Sequential Voting with Abstention

Marco Battaglini[1]

This version: May 2004

[1]Princeton University and CEPR. Part of this paper was completed while I visited the Economics Department at the Massachusetts Institute of Technology. I am grateful for the hospitality. I am grateful for helpful comments from Eddy Dekel, Thomas Palfrey, Wolfgang Pesendorfer, Michele Piccione, Thomas Romer, Tymon Tatur, the editor and two referees. Gabor Virag provided excellent research assistance. Address: Economics Department, Princeton University, 001 Fisher Hall, Princeton, NJ 08544-1021; E-mail: mbattagl@princeton.edu

Abstract

Dekel and Piccione (2000) have proven that information cascades do not necessarily affect the properties of information aggregation in sequential elections: under standard conditions, any symmetric equilibrium of a simultaneous voting mechanism is also an equilibrium of the correspondent sequential mechanism. We show that when voters can abstain, these results are sensitive to the introduction of an arbitrarily small cost of voting: the set of equilibria in the two mechanisms are generally disjoint; and the informative properties of the equilibrium sets can be ranked. If an appropriate q-rule is chosen, when the cost of voting is small the unique symmetric equilibrium of the simultaneous voting mechanism dominates all equilibria of the sequential mechanism.

Keywords: voting, elections, sequential elections, information aggregation.

JEL Classification: D72, D82.

I Introduction

Despite the fact that much attention in the theoretical literature has been focused on simultaneous voting mechanisms, many important decisions are actually taken in sequential mechanisms. This is certainly the case of presidential primaries, and of roll-call voting in legislatures; but it is often the case that even mechanisms that are supposed to be simultaneous, are not in practice, because late voters generally have access to exit polls and other information describing the choice of early voters.[1] A famous example is the 1980 Presidential election in which the news that Ronald Reagan was winning by a landslide changed expectation "suddenly and dramatically."[2] The "popular" intuition, therefore, is that because of bandwagon effects, timing matters. This view is also supported with strong historical and experimental evidence (see Hung and Plott, 2000, and Morton and Williams, 2000a and 2000b).

In an influential contribution, however, Dekel and Piccione (2000) have shown that informative symmetric equilibria of a simultaneous voting game with two options are equilibria in any sequential voting structure as well; and in unanimity games the sets of equilibria in simultaneous and sequential games are identical. As they have concluded, this result "completely demolishes any hope of obtaining strong conclusions about endogenous timing [...] and it extends the successful aggregation results of Feddersen and Pesendorfer (1997) to any sequential voting environment."[3] The result is very surprising because it runs precisely against the popular intuition and shows that the literature on informational cascades cannot be directly applied to voting games.

In this paper we argue that when voters can abstain, the equivalence between the equilibria of sequential and simultaneous voting games is sensitive to the introduction of any, even arbitrarily small, cost of voting.[4] Indeed, when this is the case *the set of equilibria of sequential and simultaneous games are generally disjoint.* What is relevant in our "non-equivalence" result, however, is not only the theoretical fact that equilibria are different; but, also, the more important implication that these differences may have a non-ambiguous impact on the ability of the election to aggregate information. Despite the fact that it is very difficult to characterize all equilibria of a sequential game (which can be asymmetric, and certainly history contingent), we present a simple, yet general result on the comparative properties of information aggregation in sequential and simultaneous voting mechanisms. *When the cost of voting is small and n is large, information aggregation is always maximized by choosing a simultaneous voting game*: in large elections, there is always a q^*-rule in correspondence of which the unique symmetric equilibrium of the simultaneous voting game dominates in information aggregation all the equilibria (even asymmetric) of a sequential election; moreover, no equilibrium of any q-rule of a sequential game dominates the simultaneous voting game with this q^*-rule.

The intuition of these results is simple. Both in a simultaneous and in a sequential election, agents evaluate the net benefit of voting for a particular alternative conditioning on the event of being pivotal. After any history of a sequential election, the expected benefit of voting for some alternative is *proportional*, but not *equal* to the net expected benefit in a simultaneous voting game. Both benefits are proportional to the expected value of the alternative conditional on being pivotal. Their factor of proportionality is the ratio of the probabilities of being pivotal after the history (in a sequential game) and ex-ante (in the simultaneous game). While the utility conditional on being pivotal is not history dependent if voters receive symmetric signals, the probability of being pivotal is always history dependent. This is irrelevant if the cost of voting is exactly zero, since the factor of proportionality is weakly positive. But when there is a cost of voting, even if arbitrarily small, the magnitude of the probability of being pivotal is important for the voter's decision, and since it affects the decision to abstain, which is history dependent, it affects the number of informative signals that can be aggregated in the election as well.

Since Dekel and Piccione (2000), the literature has not been able to reconcile the theory of rational voting in sequential and simultaneous elections with the evidence cited above on differences in voting outcomes. It has been suggested that particular behavioral assumptions on voters' preferences are needed (Callander, 2003). This is certainly a promising research line. Our results, however, contribute by explaining the differences in information aggregation between simultaneous and sequential voting mechanisms while remaining in the realm of the rational choice paradigm, and relying only on strategic abstention, which is a fact documented by robust evidence.[5]

On the normative side, our results also contribute to the debate on the optimal design of elections and their timing.[6] Scholars and prominent policy makers have argued that sequential elections give lesser known candidates a chance to emerge, and limit the advantage of strong incumbents.[7] In these cases, it is argued that sequential elections may be superior in aggregating voters' preferences, despite the risk of bandwagon effects. Based on these considerations, in 1996, the Republican Convention offered bonus delegates to states willing to schedule primaries later in the season:[8] as noted by Morton and Williams

(1999), "these changes were based on the premise that sequential primaries allow later voters to make more informed (and perhaps "better") decisions than they would in simultaneous voting."[9] Our results, however, provide conditions under which simultaneous elections allow a superior aggregation of information even when candidates are ex ante heterogeneous.

The next section presents the model. Section 3 presents the general results on the equivalence of equilibria in sequential and simultaneous voting games. In order to study information aggregation, in Section 4 we consider a setting in which voters have common values and the only purpose of the election is information aggregation: here we focus on the environment introduced by Austen-Smith and Banks (1996) and prove that a simultaneous election dominates sequential elections in the sense described above. Section 5 concludes.

II The model

In this section we present a simple model of voting which will be used throughout the analysis to compare sequential with simultaneous voting.

We assume that there are n voters who have to decide between two options: the status quo N and an alternative Y. The votes are aggregated using an anonymous and monotonic decision rule. For any such aggregation rule, we can define a real number $q \in [0,1]$ such that Y is chosen if and only if it receives at least a fraction q of cast votes.

We normalize the value of the status quo to zero. Each agent has a value $v_i \in V_i \equiv [-1,1]$ for Y. We assume that the vector of values is not observable by any agent, but each agent observes a private signal $x_i \in X_i \equiv [-1,1]$. We use the notation $X = \prod_i X_i$,

$V = \prod_i V_i$ and $\mathbf{x} = (x_1, ..., x_n)$ and $\mathbf{v} = (v_1, ..., v_n)$. The joint probability distribution of these two vectors is given by an atomless density function $f(\mathbf{v}, \mathbf{x})$. We assume that $E(v_i | x_{-i}, x_i)$ is strictly increasing in x_i for any x_{-i}.[10]

Voting games may have T periods. In each period a subset of agents has the opportunity to vote simultaneously. The agents who have the opportunity to vote are denoted by $V(t)$ and t_i is the period in which voter i has the opportunity. For simplicity, we consider sequential games in which the number of agents who have the opportunity to vote at each stage is a constant s: $s = |V(t)| = |V(t')|$. The outcome is decided at the end of the last stage, and at each stage voters can see the action taken by voters who decide in previous stages. We define a *sequential voting game* a roll-call voting game in which voters choose in a purely sequential way voting one after the other; a "hybrid" system in which voters vote in T steps is called a T-*stage* system. For simplicity, we will consider populations such that $T = \frac{n}{s}$ is an integer.

Each voter has three options: he can vote for alternative Y, vote for the status quo N or abstain, A. A strategy for a voter is therefore a function that maps the player's signal and the observed history of votes into the probability of voting and the probability of choosing Y: $\varphi_i : H_{t_i} \times X_i \rightarrow [0,1]^2$ where H_t is the set of histories of length t which describes how each agent $j \in \bigcup_{l=1}^{t-1} V(l)$ behaved. Differently from Dekel and Piccione (2000), however, an agent who decides to vote incurs a cost c which can be arbitrarily small, but strictly positive.

For notational convenience we shall denote the probability of choosing Y conditional on voting $s_i(h_{t_i}, x_i)$ and the probability of abstaining from voting $a_i(h_{t_i}, x_i)$. Given a strategy profile $\boldsymbol{\varphi} = (\varphi_1, ..., \varphi_n)$, moreover, it useful to define the expectation of v_i conditional on a history h_{t_i} and an event P: $E_{\varphi}[v_i; h_{t_i}, P]$.

III (Non-)Equivalence results

In this section we compare equilibria in sequential and simultaneous voting models. First we show that there is a sense in which the two sets of equilibria are equivalent if we assume that the cost of voting is exactly zero. We then show that this equivalence does not generalize to the case in which there is a cost of voting, even if arbitrarily small.

Dekel and Piccione (2000) have studied the equivalence of the equilibria in simultaneous and sequential voting systems in which three assumptions hold:

Assumption 1 (*Agents' Symmetry*) *For any agent* i, j, $X_i = X_j$ and $V_i = V_j$; and $f(v, x) = f(T^{ij}v, T^{ij}x)$ for any i, j, and any v, x.[11]

Assumption 2 (*Full Support*) For any v, if $f(v, x) > 0$ for some x, then $f(v, x') > 0$ for any x'.

Assumption 3 (*No Abstention*) Abstention is not allowed.

Given these conditions, they have proven that a voting profile $\boldsymbol{\varphi}^*$ is a symmetric equilibrium with informative strategies of the simultaneous game if and only if it is an equilibrium of the sequential game. *Informative strategies* are simply strategies that are responsive to private information:

Definition 1 *A strategy profile is informative if for each voter i there is at least a history h_{t_i} and some pair of signals x, x', such that after observing x the agent votes for Y with positive probability, and after observing x' votes for N with positive probability.*

The next proposition extends Dekel and Piccione's (2000) equivalence result between equilibria in sequential and simultaneous games when the agents can abstain but the cost of voting is zero. Although the argument is similar to the argument used in Dekel and Piccione (2000), it is worth noting that in their model the equivalence result does not necessarily hold with abstention, even if the cost of voting is zero. Dekel and Piccione assume a finite set of signals. In this case there may be equilibria in which agents are indifferent between voting for Y (or N) and abstaining when they condition on the pivotal event. If after these signals the agent uses a mixed strategy, voters who vote later would optimally condition their actions on the outcome of the randomization since this would reveal information on the actual signal observed by the voter who randomizes.[12] Such equilibria of the sequential game would be history dependent and could not be played in a simultaneous game.[13]

This problem does not arise in our environment since, with continuous signals, the probability that an agent is indifferent between two alternatives is zero: therefore posteriors are independent of the strategies (even mixed) used in these events. Let $\boldsymbol{\varphi}$ be a profile of informative and symmetric strategies:

Proposition 1 *Assume Assumptions 1-2 and $c = 0$. In any anonymous and monotonic decision rule with abstention, a profile $\boldsymbol{\varphi}$ is a symmetric Nash equilibrium of the simultaneous voting game if and only if every T-stage voting game has a sequential equilibrium $\varphi^T = \left(\varphi_1^T, ..., \varphi_n^T\right)$ such that for each agent i, $\varphi_i^T\left(h_{t_i}, x_i\right) = \varphi_i\left(x_i\right)$ for any h_{t_i} and x_i.*

The key assumption in this result is that the cost of voting is *exactly* zero. In this case, the expected benefit of voting Y in a simultaneous game -defined $u_i^{sim}(Y, x_i)$- is the expected value of v_i conditional on being pivotal $E_\varphi(v_i \,|Piv_i; x_i)$ times the probability of being pivotal $\Pr_\varphi(Piv_i; x_i)$.[14] In a sequential game, however, after any history h_{t_i} the benefit of voting Y for an agent in a sequential voting game -defined $u_i^{\text{seq}}(Y; h_{t_i}, x_i)$- is *proportional* to the utility in a simultaneous game, but it is not *equal* because the probability of being pivotal depends on the history of previous votes: clearly it is positive if no one has voted before, and zero if more than q voters have voted unanimously. So we have:

$$\begin{aligned} u_i^{\text{seq}}(Y; h_{t_i}, x_i) &= \Pr(Piv_i; x_i, h_{t_i}) E(v_i \,|Piv_i; x_i) \qquad (1) \\ &= \Lambda\left(x_i, h_{t_i}\right) u_i^{sim}(Y, x_i) \end{aligned}$$

where $\Lambda\left(x_i, h_{t_i}\right) = \frac{\Pr(Piv;\ x_i, h_{t_i})}{\Pr_\varphi(Piv;\ x_i)} \geq 0$. When the cost of voting is zero the fact that utilities are only proportional is strategically irrelevant: the sign of the utility of voting Y always (weakly) agree, so the optimal choice in a simultaneous game is still weakly optimal in a sequential game if the voting cost is zero.[15] In a model of voting with abstention and positive voting cost, however, the *magnitude* of the factor of proportionality $\Lambda\left(x_i, h_{t_i}\right)$ is important for the voters' choice.

As in Dekel and Piccione (2000), Proposition 1 considers informative strategies in which all agents use the signals that they receive. If we are interested in the properties of information aggregation in equilibrium, this requirement may be too strong. It might be that we are not interested in the fact that the equivalence between simultaneous or sequential voting games fails if we require that *all* agents are informative. We might be satisfied if the equivalence holds for equilibria that aggregate only some information. The next result, however, shows that even considering asymmetric equilibria in which some voters are not informative, the equivalence fails if there is any, even arbitrarily small, cost of voting:

Proposition 2 *For any $c > 0$, no strategy profile φ that is an equilibrium of a simultaneous voting game in which m agents are informative, is ever an equilibrium of any T-stage voting game with $T \geq 2\frac{n}{m}$.*

Proof. Assume that $\boldsymbol{\varphi} = \{\varphi_1(x_1, h_{t_1}), ..., \varphi_n(x_n, h_{t_n})\}$ is an equilibrium of the simultaneous game and therefore it is history independent: $\varphi_i(x_i, h_{t_i}) = \varphi_i\left(x_i, h'_{t_i}\right)$ for any h_{t_i}, h'_{t_i}. Assume also, by contradiction, that φ is an equilibrium of some T-stage voting game with $T \geq 2\frac{n}{m}$ in which at least m agents are informative. We proceed in two simple steps.

Step 1. We first show that all agents who have the opportunity to vote at T must abstain after any history and any signal. Given $T \geq 2\frac{n}{m}$, there are at least $m \geq 2\frac{n}{T}$ informative voters: so we can assume without loss of generality that there are at least $\frac{n}{T}$ informative voters who have the opportunity to vote before period T. Assume first that the number of uninformative voters who always abstain is some $k \in [0, n-m]$ and the number of uninformative voters voting Y is larger or equal to $\lceil q(n-k) \rceil - \frac{n}{T}$.[16] Then there exists a history h_T^1 in which all the informative voters are voting Y, so that the total number of voters voting Y is at least $V(Y) \geq \lceil q(n-k) \rceil$. After this history, an agent voting at stage T would never be pivotal, therefore he would abstain after any signal. Since $\boldsymbol{\varphi}$ is history independent, this implies that any agent voting at T is abstaining after any history and any signal. Assume now that the number of uninformative voters who vote for Y before T is strictly less than $\lceil q(n-k) \rceil - \frac{n}{T}$. In this case, there is a history

h_T^2 in which all the informative voters vote N. After h_T^2, the number of voters voting for N must be larger than $(1-q)(n-k)$. This proves that there is a history in which the voters at the Tth stage would not be pivotal. Because $\boldsymbol{\varphi}$ is history independent, this implies that they would always abstain.

Step 2. Consider now the induction step assuming that for some $t \in (1, T)$, all voters l who vote after t are uninformative because they always abstain. Since there are no informative agents after t and the number of informative voters is $m \geq 2\frac{n}{T}$, we can also assume without loss of generality that there are at least $\frac{n}{T}$ informative voters who have the opportunity to vote before t. Therefore an argument similar to Step 1 shows that the agents who have the opportunity to vote at t must be always abstaining. We can therefore conclude that there are no informative voters except at most in the first stage of the voting game. This is, however, impossible since in the first stage there are only $\frac{n}{T}$ voters. We therefore conclude that either in $\boldsymbol{\varphi}$ there are less than m informative voters, or the sequential voting game has less than $\frac{n}{m}$ stages. ■

The previous result is general because it holds for any equilibria, even asymmetric, independently of the order of voting, and if the cost of voting is arbitrarily small.[17] Two immediate corollaries follow from Proposition 2. For any $c > 0$, an equilibrium which aggregates the signals of at least two informative voters of a simultaneous game cannot be an equilibrium of the fully sequential game (i.e. when $T = n$). Moreover, for any $c > 0$, no equilibrium with informative strategies of the simultaneous game (i.e. when all voters are informative, $n = m$) can be an equilibrium of any T-stage voting game with more than one stage.

The bottom line of Proposition 2 is that the comparison between the properties of a simultaneous and a sequential game is difficult with abstention because there are no easy parallels between equilibria. In a sense, Proposition 1 opens a "Pandora's box:" the only way to compare the two environments is to calculate the equilibria of the two games; but this is extremely difficult since, at least in the sequential game, these are necessarily history dependent.

Without abstention, Dekel and Piccione's (2000) equivalence results can be used to draw strong conclusions on the informative content of the equilibria in a simultaneous and sequential game, showing general conditions for the two electoral systems to be informatively equivalent. Proposition 2 does not have immediate implications for informativeness, it simply says that the set of equilibria are disjoint and cannot be easily compared. In the next section, however, we show a general environment in which we can compare the informational properties of simultaneous and sequential voting games, and we can prove that, with small cost of voting, the set of equilibria in simultaneous games are generally *uniformly* more informative when the appropriate q-rule is chosen.

IV Information aggregation

As we mentioned in the Introduction, the comparative study of simultaneous and sequential elections has been motivated in the literature by the hypothesis that informational cascades may occur in sequential elections, and affect information aggregation in a predictable way. The results that we have presented in the previous section prove that strategic abstention makes the set of equilibria disjoint in the two systems, but they do not help to rank the informative properties of the equilibria. In this section, we show that *for any large n, if the cost of voting is small, information aggregation is maximized by choosing a simultaneous voting game*: there is always a q^*-rule in correspondence of which the unique symmetric equilibrium of the simultaneous voting game generally strictly dominates in information aggregation all the equilibria (even asymmetric) of the sequential system in which a given number of voters vote simultaneously at each stage; moreover no equilibrium of any q-rule of the sequential game dominates the simultaneous voting game with this q^*-rule.

To study the informative properties of the equilibria, we focus attention on the environment introduced by Austen-Smith and Banks (1996). This is a special version of the model presented in Section II in three ways. First, there are only two states of the world $\{Y, N\}$. We assume that state Y has prior probability $\pi \in (0, 1)$. Second preferences are represented by $u_i(Y, Y) = u_i(N, N) = v > 0$ and $u_i(N, Y) = u_i(Y, N) = 0$ for any $i = 1$ to n, where the first argument of u_i is the alternative selected by the outcome of the election and the second is the state of the world. Third, as also in the section on common values in Dekel and Piccione (2000), we now assume that each voter can privately observe only two signals $\{g, b\}$. We assume that signals are conditionally independent and informative, satisfying $\Pr[g; Y] = \Pr[b; N] = p \in \left(\frac{1}{2}, 1\right)$. To make this environment non trivial we make an assumption that guarantees that the outcome should depend on the signals. Define the posterior probability that the state is Y when there are k "g" signals out of n:

$$\beta_n(k) = \frac{\pi p^k (1-p)^{n-k}}{\pi p^k (1-p)^{n-k} + (1-\pi) p^{n-k} (1-p)^k}.$$

It is easy to prove that in the benchmark case with complete information, a voter would prefer alternative Y if $\beta_n(k) > \frac{1}{2}$. We assume:

Assumption 4 (*Non-Triviality*) There is a number of signals k such that $\beta_n(k) > \frac{1}{2} > \beta_n(k-1)$.

Clearly, this condition is satisfied for any generic choice of π and p when n is large enough. As in the previous Sections, we define a q-rule as the rule in which alternative Y passes if and only if it receives at least a fraction q of the cast votes. Keeping p, π and v constant (but arbitrary), a voting environment is therefore characterized by the parameters $\{c, n, q\}$. For any $\{c, n, q\}$, we can have a game $\Gamma_{sim}(c, n, q)$ in which agents vote simultaneously; or a game $\Gamma_{\text{seq}}(c, n, q, s)$ in which $s \geq 1$ voters vote simultaneously at each of the $T = \frac{n}{s}$ stages. Here too, we will consider populations of size n such that the number of stages is an integer.

In this setting, there is an unequivocal way to rank the informative value of these voting games. We say that an equilibrium of the simultaneous voting system with q-rule *weakly dominates* an equilibrium of the sequential system if and only if the probability of committing a mistake in $\Gamma_{sim}(c, n, q)$ is not larger than in $\Gamma_{\text{seq}}(c, n, q, s)$; it *strictly dominates* if the probability of a mistake is strictly smaller.[18]

The next result shows that, when n is large, simultaneous voting is superior to sequential voting if the cost of voting is small, independently of the number of voters s who vote simultaneously at each stage.

Proposition 3 *Let ε be an arbitrarily small number. There is a $\overline{n}$ such that for any $n > \overline{n}$ there are two cutoffs $c_2(n) < \varepsilon$ and $c_1(n) < c_2(n)$, and a $q^*(n) \in (0, 1)$ such that:*

i) If the $q^* (n)$-rule is adopted, the unique symmetric and informative equilibrium of $\Gamma_{sim} (c, n, q^* (n))$ weakly dominates any equilibrium of $\Gamma_{\text{seq}}(c, n, q^* (n), s)$ for any $c < c_2(n)$; and it strictly dominates for any c such that $c_1(n) < c \leq c_2(n)$.

ii) Moreover, no equilibrium of any other q-rule of the sequential game $\Gamma_{\text{seq}}(c, n, q, s)$ dominates the symmetric equilibrium of the simultaneous voting game with this $q^* (n)$-rule.

This result has an immediate implication. In order to aggregate information with large n, *there is never a rationale to choose a sequential election system if the cost of voting is small enough* (i.e., $c \leq c_2(n)$): there is always a q^*-rule, that is unbeatable by any equilibrium of any q-rule of the sequential game. Indeed, when $c \in (c_1(n), c_2(n)]$ the simultaneous system is strictly optimal, regardless of the equilibrium choice.[19]

In the $c \in (c_1(n), c_2(n)]$ interval, simultaneous voting is superior because it induces lower incentives to abstain, and therefore it elicits a larger number of informative signals. When c is very small (i.e., $c \leq c_1(n)$), however, abstention is not necessarily an issue for information aggregation even in sequential elections: this is why the two voting games may be equivalent in $(0, c_1(n)]$. Although it is reasonable to assume that the cost of voting is small, it is not plausible to assume that it is infinitesimal relative to the other parameters of the model. How much is $c_1(n)$ smaller than $c_2(n)$? Indeed if $c_1(n)$ converges to zero faster than $c_2(n)$ as $n \to \infty$, then the range $(c_1(n), c_2(n)]$ *essentially coincides with* $(0, c_2(n)]$ *in large elections.* We have:

Proposition 4 *As $n \to \infty$, $\frac{c_1(n)}{c_2(n)} \to 0$.*

These results should be interpreted in the light of the existing debate on optimal timing of elections. Two opposite camps have disputed the merits of sequential versus simultaneous elections (see Morton and Williams, 2000b, for an extensive account). One side of this dispute, as mentioned in the Introduction, claims that sequential elections give lesser known candidates a chance to emerge, and therefore limit the natural advantage of strong incumbents. In these cases, even if the prior favors the incumbent, it is argued that voters can learn from previous votes and update their posterior beliefs. This opinion is supported (and perhaps suggested) by a few famous episodes in which "dark horses" emerged in early primaries of presidential elections, as with George McGovern in 1972, Jimmy Carter in 1976, and Gary Hart in 1984; as well as by some careful econometric evidence of these learning effects (Bartels, 1986 and 1988, Norrander, 1983, and others). In an attempt to separate irrational momentum from real learning in the 1984 Democratic primaries, Bartels notes that:

> *"Although there is some room here for a momentum effect independent of real political content, it seems clear that no apolitical bandwagon hypothesis can account for the broad patterns of response to Hart's emergence".*[20]

Morton and Williams (1999), similarly, have shown with theoretical arguments and experimental evidence that in environments with symmetric information sequential elections may help voters to coordinate on superior candidates, and therefore improve the ability of the electoral system to aggregate preferences. Many policy makers have embraced these opinions: as mentioned in the Introduction, these arguments has even led the 1996 Republican Convention to explicitly discourage the concentration of primaries.

The "learning" story which is used to justify the evidence presented above implicitly assumes that naive voters always vote informatively using their signals even if it is not individually rational, or assumes symmetric information and focuses on coordination problems (as in Morton and Williams, 1999). Those who favor simultaneous elections, however, have suggested that at some stage of the election, voters would necessarily stop learning and fall trapped in bandwagon effects in which little information is aggregated. Taking the opposite position of the 1996 Republican Convention cited above, Bikhchandani et al. (1992) suggested that "the Super Tuesday, in which many southern states coordinate their primaries on the same date, was an attempt to avoid the consequences of sequential voting."[21] Although focused on individualistic institutions in which the action of a player does not directly affect other players' payoffs, the literature on herding behavior (Banerjee, 1992, Bikhchandani et al., 1992) has significantly influenced the research on voting models. Motivated by this literature, Hung and Plott (2001), for example, have produced experimental evidence showing that information cascades occur in voting models as well; and Callander (2003) has rationalized information cascades in elections with a behavioral model in which voters prefer to vote for a winner.

The important contribution of Deckel and Piccione (2000) is to clarify this debate showing that, because rational voters vote contingent on being pivotal, the literature on information cascades can not be directly applied to voting games. According to their results, however, both sides are incorrect: on the one hand, strategic voters do not learn from previous voters because they act as if they are pivotal; and on the other hand, exactly when a "bandwagon equilibrium" is possible in a sequential election, the same is also possible in a simultaneous election: so there is no reason to prefer one system to the other on this ground. Their results, therefore, seem to "demolish any hope" to solve this debate on the timing of elections remaining in standard voting theory (founded on the assumption that voters are rational and strategic).[22]

Clearly, this paper cannot solve the debate described above; our results, however, contribute to it in two ways. First, from a normative perspective, we show that it is generally possible to rank the informative properties of simultaneous and sequential voting systems in a non ambiguous way that does not depend on the particular choice of equilibria. And, second, they have a positive implication as well. As mentioned, an argument in favor of sequential elections is that they are better in aggregating information when one alternative has an initial advantage over the other. This advantage is measured in this model by the initial prior π. Propositions 3 and 4, however, are true even if the alternatives are heterogeneous ($\pi > 1/2$ or $\pi < 1/2$): in this case, too, there is a monotonic decision rule in a simultaneous election that dominates any decision rule of any equilibrium of the sequential election. Interestingly, the logic why simultaneous elections are superior to sequential elections is different from the intuition suggested by the literature on information cascades. In our model, the fact that voters do not know other voters' choices works as a commitment device to vote: in a sequential election this commitment device is absent and voters find it optimal to abstain when an interim outcome suggests that one option is winning (and so the probability of being pivotal is small). This logic is confirmed by evidence on voters' behavior, who are more likely to abstain when the probability of being pivotal is small (see Jackson, 1983, and Rosenthal and Sen, 1973).

The results presented above, however, leave a few open questions. First, given a particular q-rule (not necessarily the optimal q^*-rule): is a simultaneous or a sequential voting system better? It is very hard to obtain general results for this question. The power of Propositions 3 and 4 derives from the fact that when c is small, it is always possible to find a q-rule which guarantees that informative voting is rational in a simultaneous game; this is generally impossible, even when c is small, in a sequential voting system with large n. By choosing the q-rule, we can focus on a case in which the equilibrium of the simultaneous game has a nice property: informative voting is rational. To compare the properties of the electoral system, then, we only need to prove that for large n and small c, this is impossible in a sequential game. When q is arbitrary, informative voting is not necessarily rational in simultaneous elections; therefore, to compare the voting systems, we would need to compare the outcomes in the two sets of mixed strategies equilibria: these sets may be very complicated.

A strictly related question concerns the case when the cost of voting c is large and it is not possible to find a q-rule for which informative voting is rational in a simultaneous election. In this case the comparison between the electoral systems depends on the parameters describing the environment; but there are situations in which sequential voting may improve information aggregation. We show this fact with an example.

Example 1 *Consider an environment with the same payoffs and signal structure as described above in this section, and set* $\pi = 1/2$. *There are two voters who decide by majority rule: if they disagree, or if they both abstain, then option* Y *is chosen with probability* $1/2$. *The cost of voting is* $\frac{\alpha}{2}(2p-1)v$, *where* $\alpha < 1$ *but is close to one.*

In this case, the cost of voting is almost equal to the expected utility of choosing Y

(respectively, N) when a voter is alone and only one signal g (respectively, b) is available (i.e., $\frac{1}{2}(2p-1)v$); and there is never an equilibrium in which the two voters always vote informatively in the simultaneous election: if this were true, a voter would be pivotal precisely when the other voter has received the opposite signal; but in this case the posterior would be 1/2, the voter would be indifferent between the options, and abstention would be optimal.[23] There is, however, a symmetric equilibrium in which voters abstain with probability λ, and vote informatively with probability $1-\lambda$. In this equilibrium the expected net benefit of voting would be $\frac{\lambda-\alpha}{2}(2p-1)$.[24] Since in this case each voter must be indifferent between voting and abstaining, it must be that in a symmetric equilibrium $\lambda = \alpha$. So as $\alpha \to 1$, the probability of a mistake approaches 1/2 in the symmetric equilibrium of the simultaneous game, since the voters almost always abstain. There is, however, an equilibrium of the sequential game in which the first voter abstains and the second voter votes informatively.[25] In this case the probability of a mistake is only $1-p<1/2$, so the sequential election is better in aggregating information.[26] Intuitively, the reason why the equilibrium in the sequential election is better than the symmetric equilibrium of the simultaneous election is that it allows the voters to coordinate their actions: the first voter credibly delegates decision power to the second. In a symmetric equilibrium of a simultaneous game this coordination is lost when there is no pure strategy equilibrium (as in Proposition 3), and the equilibrium can be inefficient with large c.

Finally, how large is the advantage in information aggregation for the simultaneous voting system when the cost of voting is small? The proofs of Propositions 3 and 4 show that there is no equilibrium in the sequential game with a q-rule that is as informative as the symmetric equilibrium of the simultaneous game with a q^*-rule; but it does not provide a full characterization of all the possible equilibria in the sequential game. Because of this, we cannot prove that there are no equilibria of the sequential game that converge to the equilibrium of the simultaneous game in terms of information aggregation. However, there is no reason to believe this or, more in general, to believe that the informative properties of the two classes of equilibria are in any way comparable. The fact that voters abstain after some history clearly affects the expected value of any previous voter. In order to quantify these effects, it would be necessary to characterize not one, but all possible equilibria.

V Conclusion

This paper has shown that when voters can abstain and there is a cost to voting, the set of informative equilibria of simultaneous and sequential elections with binary choices are *disjoint*, even if the cost of voting is arbitrarily small. This may imply, as the popular intuition and history seems to suggest, that the two systems aggregate information in different ways; indeed we have shown that when the cost of voting is small there is always a q-rule in correspondence of which the simultaneous voting mechanism dominates in a strong sense a sequential mechanism from an informative point of view.

Is abstention really strategic and does it really play in reality the same role as in our argument? Jackson (1983) analyzed data on individual turnout in the 1980 Presidential election in what resembles a classical natural experiment. He assessed whether exposure to election night news reduced the likelihood of voting among those exposed. Consistently with our theory, he concludes that hearing news of the projected outcome decreased the likelihood of voting among those who had not already voted.[27] Strategic abstention, therefore seems to be an important component in voting games that are not simultaneous.

Appendix

A. Proof of Proposition 1

Let $\boldsymbol{\varphi}$ be a symmetric strategy profile in which the strategy for any agent i depends only on the agent's signal x_i. Given that all agents except at most i are following such a symmetric strategy, assume that agent i is indifferent between voting and abstaining after observing a signal x_i and conditioning on being pivotal (i.e., $E_{\boldsymbol{\varphi}}(v_i\,|Piv, x_i) = 0$). Then any agent with signal $x_i' > x_i$ or $x_i' < x_i$ would have a strict preference to vote for Y or N. Therefore the measure of voters who abstain in any symmetric equilibrium is zero. This implies that for any agent i there is only one family of pivotal events that receives positive probability: when all voters vote, and exactly $\lceil qn \rceil - 1$ vote for Y (where $\lceil x \rceil$ is the minimal integer larger or equal than x). Let us denote $\mathcal{P}_i$ this family of events in which a voter is pivotal with positive probability. By anonymity and the assumption of symmetric distribution of signals, for any $P, P' \in \mathcal{P}_i$, $E_{\boldsymbol{\varphi}}[v_i; h_{t_i}, P] = E_{\boldsymbol{\varphi}}[v_i; h_{t_i}, P'] = E_{\boldsymbol{\varphi}}[v_i; P]$ in any symmetric equilibrium $\boldsymbol{\varphi}$, since the identity of the "yes" and "no" voters is irrelevant. The utility of a voter to vote "yes" is therefore

$$u(v_i; x_i) = \Pr(\mathcal{P}_i; v_i; \boldsymbol{\varphi}) E_{\boldsymbol{\varphi}}[v_i; P].$$

Since by the full support assumption $\Pr(\mathcal{P}_i; v_i; \boldsymbol{\varphi}) > 0$, a voter i strictly prefers to vote for N if $E_{\boldsymbol{\varphi}}[v_i; P] > 0$, is indifferent if $E_{\boldsymbol{\varphi}}[v_i; P] = 0$ and strictly prefers if $E_{\boldsymbol{\varphi}}[v_i; P] < 0$. Consider now a agent in a T-stage voting game in an equilibrium which may be history dependent. His utility after a history h_{t_i} is

$$\begin{aligned} u(v_i; h_{t_i}, x_i) &= \Pr(\mathcal{P}_i; v_i, h_{t_i}, \boldsymbol{\varphi}) E_{\boldsymbol{\varphi}}[v_i; P] \\ &= \Lambda(h_{t_i})\, u(v_i; x_i) \end{aligned} \tag{A.1}$$

where $\Lambda(h_{t_i}) = \frac{\Pr(\mathcal{P}_i;\ v_i, h_{t_i}, \boldsymbol{\varphi})}{\Pr(\mathcal{P}_i;\ v_i, \boldsymbol{\varphi})} \geq 0$ for any h_{t_i} by, again, the full support assumption. To show the sufficiency part of the claim, note that:

$$\begin{aligned} u(v_i; h_{t_i}, x_i) &> 0 \Rightarrow u(v_i; x_i) > 0 \\ u(v_i; h_{t_i}, x_i) &< 0 \Rightarrow u(v_i; x_i) < 0 \end{aligned}$$

so if after history h_{t_i} the voter has strict preferences, then he would behave as if he did not know the history. However $u(v_i; h_{t_i}, x_i) = 0$ does not imply $u(v_i; x_i) = 0$ because it could be that $\Lambda(h_{t_i}) = 0$. To prove sufficiency, assume now that $\varphi = \{\varphi_i(x_i)\}_{i=1}^n$ is an equilibrium of the sequential game and: $u(v_i; h_{t_i}, x_i) = 0$, and the agent votes "no" with positive probability after history h_{t_i}; but $u(v_i; x_i) > 0$, and the voter would vote "yes" in the simultaneous game. Let h'_{t_i} be a history in which the probability of being pivotal is strictly positive. Since the strategy is history independent, it must be that $u(v_i; h'_{t_i}, x_i) \leq 0$, which implies $u(v_i; x_i) \leq 0$, a contradiction. Analogously, we can prove that if $u(v_i; h_{t_i}, x_i) = 0$ and "no" is voted with positive probability, then $u(v_i; x_i) \geq 0$. We conclude that after observing only x_i, a voter must find it optimal to vote as after $\{h_{t_i}, x_i\}$.

Consider now necessity. We have:

$$\begin{aligned}
u(v_i;x_i) &> 0 \Rightarrow u(v_i;h_{t_i},x_i) \geq 0 \\
u(v_i;x_i) &< 0 \Rightarrow u(v_i;h_{t_i},x_i) \leq 0 \\
u(v_i;x_i) &= 0 \Rightarrow u(v_i;h_{t_i},x_i) = 0.
\end{aligned}$$

Therefore any strategy that is strictly or (weakly) optimal in a simultaneous game is weakly optimal in a sequential game. ■

B. Proof of Propositions 3 and 4

To prove Propositions 3 and 4 it is useful to show a preliminary result. Let θ_1, θ_2, α_{j1}, α_{j2} $j = 1, 2$, and π be parameters with $\boldsymbol{\alpha}_1 = (\alpha_{11}, \alpha_{12})$, $\boldsymbol{\alpha}_2 = (\alpha_{21}, \alpha_{22})$, $\boldsymbol{\theta} = (\theta_1, \theta_2)$; and let v_n and k_n are sequences of positive integers indexed by n. Define $e_n^{k_n}(\boldsymbol{\alpha}_1, \boldsymbol{\theta})$ as:

$$e_n^{k_n}(\boldsymbol{\alpha}_1, \boldsymbol{\theta}) = \left(\frac{(n-1)!}{k_n!\,(n-1-k_n)!}\right)\left(\frac{\pi\alpha_{11}\theta_1^{k_n}\theta_2^{n-1-k_n} + (1-\pi)\,\alpha_{12}\theta_2^{k_n}\theta_1^{n-1-k_n}}{\pi\alpha_{11} + (1-\pi)\,\alpha_{12}}\right) \tag{B.2}$$

and $d_n^{k_n,v_n}(\boldsymbol{\alpha}_2, \boldsymbol{\theta})$ as

$$d_n^{k_n,v_n}(\boldsymbol{\alpha}_2, \boldsymbol{\theta}) = \frac{\pi\alpha_{21}\theta_1^{k_n}\theta_2^{n-1-k_n} + (1-\pi)\,\alpha_{22}\theta_2^{k_n}\theta_1^{n-1-k_n}}{\pi\theta_1^{k_n}\theta_2^{v_n+1} + (1-\pi)\,\theta_2^{k_n}\theta_1^{v_n+1}}. \tag{B.3}$$

We have:

Lemma 1 *Assume k_n is a non decreasing sequence such that $k_n \leq \frac{n-1}{2}$ $\forall n$ and $k_n \to \infty$ as $n \to \infty$; and v_n is a sequences of integers which may take value from 0 to a finite integer $\boldsymbol{\iota}$. For any $\alpha_{ji} \in (0,1)$ $\forall j, i$, $\pi \in (0,1)$, $\theta_1 \in (0,1)$, $\theta_2 \in (0,1)$ with $\theta_1 \neq \theta_2$, and either θ_1 or θ_2 is larger than $\frac{1}{2}$ and the other lower than $\frac{1}{2}$, then as $n \to \infty$:*

$$\frac{d_n^{k_n,v_n}(\boldsymbol{\alpha}_2, \boldsymbol{\theta})}{e_n^{k_n,v_n}(\boldsymbol{\alpha}_1, \boldsymbol{\theta})} \to 0$$

Proof. Assume, without loss of generality, that $\theta_1 > \frac{1}{2}$ (the case with $\theta_2 > \frac{1}{2}$ is analogous). The ratio $\frac{d_n^{k_n,v_n}(\boldsymbol{\alpha}_2,\boldsymbol{\theta})}{e_n^{k_n,v_n}(\boldsymbol{\alpha}_1,\boldsymbol{\theta})}$ can be written as:

$$A(n) \cdot \left(\frac{k_n! \, (n-1-k_n)!}{(n-1)! \left[\pi \theta_1^{k_n} \theta_2^{v_n+1} + (1-\pi) \, \theta_2^{k_n} \theta_1^{v_n+1} \right]} \right) \tag{B.4}$$

where

$$A(n) = \frac{\left[\pi \alpha_{11} + (1-\pi) \, \alpha_{12} \right] \left[\pi \alpha_{21} \left(\frac{\theta_2}{\theta_1} \right)^{n-1-2k_n} + (1-\pi) \, \alpha_{22} \right]}{\pi \alpha_{11} \left(\frac{\theta_2}{\theta_1} \right)^{n-1-2k_n} + (1-\pi) \, \alpha_{12}} > 0. \tag{B.5}$$

Since $n-1-2k_n \geq 0$, $\left(\frac{\theta_2}{\theta_1} \right)^{n-1-2k_n}$ in (B.5) is bounded above by one, it can be shown that $A(n)$ is bounded above by a constant ϕ_1, so we can focus on the second term in (B.4). Using the Stirling formula (see, for example, Durrett, §2.1, 1995) and some algebra, we have that

$$\lim_{n \to \infty} \frac{\frac{k!(n-1-k)!}{(n-1)!}}{\frac{k^k (n-1-k)^{n-1-k} \cdot \sqrt{2\pi k (n-1-k)}}{(n-1)^{n-1} \sqrt{(n-1)}}} = 1$$

From $0 < k \leq \frac{n-1}{2}$ we have that $\left(\frac{k}{n-1} \right)^k \leq \left(\frac{1}{2} \right)^k$ and $\left(1 - \frac{k}{n-1} \right)^{n-1-k} < 1$, so:

$$\begin{aligned} & \frac{k^k (n-1-k)^{n-1-k} \cdot \sqrt{2\pi k \, (n-1-k)}}{(n-1)^{n-1} \sqrt{(n-1)}} \\ = \; & \left(\frac{k}{n-1} \right)^k \left(1 - \frac{k}{n-1} \right)^{n-1-k} \frac{\sqrt{2\pi k \, (n-1-k)}}{\sqrt{(n-1)}} < \frac{\sqrt{2\pi k}}{2^k}. \end{aligned}$$

Therefore:

$$0 \leq \lim_{n \to \infty} \frac{2^k k! \, (n-1-k)!}{\sqrt{2\pi k} \, (n-1)!} \leq 1. \tag{B.6}$$

We can write:

$$
\begin{aligned}
\lim_{n\to\infty}\frac{d_n^{k_n,v_n}(\boldsymbol{\alpha}_2,\boldsymbol{\theta})}{e_n^{k_n,v_n}(\boldsymbol{\alpha}_1,\boldsymbol{\theta})} &\leq \lim_{n\to\infty}\left[\frac{\phi_1\sqrt{2\pi k_n}}{\pi 2^{k_n}\theta_1^{k_n}\left[\theta_2^{v_n+1}+\frac{1-\pi}{\pi}\left(\frac{\theta_2}{\theta_1}\right)^{k_n}\theta_1^{v_n+1}\right]}\cdot\frac{2^{k_n}k_n!\,(n-1-k_n)!}{\sqrt{2\pi k_n}\,(n-1)!}\right] \\
&= \lim_{n\to\infty}\frac{\phi_1\sqrt{2\pi k_n}}{\pi 2^{k_n}\theta_1^{k_n}\left[\theta_2^{v_n+1}+\frac{1-\pi}{\pi}\left(\frac{\theta_2}{\theta_1}\right)^{k_n}\theta_1^{v_n+1}\right]} \\
&\quad\cdot\lim_{n\to\infty}\frac{2^{k_n}k_n!\,(n-1-k_n)!}{\sqrt{2\pi k_n}\,(n-1)!} \\
&\leq \frac{\phi_1}{\pi\theta_2^{\iota+1}}\lim_{n\to\infty}\frac{\sqrt{2\pi k_n}}{(2\theta_1)^{k_n}}
\end{aligned}
$$

where the last inequality follows from (B.6), the fact that $\left(\frac{\theta_2}{\theta_1}\right)^{k_n}$ converges to zero as $n\to\infty$, that $\theta_1^{v_n+1}$ and $\theta_2^{v_n+1}$ are bounded, and $\theta_2^{v_n+1}$ is bounded below by $\theta_2^{\iota+1}$. Finally, given $\theta_1\in\left(\frac{1}{2},1\right)$, it is well known that $\lim_{x\to\infty}\frac{\sqrt{\pi x}}{(2\theta_1)^x}=0$, and therefore we have that the ratio $\frac{d_n^{k_n,v_n}(\boldsymbol{\alpha}_2,\boldsymbol{\theta})}{e_n^{k_n,v_n}(\boldsymbol{\alpha}_1,\boldsymbol{\theta})}$ converges to zero as $n\to\infty$. ■

We can now prove Proposition 3.

Proof of Proposition 3. Given Assumption 4, for any n we can find a $q^*\in(0,1)$ such that: $\beta_n(\lceil q^*n\rceil-1)<\frac{1}{2}<\beta_n(\lceil q^*n\rceil)$. Let $q_n^*=\lceil q^*n\rceil$, so that if there is no abstention Y is chosen if and only if there are at least q_n^* votes for Y. Observe, moreover, that it must be that $n-q_n^*\to\infty$ as $n\to\infty$. Denote $P_{q_n^*}$ the pivotal event in which q_n^*-1 out of n-1 voters vote for "yes" and $n-q_n^*$ out of n-1 vote for "no". By the definition of q^*, both the expected utility of voting "yes" conditional on event $P_{q_n^*}$ and a signal g, i.e. $E\left[u\,(vote\ Y)\ \middle|x=g,P_{q_n^*}\right]$, and the expected utility of voting "no" conditional on $P_{q_n^*}$ and a signal b, i.e. $E\left[u\,(vote\ N)\ \middle|x=b,P_{q_n^*}\right]$, are strictly positive and the voter would either vote according to the signal (informatively) or abstain. We consider now two cases.

Case 1: Assume first that $q_n^* \leq \frac{n+1}{2}$. We define $c_2(n)$ as:

$$c_2(n) = \min\left\{e_n^{q_n^*-1}(\boldsymbol{\alpha}, \boldsymbol{\theta}), e_n^{q_n^*-1}(\boldsymbol{\gamma}, \boldsymbol{\theta})\right\} \underline{EV},$$

where $\boldsymbol{\alpha} = \{\alpha_1, \alpha_2\}$, $\boldsymbol{\theta} = \{\theta_1, \theta_2\}$, and $\boldsymbol{\gamma} = \{\gamma_1, \gamma_2\}$ are given by: $\alpha_1 = 1-p$, $\alpha_2 = p$; $\gamma_1 = p$, $\gamma_2 = 1-p$; $\theta_1 = p$, $\theta_2 = 1-p$; and:

$$\underline{EV} = \min\left\{E\left[u\left(vote\ Y\right)\ \left|x = g, P_{q_n^*}\right.\right], E\left[u\left(vote\ N\right)\ \left|x = b, P_{q_n^*)}\right.\right]\right\}.$$

The variable $c_2(n)$ is a lower bound on the expected value of voting informatively in the simultaneous voting game ($e_n^{q_n^*-1}(\boldsymbol{\alpha}, \boldsymbol{\theta})$ and $e_n^{q_n^*-1}(\boldsymbol{\gamma}, \boldsymbol{\theta})$ are the interim probability of being pivotal when there is no abstention, after the agent observes, respectively, a "b" and a "g" signal). If $c < c_2(n)$ the voters in a simultaneous election always find it strictly optimal to vote informatively. Therefore, the simultaneous election is never strictly dominated by any sequential election (regardless of the q-rule used in the sequential election and of the equilibrium selection) since, at most, the equilibrium in the sequential cases uses the same number of informative signals as in the simultaneous game. Therefore, we only have to verify that there is a threshold $c_1(n) < c_2(n)$ such that when $c \in (c_1(n), c_2(n))$ the sequential election with q^*-rule described by $\Gamma_{\text{seq}}(c, n, q, s)$ uses a strictly lower number of signals, and therefore it is strictly dominated. Lets define:

$$z_n^1 = \min\left\{z \in \mathbb{N}^+ \text{ s.t. } z \cdot s \geq q_n^* - 1\right\} \tag{B.7}$$

and $v_n = z_n^1 s - q_n^* + 1$. It can be verified that $v_n \in [0, s]$ for any n, so v_n is a sequence of positive integers bounded above by s, moreover, since $q_n^* - 1 \leq \frac{n-1}{2}$ (this follows from $q_n^* \leq \frac{n+1}{2}$), $T \geq 2$ implies $z_n^1 s < n$ (i.e., z_n^1 is not the last stage of the voting game). Given this we can define:

$$c_1(n) = d_n^{q_n^*-1,v_n}(\boldsymbol{\alpha},\boldsymbol{\theta})\overline{EV},$$

where $\overline{EV} = \max\left\{E\left[u(vote\ Y)\ \middle| x = g, P_{q_n^*}\right], E\left[u(vote\ N)\ \middle| x = b, P_{q_n^*}\right]\right\}$. Since $q_n^* - 1 \leq \frac{n-1}{2}$, we can apply Lemma 1. Given this we know that both $\frac{d_n^{q_n^*-1,v_n}(\boldsymbol{\alpha},\boldsymbol{\theta})}{e_n^{q_n^*-1}(\boldsymbol{\alpha},\boldsymbol{\theta})}$ and $\frac{d_n^{q_n^*-1,v_n}(\boldsymbol{\alpha},\boldsymbol{\theta})}{e_n^{q_n^*-1}(\boldsymbol{\gamma},\boldsymbol{\theta})}$ converge to zero and therefore $\frac{c_1(n)}{c_2(n)}$ is lower than 1 for for n large enough. It follows that there is a $\overline{n}$ such that for $n > \overline{n}$ the set $(c_1(n), c_2(n))$ is always non empty. We now show that when $c \in (c_1(n), c_2(n))$ there is at least a history with positive probability in correspondence of which the simultaneous election strictly dominates the sequential election. Consider the sequential game, and assume that all the agents except at most an agent (say i) who votes in stage $z_n^1 + 1$ are voting informatively after a history in which the probability of being pivotal is larger than zero. Clearly if this were not the case, then the sequential game would be dominated by the simultaneous game and the result would be proven. Let us define $h^*(z_n^1)$ as a history of length z_n^1 in which $q_n^* - 1$ voters have voted "yes" and v_n have voted "no". Consider voter i who decides his action in stage $z_n^1 + 1$ after observing history $h^*(z_n^1)$ and a signal "b." The probability that this voter will be pivotal given history $h^*(z_n^1)$ is the probability that all the following voters vote "no" (conditional on history $h^*(z_n^1)$ and the observed signal). The probability of this event is $d_n^{q_n^*-1,v_n}(\boldsymbol{\alpha},\boldsymbol{\theta})$ defined in (B.3). Therefore an upperbound on the expected benefit for voting by this voter is given by $c_1(n) = d_n^{q_n^*-1,v_n}(\boldsymbol{\alpha},\boldsymbol{\theta})\overline{EV}$. After such a history, the expected utility of voting "no" for agent i at stage $z_n^1 + 1$ would be lower or equal than $d_n^{q_n^*-1}(\boldsymbol{\alpha},\boldsymbol{\theta})\overline{EV} < c$ and, therefore, he would abstain. So in an event in which history $h^*(z_n^1)$ is realized and all the agents in the stages after z_n^1 observe a "b" signal, Y would win, but N should win. In this case a simultaneous game would be strictly better.

Case 2: When $q_n^* > \frac{n+1}{2}$, the same result can be proven with a completely symmetric argument. In this case we define:

$$z_n^2 = \min\left\{z \in \mathbb{N}^+ \text{ s.t. } z \cdot s \geq n - q_n^*\right\}. \tag{B.8}$$

We then consider the history $h^*(z_n^2)$ in which $n-q_n^*$ agents have voted "no" and $z_n^2 s-n+q_n^*$ have voted "yes", and consider the decision of a voter who votes in stage z_n^2+1 after history $h^*(z_n^2)$ and observing a signal g. Since $n - q_n^* < \frac{n-1}{2}$ and $n - q_n^* \to \infty$ we can apply Lemma 1 to replicate the argument used in Case 1 above.

Finally, note now that, since both $c_1(n)$ and $c_2(n)$ converge to zero as $n \to \infty$, for any ε small, there is a n^* such that for $n > n^*$, $c_1(n) < c_2(n) < \varepsilon$. ■

We conclude with:

Proof of Proposition 4. From the definition of $c_1(n)$ and $c_2(n)$, and Lemma 1, it follows that $\frac{c_1(n)}{c_2(n)} \to 0$ as $n \to \infty$. ■

Footnotes

[1] A related but different situation occurs when many distinct elections are held in sequence. In this case too historical data on previous votes is available (and may affect voters' decisions), but here votes in previous elections do not directly affect the outcome. See McKelvey and Ordeshook (1985) for an analysis of these mechanisms.

[2] Quoted from Jackson (1983), pp. 615.

[3] Dekel and Piccione (2000), pp. 35.

[4] The observation that the act of voting is costly has a long tradition in political science: Downs (1957), Tullock (1967), Riker and Ordeshook (1968), Palfrey and Rosenthal (1983a) and (1983b).

[5] See Jackson (1983), for instance. Rosenthal and Sen (1973) provide evidence that the choice of abstention strategically depends on the probability of being pivotal.

[6] A more detailed account of this debate is presented in Section 4.

[7] Two prominent examples are George McGovern in 1970 and Jimmy Carter in 1976. See Morton and Williams (1999) and Morton (2004) for a discussion.

[8] Morton and Williams (1999), pp. 51.

[9] See Morton and Williams (1999), pp. 51.

[10] This condition is satisfied if, for example, signals are affiliated.

[11] For any n-tuple z, $T^{ij}z$ is the n-tuple obtained from z by exchanging z_i and z_j.

[12] In this case, after conditioning on pivotal events with different abstention rates, posterior beliefs may be different.

[13] To generalize the result to the case with abstention and discrete signals, therefore, we would need to make additional assumptions on the types' distribution that rule out ties in equilibrium. Note, however, that for any arbitrarily small but positive c, ties are often necessary to guarantee the existence of an equilibrium in large elections when signals are discrete.

[14] Clearly, the expected utility of voting for an alternative is the benefit of voting net of the cost of voting: $u_i^{sim}(Y, x_i) - c$. In this case, since $c = 0$, the distinction is irrelevant. The same observation is true for $u_i^{\text{seq}}(Y; h_{t_i}, x_i)$ defined below.

[15] With abstention a slightly more subtle issue needs to be taken care of regarding mixed strategies after events in which the probability of being pivotal is zero, but it does not affect the basic intuition. See the Appendix for details.

[16] $\lceil x \rceil$ is the smallest integer larger or equal than x.

[17] Note, moreover, that while we use the assumption of a continuum of types for Proposition 1, Proposition 2 holds even with discrete types.

[18] The probability of committing a mistake is the probability of choosing Y in state N plus the probability of choosing N in state Y. In our environment the distinction between errors of type I and errors of type II is not important since the utility is symmetric in the states of the world. Even if these errors were different, however, the result presented below would immediately generalize if we measured the informational performance with any weighted average of the two types of error: indeed both errors would be reduced.

[19] The sequential voting game has many equilibria which may have different informative properties.

[20] Bartels (1986), pp.21.

[21] See also Aldrich (1980) for an analysis of momentum in sequential elections.

[22] Dekel and Piccione (2000) study a model in which all voters receives signals of the same quality. This is a natural assumption in large non anonymous elections; in small elections, however, some voters may be recognized as more informed than others. Although the main idea that voters condition their action on being pivotal is still valid, information cascades are possible with asymmetric signals. A complete characterization of the equilibrium sets with asymmetric signals, therefore, seems a profitable (and still open) research question that may contribute to determine conditions under which bandwagon effects occur in small, non anonymous elections (as, for instance, in a faculty meeting).

[23] If π were different from $\frac{1}{2}$ or one expert had a more informative signal than the other, then a voter would not be exactly indifferent between the options in the pivotal event; this, however, would not be relevant for the result: if the cost of voting is large enough (as in this example), the voter would still find it optimal to abstain. Indeed, in correspondence to the pivotal event, the fact that the other voter has an opposite signal always reduces the expected benefit to vote (compared to the case in which a voter determines the outcome alone).

[24] With probability λ the other voter is abstaining and the benefit of voting (excluded the cost of voting) is $\frac{\lambda}{2}(2p-1)$; with probability $1-\lambda$, both voters are voting but when the voter is pivotal, there is no benefit of voting since the voter is indifferent between the options. For the net benefit of voting we need to subtract the cost of voting $c = \frac{\alpha}{2}(2p-1)$.

[25] If the first voter abstains, then the second voter would find optimal to vote informatively. If the first voter votes, then (in this equilibrium) the second voter assumes that the first voter has voted informatively and he would find optimal to abstain: either he would have to vote as the first voter, but this would be useless; or he has an opposite signal, and then he is indifferent between the alternatives. Given this reaction of voter 2, it is optimal for the first voter to abstain.

[26] When there are more than two voters we may have more that one symmetric equilibrium in the simultaneous game, but a similar equilibrium can be found. Note moreover that the simultaneous election has other non symmetric equilibria (as when one of the two voters abstains and the other votes informatively); still the symmetric equilibrium is dominated by the equilibrium of the sequential game.

[27] Interestingly the fact that the likelihood of turnout among Republicans fell more then among democrats, contradicts the behavioral assumption that voters like to vote for the winner.

References

[1] Aldrich, J., 1980. A Dynamic Model of Presidential Nomination Campaigns. Amer. Pol. Sci. Rev. 74, 651-669.

[2] Austen-Smith, D. and Banks J., 1996. Information Aggregation, Rationality, and the Condorcet Jury Theorem. Amer. Polit. Sci. Rev. 90, 34-45.

[3] Banerjee, A., 1992. A Simple model of Herd Behavior. Quart. J. Econ. 107, 797-817.

[4] Bartels, L., 1986. Candidate Choice and the Dynamics of the Presidential Nominating Process. Amer. J. Polit. Sci. 31:1-30.

[5] Bartels, L., 1988. Presidential Primaries and the Dynamics of Public Choice. Princeton University Press, Princeton, N.J.

[6] Bikhchandani, S., Hirshleifer D. and Welch I., 1992. A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades. J. Pol. Econ. 100, 992-1026.

[7] Callander S., 2003. Bandwagon and Momentum in Sequential Voting. Mimeo. MEDS, Northwestern University.

[8] Dekel, E. and Piccione M., 2000. Sequential Voting Procedures in Symmetric Binary Elections. J. Pol. Econ. 108, 34-55.

[9] Downs, A, 1957. An Economic Theory of Democracy. Harper and Row, New York.

[10] Durrett, R., 1995. Probability. Duxbury Press, Belmont.

[11] Feddersen, T. and Pesendorfer W., 1997. Voting Behavior and Information Aggregation in Elections with Private Information. Econometrica. 65, 1029-58.

[12] Hung A. and Plott C., 2000. Information Cascades: Replication and Extension to Majority Rule and Conformity-rewarding Institutions. Amer. Econ. Rev. 91, 1508-1520.

[13] Jackson J., 1983. Election Night Reporting and Voter Turnout. Amer. J. Polit. Sci. 27, 615-635.

[14] McKelvey R. and Ordeshook P., 1985. Sequential Elections with Limited Information. Amer. J. Polit. Sci. 29, 480-512.

[15] Morton, R., 2004. Analyzing Elections. Forthcoming. W.W. Norton, New York.

[16] Morton, R. and Williams K., 2000a. Information Asymmetries and Simultaneous versus Sequential Voting. Amer. Polit. Sci. Rev. 93, 51-67.

[17] Morton, R. and Williams K., 2000b. Learning by Voting: Sequential Choices in Presidential Primaries and Other Elections. University of Michigan Press, Ann Arbor.

[18] Norrander, B., 1993. Ideological Representativeness of Presidential Primary Voters. Amer. J. Polit. Sci. 33, 570-87.

[19] Palfrey T. and Rosenthal H., 1983a. A Strategic Calculus of Voting. Public Choice. 41, 7-53.

[20] Palfrey T. and Rosenthal H., 1983b. Voting Participation and Strategic Uncertainty. Amer. Polit. Sci. Rev. 79, 62-78.

[21] Riker, W. and Ordershook P., 1968. A Theory of The Calculus of Voting. Amer. Polit. Sci. Rev. 62, 25-42.

[22] Rosenthal, H. and Sen S., 1973. Electoral Participation in the French Fifth Republic," Amer. Polit. Sci. Rev. 67: 29-54.

[23] Tullock, G., 1967. Toward a Mathematics of Politics. University of Michigan Press, Ann Arbour.