goes to zero. To demonstrate Fact $2B$, let $\dot{v} \equiv \min\left\{0.5-\bar{k}, \bar{k}\right\}$ be the maximum possible value for v. Let $\dot{C}$ be $\frac{(2\rho-1)}{L\left(\beta,\rho,\bar{k},v\right)4\beta v}$. If $C \leqslant \dot{C}$ then $L\left(\beta, \rho, \bar{k}, \dot{v}\right) \leqslant G\left(\beta, \rho, C, \dot{v}\right)$. So, $\hat{q} = L\left(\beta, 1, \bar{k}, v\right)$ for $v \in (0, \dot{v})$. Otherwise there is $\bar{v} \in (0, \dot{v})$ such that $L\left(\beta, \rho, \bar{k}, \bar{v}\right) = G\left(\beta, \rho, C, \bar{v}\right)$. It follows that $\hat{q} = L\left(\beta, 1, \bar{k}, v\right)$ for $v \in (0, \bar{v})$ and $\hat{q} = G\left(\beta, \rho, C, v\right)$ for $v \in (\bar{v}, \dot{v})$

Proof of Fact 3. $G\left(\beta, \rho, C, v\right)$ is decreasing in C and $L\left(\beta, \rho, \bar{k}, v\right) < 1$ does not change with C. Hence, if C is sufficiently small then $\hat{q} = L\left(\beta, \rho, \bar{k}, v\right)$. Increasing q above $L\left(\beta, \rho, \bar{k}, v\right)$ does not change the chances of candidate $w \in (0,1)$ in state w.

Proof of Fact 4A. L decreases with ρ and G increases with ρ. Let $\bar{C}$ be $\frac{4\beta v}{L\left(\beta,1,\bar{k},v\right)}$. If $C < \bar{C}$ then $L\left(\beta, 1, \bar{k}, v\right) < G\left(\beta, 1, C, v\right)$. Under the assumption that $\beta < \beta_2$, it follows that $L\left(\beta, 0.5, \bar{k}, v\right) > G\left(\beta, 0.5, C, v\right) = 0$. So, there is $\bar{\rho} \in (0.5, 1)$ such that $L\left(\beta, \bar{\rho}, \bar{k}, v\right) = G\left(\beta, \bar{\rho}, C, v\right)$. It follows that $\hat{q} = G\left(\beta, \rho, C, v\right)$ for $\rho \in (0.5, \bar{\rho})$ and $\hat{q} = L\left(\beta, \rho, \bar{k}, v\right)$ for $\rho \in (\bar{\rho}, 1)$. So, $\hat{q}$ is increasing in ρ for $\rho \in (0.5, \bar{\rho})$ and $\hat{q}$ is decreasing in ρ for $\rho \in (\bar{\rho}, 1)$. If $C \geqslant \bar{C}$ then $L\left(\beta, 1, \bar{k}, v\right) \geqslant G\left(\beta, 1, C, v\right)$ and $\hat{q} = G\left(\beta, \rho, C, v\right)$ for $\rho \in (0.5, 1)$. So, $\hat{q}$ is increasing in ρ.

Proof of Fact 4B. The probability that candidate 0 wins the election ($P(0)$) is 0.5 plus 0.5 times the difference between the probability that candidate 0 is elected in state 0 and the probability that candidate 1 is elected in state 1. Hence, $P(0) =$

$$0.5 + 0.5\left(\frac{(m+l)-\beta^l}{4\beta v} - \frac{\beta^u-(m-l)}{4\beta v}\right) = 0.5 + 0.5\left(\frac{m-\beta(1-2\bar{k})}{2\beta v}\right) \quad (\Upsilon_3)$$

By definition, $m = (1-\beta)(1-\hat{q})\min\left\{\frac{\beta(1-2\bar{k})}{(1-\beta)(1-\hat{q})}, 1\right\}$. So, a marginal increase in ρ either leaves $P(0)$ unchanged (when $\beta(1-2\bar{k}) < (1-\beta)(1-\hat{q})$) or it changes $P(0)$ in the opposite direction that $\hat{q}$ changes. It follows that for $\rho \in (0.5, \bar{\rho})$, $P(0)$ is weakly decreasing in ρ (because, by Fact $4A$, $\hat{q}$ is increasing in ρ). For $\rho \in (0.5, 1)$, $P(0)$ is weakly increasing in ρ (because $\hat{q}$ is decreasing in ρ). On the other hand, if $C \geqslant \bar{C}$ then $P(0)$ is weakly decreasing in ρ (because $\hat{q}$ is increasing in ρ).

Proof of Fact 4C. Full information equivalence holds in state 0 when the fraction of partisans supporting candidate 1 is the majority of the electorate. This follows because in this event candidate 1 receives the majority of the votes regardless of the independents vote (and, therefore, regardless of what information independents posses). The probability of this event is independent of ρ. The only other event in which full information equivalence holds occurs when candidate $w \in \{0,1\}$ is elected in state w. The probability of this event is given by equation $(**)$. By equation $(**)$, this probability is increasing in ρ because both $(2\rho-1)L$ and $(2\rho-1)G$ are increasing in ρ.

Proof of Fact 5. G decreases with β. L increases with β if $\beta \leqslant \beta_1$ and decreases otherwise. If $G(\beta_1, \rho, C, v) \geqslant L(\beta_1, \rho, \bar{k}, v)$ then $\hat{q} = L(\beta, \rho, \bar{k}, v)$ for $\beta \in (0,1)$. Otherwise there is β_3, $\beta_3 < \beta_1$, such that $G(\beta_3, \rho, C, v) = L(\beta_3, \rho, \bar{k}, v)$. Hence, $\hat{q} = L(\beta, \rho, \bar{k}, v)$ if $\beta \leqslant \beta_3$ and $\hat{q}$ decreases with β if $\beta > \beta_3$ because either $\hat{q} = G(\beta, \rho, C, v)$ or $\hat{q} = L(\beta, \rho, \bar{k}, v)$ and $\beta \geqslant \beta_1$.

Proof of Fact 6. $L(\beta_1, \rho, \bar{k}, v)$ goes to 1 and $G(\beta_1, \rho, C, v)$ goes to $\frac{0.5}{C}$ as ρ goes to 0.5. So, if $C > 0.5$ then $G(\beta_1, \rho, C, v) < L(\beta_1, \rho, \bar{k}, v)$ when ρ is sufficiently close to 0.5. Hence, $\hat{q} = G(\beta, \rho, C, v)$ and $\hat{e}_i(\phi) < 1$ for β in a neighborhood of β_1 and ρ in a neighborhood of 0.5. In this neighborhood, turnout is given by $T = \beta + (1-\beta)\frac{(2\rho-1)}{4\beta C v} + \beta(1-2\bar{k})$. At β_1, an increase in β implies in a marginal increase in T given by $\frac{\partial T}{\partial \beta}(\beta_1) = 2(1-\bar{k}) - \frac{(2\rho-1)}{4Cv}\frac{1}{(\beta_1)^2}$ which is negative if ρ is sufficiently close to 0.5 because $\frac{(2\rho-1)}{(\beta_1)^2}$ goes to infinity as ρ goes to 0.5.

Proof of Facts 7A, 7B and 7C. Fix an arbitrary $\dot{\rho}$ and $\dot{v}$. G goes to infinity as β goes to zero. So, if β is small enough then $\hat{q} = \frac{2\beta v}{(1-\beta)(2\rho-1)}$ in a neighborhood of $\dot{\rho}$ and $\dot{v}$. In this neighborhood, it must be the case that $\beta \leqslant \beta_1$ and $\hat{e}_i(\phi) = \frac{\beta(1-2\bar{k})}{(1-\beta)(1-\hat{q})} < 1$. So, $T = \beta + \frac{2\beta v}{(2\rho-1)} + \beta(1-2\bar{k})$. Turnout is decreasing with ρ and $\bar{k}$ and increasing with v. Given that candidate $w \in \{0,1\}$ is elected in state w, the expected margin of victory is given by

$$EMV = \frac{0.5E\left\{(1-\beta)\hat{q}(2\rho-1) + \beta(1-2\bar{k}) - \beta(1-2\tilde{k})\right\}}{T} +$$

$$+\frac{0.5E\left\{\beta(1-2\tilde{k}) - \beta(1-2\bar{k}) + (1-\beta)\hat{q}(2\rho-1)\right\}}{T} = \frac{v(2\rho-1)}{v+(1-\bar{k})(2\rho-1)}.$$

So, the expected margin of victory is increasing with ρ and v⊡

References

[1] Austen-Smith, David and Jeffrey S. Banks. 1996. "Information Aggregation, Rationality and the Condorcet Jury Theorem." *American Political Science Review* 90(1): 34-45.

[2] Blais, A. 2000. *To Vote or Not to Vote: The Merits and Limits of Rational Choice Theory.* University of Pittsburgh Press.

[3] Bartels, L. 1996 "Uninformed Votes: Information Effects in Presidential Elections" *American Journal of Political Science*, 40 (1), 194-230.

[4] Feddersen, Timothy J. 2004. "Rational Choice Theory and the Paradox of Voting." *Journal of Economic Perspectives.* Vol 18, Winter.

[5] Feddersen, Timothy J. and Wolfgang Pesendorfer. 1996. "The Swing Voter's Curse." *American Economic Review*, 86(3): 408-24.

[6] Feddersen Timothy J. and Wolfgang Pesendorfer. 1999. "Abstention in Elections with Asymmetric Information and Diverse Preferences." *American Political Science Review* 93(2): 381-398.

[7] Feddersen, Timothy J., and Alvaro Sandroni. 2004a. "A Theory of Participation in Elections." Typescript.

[8] Feddersen, Timothy J., and Alvaro Sandroni. 2004b. "Companion to A Theory of Participation in Elections." Typescript.

[9] Grofman, Bernard and Scott Feld. 1988. "Rousseau's General Will: A Condorcetian Perspective." *American Political Science Review. 82(2): 567-76.*

[10] Lassen, David Dryer. 2005. "The Effect of Information on Voter Turnout: Evidence from a Natural Experiment." *American Journal of Political Science.* 49(1): 103-118.

[11] Martinelli, C. 2003. "Would Rational Voters Acquire Costly Information." Typescript.

[12] McClennan, Andrew. 1998. "Consequences of the Condorcet Jury Theorem for Beneficial Information Aggregation by Rational Agents." *American Political Science Review. 92(2): 413-18.*

[13] Miller, Nicholas. 1986. "Information, Electorates, and Democracy; some Extensions and Interpretations of the Condorcet Jury Theorem." In *Information Pooling and Group Decision Making,* ed. Bernard Grofman and Guillermo Owen. ed. Greenwich, CT: JAI Press.

[14] Nalebuff, N. and R. Shachar (1999) "Follow the Leader: Theory and Evidence on Political Participation," *American Economic Review,* 89-3, 525-47.

[15] Persico, N. 2004 "Committe Design with Endogenous Information." *Review of Economic Studies* 71(1): 165-194.

[16] Young, Peyton. 1988. "Condorcet's Theory of Voting." *American Political Science Review* 82(4): 1231-44.

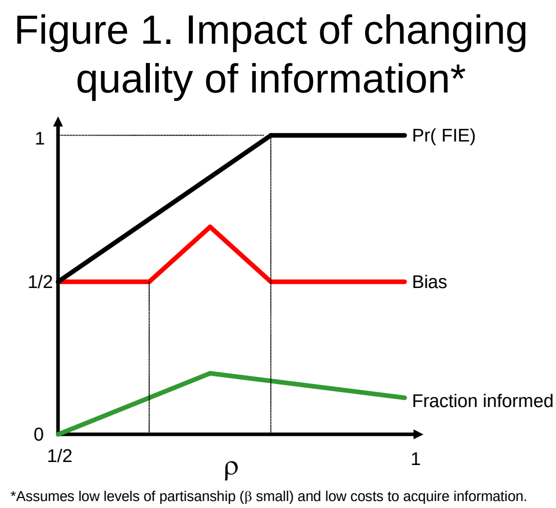
Figure 1. Impact of changing quality of information*
1
1/2
0
1/2
ρ
1
Pr(FIE)
Bias
Fraction informed
*Assumes low levels of partisanship (β small) and low costs to acquire information.

Figure 6.1: