American Economic Journal: Microeconomics 2014, 6(4): 162–179
http://dx.doi.org/10.1257/mic.6.4.162

The Strategic Dis/advantage of Voting Early†

By Eddie Dekel and Michele Piccione*

Under sequential voting, voting late enables conditioning on which candidates are viable, while voting early can influence the field of candidates. But the latter effect can be harmful: shrinking the field increases not only the likelihood that future voters vote for one's favorite candidate, but also that they vote for an opponent. Specifically, if one's favorite candidate is significantly better than all others, then early voting is disadvantageous and all equilibria are equivalent to simultaneous voting. Conversely, when some other candidate is almost as good, then any Markov, symmetric, anonymous equilibrium involves sequential voting (and differs from simultaneous voting). (*JEL* D72)

We study a simple model of voting in which voters choose the time in which they cast their votes. Our main objective is to study how the preferences of voters affect the temporal structure of voting. We highlight a particular strategic disadvantage of "early voting" in the presence of multiple candidates: narrowing down the field of competitors induces subsequent voters to vote either for early voters' preferred candidate or for her opponents, and the latter negative effect can dominate the former positive one. More specifically, we identify, in a particular parametrized model, when sequential or simultaneous voting will occur.

Several institutions present instances of sequential voting, with either exogenous or endogenous timing. In many legislative bodies, voting takes place sequentially. US presidential primaries are a particularly interesting example of sequential voting since each state decides the date of its own election.

Momentum or bandwagon effects are generally held to be the main manifestation of the difference between simultaneous and sequential voting. Such effects underscore the advantages of early voters over later voters in the determination of the election outcome. Momentum effects can arise from two distinct forces that interact in sequential voting. The first is strategic. As votes are cast sequentially, some candidates may find their chances of winning significantly reduced, and some voters may decide to shift their votes in favor of candidates that are more likely to succeed.

* Dekel: Department of Economics, Northwestern University, 2001 Sheridan Road, Evanston, IL 60208, and The Eitan Berglas School of Economics, Tel Aviv University (e-mail: dekel@northwestern.edu); Piccione: Department of Economics, London School of Economics and Political Science, Houghton Street, London WC2A 2AE, UK (e-mail: m.piccione@lse.ac.uk). We gratefully acknowledge ESRC Grant R000222823 and NSF Grant SES-1227434 for financial support. We also appreciate the excellent research assistant work from Michael Safronov and Jaber Zarezadeh. We are grateful to the referee for helpful comments and guidance.

† Go to http://dx.doi.org/10.1257/mic.6.4.162 to visit the article page for additional materials and author disclosure statement(s) or to comment in the online discussion forum.

The second is informational. Early voting can signal the quality of the candidates to later voters.

In this paper, we abstract from issues pertinent to the signaling of information about the valuation of candidates by assuming that the preferences of the voters are independent. We study an election with three candidates in which voting consists of two stages. In the first, "timing game," stage, prior to the preferences being realized, voters choose simultaneously the period in which they will cast their votes. In the second, "voting game," stage, the preferences are realized privately and voting takes place according to the order chosen in the first stage. In the voting game, voters in later periods observe the votes expressed by the voters in earlier periods. The candidate that obtains the majority of votes wins the election and ties are broken by a fair lottery.

We focus on the trade-off between two incentives that guide the strategic considerations of voters with respect to the temporal dimension of voting. On the one hand, a later voter has lower chances of wasting a vote as she acquires better knowledge about the likelihood of victory of the candidates. On the other hand, an early voter can influence the outcome of the election by affecting the probability of victory of some candidates in later periods. In particular, voting early influences the field of candidates and can keep one's candidate viable and make others unviable. We demonstrate a novel insight into the nature of sequential voting: if later voters persist in voting for their favorite candidate despite low chances of victory, early voting has a strategic disadvantage. In particular, if voters use voting strategies that we call *persistent*, that is, they vote for their preferred candidates as long as these candidates have a positive chance of winning, all equilibria in the first stage are equivalent to simultaneous voting. The reason is that, with such voting behavior, voting for one's favorite candidate, say A, in the first period can only be advantageous by making another candidate, say B, unviable and, thus, inducing voters for whom B was the favorite and A the second best to switch to A. But voters for whom B was the favorite and C the second best would also be induced to switch to C. While these two forces might seem to cancel one another, it turns out that the losses are more significant than the gains. Thus, when voters continue to vote for their favorite candidate so long as she might win, shrinking the field is detrimental and voters have no incentives to vote early.

More precisely, the paper focuses on *simple* strategies for which voting (but not necessarily timing) is pure and symmetric across voters and candidates (see Section III for a formal definition). We show (Theorem 1) that when restricting attention to simple persistent strategies, if the number of voters is at least six, then all equilibria are equivalent in terms of the outcome to one in which everyone votes in the last period. The restriction to persistent voting is consistent with equilibrium behavior if the value of every voter's second-best candidate is low compared to the favorite candidate. However, if the second-best candidate's value is close to the value of the favorite candidate, then equilibria in simple strategies cannot be persistent and must involve voting in multiple periods. (Theorem 2 and Corollary 1). Thus, we are able to link some preference environments to sequential or simultaneous voting.

The paper is organized as follows. In the next section, we briefly discuss the related literature. The model and results are contained in Sections II, III, IV, and V.

In Section VI, we discuss possible extensions. The Appendix contains some of the proofs.

I. Related Literature

The literature on sequential voting has mostly dealt with the issue of information aggregation in binary elections. Dekel and Piccione (2000) study a model in which voting is sequential and show that, in symmetric environments with two candidates, the symmetric equilibria in simultaneous voting are also equilibria in any sequential structure. Battaglini (2005) shows that, with abstention and costly voting, the above inclusion fails, and the set of equilibria in simultaneous and sequential voting can be disjoint. Fey (1996) and Wit (1997) study a two-signal, common-value environment in which herd-cascade equilibria exist. Morton and Williams (1999) find theoretical and laboratory evidence that later voters use the information transmitted by earlier voters. Callander (2007) shows the existence of "bandwagons" when the number of voters is infinite and voters have a desire to conform with the majority. Battaglini, Morton, and Palfrey (2007) compare the equity, information aggregation, and efficiency of simultaneous and sequential voting rules when voting is costly and information is incomplete. Ali and Kartik (2012) construct equilibria in which voters vote sincerely and that exhibit momentum effects. Hummel (2012) studies sequential equilibria with three candidates, where voters' preferences are correlated. First period outcomes are informative about later behavior and lead to equilibria where later voters do not vote for candidates who do poorly early, as they know their chances are slim.

Knight and Schiff (2010) study polling data from US presidential elections and find that early voters have significantly more influence than late voters. Deltas, Herrera, and Polborn (2010) provide a framework for analyzing the trade-off between learning about the candidates' quality (valence) and coordination in primary elections. They simulated various types of sequential elections using estimated structural parameters and showed that sequential elections in which candidates remain in the race yield the highest expected valence.

The theoretical literature on multi-stage voting also includes Sloth (1993), who studies sequential voting with perfect information and shows that the subgame perfect equilibria of roll-call voting are closely related to sophisticated equilibria of simultaneous voting, and Bag, Sabourian, and Winter (2009) who show that a class of voting procedures based on repeated ballots is Condorcet consistent.

II. The Model

The model has $N, N \geq 4$, voters and three candidates, denoted by A, B, C, whom voter i values as (v_A^i, v_B^i, v_C^i). We assume that each (v_A^i, v_B^i, v_C^i) is independently and identically distributed across voters with probability $f(\cdot)$ on $[0, 1]^3$.[1] The function $f(\cdot)$ is symmetric across the candidates and assigns zero probability to ties. In one

[1] The symmetry assumption is important for our results.

special case that we will consider in this paper, each voter's preferences are a random selection of permutations of $(1, x, 0)$, for fixed $x \in (0, 1)$. We will refer to this case as the x-model.

Voting consists of two stages: a timing stage and then a voting stage. In the timing stage, before the preferences are realized, voters choose simultaneously the period $t \in \{1, 2\}$ in which they will cast their votes.[2] This decision is assumed to be irrevocable. In the second, voting, stage, preferences are realized privately and voting takes place according to the order decided in the timing stage. Voters know the timing stage decisions of all voters and, at their time of voting in the second stage, know the earlier votes. The election is won by the candidate that obtains the majority of votes. Ties are decided by a fair lottery.

To define the voting stage strategies of the voters, let $\Omega = \{1, 2\}^N$ be the set of all possible timing-stage outcomes, i.e., specifications of who votes when. For $\omega \in \Omega$, let $t^i(\omega)$ be the period in which player i votes, $\mathcal{H}^2(\omega) = \{A, B, C\}^{|\{i\,:\,t^i(\omega)=1\}|}$ be the set of possible realizations of votes in period 1, and $\mathcal{H}^1(\omega)$ be the empty history. A *voting-stage* strategy for player i is a collection $s^i = \{s^i(\omega)\}_{\omega\in\Omega}$, where each $s^i(\omega)$ maps $\mathcal{H}^{t^i(\omega)}(\omega) \times [0, 1]^3$ to Δ, the set of probability distributions over $\{A, B, C\}$.

III. Equilibria

Throughout the paper we focus on equilibria that in the voting stage involves strategies that are pure and symmetric across candidates and across voters; we call these *simple* voting-stage strategies. Formally define n_J^1 to be the total number of votes received by candidate J in period 1, $J = A, B, C$. Given a profile of voting-stage strategies $\mathbf{s}$, let $\mathbf{s}(\omega)$ denote the N-tuple $(s^1(\omega), \ldots, s^N(\omega))$. The profile $\mathbf{s}$ is *symmetric across voters* if the strategies depend only on the number of voters in each period and are identical for voters in the same period, i.e., if $\mathbf{s}(\Upsilon\omega) = \Upsilon\mathbf{s}(\omega)$ for any permutation Υ and any $\omega \in \Omega$. Given a profile $\mathbf{s}$ that is symmetric across voters, let $\pi_{\mathbf{s}}^t(N^1)$ be the expected payoff of a voter of period t given that N^1 voters vote in period 1. A voting-stage strategy profile is *symmetric across candidates* if for any $\omega \in \Omega$, the strategy of each voter i maps (after the appropriate reordering) every permutation of the triple

$$(n_A^{t^i(\omega)-1}, v_A^i), (n_B^{t^i(\omega)-1}, v_B^i), (n_C^{t^i(\omega)-1}, v_C^i)$$

to an identical permutation of its image in Δ.[3]

We now provide two elementary existence results for the timing-game stage, when voting-stage strategies are symmetric across voters. Given a profile of voting-stage strategies $\mathbf{s}$, define a *timing-stage Nash equilibrium* (induced by $\mathbf{s}$) as a

[2] The number of periods is chosen solely for simplicity. See Section VI. If the choice of timing was made after preferences were realized then the game would be quite different. We think that our approach to timing is more realistic and interesting; for instance, it seems to match primaries better.

[3] Where $\left(n_J^{t^i(\omega)-1}\right)_{J\in\{A, B, C\}}$ is the null triple if i moves in period 1.

Nash equilibrium of the game in which the strategy space in the voting stage is restricted to the singleton **s**.

PROPOSITION 1: *Given any profile* **s** *of voting-stage strategies that is symmetric across voters, the game has a timing-stage Nash equilibrium in pure strategies. Moreover, if* $\pi^1_{\mathbf{s}}(X+1) > \pi^2_{\mathbf{s}}(X)$*, there exists a timing-stage Nash equilibrium in which the number of voters in the first period is strictly greater than X.*

PROOF:

To show existence, first note that, if $\pi^1_{\mathbf{s}}(N) \geq \pi^2_{\mathbf{s}}(N-1)$, the claim is trivially true. If $\pi^1_{\mathbf{s}}(1) < \pi^2_{\mathbf{s}}(0)$, it is a timing-stage Nash equilibrium for everyone to vote in the second period. Otherwise, define $\tilde{N}^1$ to be the largest N^1, such that $\pi^1_{\mathbf{s}}(N^1) \geq \pi^2_{\mathbf{s}}(N^1-1)$. It is trivial to see that $\tilde{N}^1$ voters choosing the first period is indeed a timing-stage Nash equilibrium since $\pi^1_{\mathbf{s}}(\tilde{N}^1) \geq \pi^2_{\mathbf{s}}(\tilde{N}^1-1)$ and $\pi^1_{\mathbf{s}}(\tilde{N}^1+1) < \pi^2_{\mathbf{s}}(\tilde{N}^1)$. The second part of the proposition follows by repeating the argument for $N^1 \geq X$.

REMARK 1: *For a second existence result, note that if we fix any* **s** *that is symmetric across voters, then the timing game is symmetric and, hence, has a symmetric timing-stage equilibrium (possibly in mixed strategies).*

IV. Simultaneous Voting

Our goal is to gain insights into the relationship between the preference of voters and the timing of voting. As discussed, late voters may benefit from being informed about some of the votes; and early voters may benefit from influencing the behavior of late voters. One difficulty in deriving results about the relationship between preferences and timing is the well-known feature of many voting models that voters can coordinate their behavior in an unappealing manner. To see this, observe that in the x-model a sequential equilibrium with simple voting-stage strategies and sequential voting is easily obtained for any $x \in (0, 1)$, when voters vote (i) for the preferred candidate in the first period, and (ii) for the most preferred candidate among the (possibly unique) leading candidates in the second period when $N^1 < N-1$. When $N^1 = N-1$, an optimal voting strategy for the (single) second-period voter that is symmetric across candidates is easily ascertained. Obviously, given the above voting-stage profile, it cannot be an equilibrium that all voters vote in period 2, as the most preferred candidate of a unique first-period voter wins with certainty. Existence of sequential voting in pure strategies then follows by Proposition 1, and in symmetric strategies by Remark 1. Note that this observation also provides an existence proof in the x-model for equilibria with simple second-stage strategies.

In this section, we sidestep the above problem of undesirable voting coordination rather bluntly by fixing the voting-stage strategies so that voters vote for their favorite candidate so long as it is possible she might win, and focus on equilibrium behavior in the timing stage. Formally, given a realization of votes in the first period, a candidate J is said to be *second-period viable* if J wins the election with strictly positive probability when all voters in the second period vote for J. A voting-stage strategy is said to be *persistent* if, when voting in the first period, it votes for the candidate with the highest valuation and, when voting in the second period, it votes for the *viable* candidate with the highest valuation. The game in which the set of voting strategies is restricted to persistent strategies is called a *P-voting game*. Note that persistent voting strategies are simple: pure and symmetric across voters and candidates.

The assumption that voting is persistent is consistent with equilibrium behavior when preferences are such that the difference in valuations between the most preferred and the second most preferred candidate is sufficiently large for all realizations. For example, in the x-model persistence is consistent with equilibrium behavior when x is small. Since the number of voters and candidates is finite, the set of possible realizations of votes is also finite. Hence, the increase in the probability of victory of a viable candidate from receiving one extra vote is bounded from below by a strictly positive number. By contrast for large x, persistent strategies are not consistent with equilibrium behavior; this follows because, for some realization of first-period votes, a voter whose favorite candidate is viable but behind the second-best candidate votes for his second-best candidate if x is close enough to one. The intuition is that, when x is close to one, a voter wishes to minimize the chances of winning of the least preferred candidate and can do so by voting for whoever between the other two candidates is leading. This will be shown formally in Lemma 1 in the next section.

The following example shows that persistence is not enough to rule out sequential voting. It constructs a sequential-voting equilibrium in a P-voting game that is an equilibrium and is distinct from a simultaneous-move equilibrium (which also exists in this case, and is strictly worse from the perspective of one player).

EXAMPLE 1: *Consider the x-model and suppose that* $N = 5$, *and that four voters vote in the first period. The behavior of a voter in the last period differs from the behavior in simultaneous voting and affects the outcome only when the candidate with the highest valuation obtains zero votes in the first period and the other two candidates get two votes each.*[4] *The expected utility of a first-period voter, conditional upon such realizations, is* $1/2 + x/4$.[5] *This is also the expected utility of all voters when they move simultaneously. However, the utility of the second-period voter with a persistent voting strategy is higher than in simultaneous voting, as the candidate valued x wins with certainty when the most preferred candidate is not viable. To see that this is indeed an equilibrium, note that a first-period voter is indifferent between voting in the first or in the second period as, by moving to the second period, voting will be equivalent in outcome to simultaneous voting.*[6]

[4] If some candidate receives three or more votes in the first period they would receive that whether the fifth voter votes in the first or second period, hence this voter's timing is irrelevant. If both candidates receive two votes and the fifth voter's favorite candidate is one of those, then the fifth voter would vote for that candidate regardless of whether he voted after the first four or simultaneously with them. Thus, his timing only matters as indicated in the text.

[5] The two candidates with two votes are equally likely to win; if it is the candidate for whom a first-period voter voted, then their utility is 1 (and this candidate wins with probability $1/2$) and if it is the other candidate it is equally likely to be this voter's second or least favorite candidate, yielding (with probabilty $1/2$) either x or 0 with equal probability. Hence, $1/2 + (x + 0)/4$.

[6] The proof of the following theorem will imply that there cannot exist an example in which the voters in the first-period strictly prefer voting in the first period.

The next theorem shows that this example is very special. If $N \geq 6$, sequential voting cannot be an equilibrium outcome.

THEOREM 1: *Suppose that $N \geq 6$. Any P-voting game has an equilibrium in which all voters choose to vote in the second period. Moreover, all P-voting equilibria are equivalent in outcome to this equilibrium.*

PROOF:
See Appendix.

The intuition behind this result is simple despite its long proof. The events in which a candidate ceases to be voted for are exactly those in which that candidate, say candidate C, has no chance of winning the election. Therefore, one cannot save one's favorite candidate by voting early. The question remains whether one can facilitate coordination on one's favorite candidate by voting early and making another candidate unviable. Conditional upon the event of C, say, becoming unviable, voting is effectively binary in that only two candidates can win the election. However, in a binary election, the utility of a voter is decreasing in the number of voters as her influence gets diluted. Having some voters switch from C to A or B with equal probabilities is then equivalent to increasing the number of voters in a binary election. Hence, a voter is better off by voting in later periods and allowing candidates that are not viable to receive votes.

This analysis clearly relies on the symmetry of the candidates. Nevertheless, it helps clarify and highlight the general point that decreasing the field of candidates can be disadvantageous, as it facilitates focusing not only on one's preferred candidate, but also on that candidate's opponents. While we do not develop this, it is clear that with more candidates this disadvantage would persist.

V. Sequential Voting

In this section, we will investigate when sequential voting *is* an equilibrium; naturally this will require that voters in the voting stage switch their vote away from their preferred candidate at times when she is still viable. Proposition 1 implies that a sufficient condition for the existence of equilibria with sequential voting, when voting-stage strategies are symmetric across voters and candidates, is that, if $N^1 = n_J^1 = 1$, second-period voters vote for J with a probability higher than $1/3$. To provide a sharper characterization, we focus on the x-model for the case of large x. In particular, we show that for x large all equilibria in simple strategies must involve sequential voting (and, hence, do not involve persistent strategies). After proving the result, we provide an example of a natural class of equilibria that do not exist for small x.

THEOREM 2: *Consider the x-model. If x is sufficiently close to one, there does not exist a pure-strategy sequential equilibrium with simple voting-stage strategies in which all the voters vote in the same period.*

Before turning to the proof of this theorem, we state a corollary.

COROLLARY 1: *Consider the x-model. If x is sufficiently close to one, there exists an* $\varepsilon < 1$, *such that in any (mixed strategy) sequential equilibrium with simple voting-stage strategies, the probability of all players voting in the same period is bounded above by* ε.

In particular if x is large then there is an $\varepsilon' < 1$ that bounds from above the probability of any player choosing to move in either period in all symmetric equilibria. The corollary follows because otherwise there is a sequence of simple mixed-strategy sequential equilibria whose limit involves all players playing in the same period, and one can see that this limit is inconsistent with Theorem 2. (The corollary can also be proven directly along the lines of the proof of Theorem 2 below, but instead of using the special case of Remark 2 used in the proof, using the full strength of Lemma 1 and considering all possible timing-stage outcomes.)

To prove Theorem 2, we make use of the following result. Let $r_C(n_A, n_B)$ be the probability of victory of C conditional upon the information that n_A voters have voted for A, n_B voters for B, $N^1 - n_A - n_B$ for C, and the remaining voters vote for candidate J with probability e_J. The next lemma states that if $e_A \geq e_B$, then—conditioning upon candidates having at least two more votes for A than for B—changing the vote for B into a vote for A weakly decreases the probability of victory of C, and if C's probability of winning was interior, then it strictly decreases that probability.

LEMMA 1: *Suppose that* $e_A \geq e_B$. *Then,*

(*i*) *If* $n_A - 1 > n_B$ *then* $r_C(n_A, n_B) \leq r_C(n_A - 1, n_B + 1)$

(*ii*) *If* $n_A - 1 > n_B$ *and* $0 < r_C(n_A, n_B) < 1$ *then* $r_C(n_A, n_B) < r_C(n_A - 1, n_B + 1)$.

REMARK 2: *In particular,* $r_C(2, 0) < r_C(1, 1)$, *i.e., if there is one vote each for A and B, then switching from B to A decreases the likelihood of C winning.*

PROOF:

See Appendix.

PROOF OF THEOREM 2:

If all voters vote in the same period, simple voting-stage strategies vote for the highest valued candidate. Hence, there does not exist an equilibrium in which all voters vote in the first period, since one voter would be better off moving to the second period. (To see this consider the event that votes are evenly split between one's worst and middle candidate, with either zero or one vote for one's favorite. Then, by waiting, one increases expected utility, as instead of wasting one's vote on one's favorite candidate, one increases from $1/2$ to 1 the likelihood of one's second favorite rather than one's worst candidate succeeding.) We need to show that if they all vote in the second period, one voter will move to the first period. Consider for simplicity the realization of votes $(n_A^1, n_B^1, n_C^1) = (1, 0, 0)$. There are four simple strategy profiles that might be an equilibrium in the second-period of the voting stage:

- Voters may vote for either B or C and cease voting for A. For this type of strategies and x close to one, it is optimal for a voter to switch to the first period and vote for the lowest valued candidate as, for $N \geq 4$, the probability of victory of this candidate drops to zero.
- Voters may choose persistent voting strategies. For x sufficiently close to one, a voter maximizes the expected payoff by minimizing the probability of victory of the candidate with the lowest valuation. Then, after an A vote in the first period, by Lemma 1, and in particular by Remark 2, it is a strict best reply to vote for A when valued x (assuming everyone else votes for their favorite candidate). Hence, persistent strategies are not an equilibrium.
- Third, voters with preferences $(1, x, 0)$, $(x, 1, 0)$, $(1, 0, x)$, and $(x, 0, 1)$ vote for A and voters with the remaining preferences vote for the highest valued candidate. (It is clear that that by Lemma 1 this profile of strategies is an equilibrium for x close to one, but this is not needed for our conclusion.) In this case, the probability that the second-period voters vote for A after the realization of votes $(n_A^1, n_B^1, n_C^1) = (1, 0, 0)$ is higher than with $(n_A^1, n_B^1, n_C^1) = (0, 0, 0)$. Hence, a voter with preferences $(1, x, 0)$ prefers to move to the first period and vote for A.
- All voters in the second period vote for A. Obviously, in this case a voter prefers to vote in the first period for the candidate whom he values most.

The intuition behind this theorem is as follows. When only one voter votes in the first period, if x is close to one, there does not exist an equilibrium in which voters in the second period use persistent voting strategies. Thus, the result from the preceding section does not apply. If the voters coordinate in equilibrium by ceasing to vote for the candidate receiving the first period vote, then it is optimal for one voter to move to the first period and vote for the option with zero value. If voters vote for the leading candidate with probability higher than $1/3$, it is optimal for one voter to move to the first period and vote for the option with the highest value. Thus, at least one voter moves to the first period.

Arguments similar to those used in the proof of Lemma 1 establish that, when x is close to one, following some first-period outcomes, it is not optimal to vote for the most preferred candidate in the second period. Thus, voting behavior is not equivalent to simultaneous voting.

In the x-model with x large, existence of an equilibrium with simple second-stage strategies follows from the example in the first paragraph of Section IV. The existence of equilibria that do not involve the perverse coordination that appears there follows from Proposition 1 and Remark 1 using the following profile of simple second-period strategies. These equilibria are an example of what might be called satisfying voting strategies: voting for one's favorite candidate so long as the second-best candidate has fewer votes than the favorite, and voting for the second-best otherwise. That for x large these strategies are an equilibrium for the second period of the voting game given any outcome of the timing game and any outcome of first-period voting can be shown using Lemma 2. Moreover, it is easy to see that these strategies cannot be an equilibrium for small x, as voting for the second-best candidate as required in part 2 below is not optimal for small enough x.

EXAMPLE 2:

(*i*) *If, in the first period, candidate J receives a number of votes strictly less than the other two candidates, J′ and J″, voters in the second period cease voting for J and vote for J′ or J″ with the higher valuation.*

(*ii*) *If, in the first period candidate, J is leading and the other two candidates, J′ and J″, have the same number of votes, a voter in the second period votes for J if J is valued one, for J if J is valued x and at least one other voter votes in the second period, and for J′* or *J″* with the higher valuation in the remaining cases.

(*iii*) *If, in the first period, all the candidates have the same number of votes, voters in the second period vote for the candidate with the highest valuation.*

VI. Conclusions and Extensions

The main conclusion of this paper is that for sequential voting to arise endogenously, later voters must shift votes in favor of the second-best candidates when the probability of victory of their favorite candidate is not negligible. If voters desist from voting for their favorite candidate only when her probability of victory is very small, voting early has a strategic *dis*advantage as it decreases the probability of victory of one's favorite candidate. Although our results were proved for the case of three candidates, the intuitions behind them carry over quite naturally to the case of more than three candidates. While we have not verified the formal details, we expect the above results to hold in the latter case.

Theorem 1 extends to the case of a finite but arbitrary number of periods. The proof follows from minor modifications of the arguments available in this paper and is not provided. However, when voting has more than two periods, the strategic environment is complex and can involve strategies that, at least at first glance, appear nonintuitive. For example, strategic voting for the second-best candidate can occur in some periods even when one's favorite is leading. Consider the x-model and suppose that voting takes place in three periods. Assume that $x = 1 - \varepsilon$ and that in period 1 the votes for A, B, and C are n_A^1, $n_A^1 - 1$, and $n_A^1 - 1$. Also suppose that one voter votes in period 2 and two voters vote in period 3, and that the preferences of the period 2 voter are $(1, 1 - \varepsilon, 0)$. If in period 2 the voter votes for A, the distribution of votes in period 3 is $n_A^1 + 1, n_A^1 - 1, n_A^1 - 1$, whereas voting for B yields $n_A^1, n_A^1, n_A^1 - 1$. In the latter case, the subsequent voters will vote for their best between A and B to minimize their probability of getting 0, thereby giving the period 2 voter a payoff close to 1. If the distribution of votes in period 3 is $n_A^1 + 1, n_A^1 - 1, n_A^1 - 1$ and both period 3 voters have preferences $(0, 1 - \varepsilon, 1)$ (the probability of this event is $1/36$), the period 2 voter obtains a payoff equal to 0 with probability $1/2$. Note the anti-herding feature in this example: an increase in the number of votes for B can lead a voter to vote for A instead.

Strategic voting can even induce voters to vote for their *least*-preferred candidate. Consider again the x-model for x small and that voting takes three periods. The realization of votes in the first period is 1 vote for A, $N/2 - 1$ votes for C, and none for B. There is only one voter in period 2 and he has preferences $(1, x, 0)$. If he votes for A or B, then B is still viable and, if x is small, it is an equilibrium for voters in period 3 to vote for the favorite candidate. Voting for C makes B not viable and, hence, voters will switch to the preferred candidate between A and C. This can increase the probability of A winning if N is large.

Our results raise some interesting questions for the timing of voting of states of different sizes. On the one hand, a large state alone could destabilize simultaneous voting by early voting, whereas a small state alone may be unable to affect the behavior of later voters. On the other hand, a truly large state may have lower incentives than small states to vote early in the election. A large state could have a significant influence in determining the choices of later states, but, as we have seen in Theorem 1, this can be detrimental to the probability of success of a large state's preferred candidate. A specific example demonstrating this possibility is available in the working paper version of this paper. We leave studying the case of asymmetric voters for future work.

Appendix

PROOF OF THEOREM 1:

Clearly, it is an equilibrium for all voters to vote in the last period. (One voter moving earlier cannot make any candidate unviable, hence, there is no benefit to moving earlier.) Consider an equilibrium in which voting takes place in two periods and that is not equivalent to simultaneous voting. For simplicity in notation, we will assume that if two candidates cease to be viable in period 1, voters vote for their highest valued candidate in period 2. As the outcome is thus determined in period 1, this assumption is inessential.

Consider a voter i who votes in period 1. Without loss of generality, suppose that A is the candidate having the highest valuation for i. We will first show that, for any N, i's payoff cannot decrease if he decides to vote for A in period 2. This will imply that he is at least as well off if he votes in period 2 for the best viable candidate. We will then show that, for $N \geq 6$, i's payoff must increase if he decides to vote for A in period 2.

Given a fixed number of period 1 voters, consider the events for which, if voter i deviates and votes in period 2, the voting of the other voters is affected. To define such pivotal events, let $\hat{n}_J^1$, $J = A, B, C$ denote the number of votes received by candidate J in period 1 excluding the vote of player i. Define the inequalities

$$N - \hat{n}_A^1 - \hat{n}_J^1 - 1 < \max\{\hat{n}_A^1 + 1, \hat{n}_J^1\} \tag{A1}$$

(A2) $$N - \hat{n}_A^1 - \hat{n}_J^1 \geq \max\{\hat{n}_A^1, \hat{n}_J^1\},$$

where $J \in \{B, C\}$. If (A1) and (A2) hold for $J = B$ and both A and B are viable, then period 2 voters, whose most valuable candidate is C, vote for their second-most valuable candidate (A or B) when voter i votes for A in the first period (as then C is not viable by (A1)), and vote for C when voter i votes for A in the second period (as then C is viable by (A2)). A symmetric explanation holds for $J = C$.

Take $\gamma \leq \delta - 1$ such that, for $\hat{n}_A^1 = \gamma$ and $\hat{n}_B^1 = \delta$, (A1) and (A2) hold, and both A and B are viable. Define the events

$$\mathcal{E}_1(k) = \{\hat{n}_A^1 = \gamma,\ \hat{n}_B^1 = \delta,\ k \text{ voters vote for } C \text{ in period } 2\}$$

$$\mathcal{E}_2(k) = \{\hat{n}_A^1 = \delta - 1,\ \hat{n}_B^1 = \gamma + 1,\ k \text{ voters vote for } C \text{ in period } 2\}$$

and define $\mathcal{E}(k) = \mathcal{E}_1(k) \cup \mathcal{E}_2(k)$. The following elementary fact is stated without proof.

FACT: *If* (A1) *and* (A2) *hold and A and J are viable for* $\hat{n}_A^1 = \gamma$ *and* $\hat{n}_J^1 = \delta$, *where* $\gamma \leq \delta - 1$, *then* (A1) *and* (A2) *hold and A and J are viable for* $\hat{n}_A^1 = \delta - 1$ *and* $\hat{n}_J^1 = \gamma + 1$.

We will now show that the probability of A winning conditional on $\mathcal{E}(k)$ does not decrease with k. To do so, we will replace one second-period vote for C with a vote for A or B with probability 0.5 each, and show that the probability of A winning cannot increase. In particular, we will show that, conditional on $\mathcal{E}_1(k)$, the effect of decreasing k is to weakly increase the probability that A wins. However, for the corresponding event $\mathcal{E}_2(k)$, decreasing k weakly decreases the probability that A wins by an offsetting amount. Since $\mathcal{E}_2(k)$ is at least as likely as $\mathcal{E}_1(k)$, the overall effect conditional on $\mathcal{E}(k)$ of decreasing k is to weakly decrease the probability that A wins. We will then show that for all remaining events with $\hat{n}_A^1 \geq \hat{n}_B^1$, the probability of A winning decreases as k decreases.

Let $\tilde{n}_A$ denote the total number of votes for candidate A in period 2 when the vote of voter i is not included, and N^1 the total number of voters in period 1 including voter i. First note that conditional on $\mathcal{E}(k)$ and i voting for A in either the first or the second period, A wins with probability 1 if

(A3) $$\tilde{n}_A > \frac{N - 1 - k - \hat{n}_A^1 - (N^1 - \hat{n}_B^1)}{2},$$

and wins with probability $1/2$ if

(A4) $$\tilde{n}_A = \frac{N - 1 - k - \hat{n}_A^1 - (N^1 - \hat{n}_B^1)}{2}.$$

First, suppose that $N - 1 - k - \hat{n}_A^1 - (N^1 - \hat{n}_B^1)$ is even. Note that if it is even for one component event in $\mathcal{E}(k)$, it is even for the other component as well. If k is decreased by one, the only N-tuples whose outcome is affected are those in (A4). In this case, the probability of A winning is unchanged when k is decreased by one since the voter who ceases voting for C votes for A with probability $1/2$ and B with probability $1/2$.

Now consider the case of odd $N - 1 - k - \hat{n}_A^1 - (N^1 - \hat{n}_B^1)$ and suppose that one of the voters who votes for C in period 2 switches to A or B with probability $1/2$ each. To evaluate the change in the probability of victory of A, in $\mathcal{E}_1(k)$ we need to consider N-tuples for which

$$\tilde{n}_A = \frac{N - N^1 - k + (\delta - \gamma)}{2} \quad (WL1)$$

$$\tilde{n}_A = \frac{N - N^1 - k + (\delta - \gamma)}{2} - 1 \quad (LW1)$$

and, in $\mathcal{E}_2(k)$,

$$\tilde{n}_A = \frac{N - N^1 - k - (\delta - \gamma)}{2} + 1 \quad (WL2)$$

$$\tilde{n}_A = \frac{N - N^1 - k - (\delta - \gamma)}{2} \quad (LW2).$$

For the events satisfying $WL1$ and $WL2$, if one of the k voters for C switches to voting for A with probability $1/2$ and B with probability $1/2$, the probability of A winning decreases from 1 to $3/4$. For the events satisfying $LW1$ and $LW2$, the probability of A winning increases from 0 to $1/4$.

Since the distribution of $\tilde{n}_A$ conditional upon $\mathcal{E}_1(k)$ or $\mathcal{E}_2(k)$ is binomial and symmetric around $(N - N^1 - k)/2$, we have that

$$\text{(A5)} \qquad \Pr(WL1 \,|\, \mathcal{E}(k)) = \Pr(LW2 \,|\, \mathcal{E}_2(k))$$

$$\text{(A6)} \qquad \Pr(WL2 \,|\, \mathcal{E}(k)) = \Pr(LW1 \,|\, \mathcal{E}_1(k)).$$

By the same token,

$$\text{(A7)} \qquad \Pr(LW1 \,|\, \mathcal{E}_1(k)) \geq \Pr(WL1 \,|\, \mathcal{E}_1(k)).$$

Since the symmetry of the binomial distribution also implies that the probability of $\mathcal{E}_1(k)$ cannot exceed the probability of $\mathcal{E}_2(k)$, it can be easily verified that the probability of A winning cannot increase if one of the k second period C-voters switches to A or B with probability $1/2$ each.

To see this note that the change in the probability of A winning after 1 of the k second-period voters switches to A or B with probability $1/2$ is equal to

$$\frac{1}{4}\Big[\Pr\big(LW1\,|\,\mathcal{E}_1(k)\big) - \Pr\big(WL1\,|\,\mathcal{E}_1(k)\big)\Big]\Pr\big(\mathcal{E}_1(k)\big)$$
$$+\ \frac{1}{4}\Big[\Pr\big(LW2\,|\,\mathcal{E}_2(k)\big) - \Pr\big(WL2\,|\,\mathcal{E}_2(k)\big)\Big]\Pr\big(\mathcal{E}_2(k)\big).$$

Given equations (A5), (A6), and (A7) we have

$$\Pr\big(LW1\,|\,\mathcal{E}_1(k)\big) - \Pr\big(WL1\,|\,\mathcal{E}_1(k)\big)$$
$$= -\Big[\Pr\big(LW2\,|\,\mathcal{E}_2(k)\big) - \Pr\big(WL2\,|\,\mathcal{E}_2(k)\big)\Big] > 0,$$

and, since $\Pr\big(\mathcal{E}_1(k)\big) \leq \Pr\big(\mathcal{E}_2(k)\big)$, the result follows.

In view of the Fact, to conclude that i's payoff never decreases when he votes for his highest-valued candidate in period 2, we need to consider first period realizations for which $\hat{n}_A^1 \geq \hat{n}_B^1$. In particular, we want to show that, conditional upon the event $\{\hat{n}_A^1 \geq \hat{n}_B^1$, k voters vote for C in period 2$\}$, the probability of A winning the election decreases with k.

As before, we consider two cases. If $N - 1 - k - \hat{n}_A^1 - (N^1 - \hat{n}_B^1)$ is even and k decreases by 1, the N-tuples whose outcome is affected are those in (A4). As before, if a voter who ceases voting for C votes for A with probability $1/2$, the probability of A winning is unchanged.

If $N - 1 - k - \hat{n}_A^1 - (N^1 - \hat{n}_B^1)$ is odd, decreasing k by 1 changes the outcome of the election only when

$$\tilde{n}_A = \frac{N - k - \hat{n}_A^1 - (N^1 - \hat{n}_B^1)}{2}$$

or

$$\tilde{n}_A = \frac{N - k - \hat{n}_A^1 - (N^1 - \hat{n}_B^1)}{2} - 1.$$

Since $\hat{n}_A^1 \geq \hat{n}_B^1$, the event defined by the first equality is more likely. For an N-tuple satisfying the first equality, decreasing k decreases the probability of A winning from 1 to $3/4$. For an N-tuple satisfying the second equality, decreasing k increases the probability of A winning from 0 to $1/4$. Hence, decreasing k cannot increase the probability of A winning. Since conditional upon i voting for A, the probabilities of B or C winning the election are identical, voter i is not worse off voting for A in period 2.

We conclude the proof of the theorem showing that, if $N \geq 6$, voter i is strictly better off voting in period 2 for the best viable candidate. First, suppose that N is even. If $N^1 \geq N/2 + 2$, consider the event $\{\hat{n}_B^1 = N/2, \hat{n}_C^1 = N^1 - N/2 - 1\}$. Conditional upon this event, if voter i votes for A in period 1, B wins the election with probability 1, whereas, if the second best is C, voting for C in period 2 gives C a positive probability of victory. If $N^1 = N/2 + 1$, (A1) and (A2) are satisfied for $\hat{n}_A^1 = 0$ and $\hat{n}_B^1 = N/2$. This corresponds to the case where $\gamma = 0$ and $\delta = N/2$ in the definition of $\mathcal{E}(k)$. For $k = 1$,

$$\Pr(WL1 \mid \mathcal{E}_1(1)) - \Pr(LW1 \mid \mathcal{E}_1(1)) = -\frac{1}{2}\Pr\Big(\tilde{n}_A = \frac{N}{2} - 2\Big),$$

since $\Pr(WL1 \mid \mathcal{E}_1(1)) = 0$ and $\Pr(LW1 \mid \mathcal{E}_1(1)) = \Pr(\tilde{n}_A = N/2 - 2 \mid \mathcal{E}_1(1)) = (1/2)\Pr(\tilde{n}_A = N/2 - 2)$.

$N \geq 6$ implies that $\Pr(\tilde{n}_A = N/2 - 2) > 0$ and, since the probability of $\mathcal{E}_1(1)$ is strictly smaller than the probability of $\mathcal{E}_2(1)$, voter i is strictly better off voting in period 2.

Now suppose that N is odd. If $N^1 \geq (N+5)/2$, consider the event $\{\hat{n}_B^1 = (N-1)/2, \hat{n}_C^1 = N^1 - (N-1)/2 - 1\}$. Conditional upon this event, A cannot win the election and, if voter i's second best is B, voting for B in period 2 gives B the certainty of victory, whereas voting for A in period 1 gives C a positive chance of winning the election. If $N^1 = (N+3)/2$, (A1) and (A2) are satisfied for $\hat{n}_A^1 = 1$ and $\hat{n}_B^1 = (N-1)/2$. This corresponds to the case where $\gamma = 1$ and $\delta = (N-1)/2$ in the definition of $\mathcal{E}(k)$. For $k = 2$,

$$\Pr(WL1 \mid \mathcal{E}_1(k)) - \Pr(LW1 \mid \mathcal{E}_1(k)) = -\frac{1}{4}\Pr\Big(\tilde{n}_A = \frac{N-5}{2} - 1\Big),$$

since $\Pr(WL1 \mid \mathcal{E}_1(2)) = 0$ and $\Pr(LW1 \mid \mathcal{E}_1(2)) = \Pr(\tilde{n}_A = (N-5)/2 - 1 \mid \mathcal{E}_1(2))$ which, in turn, equals $\Pr(\tilde{n}_A = (N-5)/2 - 1)/4$.

When $k = 2$, $\Pr(\tilde{n}_A = (N-5)/2 - 1) > 0$. But k can be equal to 2 if there are at least two voters (excluding i) in period 2, that is, $N \geq 7$.

PROOF OF LEMMA 1:

We first state and prove an intermediate lemma.

LEMMA 2: *For nonnegative integers* μ_1, μ_2 *and* $\bar{\mu}$, *if* $\mu_2 \geq \mu_1$, $\mu_2 + \mu_1 \geq \bar{\mu} + 1$, *and for e such that* $1/2 \leq e \leq 1$,

$$\sum_{\mu=\mu_1}^{\mu_2} (1-e)^{\mu} e^{\bar{\mu}-\mu} \binom{\bar{\mu}}{\mu} < \sum_{\mu=\mu_1-1}^{\mu_2-1} (1-e)^{\mu} e^{\bar{\mu}-\mu} \binom{\bar{\mu}}{\mu}.$$

PROOF OF LEMMA 2:

The difference of the left- and right-hand sides simplifies to

$$-(1-e)^{\mu_1-1} e^{\bar{\mu}-\mu_1+1} \binom{\bar{\mu}}{\mu_1 - 1} + (1-e)^{\mu_2} e^{\bar{\mu}-\mu_2} \binom{\bar{\mu}}{\mu_2}.$$

Note that

$$\frac{(1-e)^{\mu_1-1}e^{\bar{\mu}-\mu_1+1}}{(1-e)^{\mu_2}e^{\bar{\mu}-\mu_2}} = \left(\frac{e}{1-e}\right)^{-\mu_1+1+\mu_2} \geq 1.$$

If $\mu_1 \geq \bar{\mu}/2$, the claim is obviously true as $\mu_2 \geq \mu_1$ because $\binom{\bar{\mu}}{l}$ is decreasing in l for $l \geq \bar{\mu}/2$. If $\mu_1 < \bar{\mu}/2$, then $\bar{\mu} - (\mu_1 - 1) > \bar{\mu}/2$, and, since $\mu_2 > \bar{\mu} - (\mu_1 - 1)$ by assumption and $\binom{\bar{\mu}}{\bar{\mu}-(\mu_1-1)} = \binom{\bar{\mu}}{\mu_1 - 1}$, the claim follows using the same argument.

To prove Lemma 1, define $q_C(\underline{n}_B, n_C)$ to be the (probability) that C wins conditional upon exactly n_C voters voting for C, at least $\underline{n}_B$ for B, and at least $M - \underline{n}_B$ for A, where $0 \leq \underline{n}_B < M/2 - 1$.

Obviously, if $q_C(\underline{n}_B, n_C) = 1$, then $q_C(\underline{n}_B + 1, n_C) = 1$ for $0 \leq \underline{n}_B < M/2 - 1$.

Now we will show that, if $0 < q_C(\underline{n}_B - 1, n_C) < 1$, then $q_C(\underline{n}_B, n_C) > q_C(\underline{n}_B - 1, n_C)$ for $1 \leq \underline{n}_B \leq M/2$. For this purpose, define $e = e_A/(e_A + e_B)$. The proof proceeds by showing this for various cases of n_C and M, and concludes by taking expectations over n_C.

First suppose that $n_C = N/3$. Note that if $0 < q_C(\underline{n}_B - 1, n_C) < 1$, then $N/3 \geq M - \underline{n}_B$, as the latter is the number of votes for A, so otherwise C looses to $q_C = 1$. Now observe that

$$q_C\left(\underline{n}_B, \frac{N}{3}\right) = \frac{1}{3}e^{\frac{N}{3}-M+\underline{n}_B}(1-e)^{\frac{N}{3}-\underline{n}_B}\binom{\frac{2N}{3}-M}{\frac{N}{3}-\underline{n}_B}.$$

Straightforward algebra shows that this expression is increasing in $\underline{n}_B$ for $0 \leq \underline{n}_B \leq M/2$.

Now consider the case that $n_C = N/2$. Then $q_C(0, n_C) < 1$ and $q_C(1, n_C) = 1$.

Next, if $n_C = (N-1)/2$ and $M = 2$, then

$$q_C(1, n_C) = 1 - \frac{1}{2}e^{\frac{N-3}{2}} - \frac{1}{2}(1-e)^{\frac{N-3}{2}}$$

$$q_C(0, n_C) = 1 - e^{\frac{N-3}{2}} - \frac{1}{2}\cdot\frac{N-3}{2}e^{\frac{N-5}{2}}(1-e).$$

As $N \geq 5$, it is easy to see that $q_C(0, n_C) < q_C(1, n_C)$.

One more special case to consider is when $n_C = (N-1)/2$ and $M > 2$. Then,

$$q_C(1, n_C) = 1 - \frac{1}{2}e^{\frac{N+1}{2}-M}$$

$$q_C(0, n_C) = 1 - e^{\frac{N+1}{2}-M} - \frac{1}{2}\cdot\left(\frac{N+1}{2} - M\right)e^{\frac{N-1}{2}-M}(1-e).$$

Again, it is straightforward to verify that, as $N \geq 5$, $q_C(0, n_C) < q_C(1, n_C)$. Since $q_C(2, n_C) = 1$, the claim again holds.

Finally suppose that $N/3 < n_C \leq (N-2)/2$. Then,

$$q_C(\underline{n}_B, n_C) = \sum_{l=N-2n_C-\underline{n}_B}^{n_C-\underline{n}_B} \alpha_l \beta_l \binom{N-M-n_C}{l},$$

where $\beta_l = e^{N-M-n_C-l}(1-e)^l$, $\alpha_l = 1/2$ for $l = N - 2n_C - \underline{n}_B$ or $l = n_C - \underline{n}_B$, and $\alpha_l = 1$ for $N - 2n_C - \underline{n}_B < l < n_C - \underline{n}_B$. (The parameter l denotes the number of votes for A.)

Lemma 2 implies that

$$\sum_{l=N-2n_C-\underline{n}_B}^{n_C-\underline{n}_B-1} \beta_l \binom{N-M-n_C}{l} < \sum_{l=N-2n_C-\underline{n}_B-1}^{n_C-\underline{n}_B-2} \beta_l \binom{N-M-n_C}{l}$$

and

$$\sum_{l=N-2n_C+1-\underline{n}_B}^{n_C-\underline{n}_B} \beta_l \binom{N-M-n_C}{l} < \sum_{l=N-2n_C-\underline{n}_B}^{n_C-\underline{n}_B-1} \beta_l \binom{N-M-n_C}{l}.$$

(The conditions on the lower and upper bounds of the summations required for Lemma 2 follow since $\underline{n}_B < M/2 - 1$.) Since

$$2q_C(\underline{n}_B, n_C) = \sum_{l=N-2n_C-\underline{n}_B}^{n_C-\underline{n}_B-1} \beta_l \binom{N-M-n_C}{l} + \sum_{l=N-2n_C-\underline{n}_B+1}^{n_C-\underline{n}_B} \beta_l \binom{N-M-n_C}{l},$$

it is easy to verify that the above inequalities imply that $q_C(\underline{n}_B, n_C) - q_C(\underline{n}_B + 1, n_C) < 0$. The proof is then completed by taking expectations over n_C.

REFERENCES

Ali, S. Nageeb, and Navin Kartik. 2012. "Herding with Collective Preferences." *Economic Theory* 51 (3): 601–26.

Bag, Parimal K., Hamid Sabourian, and Eyal Winter. 2009. "Multi-Stage Voting, Sequential Elimination and Condorcet Consistency." *Journal of Economic Theory* 144 (3): 1278–99.

Battaglini, Marco. 2005. "Sequential Voting with Abstention." *Games and Economic Behavior* 51 (2): 445–63.

Battaglini, Marco, Rebecca Morton, and Thomas Palfrey. 2007. "Efficiency, Equity, and Timing of Voting Mechanisms." *American Political Science Review* 101 (3): 409–24.

Callander, Steven. 2007. "Bandwagons and Momentum in Sequential Voting." *Review of Economic Studies* 74 (3): 653–84.

Dekel, Eddie, and Michele Piccione. 2000. "Sequential Voting Procedures in Symmetric Binary Elections." *Journal of Political Economy* 108 (1): 34–55.

Deltas, George, Helios Herrera, and Mattias K. Polborn. 2010. "A Theory of Learning and Coordination in the Presidential Primary System." http://tintin.hec.ca/pages/helios.herrera/primaries.pdf.

Fey, Mark. 1996. "Informational Cascades, Sequential Elections, and Presidential Primaries." Paper presented at the Annual Meeting of the Midwest Political Science Association, Chicago, April 18–20.

Hummel, Patrick. 2012. "Sequential Voting in Large Elections with Multiple Candidates." *Journal of Public Economics* 96 (3–4): 341–48.

Knight, Brian, and Nathan Schiff. 2010. "Momentum and Social Learning in Presidential Primaries." *Journal of Political Economy* 118 (6): 1110–50.

Morton, Rebecca B., and Kenneth C. Williams. 1999. "Information Asymmetries and Simultaneous versus Sequential Voting." *American Political Science Review* 93 (1): 51–67.

Sloth, Birgitte. 1993. "The Theory of Voting and Equilibria in Noncooperative Games." *Games and Economic Behavior* 5 (1): 152–69.

Wit, Jorgen. 1997. "Herding Behavior in a Roll-Call Voting Game." Unpublished.

This article has been cited by:

1. Marina Agranov. 2016. Flip-Flopping, Primary Visibility, and the Selection of Candidates. *American Economic Journal: Microeconomics* **8**:2, 61-85. [Abstract] [View PDF article] [PDF with links]