Primaries

Current U.S. presidential nomination process consists of a series of elections (primaries) held over several months.

Sequential nature of primaries has produced presidential nominees most experts feel would not have won nomination with simultaneous voting. (Carter 1976, Kerry 2004)

Knight & Schiff (2007) estimate that voters in early primary states have up to 20 times more influence in the selection of candidates than later voters.

Despite significance, the current primary system has developed through tradition, political maneuvering, and sheer accident rather than central party planning.
Motivation

- There is evidence that political parties wish to regulate the general structure of the primary calendar.
  - Both parties moved to prevent Florida and Michigan from holding early primaries in 2008.
  - In same year, Nevada and South Carolina were encouraged to hold early primaries to increase minority influence.
- Despite such minor interventions, primary calendar has never been fully controlled by the national parties.
- **Question:** *If a party could make all decisions regarding primary dates centrally, what sequence of primaries should it use to select its nominee?*
Introduction

Objective of this paper

- I focus on the problem faced by a political party which wants to choose the optimal voting sequence to select its presidential nominee.

- Assuming that each primary contains a candidate which is of the highest quality (meaning best for the party—best chance to win general election and serve party interests), party wishes to choose sequence of primaries most likely to select that candidate.

- I consider a social learning environment in which a sequential election may generate voter herding.

- When and how can the party benefit from voter herding to maximize the probability of selecting the highest quality candidate?
This paper *normatively* frames an isolated party’s problem of choosing a temporal structure for its primaries.

Party can control momentum effects and guarantee herding will be ex ante beneficial by choosing:

- whether to hold primaries sequentially
- and if so, the actual sequence in which states vote
Related literature

- Models of social learning and information cascades

- Temporal structure of elections

- Momentum in presidential primaries
  - Klumpp & Polborn (2006): normative model of primaries (models only candidate campaigning; abstracts from voters).
Primary election with two candidates \((c = A, B)\) from the same political party.

Nature determines which candidate \(c^{HQ} \in \{A, B\}\) is of highest quality, each with probability 1/2.

The election is contested in states \(s \in S = \{1, \ldots, S\}\).

Each state \(s\) holds an election which determines \(d_s \in \mathbb{N}\) delegates (by proportional representation).

Candidate with majority of all delegates wins party nomination.
Voter preferences

- Voting is costless for all voters.
- There are three types of voters in each state:
  - *loyal voters* of candidate A
  - *loyal voters* of candidate B
  - *uncommitted voters*
Voter preferences

- *Loyal voters* are characterized by a rigid preference for one candidate and always vote for that candidate.

- *Uncommitted voters* get utility one when voting for $c^{HQ}$ and zero otherwise.
  - However, the true identity of $c^{HQ}$ is never observed.
  - Thus *expected* utility of uncommitted voters is analyzed.
Information structure

- An uncommitted voter $i$ in state $s$ receives a private noisy signal $\theta_{i,s} \in \{A, B\}$ regarding which candidate is highest quality, where

$$\Pr\left\{ \theta_{i,s} = c^{HQ} \right\} = q_s > 1/2$$

- Elections in different states may be held on different dates. Let $\tau(s) \in \{1, \ldots, T\}$ denote the date on which state $s$ votes.
- When an election is held in state $s$, voters in all future states observe the number of votes cast for each candidate in that election, $(v_s^A, v_s^B)$.
- Those voting on date $t$ do so having observed the voting history

$$\mathcal{H}_t = \left\{ \left\{ v_{s'}^A, v_{s'}^B \right\}_{s'} : \tau(s') < t \right\}.$$
Uncommitted voters’ expected utility

- Uncommitted voters are Bayesian.
- Expected utility of an uncommitted voter voting for candidate $c$ on date $t$ is the posterior probability (as of date $t$) that $c = c^{HQ}$.
- This probability is conditional on private signal $\theta_{i,s}$, private signal quality $q_s$, and observed voting history $\mathcal{H}_{\tau(s)}$.
- Formally,

$$U_{i,s}(c) = \Pr \left\{ c = c^{HQ} \mid \theta_{i,s}, q_s, \mathcal{H}_{\tau(s)} \right\}.$$

- The uncommitted voter chooses $c \in \{A, B\}$ to maximize $U_{i,s}(c)$. 

Composition of electorate

- The composition of the electorate is determined stochastically.
- Voters (and the party) observe only the distribution from which this composition is drawn.
- Each voter in state \( s \) receives a preference shock so that she is a:
  - loyal voter of candidate \( A \), with prob. \( p_s^A \)
  - loyal voter of candidate \( B \), with prob. \( p_s^B \)
  - uncommitted voter, with prob. \( 1 - p_s^A - p_s^B \).
Political party’s problem

- The political party is faced with choosing an *optimal temporal structure* for its primaries.
- Structure cannot be specific to a particular election cycle or candidates.
  - Party does not know true identity of $c^{HQ}$. It has the same (uniform) prior for $c^{HQ}$ as uncommitted voters.
  - Party cannot observe the exact number of loyal voters and uncommitted voters in the electorate.
- The party’s objective is to choose the ordered partition of the set of states which maximizes the probability of nominating the highest quality candidate.
Bayesian updating of uncommitted voters

- On date $t = 1$, uncommitted voters have only private information on which to base their vote. They hence vote in accordance with private signal $\theta_{i,s}$.
- We introduce the notation
  \[
  \gamma_s^{c_{HQ}} \equiv \Pr\{\text{random state } s \text{ voter votes for } c_{HQ}\} \\
  = p_s^{c_{HQ}} + \left(1 - p_s^A - p_s^B\right)q_s
  \]
- The distribution of votes for candidate $c_{HQ}$ in state $s$ in the first period is a binomial with $n_s$ trials and probability $\gamma_s^{c_{HQ}}$:
  \[
  \nu_s^{c_{HQ}} \sim \text{Binomial}\left(n_s, \gamma_s^{c_{HQ}}\right).
  \]
Bayesian updating of uncommitted voters

- Uncommitted voters use this fact to calculate their expected utility $U_{i,s}(c) = \Pr \left\{ c = c^{HQ} \mid \theta_{i,s}, q_s, \mathcal{H}_{\tau(s)} \right\}$, which yields the following result:

Lemma

If $n_1$ voters all characterized by $(p^A_1, p^B_1, q_1)$ vote in the first period, with $\nu^A_1$ of them voting for candidate A, then the optimal voting decision for a second period uncommitted voter with private signal $\theta_{i,2}$ and signal quality $q_2$ is given by

$$v^*_i, 2 = \begin{cases} 
A & \text{for } \nu^A_1 > \Omega^A(n_1) \\
\theta_{i,2} & \text{for } \Omega^B(n_1) \leq \nu^A_1 \leq \Omega^A(n_1) \\
B & \text{for } \nu^A_1 < \Omega^B(n_1),
\end{cases}$$
Lemma

where

\[
\Omega^A(n_1) \equiv \Omega^A(n_1; q_2) = \frac{n_1 \ln \left( \frac{\gamma_1^B}{1-\gamma_1^A} \right) + \ln \left( \frac{q_2}{1-q_2} \right)}{\ln \left( \frac{\gamma_1^A}{1-\gamma_1^B} \right) + \ln \left( \frac{\gamma_1^B}{1-\gamma_1^A} \right)},
\]

\[
\Omega^B(n_1) \equiv \Omega^B(n_1; q_2) = \frac{n_1 \ln \left( \frac{\gamma_1^B}{1-\gamma_1^A} \right) - \ln \left( \frac{q_2}{1-q_2} \right)}{\ln \left( \frac{\gamma_1^A}{1-\gamma_1^B} \right) + \ln \left( \frac{\gamma_1^B}{1-\gamma_1^A} \right)} < \Omega^A(n_1).
\]

and \( \gamma_s^c \equiv p_s^c + (1 - p_s^A - p_s^B) q_s \).
When $p_1^A = p_1^B$
When $p_1^A = p_1^B$

(increasing $q_{s'}$)
When $p_1^A > p_1^B$

Figure 3
When $p_1^A > p_1^B$

(increasing $q_{s'}$)
Simultaneous election is best when equal loyal support

- We find that a simultaneous election is best when $p^A = p^B$. In fact, we prove a stronger result.

**Theorem**

Suppose that $p^A = p^B$. For any number of total voters $N$, any signal quality $q > 1/2$ and any possible first period pair $(n_1, v_1^A)$ of number of votes $n_1$ of which $v_1^A$ are cast for candidate A, the probability of selecting the highest quality candidate is at least as high when $(n_1, v_1^A)$ is NOT revealed to the remaining $N - n_1$ voters as when it is revealed.

**Corollary**

When voters across states are ex-ante homogenous and $p^A = p^B$, a simultaneous election outperforms any sequential election.
Intuition: simultaneous election best with equal loyal support

- Sequential election carries an inherent risk for the party.
- Info from early states alone provides a less accurate estimate of $c^{HQ}$ than all states together, so voters may herd to the wrong candidate.
- This phenomenon similar to information cascade literature (i.e. Banerjee (1992)), in which herding is a negative result.
- When candidates have equal expected loyal support, the party is not willing to take this risk.
- In this case, the political party does best by allowing all private information from voters to aggregate and determine the winner without momentum effects.
Sequential election is best with sufficient loyal voter imbalance or low information quality

Lemma

Suppose \( p^A > p^B \). Then:

1. For any \( q > 1/2 \), there exists \( N^*(q) > 0 \) such that a sequential election outperforms simultaneous voting iff \( N > N^*(q) \).

2. For any \( N > \frac{1-2(p^A)^2 + p^A - p^B + 2p^A p^B}{(1-p^A-p^B)(p^A-p^B)} \), there exists \( q^*(N) > 1/2 \) such that a sequential election outperforms simultaneous voting iff \( q < q^*(N) \).

Herding optimal if loyal voter imbalance is large, private information is poor, or total population is large.
Intuition: sequential election best with loyal voter imbalance and low information quality

- A sequential election yields social learning among uncommitted voters, resulting in voter herding.
- This herding can generate enough votes for \( c^{HQ} \) to overcome the imbalance in loyal voters.
- When private information is not of high enough quality, simultaneous voting does not induce a strong enough result to overcome this imbalance.
- Voter herding in a sequential election serves to *compensate* for the imbalance in loyal voters.
- This result contrasts social learning literature, in which herding is inefficient in the ex ante welfare sense.
Question: Given that the party uses a sequential election, what sequence of states is best?

To answer, we allow states voting in different periods to have different characteristics.

A state $s$ is characterized by

- $d_s$: number of delegates at stake in state primary (proportional to population)
- $(p_s^A, p_s^B)$: expected fraction of voters who are loyal voters of candidates $A, B$ respectively
- $q_s$: expected fraction of uncommitted voters who receive correct signal regarding cand. quality
- $\lambda_s$: voter diversity (low correlation between voter preference shocks)
Optimal sequence of states

In a sequential election, it is optimal for the states voting early in the primary season to be those which:

1. are smaller
   - By having fewer voters vote early, the party is accepting the following trade-off: lower probability that $c^{HQ}$ will be correctly identified by early voting results, but more late period voters to benefit from social learning, and hence higher probability of pivotal herding.

2. have fewer loyal voters

3. have better informed voters
   - Both fewer loyal voters and better informed voters make an early state’s results more informative to future states.

4. display more voter diversity
   - Higher voter diversity makes a state’s results more informative given its size.
Primary calendar in recent elections

- There has been significant movement in states’ primary dates in recent elections.
- Is there a systematic pattern when comparing early states to late states over the last three election cycles?
- What demographic characteristics are correlated with states’ primary dates?
Observed patterns: early vs. late states in 2000-2008 Democratic primaries

<table>
<thead>
<tr>
<th></th>
<th>2000</th>
<th>2004</th>
<th>2008</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>early</td>
<td>late</td>
<td>early</td>
</tr>
<tr>
<td># 2008 delegates</td>
<td>1700</td>
<td>2428</td>
<td>2435</td>
</tr>
<tr>
<td>diversity</td>
<td>.472</td>
<td>.424</td>
<td>.454</td>
</tr>
<tr>
<td></td>
<td>(.056)</td>
<td>(.055)</td>
<td>(.059)</td>
</tr>
<tr>
<td>BA education</td>
<td>.290</td>
<td>.256</td>
<td>.283</td>
</tr>
<tr>
<td></td>
<td>(.043)</td>
<td>(.055)</td>
<td>(.055)</td>
</tr>
<tr>
<td>union</td>
<td>.145</td>
<td>.113</td>
<td>.134</td>
</tr>
<tr>
<td></td>
<td>(.055)</td>
<td>(.055)</td>
<td>(.057)</td>
</tr>
</tbody>
</table>

$p \leq 0.01 : ***$ 

$p \leq 0.05 : **$ 

$p \leq 0.10 : *$
Some observations regarding 2008 calendar

- Rhode Island, Montana: voted late but are small, have high education and high diversity, optimal to vote early.
- Less diverse states (such as Arkansas and Alabama) moved to earlier dates, more diverse states (such as Ohio and Rhode Island) moved to later dates.
- States with high education levels (such as Maryland and Washington) moved to later dates.
- Florida and Michigan: large states with average diversity and education levels, not optimal to vote early.