A Theory of Participation in Elections

By TIMOTHY FEDDERSEN AND ALVARO SANDRONI*

The probability that a single vote is pivotal is negligible in a large election. Hence, small voting costs should dissuade turnout (Anthony Downs, 1957). Yet, significant turnout is often observed. The voting literature has worked around the "paradox of not voting," either by eliminating voters as strategic actors or by assuming that the decision to vote is independent of other strategic choices. Both approaches are problematic. Eliminating voters as strategic actors flies in the face of considerable evidence that voting behavior is strategic. For example, voters seem to condition their choice on the viability of candidates (Paul R. Abramson et al., 1992). Similarly, voting behavior under plurality rule is broadly consistent with the predictions of strategic voting (Gary C. Cox, 1997). Turnout is also correlated with education and income levels (Raymond E. Wolfinger and Steven J. Rosenstone, 1980). This is consistent with game-theoretic models that show that information levels influence turnout (Feddersen and Wolfgang Pesendorfer, 1996, 1997).

Alternatively, assuming that voters behave strategically while in the voting booth, but not when deciding to vote, also seems to contradict empirical data. William H. Riker and Peter C. Ordeshook (1968) find that turnout is inversely related to voting costs, and André Blais (2000) finds that closeness of elections influences turnout. This suggests that voters take the costs and benefits of participation into account.

Given the extensive evidence of strategic voter behavior, it is unsettling that so far there is not a canonical rational choice model of voting in elections with costs to vote. This paper aims at providing such a model.

There is considerable evidence that voters are motivated by a sense of civic duty (Blais, 2000). In this paper, we offer a model in which agents are motivated to vote out of a sense of ethical obligation. The starting point for our model is work by John C. Harsanyi (1977, 1980, 1992). Harsanyi (1977) considers a general game-theoretic model in which people receive a payoff from acting ethically. He assumes that a fraction of the population are "rule utilitarians." A rule utilitarian is an agent who receives a payoff for acting according to a strategy that maximizes social welfare (the sum of utilities), if everyone acts according to it. Harsanyi (1980) illustrates his idea with an example of costly voting. There are two candidates, one of whom is assumed to maximize social welfare if elected. Harsanyi assumes that a fixed fraction of the population votes for the socially inferior candidate. Given that the payoff for acting ethically is higher than the cost to vote, Harsanyi's model provides a microlevel explanation for turnout that depends not only on the relative magnitude of the cost to vote and the payoff for acting ethically, but also upon the level of support for the inferior candidate.

Harsanyi's example relies on the assumption that a substantial fraction of the voters are voting for an inferior candidate. If all rule utilitarians agree on which candidate is best, however, then it's not clear why an inferior candidate should receive any votes. In this paper, we introduce diversity into Harsanyi's framework.

We model a large election with two candidates. Each agent may vote for candidate 1, vote for candidate 2, or abstain. Voting costs vary within the population, and a single vote is never pivotal. The winner of the election is the candidate who receives the majority of votes cast. Like Harsanyi, we assume that some agents care about how they should behave and that agents have preferences over the candidates and the cost of the election. Unlike Harsanyi, voters'

* Feddersen: Kellogg School of Management, MEDS Department, Northwestern University, 2001 Sheridan Road, Evanston, IL 60208 (e-mail: tfed@kellogg.northwestern.edu); Sandroni: Kellogg School of Management, MEDS Department, Northwestern University, 2001 Sheridan Road, Evanston, IL 60208 (e-mail: sandroni@kellogg.northwestern.edu). We thank many audiences and two anonymous referees for their comments. Sandroni gratefully acknowledges financial support from the National Science Foundation grant SES 0109650. All errors are our own.

preferences are not necessarily related to social welfare and are not identical across all voters.[1]

There are two types of agents in our model: those who prefer candidate 1 (type 1) and those who prefer candidate 2 (type 2). We assume that there are more type 2 agents than type 1. Hence, we call type 2 agents the majority and type 1 agents the minority. Fixing the probability of winning for each candidate, all agents prefer to minimize the cost of the election.

In our model, each agent has an action he *should* take and receives utility from taking this action. Given a preference type, a *rule* defines a cut-off point such that agents with voting costs below this threshold should vote for their favored candidate. Agents with voting costs above the threshold should abstain. We say that agents who act as they should are "doing their part." Some agents (called *ethicals*) receive a payoff for doing their part which is sufficiently high so they prefer to act as they should. Other agents (called *abstainers*) receive no payoff for doing their part and, therefore, always abstain.

The central question in our model is which rules will ethical agents determine they must follow. Agents face a basic trade-off when they evaluate the merits of different rules. If followed, a rule that directs more agents of a given type to vote produces a higher chance that their favored candidate is elected, but at a higher cost of voting. Ethical agents take as given the behavior of abstainers and ethical agents of the other type when deciding which rule should be followed. Under this assumption, each voter independently decides that the right rule is the one that produces the best social outcome (given his preferences). We say that a behavioral rule profile is *consistent* if it defines rules that each type of agent decides he must follow, given a proper anticipation of the behavior of other types of agents.

Our model delivers comparative statics results on turnout, winning probabilities, and margin of victory as a function of parameters such as the level of disagreement within the electorate, the importance of the election, voting costs, and the payoff for doing one's part. The level of disagreement is measured by the size of the minority. The importance of the election is the weight in voters' preferences given to who wins the election relative to the voting costs. Therefore, in addition to the comparative statics results, welfare comparisons can be conducted because the model relates endogenous variables such as turnout and margin of victory to agents' preferences.

When the size of the minority becomes arbitrarily small (and all agents are assumed to be rule utilitarians) our model becomes like Harsanyi's: there is essentially only one ethical type. In that case, turnout goes to zero. This result shows the importance of heterogeneity of preferences as a factor in explaining turnout. It also illustrates why Harsanyi's model does not explain turnout.

In one respect, our model is similar to the earlier model proposed by Riker and Ordeshook (1968). They analyze a model of participation in which agents receive a "duty" payoff when they vote for their preferred candidate. Ethical agents in our model also receive the equivalent of a duty payoff when they act according to the best rule. In Riker and Ordeshook's model, however, the action an agent should take is determined exogenously. In contrast, the action an agent should take is determined endogenously in our model. It may be (and often is) the case that some agents receive an additional payoff for voting, while others receive a payoff for *not voting*. This follows because some agents will reason they should vote, while other will decide that they should not. More importantly, our comparative statics results differ from those of Riker and Ordeshook. For example, in Riker and Ordeshook, turnout levels are fixed and independent of the relative size of the minority. In contrast, our model predicts that turnout will be higher among the minority than the majority, yet the majority's preferred candidate wins with probability greater than a half.

It is also instructive to compare our results to those in the model of Thomas R. Palfrey and Howard Rosenthal (1983, 1985). Their model relies on events of insignificant likelihood, such as the chances of being pivotal in a large election. These models provide useful insights, but they cannot explain large-scale turnout with reasonable voting costs. Moreover, these results are unlikely to be robust. Other factors, such as

[1] Hence, voters are not necessarily rule utilitarians in our model.

a marginal desire to vote for one's favorite candidate, may radically alter the predictions of such models. Conversely, our results would not be significantly altered if there is a small, but positive, chance that a vote is pivotal.

In both Palfrey and Rosenthal (1983, 1985) and in our model, there is a negative correlation between turnout and margin of victory. Our result follows from the fact that when the size of the minority is small, agents in the majority group can win the election with high probability, even with low turnout. In a consistent rule profile, only a small fraction of a large majority will participate, but this fraction is sufficiently large to ensure a high chance of victory. So, expected turnout will be small and expected margin of victory large. As the size of the minority increases, the preferred rule of the majority requires a large turnout. When the group sizes are equal, the expected margin of victory is minimized and expected turnout maximized. Turnout and margin of victory are inversely related when the level of disagreement within the electorate changes, but these variables need not be inversely related when other parameters of the model vary. For example, we show a positive relationship between expected turnout and margin of victory as the importance of the election changes.

In Section I we present our model, and in Section II we show the comparative statics the model delivers. In Section III we justify our modelling choices, analyze the robustness of our results, and discuss some related work. The paper is concluded in Section IV.

I. The Basic Model

We model an election with two candidates, 1 and 2, and a continuum of voters who must either vote for candidate 1, vote for candidate 2, or abstain. Let A be the set of these three actions. The election is decided by majority rule.

Each agent has a cost of voting given by $\bar{c} > 0$ multiplied by an independent uniformly distributed random variable over the interval (0, 1). The cost of voting can be thought of as, for example, the time spent in the process of voting. We assume that each agent's cost of voting is independent of any other random variable in this model. Each agent knows her own realized voting costs, but not the realization of other agents' voting costs.

Agents have preferences about which candidate wins and the social cost of the election. These preferences reflect not only agents' self-interest but also their religious, ethical, or philosophical perspectives. For example, they might be associated with a notion of social welfare (utilitarianism), a concern for distributive justice, or support for human rights.[2]

There are two types of agents. Type 1 agents prefer candidate 1 and type 2 agents prefer candidate 2. Preferences must reflect choices, but no single agent decides who will be elected. Hence, preferences over social outcomes reflect the choices the agent would make *if* he were a social planner and could make such decisions. We assume that all agents prefer the social cost of voting to be minimized. So, if the agent were the social planner, then, holding constant the probability that candidate 1 wins the election, he prefers low turnout to minimize the social costs of voting.[3] Formally, type 1 and 2 agents have a utility function given by

$$(1) \quad wp - \phi \qquad \text{and} \qquad w(1-p) - \phi,$$

respectively, where p is the probability that candidate 1 wins the election, ϕ is the expected social cost of voting, and $w \in \mathbb{R}_+$ is a parameter of the model we call the *importance of the election*.[4]

[2] See John E. Roemer (1996) for an exposition of these ideas.

[3] Readers may wonder if this assumption is at odds with the fact that people worry about low turnout. Expressions of concern about low turnout need not represent a concern for turnout, per se. Instead, people may be concerned by what low turnout signals about society, e.g., lack of civic mindedness. If the winner of the election is fixed (independently of the votes) along with the level of civic mindedness, then a desire to maximize turnout seems hard to justify. Even so, including agents who prefer high turnout (or are indifferent about turnout) is not problematic. We only require that *some* agents prefer to minimize social costs. We do not include agents who prefer high turnout because in effect we would be assuming what we are trying to explain.

[4] Holding everything else constant, the higher the value of w, the higher the expected social cost that agents would be willing to trade for an increase in the chances that their favored candidate wins. Thus, w parameterizes the importance of the difference between the two candidates relative to the social cost of voting.

The fraction of type 1 agents in the electorate is $k \in (0, \frac{1}{2}]$. So, type 1 agents are a minority. The parameter k may be thought of as indicating the *level of disagreement* within the electorate. When k is small, almost everyone agrees that candidate 2 is preferred to candidate 1. When k is close to 0.5 the society is nearly divided on the question of which candidate is preferable.

The model as defined so far is standard. In voting games, as above, with a continuum of agents and costly voting there are generically no equilibria in which a positive fraction of the population participates. We now alter this standard game. We assume that each agent has a rule that he understands he should follow. If this agent acts according to this rule, then we say that the agent is "doing his part." We assume that some agents derive utility from doing their part.

Let a *rule* profile be cutoff points $\sigma_i \in [0, 1]$, $i \in \{1, 2\}$, which specify that type i agents with costs below $\sigma_i \bar{c}$ should vote for i and type i agents with costs above $\sigma_i \bar{c}$ should abstain. Some agents (called *ethical* agents) receive a payoff $D > \bar{c}$ for doing their part and, therefore, always do so. Other agents (called *abstainers*) receive zero payoff for doing their part. They always prefer to abstain. The fraction of ethical agents in each type group, $\tilde{q}_1$ and $\tilde{q}_2$, are independent and uniformly distributed over [0, 1].

In Riker and Ordeshook (1968), the set of agents who understand they should vote is exogenously determined. Unlike Riker and Ordeshook (1968), we assume that the agents determine their best rule as follows: taking the behavior of abstainers and ethical agents of type $j \neq i \in \{1, 2\}$ as given, each ethical agent type i independently considers what would occur if they (i.e., *ethicals* of type i) all follow rule σ_i. The rule σ_i^* that produces the best social outcome (for type i) is the one that each type i agent reasons is the rule he should follow.[5]

Agents face a trade-off when determining which rule to follow. A higher cutoff σ_i implies a higher chance that his favored candidate is elected, but also a higher social cost. We now formalize this trade-off. Assume that ethical agents follow the rule profile (σ_1, σ_2). The expected social cost of voting is

$$\phi(\sigma_1, \sigma_2) \equiv \bar{c}\left(kE(\tilde{q}_1)\int_0^{\sigma_1} x\partial x + (1-k)E(\tilde{q}_2)\int_0^{\sigma_2} x\partial x\right)$$

$$= \left(\frac{\bar{c}}{4}\right)(k(\sigma_1)^2 + (1-k)(\sigma_2)^2).$$

Candidate 1 is elected if he receives the majority of votes. This occurs if

$$k\tilde{q}_1\sigma_1 \geq (1-k)\tilde{q}_2\sigma_2 \Leftrightarrow \frac{\tilde{q}_2}{\tilde{q}_1} \leq \frac{k\sigma_1}{(1-k)\sigma_2}.$$

So, candidate 1 is elected with probability

$$p(\sigma_1, \sigma_2) \equiv F\left(\frac{k\sigma_1}{(1-k)\sigma_2}\right),$$

where F is the cumulative distribution function of $\tilde{q}_2/\tilde{q}_1$.

Given agents' preferences in (1), it follows that if ethical agents act according to the rule profile (σ_1, σ_2), then the induced payoffs for agents type $i \in \{1, 2\}$ are

$$R_1(\sigma_1, \sigma_2) \equiv wp(\sigma_1, \sigma_2) - \phi(\sigma_1, \sigma_2);$$

$$R_2(\sigma_1, \sigma_2) \equiv w(1 - p(\sigma_1, \sigma_2)) - \phi(\sigma_1, \sigma_2).$$

So, when evaluating the merits of different behavioral rules, the costs that agents take into account are those of the entire society. Stephen Coate and Michael Conlin (2004) show an alternative version of our model where agents consider only the voting costs of their group. Both alternatives are reasonable and produce similar models. Our choice is closer in spirit to the rule-utilitarian approach that inspired this paper. In this approach, agents take into account the welfare of the entire society when reasoning

[5] Hence, we can subsume the Riker and Ordeshook (1968) model by considering preferences such that agents will understand they should always vote, no matter the cost.

what they should do (although agents might disagree on which policies are best).

DEFINITION 1 (*Consistency requirement*): *The pair* $(\sigma_1^*, \sigma_2^*) \in (0, 1] \times (0, 1]$ *is a consistent rule profile if*

$$(2) \quad R_1(\sigma_1^*, \sigma_2^*) \geq R_1(\sigma_1, \sigma_2^*) \quad \text{for } \sigma_1 \in [0, 1];$$

$$(3) \quad R_2(\sigma_1^*, \sigma_2^*) \geq R_2(\sigma_1^*, \sigma_2) \quad \text{for } \sigma_2 \in [0, 1].^6$$

If a rule profile is not consistent, then at least one agent must conclude that the ethical agents of his type should follow an alternative rule and, thereby, achieve a better outcome. Conversely, in a consistent rule profile, no agent concludes that the ethical agents of his type can achieve a better outcome by following an alternative rule.

An agent may take a costly action even though he understands that this single action has no effect on the final outcome and, hence, does not benefit anyone. They take these costly actions because they feel morally obligated to do their part. The right behavioral rules are determined by the cutoff points $\bar{c}\sigma_1^*$ and $\bar{c}\sigma_2^*$. Ethical agents of type $i \in \{1, 2\}$ understand they should vote for i when their voting cost is below $\bar{c}\sigma_i^*$ and abstain when their voting cost is above $\bar{c}\sigma_i^*$.

The consistent rules (σ_1^*, σ_2^*) are determined so that each voter correctly anticipates behavior. No agent is surprised by the observed behavior of any other agent. Hence, the cutoff points $\bar{c}\sigma_1^*$ and $\bar{c}\sigma_2^*$ determine how agents understand they should behave, and also how agents will behave.[7]

[6] The level of turnout for each group type must be strictly positive. Otherwise, one group can win the election at a infinitesimally small social cost.

[7] If the payoff D for doing one's part is positive, but smaller than the maximum cost $\bar{c}$, then when rules are evaluated agents must take into account that some rules might not be followed. We refer the reader to Feddersen and Sandroni (2005) for a discussion on this issue. This paper is available at https://kellogg.northwestern.edu/faculty/fedderse/homepage/papers/dutytechnicalb.pdf.

The consistent rule profile (σ_1^*, σ_2^*) can be derived, in closed-form solution, as a function of the parameters $(k, w, \bar{c}, D)$. Let the cumulative distribution and density function of $\tilde{q}_1/\tilde{q}_2$ (F and f, respectively) be given by

$$F(z) = \frac{z}{2} \quad \text{if } z \leq 1; \quad f(z) = \frac{1}{2} \quad \text{if } z \leq 1;$$

$$F(z) = 1 - \frac{1}{2z} \quad \text{if } z \geq 1. \quad f(z) = \frac{1}{2z^2} \quad \text{if } z \geq 1.$$

Let $\bar{k} \equiv (k/1 - k)$. The first-order conditions of the maximization problem (2) are

$$(4) \quad wf\left(\bar{k}\frac{\sigma_1}{\sigma_2}\right)\bar{k}\frac{1}{\sigma_2} - k\left(\frac{\bar{c}}{2}\right)\sigma_1 \begin{cases} = 0 & \text{if } \sigma_1 \in (0, 1) \\ \geq 0 & \text{if } \sigma_1 = 1. \end{cases}$$

Note that $f(z) = f(1/z)1/z^2$. So, the first-order conditions of the maximization problem (3) can be written as

$$(5) \quad wf\left(\bar{k}\frac{\sigma_1}{\sigma_2}\right)\frac{\bar{k}\sigma_1}{(\sigma_2)^2} - (1 - k) \times \left(\frac{\bar{c}}{2}\right)\sigma_2 \begin{cases} = 0 & \text{if } \sigma_2 \in (0, 1) \\ \geq 0 & \text{if } \sigma_2 = 1. \end{cases}$$

With some algebra, it can be shown that the solution to (4) and (5) is given by Table 1.

II. Properties of Consistent Profiles

The final outcome of a large democratic election, where each agent independently decides to vote or to abstain in the absence of incentives such as bribes or pressure, depends on the fraction of agents who support each candidate, k and $1 - k$, and also on the fractions $\tilde{q}_1\sigma_1^*$ and $\tilde{q}_2\sigma_2^*$, which is comprised of ethicals of each preference type who understand they should vote. Thus, a fraction $k\tilde{q}_1\sigma_1^*$ of the electorate will vote for 1, a fraction $(1 - k)\tilde{q}_2\sigma_2^*$ of the electorate will

TABLE 1—CONSISTENT PROFILES IN CLOSED FORM

$\sigma_1^* = \sqrt{\frac{w}{\bar{c}}}\frac{1}{\sqrt[4]{k(1-k)}}, \sigma_2^* = \sqrt{\frac{w}{\bar{c}}}\sqrt[4]{\frac{k}{(1-k)^3}}$	if	$\frac{\bar{c}}{w} > \frac{1}{\sqrt{k(1-k)}}$;
$\sigma_1^* = 1, \sigma_2^* = \sqrt[3]{\frac{wk}{\bar{c}(1-k)^2}}$	if	$\frac{k}{(1-k)^2} < \frac{\bar{c}}{w} \leq \frac{1}{\sqrt{k(1-k)}}$;
$\sigma_1^* = \sigma_2^* = 1$	if	$\frac{\bar{c}}{w} \leq \frac{k}{(1-k)^2}$.

vote for 2, and all others will abstain. Hence, the expected total turnout is

$$T \equiv E(k\sigma_1^*\tilde{q}_1 + (1-k)\sigma_2^*\tilde{q}_2)$$

$$= 0.5(k\sigma_1^* + (1-k)\sigma_2^*).$$

The expected margin of victory is

$$MV \equiv E\left|\frac{(1-k)\tilde{q}_2\sigma_2^* - k\tilde{q}_1\sigma_1^*}{(1-k)\tilde{q}_2\sigma_2^* + k\tilde{q}_1\sigma_1^*}\right| = \frac{k\sigma_1^*}{(1-k)\sigma_2^*}$$

$$\times\left(2\ln 2 - 1 - \ln\left(1 + \frac{(1-k)\sigma_2^*}{k\sigma_1^*}\right)\right)$$

$$+ \frac{(1-k)\sigma_2^*}{k\sigma_1^*}\ln\left(1 + \frac{k\sigma_1^*}{(1-k)\sigma_2^*}\right).$$

The probability of victory for candidate 2 (supported by the majority) is

$$PV \equiv F\left(\frac{\sigma_2^*(1-k)}{\sigma_1^* k}\right).$$

The expected margin of victory and the probability of victory for candidate 2 are both increasing functions of $[(1-k)\sigma_2^*/k\sigma_1^*] \geq 1$. Hence, the comparative statics results for them are identical.

It follows from the equations above (and the closed-form solutions for σ_1^* and σ_2^* in Table 1) that T, MV, and PV can be obtained as a function of the parameters (k, w, $\bar{c}$, D). Therefore, a formal derivation of the properties of consistent profiles can be obtained by direct investigation of these functions. We omit this derivation and report the results.

PROPERTY 1: *The expected fraction of agents in the majority group who vote* $(0.5\sigma_2^*)$ *is smaller than the expected fraction of agents in the minority group who vote* $(0.5\sigma_1^*)$. *However, the total expected turnout of the majority* $(0.5(1-k)\sigma_2^*)$ *is greater than the total expected turnout of the minority* $(0.5k\sigma_1^*)$.

Figure 1 illustrates Property 1.

The idea behind the proof of property 1 is as follows: From equations (4) and (5), it is immediate that changes in participation rates (σ_1 and σ_2) lead to a ratio of marginal benefits (and also a ratio of marginal costs) given by

$$\frac{w\dfrac{\partial p(\sigma_1, \sigma_2)}{\partial \sigma_1}}{w\dfrac{\partial p(\sigma_1, \sigma_2)}{\partial \sigma_2}} = \frac{\sigma_2}{\sigma_1} \tag{6}$$

and

$$\frac{\dfrac{\partial \phi(\sigma_1, \sigma_2)}{\partial \sigma_1}}{\dfrac{\partial \phi(\sigma_1, \sigma_2)}{\partial \sigma_2}} = \frac{k\sigma_1}{(1-k)\sigma_2}.$$[8]

The ratio of marginal benefits depends on participation rates, while the ratio of marginal costs depends on expected turnout. So, assume, by contradiction, that, in a consistent rule profile, the participation rate of the majority is equal to (or higher than) the participation rate of

[8] Here, we assume an interior solution to (4) and (5).

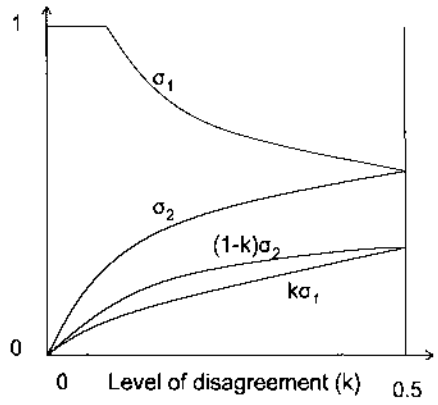

FIGURE 1. EXPECTED FRACTION OF TURNOUT IN EACH GROUP AS A FUNCTION OF THE LEVEL OF DISAGREEMENT (k)

Note: The figure holds for the case in which D is greater than $\bar{c}/2$ and $\bar{c}/w > 2$.

the minority (i.e., $\sigma_2^* \geq \sigma_1^*$). It follows from (6) that, when evaluating rules, a marginal increase in their participation rate is more costly and equally (or less) beneficial for agents in the majority than in the minority. This is impossible in a consistent profile. For both sides, marginal costs and benefits must be equal.

Assume, by contradiction, that the minority has an equal (or greater) chance of winning the election (i.e., $k\sigma_1^* \geq (1 - k)\sigma_2^*$). By (6), the marginal benefit of increasing participation rates is greater for the majority, but the marginal costs are equal (or smaller for the majority). This is impossible in a consistent profile.

In our model, a single vote cannot change election outcomes. Nevertheless, there is a "pivotality" aspect to the model. The benefits for the minority of increasing their participation rate from σ_1 to $\sigma_1 + \Delta$ is (w multiplied by) the probability that such an increase would alter the outcome of the election in favor of the minority. So, increasing participation rates (by Δ) is beneficial only if the group in question is behind in votes, but the election is sufficiently close to a tie, so that the difference in votes is smaller than the extra turnout. Otherwise, the extra turnout is wasted. The expected extra turnout of the minority ($E(\Delta k\tilde{q}_1)$) is smaller than the expected extra turnout of the majority ($E(\Delta(1 - k)\tilde{q}_2)$). Increasing turnout is relevant only when the election is close to a tie, however. Conditional on the election being tied, the expected extra turnout for the minority is $E(\Delta k\tilde{q}_1 | (1 - k)\sigma_2\tilde{q}_2 = k\sigma_1\tilde{q}_1)$, and $E(\Delta(1 - k)\tilde{q}_2 | (1 - k)\sigma_2\tilde{q}_2 = k\sigma_1\tilde{q}_1)$ for the majority. Like the ratio of marginal benefits in (6), the ratio of these conditional expected values is σ_2/σ_1 (independently of k). The intuition on the ratio of marginal costs in (6) is immediate. It is more costly for the majority than for the minority to increase their participation rates, because the unconditional expected extra turnout is greater for the majority.[9]

In the Palfrey and Rosenthal (1983, 1985) model, agents consider the effect of changing only one vote, which is beneficial only when the election is tied or the voters' preferred candidate is behind by one vote. In our model, agents consider the effect of changing many votes, which is also relevant only when their preferred candidate is tied or slightly behind in votes. In spite of this similarity, the conclusions of these two models differ significantly. In the Palfrey and Rosenthal model, the minority may be just as likely to win a large election as the majority (even if the majority is overwhelmingly large).[10]

It is also useful to contrast the properties of our model with the results that would be obtained in a purely decision-theoretic model along the lines of Riker and Ordeshook (1968), where it is exogenously determined that agents understand they should vote. In such a model, turnout is determined by the fraction of the electorate with cost to vote less than D. So, the participation rate of the majority and the minority is $min\{(D/\bar{c}), 1\}$. Hence, unlike our model, the participation rate of the minority is identical to that of the majority. This shows that the differences in turnout between the majority and the minority are related to the fact that agents endogenously determine how they should behave.

PROPERTY 2: *Expected turnout is strictly positive, and converges to zero as the level of disagreement goes to zero.*

[9] We thank an anonymous referee for this point.

[10] See Colin M. Campbell (1999), Tilman Borgers (2004), Stefan Krasa and Mattias Polborn (2004), and Curtis R. Taylor and Huseyin Yildiram (2005).

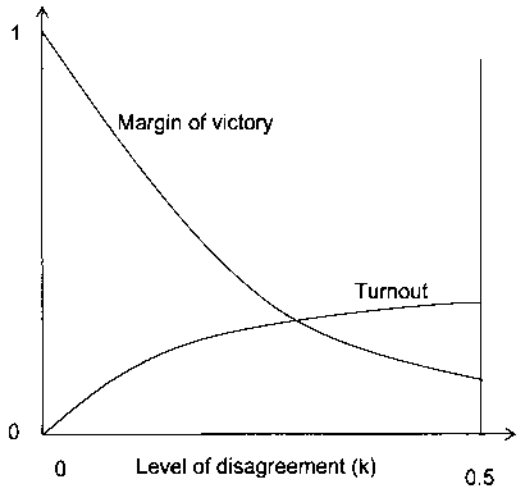

FIGURE 2. EXPECTED MARGIN OF VICTORY AND TURNOUT AS A FUNCTION OF THE LEVEL OF DISAGREEMENT

Property 2 shows that turnout is not simply a consequence of our assumption that some agents receive positive payoff for doing their part. It also depends upon the level of disagreement in the electorate. The intuition for property 2 is that, as group 2 becomes an overwhelming majority, they can vote at a low level and win the election with high probability. Properties 2 and 3 are illustrated in Figure 2.

PROPERTY 3: *Expected turnout is increasing in the level of disagreement, while the expected margin of victory is decreasing in the level of disagreement.*

Property 3 illustrates how our model can produce an inverse correlation between margin of victory and turnout, which has often been the focus of study in the empirical literature (see John C. Matsusaka, 1991; Ron Shachar and Barry Nalebuff, 1999; Blais, 2000).

The intuition behind Property 3 is as follows: assume that the majority is overwhelmingly large. Then, the majority group can set their voting threshold at a low level and still win the election with high probability. At this point, their marginal benefit (for increasing turnout even further) is low. Marginal costs are also low because increasing the voting threshold induces additional low-cost agents to vote. So, in a consistent rule profile, the majority turnout must be appropriately chosen at a low level. This avoids waste in voting efforts. Consequently, the turnout of the election is small. Given that the majority wins the election with high probability, the expected margin of victory must be large.

A more formal way of seeing this point is as follows: from equation (6) (and the fact that marginal costs and benefits equal), we obtain

$$\left(\frac{\sigma_2^*}{\sigma_1^*}\right)^2 = \frac{k}{(1-k)}. \tag{7}$$

Thus, when the level of disagreement goes to zero ($k \to 0$), the turnout of the majority goes to zero ($\sigma_2 \to 0$). On the the hand, consider groups of equal size (i.e., $k = \frac{1}{2}$). By (7), the participation rates of both groups are the same, $\sigma_2^* = \sigma_1^*$. An intuition is as follows: assume, by contradiction, that the groups choose different voting thresholds. The group with a larger threshold has higher marginal cost for increasing turnout (because it requires additional higher-cost agents to vote) and a lower marginal benefit for increasing turnout (because this group has a higher probability of winning the election). This is not possible in a consistent rule profile because of the equality between marginal costs and benefits. It follows that both sides choose the same turnout level and, so, the expected margin of victory must be low. Expected turnout must be large because decreasing turnout decreases marginal costs but does not decrease marginal benefits. Hence, a low turnout level makes marginal benefits greater than marginal costs.

In the model of Riker and Ordeshook (1968), the expected turnout does not change with the level of disagreement. An inverse relationship between turnout and margin of victory seems inconsistent with their model.

PROPERTY 4: *Expected turnout and margin of victory are increasing in the importance of election (w) and decreasing in the average voting cost ($\bar{c}/2$).*

The property that expected turnout increases with the importance of the election is consistent with the evidence that turnout is higher for presidential elections than state elections (see

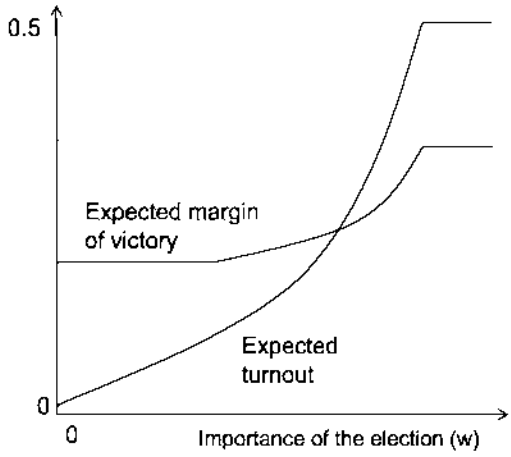

FIGURE 3. TURNOUT AND MARGIN OF VICTORY AS A FUNCTION OF THE IMPORTANCE OF THE ELECTION

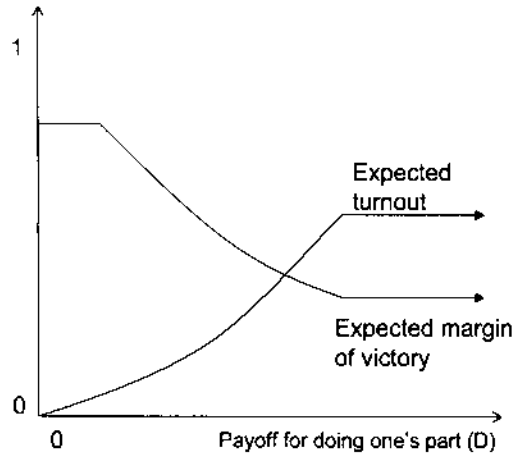

FIGURE 4. TURNOUT AND MARGIN OF VICTORY AS A FUNCTION OF THE PAYOFF FOR DOING ONE'S PART

Ruy Teixeira, 1987). The margin of victory is also increasing in the importance of the election. When the importance of the election is very high, both groups participate at maximum levels. Hence, the margin of victory is determined by the size of the groups and, therefore, is high. When the importance of the election decreases, turnout also decreases because the marginal benefit of participation decreases and the marginal costs remain the same. This decreases the chances that the candidate favored by the majority wins the election (hence, decreases the expected margin of victory) because the fraction of the minority who vote is higher than the majority. In expectation, the election becomes closer to a tie when both parties are not in full turnout than when they are.

The inverse relationship shown in Figure 2 obtains as the level of disagreement changes. Property 4 shows, however, that changes in the importance of the election can produce a positive correlation between margin of victory and turnout.

Property 4 contrasts with the results in the decision theoretic literature. In Riker and Ordeshook, neither turnout nor expected margin of victory changes when the importance of the election changes. Our model shows that both the expected turnout and the margin of victory are related to the importance of the election. Property 4 is illustrated in Figure 3.

When voting costs increase, the minority may be better off because of an increase in their chances of winning. The majority is never better off.

PROPERTY 5: *Expected turnout is weakly increasing in the payoff for doing one's part (D) while the expected margin of victory is weakly decreasing in D.*

Property 5 shows an inverse correlation between the expected margin of victory and turnout as D changes.[11] In the decision-theoretic literature, the D term and the cost to vote ($\bar{c}$) are somewhat interchangeable because behavior depends only on the ratio $D/\bar{c}$. In our setting, these variables have distinct effects. Our model delivers a positive correlation between expected turnout and margin of victory as the expected cost to vote changes, and a negative correlation as the payoff for doing one's part changes. Property 5 is illustrated in Figure 4.

To provide an intuition for the result that the expected margin of victory is decreasing in D, consider the case in which D is very small relative to $\bar{c}$. The payoff for doing one's part is small and so, in a consistent rule profile, only

[11] The comparative statics results on D require us to consider the case $D < \bar{c}$.

ethical agents with low costs vote. The cut-off point that determines when agents should abstain is given by D for both parties. It cannot be optimal to reduce turnout because the reduction in total voting costs would be small. As D increases, turnout increases. Then, the fraction of the minority who vote is no longer equal to the fraction of the majority who vote. The fraction of the minority who vote becomes greater than the fraction of the majority who vote. This decreases the chances that the candidate supported by the majority wins. It also decreases the expected margin of victory.

III. General Model and Results

Behavior motivated by moral considerations is fairly novel in formal models, both in political science and in game theory. Therefore, it might be useful to review our modelling choices and the robustness of our results. Some of our modelling choices have little or no impact on the results. Perhaps agents should receive a payoff for doing their part only when they take a costly action. This would prevent abstainers from receiving this payoff. In our model, however, these agents abstain with or without this proviso. Other assumptions are simplifying. For example, there is no loss of generality in restricting behavioral rules to be based on cut-off points. This is formally demonstrated in our companion paper, Feddersen and Sandroni (2005), where we allow agents to consider arbitrary rules (i.e., any mapping from costs to actions). The intuition behind this result is that the least costly way to achieve a given level of turnout is to have low-cost agents vote and high-cost agents abstain.

One of our main assumptions is that when ethical agents of a given type consider which rule to follow they take as given the behavior of abstainers and ethical agents of a different type. In Feddersen and Sandroni (2005), we provide a foundation for this assumption. We consider several ways in which agents could evaluate rules. At one extreme, agents could take as given the behavior of all other agents. At the other extreme, agents could consider rules that apply to everyone. We show that if agents are to evaluate rules under correct assumptions, then we must assume that agents take as given the behavior of abstainers and ethical agents of a different type.

In our companion paper, we show the existence and uniqueness of a consistent rule profile for a more general class of elections than considered here. We also analyze the robustness of our comparative statics results when the fraction of voters of each type is allowed to be random, and when assumptions of uniform distributions for voting costs and fractions of ethical agents are relaxed. Some properties (like 1 and 2) are quite general, but not every aspect of our comparative statics results generalize. In contrast with Property 4, we provide an example in which expected turnout decreases as the disagreement level increases from one point to another. Expected turnout cannot, however, monotonically decrease in the whole range of disagreement levels because it approaches zero when the level of disagreement vanishes.

A. *Related Work*

Coate and Conlin (2005) structurally estimated our model using data from liquor referenda in Texas. Their results show that the model can explain the wide variation in the turnout of different communities. Moreover, the model fits the data better than do reasonable alternatives, such as the intensity model that postulates that people are more likely to vote when they feel more strongly about the issue.

Models by Rebecca B. Morton (1991), Shachar and Nalebuff (1999), and Carole J. Ulhaner (1989) assume that elites produce participation by providing incentives for voting.[12] The elite-based models do not make explicit the underlying mechanism generating turnout. Perhaps elites make the voting costs of a part of the electorate negative either by bribing or by pressuring voters. Naturally, in this case there is no "paradox of not voting." As Shachar and Nalebuff (1999) note, an important component of the power of elites comes from their ability to persuade members of their party to vote. The difficulty is that without changing the costs to vote,

[12] See Feddersen (2004) for a review of group-based versus game-theoretic models of turnout.

standard models cannot explain how agents can be persuaded to vote.

Our model can deliver the microfoundations for persuasive elites because agents care about how they should behave and, therefore, elites might provide credible information about the importance of the election and thereby increase turnout. Large-scale turnout may still occur, however, in the absence of elites. In this paper, we focus on a basic model without elites and several other important institutional details. We leave for future work a model in which elites communicate with the electorate.

IV. Conclusion

Participation in large democratic elections is influenced by voters' sense of civic duty. This motivates a basic model in which agents endogenously determine how they should behave. Our model delivers testable implications and predicts variations in expected turnout and margin of victory as a function of costs to vote, level of disagreement within the electorate, importance of the election, and agents' incentives to do their part. Our results also show that costly voting and strategic considerations may coexist in a formal model. This provides a solid foundation for a theory of costly voting.

REFERENCES

Abramson, Paul R; Aldrich, John R.; Paolino, Philip and Rhode, David R. " 'Sophisticated' Voting in the 1988 Presidential Primaries." *American Political Science Review*, 1992, *86*(1), pp. 55–69.

Blais, André. *To vote or not to vote: The merits and limits of rational choice theory*. Pittsburgh: University of Pittsburgh Press, 2000.

Borgers, Tilman. "Costly Voting." *American Economic Review*, 2004, *94*(1), pp. 57–66.

Campbell, Colin M. "Large Electorates and Decisive Minorities." *Journal of Political Economy*, 1999, *107*(6), pp. 1199–1217.

Coate, Stephen and Conlin, Michael. "A Group Rule–Utilitarian Approach to Voter Turnout: Theory and Evidence." *American Economic Review*, 2004, *94*(5), pp. 1476–1504.

Cox, Gary W. *Making Votes Count: Strategic Coordination in the World's Electoral Systems*. Cambridge: Cambridge University Press, 1997.

Downs, Anthony. *An economic theory of democracy*. New York: Harper and Row, 1957.

Feddersen, Timothy J. "Rational Choice Theory and the Paradox of Not Voting." *Journal of Economic Perspectives*, 2004, *18*(1), pp. 99–112.

Feddersen, Timothy J. and Pesendorfer, Wolfgang. "The Swing Voter's Curse." *American Economic Review*, 1996, *86*(3), pp. 408–24.

Feddersen, Timothy J. and Pesendorfer, Wolfgang. "Voting Behavior and Information Aggregation in Elections with Private Information." *Econometrica*, 1997, *65*(5), pp. 1029–58.

Feddersen, Timothy J. and Sandroni, Alvaro. "The Calculus of Ethical Voting." Unpublished Paper, 2005.

Harsanyi, John C. "Morality and the Theory of Rational Behavior." *Social Research*, 1977, *44*(4), pp. 623–56.

Harsanyi, John C. "Rule Utilitarianism, Rights, Obligations and the Theory of Rational Behavior." *Theory and Decision*, 1980, *12*(2), pp. 115–33.

Harsanyi, John C. "Game and Decision Theoretic Models in Ethics," in Robert Aumann and Sergiu Hart, eds., *Handbook of game theory with economic applications*. Vol. I. Amsterdam: Elsevier Science, North-Holland, 1992, pp. 669–707.

Krasa, Stefan and Polburn, Mattias. "Is Mandatory Voting Better Than Voluntary Voting?" Unpublished Paper, 2004.

Matsusaka, John G. "Election Closeness and Voter Turnout: Evidence from California Ballot Propositions." *Public Choice*, 1993, *76*(4), pp. 313–34.

Morton, Rebecca B. "Groups in Rational Turnout Models." *American Journal of Political Science*, 1991, *35*(3), pp. 758–76.

Palfrey, Thomas R. and Rosenthal, Howard. "A Strategic Calculus of Voting." *Public Choice*, 1983, *41*(1), pp. 7–53.

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

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

Roemer, John E. *Theories of distributive justice.* Cambridge, MA: Harvard University Press, 1996.

Shachar, Ron and Nalebuff, Barry. "Follow the Leader: Theory and Evidence on Political Participation." *American Economic Review*, 1999, *89*(3), pp. 525–47.

Taylor, Curtis and Yildirim, Huseyin. "Public Information and Electoral Bias." Duke University, Department of Economics, Working Papers: No. 05-11, 2005.

Teixeira, Ruy A. *Why Americans don't vote: Turnout decline in the United States: 1960–1984.* New York: Greenwood Press, 1987.

Uhlaner, Carole J. "Rational Turnout: The Neglected Role of Groups." *American Journal of Political Science*, 1989, *33*(2), pp. 390–422.

Wolfinger, Raymond E. and Rosenstone, Steven J. *Who votes?* New Haven: Yale University Press, 1980.