Why Does New Hampshire Matter — Simultaneous v.s. Sequential Election with Multiple Candidates

Pei-yu Melody Lo

Abstract

In a multi-candidate election, a voter may prefer to vote for his second choice in order to defeat his least favorite candidate. I study a model in which voters know their own preference but infer support of each candidate from a private signal. I show that if private signals are sufficiently precise, an equilibrium exists and is unique in the limit as the size of the electorate increases. In this unique equlibrium, supporters of a candidate vote more strategically when their hate for their worst choice becomes more dominant, but they vote less strategically when the same is true for the opposite camp. Using this property, I show that in sequential primaries, winning early primaries improves a candidate's chance of winning later primaries by making his supporters in later primaries vote more sincerely. I show that sequential primaries help with coordination, but puts more weight on the preference of the median voters in the early primaries. When voters' worry enough about defeating their worst choice, sequential primaries are better at aggregating preferences than simultaneous primaries.

1 Introduction

The outcomes of early elections play an out-of-proportion role in the US Presidential primary. Adam (1987) reports that the 1984 New Hampshire primary got nearly 20% of the season's coverage in ABC,CBS, NBC and the New York Times, even though New Hampshire accounts for only 0.4% of the US population, and only four votes out of 538 electoral votes in the presidential election. In the 1980 Republican primaries, George Bush and Ronald Reagan spent about 3/4 of their respective campaign budgets in early primary states, which account for much less than a fifth of the votes in the Republican convention in 1980 (Malbin, 1985). The emphasis on winning early primaries may come from the widely-held belief that early winners gain "momentum" due to the sequential nature of the election.

However, recent primaries have become more "front-loaded" into the early weeks. California has recently passed a legislation to move forward its primary to Feb. 5, 2008, only after 4 other primaries held in January. The media in general views this as "selfish" behavior on the part of those states. It has been argued that a more front-loaded primary system makes it more important for candidates to raise a lot of money early (William Schneider, 1997) and a more front-loaded 2008 primary gives well-established candidates an advantage. On what ground do these assertions stand? And if they are true, through what channel does the timing structure affect the voting outcomes?

Existing literature that study sequential elections has for the most part restricted attention to contests between two candidates. However, there are usually many candidates in a presidential primary. For example, Sen. Hillary Rodham Clinton of New York, Sen. Barack Obama of Illinois and former senator John Edwards of North Carolina, are all considered front runners in the 2008 primary for the Democratic party. With only two candidates, voters simply vote for their preferred candidate. In a multi-candidate contest, however, some voters have to vote strategically for their second choice if they believe their most preferred candidate has a smaller probability of being in a close race. Therefore, voters' beliefs about relative popularity of every candidate, and the relative likelihood of different pivotal events, play an important role in their decision.

Given this element of coordination in multicandidate contest under plurality rule, it is not surprising that with common knowledge assumption of the electoral situation, the voting outcome involves either a complete success or failure of coordination. Duverger's Law (see Riker 1982) asserts that "plurality rule brings about and maintains two-party competition", because only two candidates should be expected to get any vote. This represents complete success of coordination. Most of the literature focuses on these "Duvergerian" equilibria, but offers no formal theory as to which two candidates should be considered "serious" contenders. In addition, it cannot explain the incomplete coordination observed in many multicandidate election outcomes. For example, in the 1970 New York senatorial election, even the trailing candidate among the three got more than 24% of the votes, and the winner gets only 2% more votes than the second.

Moreover, common knowledge of the electoral situation seems a very strong assumption. The 1997 British Election Survey indicates that about two-third of voters who expected their preferred party to come second actually found that it came third (Fisher, 2000). There was clearly lack of common knowledge among voters as to the identities of the first and second place winner, which is inconsistent with that literature.

This paper presents a model of preference aggregation in a multi-candidate election that features a candidate who is "a common second choice" for supporters of the other two extreme candidates. One interpretation of "the common second choice" is a candidate that's widely known and considered a "safe option". Voters in the model know their preferences over the candidates, but have only imperfect information about the distribution of preferences in the electorate. Supporters of an extreme candidate have an incentive to coordinate with supporters of the "common second choice" against their least favorite candidate. Relaxing common knowledge assumption enables meaningful analysis of this coordination effect.

I show that an equilibrium exists and is unique in the limit as the size of the electorate increases. In this unique equilibrium, coordination incentive among supporters of an extreme candidate is stronger when their love-hate ratio for that candidate is smaller, when love-hate ratio for the opposing extreme candidate is higher, or when the prior belief of the share of supporters of the extreme candidate is smaller. In addition, in those situations, there is excess coordination in that the "common second choice" wins too often, i.e. sometimes "the common second choice" wins even though the median voter favors one of the extreme candidates.

I then study an election that involves voting in three states (electorates) in which the candidate winning the most states wins the election. This is close to Republican primaries. I compare voting behavior and outcomes under simultaneous and sequential primaries. When hate is stronger than love, in the last state, supporters of the extreme candidate that has not garnered any victory always vote for the "common second choice". Thus the equilibrium exhibit winnowing down of front runners. In addition, a victory by one extreme candidate in the first state boosts the morale of her supporters in the second state and results in more sincere voting behavior by her supporters and higher chance of winning in the second state. I show that when love-hate ratio is moderately small, or when the ex ante share of moderate voters is big (eg. larger than $\frac{1}{2}$), a sequential election reduces excess coordination motive in the first state as compared to the outcome under simultaneous primaries and reduces the ex ante probability that the candidate winning that state is not the median voter's first choice.

In addition to comparing voting behavior, I can also compare voting outcome between simultaneous and sequential election. Even if sequential election does not make extreme voters in the first state more aggressive, if the median voter in the first state is extreme, then if love-hate ratio is moderate, she prefers a sequential election to a simultaneous election because she can affect voting outcome in other states toward her favorite candidate. If the median voter in the first state is moderate, then she also prefers sequential election if the probability that an extreme voter wins her state is at least 70% of that of the share of extreme voters.

I can also compare voting outcome across voting order in a sequential election. If the median voter is extreme, then she always prefers voting earlier, i.e. first rather than second. If the median voter is moderate, then she prefers that her state votes first if the other state that her state swaps voting order with is "moderate" state, one in which love-hate ratio of extreme voters is small, or ex ante share of extreme voters is small. This is because when the other state is "moderate", voting first makes its extreme voters more aggressive, which is bad from a moderate voter's point of view.

2 Literature Review

Dekel and Piccione (2000) and Ali and Kartik (2006) both study sequential elections between two candidates in which some voters have only imperfect information about their own preference over the two candidates. Dekel and Piccione (2000) show that any outcome of a voting equilibrium in a simultaneous election is also an equilibrium outcome of a sequential election with any timing structure. Ali and Kartik (2006), on the other hand, construct a Perfect Bayesian equilibrium in which "herding," i.e. voting according to the history of vote counts so far and disregarding one's own information, happens with positive probability. This suggests that in a race between two candidates, a simultaneous election can be (but is not necessarily due to multiplicity of equilibria) more efficient in gathering information than a sequential election.

Myerson and Weber (1993) and Myerson (2002) both assume common knowledge of the preference distribution of the electorate, and show that under plurality rule, for any pair of candidates in a "three-horse race", there exists an equilibrium in which only this pair are considered "serious" and get any vote. Myerson (2002) call these discriminary equilibria because labeling of the candidates matter as to whether they have positive probability of winning. They argue that "a large multiplicity of equilibria creates a wider scope for focal manipulation by political leaders."

Myerson and Weber (1993) also show via an example the existence of a "non-Duvergerian" equilibria in which a group of voters fail completely to coordinate to avoid the worst outcome, and the two losers exactly tie. They conjecture that some additional assumption of dynamic stability or persistence may be used to eliminate these "non-Duvergerian" equilibria.

This paper is most closely related to Myatt (2007), which studies simultaneous elections under plurality rule in which one candidate (the conservative status quo) has a commonly known fixed fraction ($< \frac{1}{2}$) of supporters, while the rest of the electorate share the distaste of the status quo but disagree on which of the other two (liberal) candidates is optimal. This assumption effectively reduces an election under plurality rule with three candidates to one under qualified-majority rule between two candidates. Essentially, the (liberal) voters have to coordinate behind the two (liberal) candidates. They relax the common knowledge assumption by assuming that each voter gets an imperfect signal about the preference distribution of the electorate (as evident in the UK General Election of 1997). They construct a uniqe symmetric equilibrium that is consistent with the 1970 New York Senatorial election, which displays limited strategic voting and incomplete coordination. However, the assumption of a fixed and commonly known support for one candidate does not seem to fit US Presidential primaries.

It is difficult to characterize equilibria in a large election because probability ratios of close-race events between different pairs of candidates can be quite intractable. Myatt (2007) develops the solution concept of *strategic-voting equilibrium* for large elections, which can be viewed as a Bayesian Nash equilibrium with a continuum of voters. It facilitates the calculation through law of large numbers arguments. Myerson (2000), on the other hand, tackels this issue by assuming population uncertainty. They assume that voter turnout follows a Poisson process with a commonly known preference distribution. The feature of Poisson process that an individual voter's belief about the behavior of the electorate does not depend on his own preference type facilitates comparison of limiting probabilities of different pivotal events as the size of the electorate goes to infinity.

On relaxing common knowledge assumptions in voting situations, Feddersen and Pesendorfer (1997, 1998) use a common value model for jury decision making. In their model, each juror decides on one of two votes based on a private signal about the defendant's guilt and aims to convict the guilty and acquit the innocent. Thus other jurors' information matters even for a juror's own preference over outcomes. Each juror infers about the merits of his two actions from an assessment of the information possessed by others conditional on his vote being pivotal. Therefore, if other jurors respond a lot to their signals, a juror may have an incentive to disregard his own signal because the information contained in the pivotal event outweighs his own information. This is why bandwagon effects may arise in sequential elections with two candidates in Ali and Kartik (2006). However, since there are only two outcomes, the coordination effect in multicandidate contests is not present in these models.

3 A Multicandidate Contest in One State

3.1 The Model

Voting Rule Three candidates L, M, R compete in a simultaneous election. The number of voters is a random variable that follows a Poisson distribution with mean n. We use G_n to denote the voting game with expected voter turnout n. Each voter has to voter for exactly one candidate. The winner of this election is the candidate with the most votes. In case of a tie, each of the candidate with the most votes wins with equal probability.

Preferences A voter can be of three preference types. A type r voter favors candidate R, and prefers R to M to L, a type l voter prefers candidate L to M to R, while a type m voter prefers candidate M the most and is indifferent between R and L. Candidate M can be thought of as the common second choice. He is not the most hated candidate for any voter. On the other hand, candidate R and L are loved the most by some but hated the most by some others. We can think of candidate R and M as extreme candidates and their supporters as extreme voters. A voter's payoff when his most hated candidate wins is normalized to 0. A voter who favors candidate c receives payoff V_c if his favorite candidate wins, and v_c if his second choice wins, where $V_c > v_c > 0$ for $c \in \{R, L\}$ and $V_m > v_m = 0$. We define $V_c - v_c$ to be a type c voter's love, which is his payoff difference between a win by his favorite candidate and a win by his second choice. We define v_c to be his hate, which is his payoff difference between his second choice and his worst choice. For $c \in \{R, L\}$, define $\phi_c = \frac{V_c - v_c}{v_c}$, which is type c voter's love-hate ratio. This love-hate ratio captures the importance of coordinating with the moderate voters. The smaller the love-hate ratio is, the more fearful the voter is of a win by his worst choice, and thus the more important it is to coordinate with the m voters to defeat L. Define $u_c = \log(2\phi_c)$.

All voters with the same preference type shares the same love-hate ratio. As will be shown later, a voter's behavior depends on his preference only through this love-hate ratio. One can think of m voters as having such a large love-hate ratio that it is always weakly dominant for them to vote sincerely. The love-hate ratios of each preference type are common knowledge, but a voter's preference is his own private information.

Electoral Preference Distribution A voter in the electorate is type r with probability $F(\eta - \theta)$, type l with probability $F(-\eta - \theta)$ and type m with probability $1 - F(\eta - \theta) - F(-\eta - \theta)$, where F is the cumulative distribution function for Laplace distribution with mean 0 and scale parameter 1, denoted by *Laplace* $(0, 1)$. So

$$F(x) = \begin{cases} \frac{1}{2}e^{x} & \text{if } x \leq 0 \\ 1 - \frac{1}{2}e^{-x} & \text{if } x > 0 \end{cases}.$$

θ is an exogenously given parameter of the model. We can think of voters as living on a real line as depicted in figure __. An r voter lives above θ, and an l voter lives below $-\theta$, and m voters live in $(-\theta, \theta)$. Voter location follows Laplace distribution with mean η and scale parameter 2. We can think fo η as the location of the median voter. It also completely determines the preference distribution of the electorate. Because preference is single-peaked, the median voter's favorite candidate is the Condorcet winner. We assume throughout that $e^{-\theta} < \frac{1}{2}$. This ensures that for each candidate, there are electoral preference situations where the majority of the electorate support that candidate.

Information To capture uncertainty about preference distribution of the electorate, I assume that no one knows the location of the median voter, but a voter with belief type z believes that η follows distribution with density function $h(\eta; z) = \frac{f(\alpha(\eta - \eta_0)) f(z - \eta)}{\int_{-\infty}^{\infty} f(\alpha(\zeta - \eta_0)) f(z - \zeta) d\zeta}$ where $f(.)$ is the density function of *Laplace* $(0, 1)$, and $\eta_0 \in \mathbb{R}$ and $\alpha > 0$ are exogenous parameters of the model. Thus, a voter with belief type z believes that the likelihood ratio of the event that $\eta = \eta_1$ v.s. the event that $\eta = \eta_2$ is

$$\frac{f(z - \eta_1) f(\alpha(\eta_1 - \eta_0))}{f(z - \eta_2) f(\alpha(\eta_2 - \eta_0))}.$$

The parameter η_0 captures some common prior belief about the preference distribution, and $\alpha > 0$ captures the strength of this prior belief.

We assume that conditional on the true median voter location η, the belief type z of a voter follows *Laplace* $(\eta, 1)$, and is independent of voter preference type. That is, a voter's belief about the preference distribution of the electorate is independent of his own preference type. This assumption is made for tractibility reason.

There are two stories that can deliver this. One is that an r voter and an l voter hold different priors that exactly cancel out the inference they draw from their own preference type. They are pessimistic about the size of the fellow supporters of their favorite candidate and this pessimism exactly cancels out the favorite news they draw from being a supporter of this candidate.

Another story is that voters hold a common prior that η follows $Laplace\left(\eta_0, \alpha\right)$ but do not update based on their own preference. That is, voters believe that the preference distribution among all other voters are determined by the parameter η; they hold the same prior about η, but do not believe that their own preference is generated by the same distribution determined by η. They form their posterior belief about the preference distribution among all other voters by the common prior and an additional signal z which is independently drawn from $Laplace\left(0, 1\right)$.

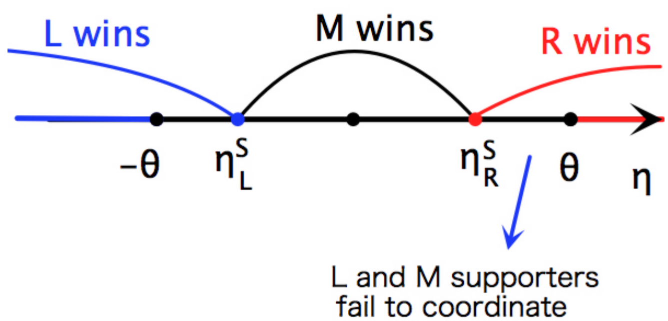
From the expression of the likelihood ratio when voter belief type is z, we can see that the precision of the private signal is held fixed in this model, instead of being allowed to get arbitrarily precise. In a sense, the precision of private signal is captured by the parameter θ, because when $|\eta| \leq \frac{\theta}{2}$, type m voters are expected to be the majority, and it is weakly dominant for them to vote sincerely, so most likely M is going to win and no voter is pivotal. An extreme voter will be weighing likelihood ratios of an event where η is in some set above $\frac{\theta}{2}$ and an event where η is in some set below $-\frac{\theta}{2}$. Thus a larger θ implies that voters will be gauging the likelihoods of two disjoint subsets of states that are further apart, and thus will be more confident whether $\eta > \frac{\theta}{2}$ or $\eta < -\frac{\theta}{2}$. [??? needs more work]

3.2 Sincere Voting and Coordination Failure

If every voter simply votes for his favorite candidate, then in a large election, vote share of candidate C, denoted by $p_c\left(\eta\right)$, is almost equal to the probability that voter is of type c. If $2F\left(-\eta\right) < \frac{2}{3}$, when the share of two exreme voters are equal to each other, it is smaller than the share of moderate voters. Because $p_R\left(\eta\right)$ increases with η, and $p_R\left(\theta\right) = \frac{1}{2}$, the median voter preferes R to M if and only if $\eta > \theta$. However, when η is close to θ but smaller than θ, R still gets almost half of the votes, while M and L share the other half. Thus R wins the election even though the median voter is moderate and the majority prefer M to R. This happens because left-wing voters and moderates fail to coordinate with each other and support M together against R. I call this cross-camp coordination failure.

3.3 Equilibria

Here I investigate equilibrium voting behavior, and how it changes with voters' love-hate ratios.



3.3.1 Strategies

Voter i's pure strategy v_i is a mapping from his preference-belief pair $(o, z) \in \{r, m, l\} \times \mathbb{R}$ to a candidate in $\{R, M, L\}$. A voter votes sincerely if he simply casts his ballot for his most preferred candidate regardless of his belief about electoral preference distribution.

A strategy profile is symmetric if voting behavior depends only on a voter's preference type and belief type, but not on the voter's identity. We look at symmetric weakly undominated strategy profiles. This rules out equilibria where a voter is never pivotal and thus any action is a best response. Because it is weakly dominant for a type m voter to vote sincerely, and it is weakly dominated for an extreme voter to vote for his worst choice, a voting strategy profile can be represented by $(v_r(z), v_l(z)) : \mathbb{R} \to [0, 1]$, where $v_c(z)$ is the probability a type c voter votes for his favorite candidate.

Given the voting strategy $v = (v_r, v_l)$ everyone adopts, and the median voter location η, a randomly chosen voter will vote for candidate C with probability

$$\begin{aligned} p_R(\eta; v) &= F(\eta - \theta) E_z [v_r(z) | \eta] \\ p_L(\eta; v) &= F(-\eta - \theta) E_z [v_l(z) | \eta] \\ p_M(\eta; v) &= 1 - p_R(\eta; v) - p_L(\eta; v). \end{aligned}$$

We call $p(\eta; v) = (p_R, p_M, p_L)(\eta; v)$ the expected vote share function. We suppress the dependence on v when it is clear.

Given any two candidates c_1, c_2, there is an equilibrium in which every voter votes for the one in $\{c_1, c_2\}$ that she prefers. In such an equilibrium, the third candidate never gets any vote, and thus is irrelevant. The election is reduced to a binary choice. One can say that the two candidates c_1 and c_2 are the focal points of the election. However, the model cannot answer the question of how front runners are chosen or have any meaningful comparative statics. In order to model how the identity of the front runners change with the underlying electoral situation, we restrict attention to strategy profiles where every candidate has positive probability of winning ex ante. In a large election, the actual vote share will be close to the expected vote share with high probability by law of large numbers. Thus, with probability close to1, the winner will be the candidate with the highest expected vote share. Thus, if each candidate's ex ante probability of winning remains positive as electorate size increases, the strategy profile must exist multi-candidate support as defined as follows.

Definition 1 *A strategy profile has multi-candidate support if, for each candidate, there exist some electoral preference distribution η where this candidate has the highest expected vote share.*

3.3.2 Best Responses

Consider a type r voter's payoff given that everyone else adopts voting strategy v. Let x_j denote the number of votes candidate j gets from everyone else. Then (x_R, x_M, x_L) is a vector of random variables whose distribution depend on the voting strategy v adopted by everyone else and electoral preference distribution η. Given η, (x_R, x_M, x_L) follow independent Poisson distributions with mean $(np_R(\eta; v), np_M(\eta; v), np_L(\eta; v))$.

A voter with preference type r will choose between voting for R and voting for his second choice M. Cosider the vote distribution without this voter's ballot. If his first and second choice either ties or is only one vote apart, he gains his love by voting sincerely [strictly prefers to vote for his first choice[. If his first and worst choice ties, or if his first choice is one vote behind, then by voting for his favorite candidate, he not only gains his love but also avoids his hate [his preference for voting sincerely is even stronger]. But if his second choice either ties with the worst choice, or is one vote behind, then he prefers to vote for his second choice to prevent his worst candidate from winning and avoid his hate. A voter does not know whether and which tie event will obtain, but forms beliefs about their probability based on the expected vote share and his belief type z. Let $\Pr\{H|\eta\}$ denote the probability of an event H when the median voter location is η. Then a voter with belief type z believes that event H happens with probability $E[\Pr\{H|\eta\}|z]$.

Let $\Pr\{Piv_+^r|\eta\}$ denote the probability that an additional vote for r's favorite candidate R rather than the second choice M will cause the winner to change from candidate M to R. Let $\Pr\{Piv_{RL}^r|\eta\}$ to be the probability that sincere voting rather than defensive voting will change the winner from r's worst choice L to r's favorite candidate R. When this is the case, by voting sincerely, the voter gains his love and also avoids suffering his hate. Let $\Pr\{Piv_-^r|\eta\}$ be the probability that sincere voting rather than defensive voting will backfire and cause the winner to change from the second choice M to the worst choice L. Let Cause the winner to change from the worst choice $-C$ to the favorite candidate C, cause the winner to change from the second choice M to the worst choice

$-C$. If a type c voter casts his vote for his favorite candidate instead of voting defensively for his second choice, the winner switches from M to C with probability $\Pr\left\{Piv_+^c|\eta\right\}$, and from M to $-C$ with probability $\Pr\left\{Piv_-^c|\eta\right\}$, and from worst $-C$ to best C with probability $\Pr\left\{Piv_{RL}^c|\eta\right\}$. Thus by voting sincerely, the expected gain to a type c voter with belief type z is

$$\left(E\left[\Pr\left\{Piv_+^c|\eta\right\}|z\right]+E\left[\Pr\left\{Piv_{RL}^c\right\}|\eta\right]\right)(V_c-v_c)-\left(E\left[\Pr\left\{Piv_-^c|\eta\right\}|z\right]-E\left[\Pr\left\{Piv_{RL}^c|\eta\right\}\right]\right)v_c.$$

So $\Pr\left\{Piv_+^c|\eta\right\}+\Pr\left\{Piv_{RL}^c\right\}$ is the probability that voter 0 gains his love by voting sincerely, while $E\left[\Pr\left\{Piv_-^c|\eta\right\}|z\right]-E\left[\Pr\left\{Piv_{RL}^c|\eta\right\}\right]$ is the net probability that voter 0 suffers his hate by voting sincerely. So a voter of type (c,z) prefers to vote sincerely if and only if his love-hate ratio is higher than the likelihood ratio of hate-avoiding events to love-enhancing events:

$$\phi_c=\frac{V_c-v_c}{v_c}\geq\frac{E\left[\Pr\left\{Piv_-^c|\eta\right\}|z\right]-E\left[\Pr\left\{Piv_{RL}^c|\eta\right\}\right]}{E\left[\Pr\left\{Piv_+^c|\eta\right\}|z\right]+E\left[\Pr\left\{Piv_{RL}^c\right\}|\eta\right]}.$$

We will establish that a multi-candidate equilibrium exists and is unique in the limit.

Definition 2 *A limit equilibrium is the limit of some sequence* $\{(v_{n_k,r},v_{n_k,l})\}_{k=1}^{\infty}$*, where* v_{n_k} *is an equilibrium when expected voter turnout is* n_k*.*

We show that when θ is sufficiently large and α is sufficiently small, equilibrium with multi-candidate support exists and is unique in the limit. That is, there exists a strategy profile with multi-candidate support which is the limit a sequence of equilibria, and is the limit of every sequence of equilibria.

Write

$$\underline{\theta}=\max\left\{\frac{u_r}{2},\frac{u_l}{2},\frac{3}{4}(u_r+u_l)\right\}-\log\left(1-\frac{1-\frac{\sqrt{1-2e^{-2\theta}}}{2}}{2}\right).$$

Theorem 1 *For all* $\theta>\underline{\theta}$*, there exists* $\overline{\alpha}>0$ *such that for all* $\alpha<\overline{\alpha}$*, there exists a unique limit equilibrium with multi-candidate support. In this unique limit equilibrium,* $v_r^*=\mathbb{I}\left\{z\geq a^*-\frac{u_r}{2}\right\}$ *and* $v_l^*=\mathbb{I}\left\{z\leq a^*+\frac{u_l}{2}\right\}$*.*

This unique limit must be in cutoff strategies. If the supporters of candidate $C\in\{R,L\}$ has higher love-hate ratio, then in equilibrium they vote more sincerely, while voters in the opposing camp $-C$ votes more defensively, and which causes candidate C to win with lower probability.

3.3.3 Limit Probability of Pivotal Events

For a supporter of candidate C, voting sincerely rather than defensively causes the winner to switch from M to C when C and M either tie or are one vote away from a tie; it causes the winner to switch from M to the worst choice $-C$ when either M and $(-C)$ tie or M is one vote behind $-C$ as the winner. The probability that candidate 2 is d votes ahead of candidate 1, given expected vote share (p_1, p_2, p_3), is

$$\sum_{k \in \mathbb{N}, \min\{k, k+d\} \geq 0} \sum_{j=0}^{\infty} \frac{e^{-n} (p_1)^k (p_2)^{k+d} (p_3)^j}{k! (k+d)! (j!)}.$$

Define as in Myerson (2006) the magnitude of an event $\{(x_1, x_2, x_3)\}$ given the mean (np_1, np_2, np_3) as

$$\mu((x_1, x_2, x_3) | np) = \sum_{i=1}^{3} \frac{x_i}{n} \left(1 - \log \frac{x_i}{np_i}\right).$$

Then the probability of $\{(x_1, x_2, x_3)\}$ given mean np is

$$\Pr\{x | np\} = \frac{e^{n\mu(x|np)}}{\Pi_{i=1}^{3} \left(\iota(x_i) \sqrt{2\pi x_i + \frac{\pi}{3}}\right)}.$$

Consider an event H. Call x^* a maximizer of the event H and x a near-maximizer of H given np if

$$x^* \in \arg \max_{x \in Co(H)} \mu(x | np)$$

and

$$x \in \arg \max_{Co(H)} \mu(x | np).$$

Myerson (2006) show that as $n \to \infty$, the probability of an event H will concentrate around the near maximizer of the event.

magnitude of a pivotal event $\{(k, k+d, j) : k, d \in \mathbb{N}\}$

$$\mu(\{(k, k+d, j) : k, d \in \mathbb{N}\} | np) = -\left(\sqrt{p_1} - \sqrt{p_2}\right)^2 + ..$$

magnitude of $\{(k, k+d, j) : k, d \in \mathbb{N}, k \geq j\}$ when the constraint is binding,

$$-1 + 3 (p_1 p_2 p_3)^{\frac{1}{3}}$$

and when it is not binding.

Lemma 3.1 *(Upper bound and lower bound on probability a pivotal event)*

So if the magnitude of an event is negative, its probability vanishes. In addition, if two events have different magnitude, as $n \to \infty$, the one with larger magnitude will become infinitely more likely than the event with smaller magnitude.

By assumption on θ and because it is weakly dominant for type m voters to vote sincerely, M's expected vote share is highest among all when η is $\left(-\frac{\theta}{2}, \frac{\theta}{2}\right)$. It is straightforward to check that there exists some $\varepsilon > 0$ and $N > 0$ such that given any voting strategy, the magnitude of an event where the two extreme candidates are one vote apart is less than $-\varepsilon$ if the expected turnout is $n > N$. (Claim 6 in "condition on p for uniform converge to an expression with magnitude.tex")

Because v is assumed to have multi-candidate support, there must exist some η such that $p_R(\eta; v) \geq, \max\{p_M(\eta; v), p_L(\eta; v)\}$. This is only possible for $\eta > \frac{\theta}{2}$ because for $\eta < \frac{\theta}{2}$, candidate M has more supporters than R, and thus definitely more votes in expectation than R. By continuity of $p_R(.; v)$, there exists $\eta_R > \frac{\theta}{2}$ such R and M tie as leader in expected vote share. At such η_R, the magnitude is 0, which is the maximum possible magnitude, and probability of a tie becomes 1. Applying the magitude comparison, as $n \to \infty$, probability of an $R - M$ tie will concentrate around such η_R's. And for any interval I of electorate situations that contain at least one such solution η_R, conditional on $\eta \in I$, the event of an $R - M$ tie also becomes infinitely more likely than the event where R and L are one vote apart.

Thus the likelihood ratio of hate-love events become close to the likelihood ratio of an $L - M$ ie over an $R - M$ tie. The next lemma shows that if v is a limit equilibrium, then there will be exactly one solution $\eta_R(v)$ at which R and M tie as the expected vote share leader.

Recall that the higher the belief type z is, the higher the likelihood ratio is between a r-dense world and a l-dense world. For an r voter, the higher her belief type is, the more optimistic she is about the share of r voters in the electorate. For an l voter, the lower her belief type z is, the more optimistic she is about the share of l voters in the electorate. A natural strategy would be a cutoff strategy $(z^*(r), z^*(l))$ where a voter votes for her favorite candidate if she is sufficiently optimistic about the size of her camp, i.e. if $z \geq z^*(r)$ for an r voter, and if $z \leq z^*(l)$ for an l voter.

Figure depicts an expected vote share function given a cutoff strategy. The higher the electoral state η is, the more likely a voter supports R, and the more likely an r voter votes for R. Thus candidate R's expected vote increases with elecotral state η. On the other hand, candidate M's expected vote share peaks when electoral state is intermediate. In those states, M has the most supporters, and M also get quite some votes from the extreme voters who are pessimistic about their favorite candidates. When R and M are the expected front runners, R's expected vote share lead increases as the electoral condition becomes more favorable to R. Same applies to L. As a result, there is a pair of cutoffs (η_R, η_L) on the electoral condition such that candidate R has the highest expected vote share if the electoral condition is higher than η_R, L has the highest expected vote share if $\eta < \eta_L$, and M does for $\eta \in (\eta_L, \eta_R)$.

We call expected vote share functions with the aforementioned properties *regular.*

Definition 3 *Say that an expected vote share function* $p : \mathbb{R} \to \Delta\{R, M, L\}$ *is regular if, for* $c, -c \in \{R, L\}$ *and* $c \neq -c$, $p_c(\eta) = p_M(\eta) \geq p_{-c}(\eta)$ *has a unique solution* η_c, *and* $p'_R(\eta) - p'_M(\eta) > 0$ *for all* $\eta \geq$ *and* $p'_L(\eta) - p'_M(\eta) < 0$ *for all* η. *A voting strategy is regular if it generates a regular expected vote share function.*

The following lemma implies that, for the purpose of finding all limit equilibria, we can focus on voting strategies generating regular vote share functions.

Lemma 3.2 *If* (v_r^*, v_l^*) *is the limit of a sequence of equilibria and has multi-candidate support, then* (v_r^*, v_l^*) *must be regular.*

Sketch of the proof. The key step is to show that for a supporter of an extreme candidate C, there exists some belief type $\hat{z}_c$ such that a type c voter will vote sincerely if he is more optimistic than $\hat{z}$.

We first see that for $\eta \in \left(-\frac{\theta}{2}, \frac{\theta}{2}\right)$, candidate M has the highest expected vote share. Let $\underline{\eta_C^*}$ and $\overline{\eta_C^*}$ denote the smallest and the largest in $\{\eta : p_C(\eta; v^*) = p_M(\eta; v^*) \geq p_{-C}(\eta; v^*)\}$. Then $\underline{\eta_R^*} > \frac{\theta}{2} > -\frac{\theta}{2} > \underline{\eta_L^*}$. We first observe that for $z \in \left(-\frac{\theta}{2}, \underline{\eta_R^*} - \delta\right)$ for any $\delta > 0$, and for voting strategy v_n sufficiently close to v, the likelihood ratio of hate-avoiding event to love-enhancing event is strictly decreasing in voter's belief type z and has a slope less than -1. This is because the love-enhancing events concentrate around certain electoral situations $\eta \geq \underline{\eta_R}(v_{n_k}) \geq \underline{\eta_R^*} - \delta$, while hate-avoiding events concertrate some $\eta \leq \underline{\eta_L^*}$. Thus as $z \in \left(\underline{\eta_L^*}, \underline{\eta_R^*} - \delta\right)$ increases, the it moves closer to love-enhancing events and away from hate-avoiding events, making the love-enhancing event more likely and the hate-avoiding event less likely. In addition, because $p_R\left(\underline{\eta_R^*}; v^*\right) = p_M\left(\underline{\eta_R^*}; v^*\right)$ and because $\underline{\eta_R^*} > \frac{\theta}{2}$, in v^*, r voters must vote for R with positive probability for belief types $z \in \left(-\frac{\theta}{2}, \underline{\eta_R^*}\right)$. This must be true v_n close to v^*.

We then see that of Because the limit v has multi-candidate support, there

1.

■

In a large election, the actual vote share will be close to the expected vote share. Consider a regular expected vote share function with critical states (η_R, η_L). Thus for $\eta > \eta_R$, it is most likely that R will win, while for $\eta \in (\eta_L, \eta_R)$, it is most likely that M will win. When $\eta = \eta_R$, it is a knife-edge race between R and M and it is when an R-M tie is most likely going to happen. Similarly, an L-M tie is most likely going to happen at $\eta = \eta_L$. When candidate R and L have the same expected vote share, they are tying for the biggest loser, not the winner, and thus it is unlikely that the two extreme candidates will tie for the winner at any electoral condition η. Thus the ratio $\frac{E[\Pr\{Piv_{RL}^c|\eta\}|z]}{E\left[\Pr\left\{Piv_+^c|\eta\right\}|z\right]} \to 0$ as expected voter turn out goes to infinity. Therefore, the likelihood ratio of hate - love events approaches the likelihood ratio of an LM tie over an RM tie.

3.3.4 Limit Best Response Correspondence

A voter will be weighing his love-hate ratio with the likelihood ratio of the event that η is close to η_R with the event that η is close to η_L. Even though the probability of either an R-M tie or an M-L tie goes to 0 except at $\eta = \eta_R$ or η_L, and the probability that the electoral condition is exactly equal to η_R or η_L is 0 given any belief type z, the likelihood ratio of the two events actually converge as $n \to \infty$. Because a C-M tie happens most likely around η_C, the limit of this ratio depends on the likelihood ratio of the two electoral states η_R and η_L, and also on how fast the distance between the two front runners' expected vote share diverges from 0 as the electoral condition η goes away from η_c. Lemma shows that the ratio converges to for all voting strategies generating regular vote share functions.

Though at all $\eta \neq \eta_R$, probability of near tie events goes to 0,

Let $\Pi_\delta := \{p \in \Delta \{R, M, L\} : p_c \geq \delta \text{ for all } c\}$.

Lemma 3.3 *Given any $\delta > 0$,*

$$\frac{n \Pr \{(k, k+d, j) : k, j \in \mathbb{N} | np\}}{\frac{\sqrt{n} e^{-n\left(\sqrt{p_1(\eta)} - \sqrt{p_2(\eta)}\right)^2}}{\sqrt{2\pi}(4 p_R(\eta) p_M(\eta))^{\frac{1}{4}}}}$$

goes to 1 *uniformly among all $p \in \Pi_\delta$.*

Because an $R-L$ tie becomes infinitely less likely than either an $R-M$ tie or an $M-L$ tie, the likelihood ratio of hate-love event approaches the likelihood ratio of $M-L$ tie over $R-L$ tie. Because the probability of a tie between C and M will concentrate on a neighborhood of η_C, the limit of the likelihood ratio will depend on the likelihood ratio of the two electorate situations η_L and η_R, and how fast the expected votes of the front runners diverge when η diverges from η_C.

Define

$$\tau(v) = \frac{\eta_R(v) + \eta_L(v)}{2} + \frac{1}{2}\left(\log \frac{|p'_R(\eta_R(v); v) - p'_M(\eta_R(v); v)|}{|p'_L(\eta_L(v); v) - p'_M(\eta_L(v); v)|} + \alpha\left(|\eta_R(v) - \eta_0| - |\eta_L(v) - \eta_0|\right)\right).$$

Lemma 3.4 *Given any v generating a regular vote share function, $\psi_n(z; v)$ converges to $(\psi_r, \psi_l)(z; v)$ where*

$$\psi_r(z; v) = \begin{cases} -2(\eta_L - \tau(v)) & \text{if } z < \eta_L \\ -2(z - \tau(v)) & \text{if } z \in [\eta_L, \eta_R] \\ -2(\eta_R - \tau(v)) & \text{if } z > \eta_R \end{cases}$$

and $\psi_l(z; v) = -\psi_r(z; v)$. Convergence is uniform over z on a bounded interval.

Figure 1 depicts the limit of the likelihood ratio of hate-love events as a function of belief type z. For z between η_L and η_R, when belief type z increases, an $R-M$ tie becomes more likely while an $L-M$ tie becomes less likely. Thus the likelihood ratio of hate -love events decreases. But when both η_L and η_R lie on the same side of the belief type z, their likelihood ratio does not change with z. It immediately follows that either an extreme voter's limit best response is a cutoff strategy that votes sincerely for belief types more optimistic than the

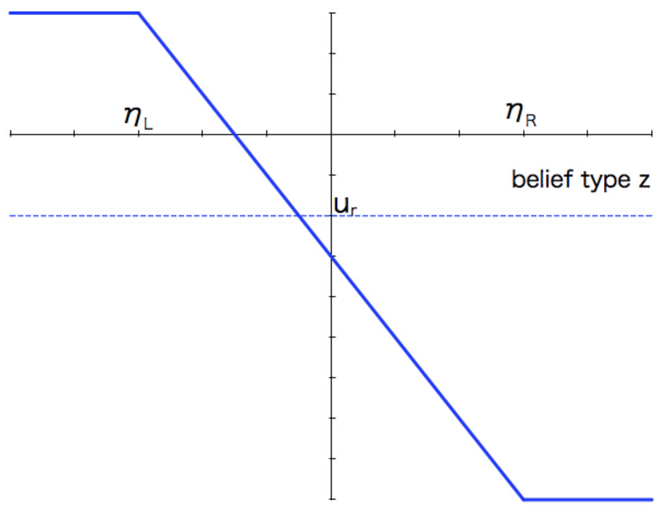


Figure 1: Limit Likelihood Ratio Function of Hate-Love Events

threshold, or an extreme voter always weakly prefer (in the limit) one candidate regardless of his belief.

The limit pivotal ratio is very simple. It is linear for belief types between η_L and η_R and is flat for more extreme belief types. Thus, for a voter that favors either R or L, her limit best response when everyone else uses the same regular strategy is either a cutoff strategy, or she always weakly prefers to vote for one candidate no matter what her belief type is. Define Σ to be the set of regular voting strategies. Let $BR_\infty : \sum \to 2^{\Sigma}$ denote the limit best response correspondence.

Lemma shows that a limit equilibrium must be a limit best response to itself. Thus we can proceed by finding all fixed points to the limit best response correspondence.

Lemma 3.5 *If v is the limit of a sequence of equilibria where every candidate wins ex ante with probability at least ε for some $\varepsilon > 0$, then v must be a limit best response to itself a.e..*

3.3.5 Unique Fixed Point to the Limit Best Response Correspondence

Theorem 2 *BR_∞ has a unique fixed point v^*.*

We first show that if v^* is a limit best response to itself, then an extreme voter does not weakly prefer one candidate at all belief types. That is, $\psi_r(\eta_L(v^*); v^*) > u_r > \psi_r(\eta_R(v^*); v^*)$ if v^* is a fixed point to BR_∞.

Lemma 3.6 *If* $\theta > \max\left\{\frac{u_r}{2}, \frac{u_l}{2}\right\} - \log\left(1 - \frac{1-\sqrt{1-2e^{-2\theta}}}{2}\right)$ *and* $\theta > \frac{3}{4}(u_r + u_l) - \log 2$*, then a fixed point to* BR_∞ *must be a cutoff strategy* $\left(a^* - \frac{u_r}{2}, a^* + \frac{u_l}{2}\right)$ *where the cutoff is the unique solution of* z *to* $\psi_c(z; v^*) = u_c$*.*

Proof. Suppose v has multi-candidate support and is a limit best response to itself. First we show that $\psi_r(z; v) < u_r$ for $z \geq \eta_R(v)$. Suppose not. Then $\psi_r(\eta_R(v); v) = u_r$ for v to have multicandidate support. Thus $v_r(z) = 0$ for all $z < \eta_R(v)$. Then

$$
\begin{aligned}
& 2p_R(\eta_R(v); v) + p_L(\eta_R(v); v) - 1 \\
\leq\ & \max\left\{\left(1 - \frac{1}{2}e^{\theta - \eta_R(v)}\right), \frac{1}{2}\right\} + \frac{1}{2}e^{-\theta - \eta_R(v)} - 1 \\
<\ & 0,
\end{aligned}
$$

contradiction to the definition of $\eta_R(v)$.

Next we show that $\psi_r(z; v) > u_r$ for all $z \geq \eta_L(v)$. It follows immediately by noting that $\eta_R(v) - \eta_L(v) \geq \eta_R^{\sin} - \eta_L^{\sin} = 2\log\frac{e^\theta + \sqrt{e^{2\theta} - 2}}{2}$, and thus for all $z \leq \eta_L(v)$,

$$
\begin{aligned}
\rho_r(z; v) &= \rho_r(\eta_L(v), v) \\
&= \eta_R(v) - \eta_L(v) + \alpha\left(\frac{\eta_R + \eta_L}{2} - \eta_0\right) \\
&\geq 2\log\frac{e^\theta + \sqrt{e^{2\theta} - 2}}{2} + \alpha\left(\frac{\eta_R + \eta_L}{2} - \eta_0\right) \\
&> u_r
\end{aligned}
$$

by the assumption on θ. Similarly, $\psi_l(z; v) > u_l$ for all $z \geq \eta_R(v)$.

Thus v must be a cutoff strategy of the form $\left(a - \frac{u_r}{2}, a + \frac{u_l}{2}\right)$. But a must be a fixed point of β. ■

Thus if v is a limit best response to itself, then an

$$
\begin{aligned}
v_r(z) &= 1_{\{z:\psi_r(z;v) \leq u_r\}} = 1_{\left\{z:z \geq \tau(v) - \frac{u_r}{2}\right\}} \\
v_l(z) &= 1_{\{z:\psi_l(z;v) \geq u_l\}} = 1_{\left\{z:z \leq \tau(v) + \frac{u_l}{2}\right\}}.
\end{aligned}
$$

So it must be in cutoff strategies of the form $\left(a - \frac{u_r}{2}, a + \frac{u_l}{2}\right)$. Thus the problem is reduced to finding a fixed point a^* to the function

$$
\tilde{\tau}(a) = \tau\left(a - \frac{u_r}{2}, a + \frac{u_l}{2}\right)
$$

where

$$
a^* - \frac{u_r}{2}, a^* + \frac{u_l}{2} \in \left(\eta_L\left(a^* - \frac{u_r}{2}, a + \frac{u_l}{2}\right), \eta_R\left(a^* - \frac{u_r}{2}, a + \frac{u_l}{2}\right)\right) \quad (1)
$$

Lemma 3.7 *If* $\left(a^* - \frac{u_r}{2}, a^* + \frac{u_l}{2}\right)$ *is a limit best response to itself, then*

$$\begin{aligned}\eta_R\left(a^* - \frac{u_r}{2}, a^* + \frac{u_l}{2}\right) &> \max\left\{\theta, a^* - \frac{u_r}{2}, a^* + \frac{u_l}{2}\right\} \\ \eta_L\left(a^* - \frac{u_r}{2}, a^* + \frac{u_l}{2}\right) &< \min\left\{-\theta, a^* + \frac{u_l}{2}, a^* - \frac{u_r}{2}\right\}.\end{aligned}$$

If $\eta_R\left(z^*(r), z^*(l)\right) > \max\left\{\theta, z^*(r), z^*(l)\right\}$, then η_R is thus the solution to

$$\begin{aligned}1 &= 2p_R\left(\eta; z^*(r), z^*(l)\right) + p_L\left(\eta; z^*(r), z^*(l)\right) \qquad (2) \\ &= 2\left(1 - \frac{1}{2}e^{\theta-\eta}\right)\left(1 - \frac{1}{2}e^{z^*(r)-\eta}\right) + \frac{1}{2}e^{-\theta-\eta}\frac{1}{2}e^{z^*(l)-\eta}.\end{aligned}$$

Thus

$$\eta_R\left(z^*(r), z^*(l)\right) = \log\left(\frac{e^\theta + e^{z^*(r)}}{2} + \frac{\sqrt{e^{2\theta} + e^{2z^*(r)} - e^{-\theta+z^*(l)}}}{2}\right).$$

Similary, if $\eta_L\left(z^*(r), z^*(l)\right) \leq \min\left\{-\theta, z^*(r), z^*(l)\right\}$, then

$$\eta_L\left(z^*(r), z^*(l)\right) = -\log\left(\frac{e^\theta + e^{-z^*(l)}}{2} + \frac{\sqrt{e^{2\theta} + e^{-z^*(l)} - e^{-\theta-z^*(r)}}}{2}\right).$$

And

$$\begin{aligned}&p_R'\left(\eta_R\right) - p_M'\left(\eta_R\right) \\ =\ &2F\left(\eta_R - z^*(r)\right)f\left(\eta_R - \theta\right) + 2f\left(\eta_R - z^*(r)\right)F\left(\eta_R - \theta\right) \\ &-f\left(-\theta-\eta\right)F\left(z^*(l) - \eta\right) - F\left(-\theta-\eta\right)f\left(z^*(l) - \eta\right) \\ =\ &-4p_R\left(\eta_R\right) + 2F\left(\eta_R - z^*(r)\right) + 2F\left(\eta_R - \theta\right) - 2p_L\left(\eta_R\right) \\ =\ &2\left(1 - \frac{1}{2}e^{z^*(r)-\eta_R}\right) + 2\left(1 - \frac{1}{2}e^{\theta-\eta_R}\right) - 2 \\ =\ &1 - e^{z^*(r)-\eta_R} + 1 - e^{\theta-\eta_R} \\ =\ &e^{-\eta_R}\left(2e^{\eta_R} - e^{z^*(r)} - e^\theta\right) \\ =\ &e^{-\eta_R}\left(\sqrt{e^{2\theta} + e^{2z^*(r)} - e^{-\theta+z^*(l)}}\right).\end{aligned}$$

Likewise,

$$p_L'\left(\eta_L\right) - p_M'\left(\eta_L\right) = -e^{\eta_L}\left(\sqrt{e^{2\theta} + e^{-z^*(l)} - e^{-\theta-z^*(r)}}\right).$$

Thus

$$\begin{aligned}&\tau\left(z^*(r), z^*(l)\right) \\ =\ &\frac{\eta_R\left(z^*\right) + \eta_L\left(z^*\right)}{2} - \frac{1}{2}\left(\begin{array}{c}\log\frac{\left|p_R'\left(\eta_R(z^*);z^*\right) - p_M'\left(\eta_R(z^*);z^*\right)\right|}{\left|p_L'\left(\eta_L(z^*);z^*\right) - p_M'\left(\eta_L(z^*);z^*\right)\right|} \\ +\alpha\left(\left|\eta_R\left(z^*\right) - \eta_0\right| - \left|\eta_L\left(z^*\right) - \eta_0\right|\right)\end{array}\right) \\ =\ &\left(\eta_R\left(z^*\right) + \eta_L\left(z^*\right)\right) - \frac{1}{2}\log\frac{\sqrt{e^{2\theta} + e^{2z^*(r)} - e^{-\theta+z^*(l)}}}{\sqrt{e^{2\theta} + e^{-z^*(l)} - e^{-\theta-z^*(r)}}} \\ &+\frac{1}{2}\alpha\left(\left|\eta_R\left(z^*\right) - \eta_0\right| - \left|\eta_L\left(z^*\right) - \eta_0\right|\right).\end{aligned}$$

Using equation (2), we get

$$\begin{aligned}
\frac{\partial \eta_R(z^*)}{\partial z^*(r)} &= -\frac{2\frac{\partial p_R(\eta;z^*)}{\partial z^*(r)} + \frac{\partial p_L(\eta;z^*)}{\partial z^*(r)}}{2p_R'(\eta_R(z^*), z^*) + p_L'(\eta_R(z^*), z^*)} \\
&= -\frac{-2f(\eta_R - z^*(r)) F(\eta_R - \theta) + 0}{e^{-\eta_R}\left(\sqrt{e^{2\theta} + e^{2z^*(r)} - e^{-\theta+z^*(l)}}\right)} \\
&= \frac{2(1 - F(\eta_R - z^*(r))) F(\eta_R - \theta)}{e^{-\eta_R}\left(\sqrt{e^{2\theta} + e^{2z^*(r)} - e^{-\theta+z^*(l)}}\right)} \\
&= \frac{2F(\eta_R - \theta) - 2p_R(\eta_R(z^*); z^*)}{e^{-\eta_R}\left(\sqrt{e^{2\theta} + e^{2z^*(r)} - e^{-\theta+z^*(l)}}\right)} \\
&= \frac{2\left(1 - \frac{1}{2}e^{\theta-\eta_R}\right) - (1 - p_L(\eta_R(z^*); z^*))}{e^{-\eta_R}\left(\sqrt{e^{2\theta} + e^{2z^*(r)} - e^{-\theta+z^*(l)}}\right)} \\
&= \frac{1 - e^{\theta-\eta_R} + p_L(\eta_R(z^*); z^*)}{e^{-\eta_R}\left(\sqrt{e^{2\theta} + e^{2z^*(r)} - e^{-\theta+z^*(l)}}\right)} \\
&= \frac{\frac{e^{z^*(r)} - e^{\theta}}{2} + \frac{\sqrt{e^{2\theta} + e^{2z^*(r)} - e^{-\theta+z^*(l)}}}{2}}{\sqrt{e^{2\theta} + e^{2z^*(r)} - e^{-\theta+z^*(l)}}} + \frac{\left(\frac{e^{z^*(r)} + e^{\theta}}{2} + \frac{\sqrt{e^{2\theta} + e^{2z^*(r)} - e^{-\theta+z^*(l)}}}{2}\right) p_L(\eta_R(z^*); z^*)}{\sqrt{e^{2\theta} + e^{2z^*(r)} - e^{-\theta+z^*(l)}}} \\
&= \frac{1}{2}\left(1 + \frac{e^{z^*(r)} - e^{\theta}}{\sqrt{e^{2\theta} + e^{2z^*(r)} - e^{-\theta+z^*(l)}}}\right) + \frac{1}{2}\left(1 + \frac{e^{z^*(r)} + e^{\theta}}{\sqrt{e^{2\theta} + e^{2z^*(r)} - e^{-\theta+z^*(l)}}}\right) p_L(\eta_R(z^*); z^*).
\end{aligned}$$

Using the same method, we get

$$\frac{\partial \eta_R(z^*)}{\partial z^*(l)} = -\frac{1}{2}\left(1 + \frac{e^{\theta} + e^{z^*(r)}}{\sqrt{e^{2\theta} + e^{2z^*(r)} - e^{-\theta+z^*(l)}}}\right) p_L(\eta_R(z^*)).$$

Because $-z^*(r) + z^*(l) = \frac{u_r + u_l}{2} < \theta$, $\sqrt{e^{2\theta} + e^{2z^*(r)} - e^{-\theta+z^*(l)}} > \left|e^{z^*(r)} - e^{\theta}\right|$,
and thus

$$\frac{d\eta_R\left(a - \frac{u_r}{2}, a + \frac{u_l}{2}\right)}{da} = \frac{1}{2}\left(1 + \frac{e^{z^*(r)} - e^{\theta}}{\sqrt{e^{2\theta} + e^{2z^*(r)} - e^{-\theta+z^*(l)}}}\right) \in (0, 1).$$

We obtain

$$\begin{aligned}
\frac{\partial \eta_L(z^*)}{\partial z^*(l)} &= \frac{1}{2}\left(1 + \frac{e^{-z^*(l)} - e^{\theta}}{\sqrt{e^{2\theta} + e^{-z^*(l)} - e^{-\theta-z^*(r)}}}\right) \\
&\quad + \frac{1}{2}\left(1 + \frac{e^{\theta} + e^{-z^*(l)}}{\sqrt{e^{2\theta} + e^{-z^*(l)} - e^{-\theta-z^*(r)}}}\right) p_R(\eta_L(z^*))
\end{aligned}$$

and

$$\frac{\partial \eta_L(z^*)}{\partial z^*(r)} = -\frac{1}{2}\left(1 + \frac{e^{\theta} + e^{-z^*(l)}}{\sqrt{e^{2\theta} + e^{-z^*(l)} - e^{-\theta-z^*(r)}}}\right) p_R(\eta_L(z^*)).$$

With some algebra, we get

$$\tau'(a) = \frac{1}{2}\left\{\begin{array}{c} \frac{1}{2}+\frac{e^{a-\frac{u_r}{2}}-e^{\theta}}{\sqrt{e^{2\theta}+e^{2a-u_r}-e^{-\theta+a+\frac{u_l}{2}}}}\left(1-\frac{\frac{e^{a-\frac{u_r}{2}}+e^{\theta}}{2}}{\sqrt{e^{2\theta}+e^{2a-u_r}-e^{-\theta+a+\frac{u_l}{2}}}}\right) \\ +\frac{1}{2}+\frac{e^{-z^*(l)}-e^{\theta}}{\sqrt{e^{2\theta}+e^{-2a-u_l}-e^{-\theta-\left(a-\frac{u_r}{2}\right)}}}\left(1-\frac{\frac{e^{\theta}+e^{-z^*(l)}}{2}}{\sqrt{e^{2\theta}+e^{-2a-u_l}-e^{-\theta-\left(a-\frac{u_r}{2}\right)}}}\right) \end{array}\right\}$$
$$+\frac{1}{2}\alpha\left(\frac{d}{da}\left|\eta_R\left(a-\frac{u_r}{2},a+\frac{u_l}{2}\right)-\eta_0\right|-\frac{d}{da}\left|\eta_L\left(a-\frac{u_r}{2},a+\frac{u_l}{2}\right)-\eta_0\right|\right).$$

It suffices to find an uniform upper bound below 1 for the term in $\{\}$. If either $z^*(r) < \theta$, or $z^*(l) > -\theta$, then the term with $\{\}$ is less than $\frac{3}{4}$. Because the two thresholds are moved up and down by the same number a, when $z^*(r) \geq \theta$ and $z^*(l) \leq -\theta$, it must be the case that $z^*(r) \in \left[\theta, -\theta - \frac{u_r+u_l}{2}\right]$ and $z^*(l) \in \left[\theta + \frac{u_r+u_l}{2}, -\theta\right]$. Because the expression is continuous in z^* and belongs to $(0,1)$, it must be no bigger than the maximum of this expression on this compact subset.

We conclude the proof showing that the unique fixed point to $\tilde{\tau}(.)$ satisfies inequality (1).

3.3.6 Existence of A Converging Sequence of Equilibria

Lastly, I construct a sequence of finite equilibria that converge pointwise to the unique fixed point v^*. To do so, I show that for sufficiently high n, the best response correspondence in G_n has a fixed point in strategies in a neighborhood of v^*.

Theorem 3 *There exists a sequence* $\{(v_{n,r}, v_{n,l})\}_{n=N}^{\infty}$ *which converge to* v^* *pointwise where* v_n *is an equilibrium in the game* G_n*.*

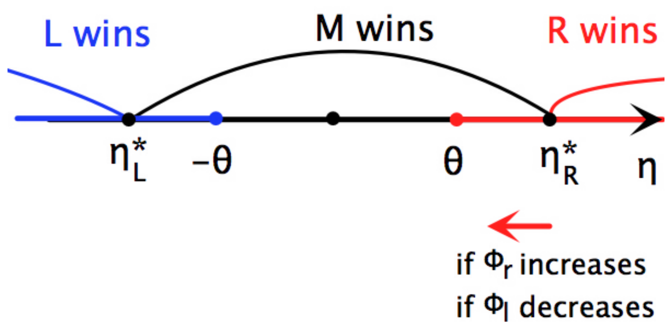
3.4 Comparative Statics

Let $\eta_c^*(u_r, u_l, \theta, \alpha) = \eta_c(a^*(u_r, u_l, \theta, \alpha))$ for $c \in \{R, L\}$. Here we analyse the comparative statics of the unique limit equilibrium.

3.4.1 Comparative Statics of Strategic Voting Equilibrium

We first see that when $\eta_0 = 0$, $a^{**}(u_r - u_l) < 0$. So the camp that has stronger love-hate ratio, i.e. the camp that is less afraid of cross-camp coordination failure, will vote more sincerely in equilibrium, and which gives an additional advantage to them, which in turn make them vote more sincerely. Because η_0 just moves the limit likelihood ratio function up and down, $\frac{\partial a^*}{\partial \eta_0} < 0$.

Proposition 1 *If* $\theta > \underline{\theta}$ *and* $\alpha < \overline{\alpha}$, *then* $\frac{\partial\left|\eta_j^*(u_r,u_l,\theta,\alpha)\right|}{\partial u_j} < 0$ *and* $\frac{\partial\left|\eta_j^*(u_r,u_l,\theta,\alpha)\right|}{\partial u_k} > 0$ *for* $j \neq k$ *and* $j,k \in \{L, R\}$*.*



η_R^* decreases with right wing voters' love-hate ratio u_r and increases with left-wing voters love-hate ratio u_l. In other words, the prior probability that R wins the election increases with u_r and decreases with u_l. This is true for preference intensities that are not very strong nor too weak.

When left-wing voters' love-hate ratio u_l goes up, there are two off-setting effects. First, this will increase the information threshold for right-wing voters and thus decrease the probability that a right wing voter votes for R by increasing the fixed point a^*. On the other hand, given the same a, this will decrease the information threshold for left-wing voters, and this will also decrease equilibrium a^*. A stronger left-wing force will eat into the voter base for M, and improves the prospect of R w.r.t. M. When u_l is not too big , the former force dominates.

$\eta_R^*(u, u, \theta, \alpha)$ is decreasing in u and increasing in θ. The ex ante probability that over-coordination occurs, i.e. the ex ante probability that $\eta \in (\theta, \eta_R^*)$ or $(-\eta_L^*, -\theta)$, decreases with θ. Here θ should be viewed as precision of private information.

4 Sequential v.s. Simultaneous Election

4.1 Model

The electorate consists of three states, state 1,2,3, or say, NH, MI and CA. The candidate that wins most states wins the election. In case of a tie between 2 or 3 candidates, the winner is determined by a random draw among those that tie for the first place. The winner within a state is determined also by plurality rule as described in the previous section. Voter i is state k is right-wing with probability $F(\eta_k - \theta_k)$ and left-wing with probability $F(-\eta_k - \theta_k)$. Every voter shares the same prior that η_k's follow *i.i.d. Laplace* $(0, \alpha_k)$. In addition to the common prior, voter i in state k obtains an additional signal $\hat{\eta}_i$ about η_k where $\hat{\eta}_i \sim$ *Laplace* $(\eta_k, 1)$. The independence of η_k's across states implies that there is no learning when voting takes place sequentially. This allows me to focus on the coordination effect of sequential voting. θ_k's and α_k's are common knowledge among voters in every state. Let G_k be the prior distribution of η_k.

Let v_{oc} denote the payoff to voter of ideology type o in state k when candidate c wins the election. We will look at the symmetric case where $v_{rR} = v_{lL} > v_{rM} = v_{lM} > v_{rL} = v_{lR}$ and $v_{mM} > v_{mL} = v_{mR}$. Define

$$\phi_k = \frac{v_{rRk} - v_{rMk}}{v_{rMk} - v_{rLk}}$$

and $u = \log 2\phi_k$. We call ϕ_k the extreme voters' love-hate ratio for their favorite candidate. Because right wing and left-wing voters both have love-hate ratio ϕ_k, the threshold a^* is 0 no matter how big ϕ_k is.

4.2 Sequential Election

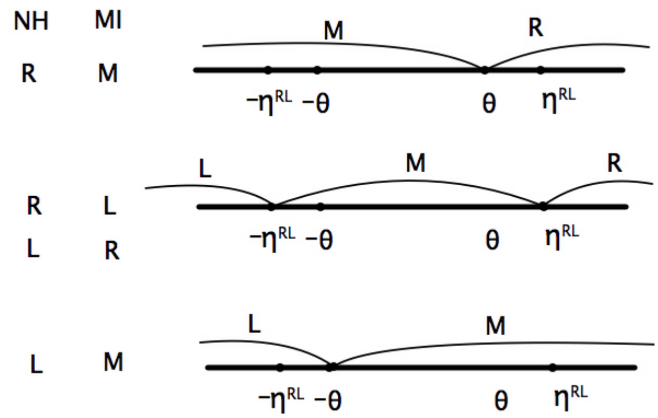
This section analyzes equilibria in a sequential election and illustrate the coordination effect. We only look at the election where $\phi < \frac{1}{2}$. In such elections, coordination is important because the payoff difference between the second and the least favorite candidate is more than twice that of the first and the second favorite candidate.

4.2.1 Voting in the last state, CA

It is weakly dominant for a moderate voter to vote for M. Given any voting strategy in which m always votes for M, the probability that candidate R ties with L vanishes more quickly than the probability that candidate L ties with M. Therefore, voter i only weighs between the probability of an $R-M$ tie and the probability of an $M-L$ tie.

When candidate L and candidate M each wins one state, then a right wing voter's payoff when candidate c wins the third state is given by $U_{rR} = \frac{v_{rR}+v_{rM}+v_{rL}}{3}$, $U_{rM} = v_{rM}$ and $U_{rL} = v_{rL}$. When $\phi < 1$, $U_{rM} - U_{rL} = \frac{(v_{rR}-v_{rM})-(v_{rM}-v_{rL})}{3} < 0$. Therefore, in both an $R-M$ tie and an $M-L$ tie, a right wing voter prefers to vote for M. Therefore, in all weakly undominated equilibria, a right wing voter votes for M. Thus the last state is a runoff between L and M. L wins the last state and the election if $\eta_3 < -\theta$ and M wins the last state and the election if $\eta_3 > -\theta$.

When candidate L and R each wins one state, $U^{LR}_{rR} - U^{LR}_{rM} = \frac{u_{rR}-u_{rM}+u_{rR}-u_{rL}}{3}$ and $U_{rM} - U_{rL} = \frac{u_{rR}-u_{rL}+u_{rM}-u_{rL}}{3}$. Therefore, the love-hate ratio for the last-state election, denoted by ϕ^{RL}, is equal to 1. Thus, the equilibrium in the subgame after R and L split the first two states gives rise to the two cutoff points $\eta^*_R(1, 1, \alpha, \theta)$ and $\eta^*_L(1, 1, \alpha, \theta)$. Because $\theta > \frac{1+1}{4}$, $\eta^*_R(1, 1, \alpha, \theta) > \theta$.



4.2.2 Voting in the second state, MI.

In this section we will show how the cutoff points on η_{MI} for different voting outcomes in state 2 depends on the voting outcome in New Hampshire, the first primary. In particular, we will show that when love-hate ratio for the overall election is moderate, probability that candidate R wins Michigan increases as the winner of New Hampshire changes from L to M to R. In particular, we will analyze how η_R^h changes with h, where $h \in \{R, M, L\}$ is winner in New Hampshire and η_R^h is the lower bound on η_2 for candidate R to win the second state.

Given the voting outcome $h \in \{R, M, L\}$ of state 1, the final election outcome depends on the voting outcome of state 2 and the electoral preference of state 3, η_3. Figure illustrates how the election outcome depends on the voting outcomes of the first two states and η_3.

Consider the voting game in state 2 after candidate R wins the first state. State 2's voting outcome is pivotal only when M will win state 3, i.e. $\eta_3 < \theta$. Therefore, we get $U_{rR} - U_{rM} = G_{CA}(\theta_{CA})(v_{rR} - v_{rM})$, where G_{CA} is the cumulative distribution function of the prior on η_{CA}. But the payoff difference when M wins state 2 v.s. when L wins state 2 gets even smaller. Therefore, we get

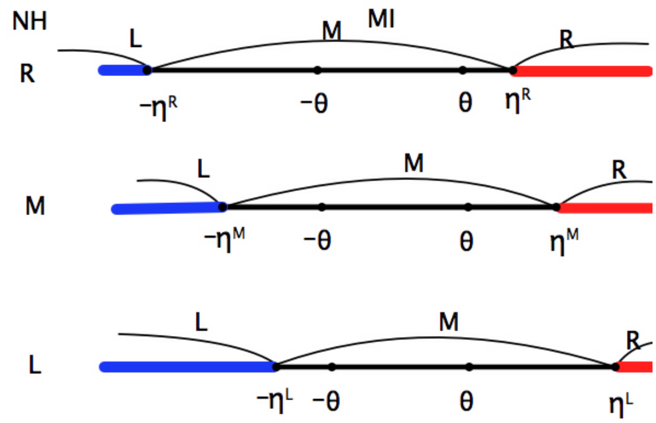
$$\begin{aligned}
\phi_r^R &= \frac{G(\theta)\phi}{\frac{1}{2}-\frac{G(-\theta,\theta)}{2}\phi-\frac{G(-\theta,\theta)}{6}(1-\phi)-\frac{G\left(\theta,\eta_R^{RL}\right)}{3}(1-\phi)} \\
&= \frac{G(\theta)\phi}{\frac{1}{2}\left(1-\frac{1}{3}G(-\theta,\theta)-\frac{2G\left(\theta,\eta_R^{RL}\right)}{3}\right)-\frac{1}{3}\left(G(-\theta,\theta)-G\left(\theta,\eta_R^{RL}\right)\right)\phi} \\
&= \frac{G(\theta)\phi}{\frac{1}{3}\left(\frac{3}{2}-\frac{G(-\theta,\theta)}{2}-G\left(\theta,\eta_R^{RL}\right)\right)-\frac{1}{3}\left(G(-\theta,\theta)-G\left(\theta,\eta_R^{RL}\right)\right)\phi} \\
&= \frac{G(\theta)\phi}{\frac{1}{3}\left(1+G\left(\eta_R^{RL},\infty\right)\right)-\frac{1}{3}\left(G(-\theta,\theta)-G\left(\theta,\eta_R^{RL}\right)\right)\phi} \\
&= \frac{G(\theta)\phi}{\frac{1}{3}\left(2G(\theta,\infty)+G\left(\theta,\eta_R^{RL}\right)+G\left(\eta_R^{RL},\infty\right)\right)+\frac{1}{3}\left(G(-\theta,\theta)-G\left(\theta,\eta_R^{RL}\right)\right)(1-\phi)} \\
&= \frac{G(\theta)\phi}{G(\theta,\infty)+\frac{1}{3}\left(G(-\theta,\theta)-G\left(\theta,\eta_R^{RL}\right)\right)(1-\phi)},
\end{aligned}$$

where $\phi=\phi_{MI}$ is the inherent love-hate ratio of extreme voters in Michigan and $G=G_{CA}$, $\theta=\theta_{CA}$. So a win by R boosts the love-hate ratio of right-wing voters in the second state. The ratio $\frac{\phi_r^R}{\phi}$ is higher the weaker the general love-hate ratio is, and the less likely an extreme candidate will win state 3. Because the game is symmetric, $\phi_l^L=\phi_r^R$. ϕ_r^R is different from the payoff difference ratio in a simultaneous election conditional on one state being taken by candidate R. Conditional on one state being R, a R-win or an M-win makes a difference when state 3 is taken by either M or L. But in a sequential election, L never wins state 3 if R wins state 1 and state 2 is taken by either R or M. In other words, voting outcome in the first two states can change a left-wing state from being taken by L to being taken by M.

Consider the voting game in state 2 after L wins the first state. We get that

$$\begin{aligned}
\phi_r^L &= \frac{\frac{1}{2}\phi-\frac{G(-\theta,\theta)}{2}+\frac{G(-\theta,\theta)}{6}(1-\phi)+\frac{G\left(\theta,\eta_R^*(1,1)\right)}{3}(1-\phi)}{G(\theta)} \\
&= \frac{1}{3}\frac{\left(1+G\left(\eta_R^{RL},\infty\right)\right)\phi-\left(G(-\theta,\theta)-G\left(\theta,\eta_R^{RL}\right)\right)}{G(\theta)}.
\end{aligned}$$

If M wins the first state, $\phi_r^M=\phi$. This is because L will not win California unless L wins Michigan. Thus when comparing expected payoff from R being the winner in Michigan and expected payoff from M being the winner, a voter does not need to consider a $M-L$ tie in other states. In other words, when Michigan's vote matters, winner in Michigan is winner of the election. Because whether R or M wins Michigan matters when at least half of the population in California is left-wing, and whether L or M wins MI matters when California is right-wing. That the prior probability of California being left-wing or right-wing implies that within-state love-hate ratio is equal to inherent pereference intensity.



Therefore, we see that ϕ_r^h increases as h changes from L to M to R. Right wing voters' love-hate ratio for the voting outcome in the second state is higher the closer the voting outcome in the first state is to their preferred choice. This is not surprising because when $\phi < 1$, $U_R - U_M$ is highest when there is an $R-L$ tie, second when there is an $R-M$ tie, but negative conditional on an $M-L$ tie. Conditional on R winning NH, an $M-L$ tie between NH and CA is ruled out, therefore boosting the payoff difference between a victory by R and a victory by M, while reducing the payoff difference between a victory by M and a victory by L. If L wins the first state, then voting for R is very risky: it is good in an $R-L$ tie but bad in an $M-L$ tie. So $\eta_R^L = \infty$ if

$$\phi \leq \frac{G\left(-\theta,\theta\right) - G\left(\theta,\eta_R^*\left(1,1\right)\right)}{1 + G\left(\theta,\infty\right) - G\left(\theta,\eta_R^*\left(1,1\right)\right)}$$

This term is increasing in θ. It is sufficient if $\phi < G_k\left(-\theta_k,\theta_k\right)$, the prior probability that a voter in California is moderate.

Proposition 2 *If* $u < 0$,$\alpha < \frac{1}{4}$, *and* $\theta > \min\left\{-\frac{u}{2} + \log 2, \frac{3}{2}\right\}$, *then* $\eta_R^R\left(u,\theta,\alpha\right) < \eta_R^M\left(u,\theta,\alpha\right) < \eta_R^L\left(u,\theta,\alpha\right)$.

This follows immediately from Proposition 1 because right wing voters' within-state love-hate ratio increases while left-wing voters' within-state love-hate ratio decreases as the winner of New Hampshire changes from L to M to R.

Note that η_R^R is still greater than θ. So the within-camp coordination problem still exists in the primary of Michigan. But this problem is less severe when the the camp's favorite candidate wins New Hampshire and more severe when the camp's worst enemy wins New Hampshire. If we define $\frac{\phi^h}{\phi}$ as the degree of sensitivity of Michigan's within-state love-hate ratio w.r.t. history, then the following proposition says that the larger moderate population in California is, the more sensitive MI's love-hate ratio is to history. On the other hand, the stronger MI extreme voters' inherent love-hate ratio is, the more sensitive their within-state love-hate ratio is to good news, i.e. to the history where their favorite candidate wins NH, but the less sensitive their within-state love-hate ratio is to bad news, i.e. to the history where their worst enemy wins N.

Lemma 4.1 $\log \frac{\phi^R}{\phi}$ *increases with both* $G_{CA}(\theta_{CA})$ *and* ϕ*, while* $\left|\log \frac{\phi^L}{\phi}\right|$ *increases with* $G_{CA}(\theta_{CA})$ *but decreases with* ϕ*.*

4.2.3 Voting in the first state, NH

Because the final outcome depends on the voting result in MI and CA, eg. when M and L splits MI and CA, victory by R in NH results in a random draw between all three candidates, while victory by M in NH results in a solid victory by M in the final election, the expected payoff difference between a victory by R and M in NH is a linear combination of the payoff difference from final election outcome between R and M and M and L, i.e. $v_R - v_M$ and $v_M - v_L$. Roughly speaking, a victory by R instead of M may change the final winner from R to M, from M to L, from L to M. Because R and L are symmetric in every state,

$$\phi_r^{\emptyset} = \frac{\phi - c^{\emptyset}}{1 - c^{\emptyset}\phi}$$

where

$$\begin{aligned} c^{\emptyset} &= \frac{\frac{1}{3}P(m)F\left(-\infty,\eta_L^R\right) - \frac{1}{3}F\left(\theta,\eta_R^*(1,1)\right)F\left(\eta_L^R\right) - F(-\theta)F\left(\eta_L^R,\eta_L^M\right)}{\begin{array}{c} P(m)F\left(\eta_R^R,\infty\right) + P(m)P(r) + 2P(r)P(l) + \frac{1}{3}P(m)F\left(\eta_L^R\right) \\ -\frac{1}{3}F\left(\theta,\eta_R^*(1,1)\right)F\left(\eta_L^R\right) + F\left(\eta_R^R,\eta_R^M\right)P(r) \end{array}} \\ &= \frac{1}{2}\frac{\frac{2}{3} - \frac{2}{3}\frac{F(\theta,\eta_R^*(1,1))}{P(m)} - 2\frac{F\left(\eta_R^M,\eta_R^L\right)}{P^L(R)}\frac{P(r)}{P(m)}}{\frac{P^R(R)}{P^L(R)} + \frac{P(r)}{P^L(R)} + \frac{1}{2} - \frac{1}{6} - \frac{1}{3}\frac{F\left(\theta,\eta_R^*(1,1)\right)}{P(m)} + 2\frac{P(r)}{P^L(R)}\frac{P(r)}{P(m)} + \frac{F\left(\eta_R^R,\eta_R^M\right)}{P^L(R)}\frac{P(r)}{P(m)}}. \end{aligned}$$

Because $\phi < 1, \phi_r^{\emptyset}$ is decreasing in $c^{\emptyset}$.

Outcome in the first state can change outcome in the second state and/or outcome in state 3. The reason a right-wing voter may strategically vote for M instead of her favorite candidate R is for fear of a tie between M and L and getting L elected instead of M in that situation. Roughly speaking, M and L tie in the overall election when one of the other two states is moderate and the other is left-wing. But when R wins the first state, and M wins the second state, no one votes for L in the third state and M will win the third state and the final election even if the median voter in CA is left-wing.

When $\phi < 1$, an extreme voter worry quite a lot about failing to coordinate with a moderate state and letting L win the election. Note that for the second state, after one victory by M, a victory by R ensures that L cannot win the election. Therefore, for NH, if the effect from changing MI from L to M is small, then the love-hate ratio for NH voters is smaller than that for MI voters when M wins NH. That is, $\eta_R^{\emptyset} > \eta_R^M$. But if the effect of changing MI from L to M is big, then $\eta_R^{\emptyset} < \eta_R^M$.

4.2.4 Why Does New Hampshire want to vote first?

Does a median voter in NH prefer to vote first or second in a sequential primary? That is, does the median voter in NH prefer to vote first or to switch order with Michigan? This depends on the distribution of preferences in NH and MI. If the median voter in NH and MI are both right-wing, then NH's median voter weakly prefers the more aggressive state to vote first. If NH and MI are ex ante identical, conditional on the super majority in both states being of the same camp, payoff does not depend on whether NH swaps order with MI. If NH is only mildly right-wing and MI is moderate, then whether NH votes first or after knowing that M wins MI may change the identity of the winner in NH. More specifically, if η^{NH} is between η_R^M and $\eta_R^{\emptyset}$ and M wins MI regardless of order, then voting first makes right-wing voters behave more aggressively if $\eta_R^{\emptyset} < \eta_R^M$, while voting after being assured that M wins MI makes them behave more aggressively if $\eta_R^{\emptyset} > \eta_R^M$. Thus, conditional on NH being mildly right-wing and MI being moderate, median voter in NH prefers to vote first if and only if $\eta_R^{\emptyset} < \eta_R^M$.

However, if MI is of the opposite camp from NH, then the median voter in NH definitely prefers to vote first. This is because the winner in MI will be M instead of L if NH votes first and R wins NH, which makes the final election outcome more favorable to NH's median voter, or because NH is not right wing enough and thus voting after MI implies voting after knowing that L has won MI, which makes right wing voters in NH more conservative and results in a victory by M instead of R in NH even though the super majority in NH prefers R.

Thus if $\eta_R^{\emptyset} > \eta_R^M$, conditional on the median voter in NH being right wing, whether she prefers that NH votes first or MI votes first depends on the relative probability of $\left\{\left(\eta^{NH}, \eta^{MI}\right) \in \left(\eta_R^M, \eta_R^{\emptyset}\right) \times \left(\eta_L^{\emptyset}, \eta_R^R\right)\right\}$ and $\left\{\left(\eta^{NH}, \eta^{MI}\right) \in \left(\eta_R^{\emptyset}, \infty\right) \times \left(\eta_L^{\emptyset}, \eta_L^R\right) \cup \left(\eta_R^{\emptyset}, \eta_R^L\right) \times \left(\eta_L^R,\right.\right.$

If the median voter in NH is moderate, then they prefer to vote first if and only if knowing that M wins NH tampers the behavior of extreme voters in MI and makes a victory by M more likely. That is, if median voter in NH is moderate, she prefers to vote first if and only $\eta_R^{\emptyset} < \eta_R^M$.

Thus we can conclude that if $\eta_R^{\emptyset} < \eta_R^M$, expected payoff from voting first is weakly higher than that from voting second for any η^{NH}. That is, voting first is unambigiously better for voters in New Hampshire when $\eta_R^{\emptyset} < \eta_R^M$. $\eta_R^{\emptyset} < \eta_R^M$ when the effect of influencing the second state's behavior is sufficiently large. Thus voting first is unambigiously better than voting second if inherent love-hate ratio is small. For example, when ϕ is smaller than the probability of a moderate voter in California, winnowing happens in the second state and thus winning the first state is necessary for an extreme candidate to win the election. Therefore, voting first is better than second.

4.3 Simultaneous (Front-loaded) Election

The payoff difference to voter i in state k when candidate c wins state k v.s. candidate c' depends on how the voting outcome in state k affects the election outcome. We will focus on symmetric equilibria in which every voter in very state use the same voting strategy. Suppose voters in the other two states use voting strategy s such that R wins state k if $\eta_k > \tilde{\eta}$ and L wins state k if $\eta_k < -\tilde{\eta}$. Then the probability that R wins state k is $G(-\tilde{\eta})$. Denote by $p^F(c)$ the probability that candidate c wins a state. This vector of probabilities depend on the voting strategy s employed and is determined by $\tilde{\eta}$.

$$\begin{aligned} U_R^F - U_M^F &= P^F(R) P^F(M)(v_{cR} - v_{cM}) + P^F(R) P^F(L) \frac{(v_{cR} - v_{cM}) + (v_{cR} - v_{cL})}{3} \\ &\quad + P^F(M) P^F(L) \frac{(v_{cR} - v_{cM}) - (v_{cM} - v_{cL})}{3} \\ &= \left(P^F(R) P(M) + \frac{2}{3} P(R) P(L) + \frac{1}{3} P(M) P(L) \right) (v_{cR} - v_{cM}) \\ &\quad - P^F(L) \left(P^F(M) - P^F(R) \right) (v_{cM} - v_{cL}) . \end{aligned}$$

Because the game is symmetric and we are looking for symmetric equilibria, $P^F(R) = P^F(L)$ and we get

$$\begin{aligned} \phi^F &: = \frac{U_R^F - U_M^F}{U_M^F - U_L^F} \\ &= \frac{\phi - c^F}{1 - c^F \phi} \end{aligned}$$

where

$$\begin{aligned} c^F &= \frac{\frac{2}{3} P^F(M) P^F(L) - \frac{2}{3} P^F(R) P^F(L)}{2 P^F(R) P^F(M) + 2 \times \frac{2}{3} P^F(R) P^F(L) + 2 \times \frac{1}{3} P^F(M) P^F(L)} \\ &= \frac{\frac{1}{3} \left(P^F(M) - P^F(R) \right)}{\frac{4}{3} P^F(M) + \frac{2}{3} P^F(L)} \\ &= \frac{1}{2} \frac{1 - \frac{P^F(R)}{P^F(M)}}{2 + \frac{P^F(L)}{P^F(M)}} . \end{aligned}$$

Note that c^F is a function of $\tilde{\eta}$, and thus u^F is a function of u and $\tilde{\eta}$.

Given that voters in the other two states use symmetric voting strategy v characterized by $\tilde{\eta}$, love-hate ratio for voting outcome of the state is given by $u^F(u, \tilde{\eta})$. Because the game within the state is symmetric, $a^* = 0$. In this equilibrium, an extreme voter votes for her favorite candidate if her signal $\hat{\eta}_i > -\frac{u^F(u,\tilde{\eta})}{2}$. Note that when $\phi < 1$, $\phi^F(\phi, \tilde{\eta}) < \phi$ if and only if $\frac{P^F(R)}{P^F(M)} < 1$. Therefore, in a symmetric equilibrium, the cutoff for R to win a state is $\eta_R^F(u, \theta, \alpha) > \eta_R^*(u, u, \theta, \alpha)$. Define $\eta^F(\tilde{\eta}; u, \theta, \alpha) = \eta_R^*\left(u^F(u, \tilde{\eta}), u^F(u, \tilde{\eta}), \theta, \alpha\right)$.

$\eta^F(\tilde{\eta})$ is increasing for $\tilde{\eta} \geq \eta_R^*(u, u, \theta, \alpha)$ and $\eta^F(\eta_R^*(u, u, \theta, \alpha)) > \eta_R^*$. Define the fixed point to be ∞ when $\eta^F(\tilde{\eta}) > \tilde{\eta}$ for all $\tilde{\eta} > \eta_R^*(u, u, \theta, \alpha)$. Then η_R^F is a fixed point of the function. $\eta_R^F = \infty$ is a simultaneous voting equilibrium in which all voters vote for M.

4.4 Comparison between Sequential and Simultaneous Election

4.4.1 Voting Behavior in State 1 (NH) under sequential and front-loaded election

Comparing $\eta_R^{\emptyset}$ and η_R^F is equivalent to comparing $c^{\emptyset}$ and c^F. When $\phi < 1$, $\eta_R^{\emptyset} < \eta_R^F$ if and only if $c^{\emptyset} < c^F$.

Proposition 3 *For θ big enough (for example when $P(r) < \frac{1}{4}$), or u small enough, voters in state 1 behave more aggressively under a sequential election than under a simultaneous election.*

It amounts to finding conditions under which $c^{\emptyset} < c^F$. If we ignore the effect of affecting other states voting behavior, we will be comparing c^F with

$$\frac{\frac{1}{3}P(m)\,F\left(-\infty, \eta_L^R\right)}{P(m)\,F\left(\eta_R^R, \infty\right) + P(m)\,P(r) + 2P(r)\,P(l) + \frac{1}{3}P(m)\,F\left(\eta_L^R\right)}.$$

Because NH voters choose without knowing the voting results of MI and CA in both systems, expected payoff difference between R-victory and M-victory depends on probability of $M-L$ tie, $R-L$ tie and $R-M$ tie in MI and CA. Payoff difference between an R-NH and M-NH is largest when R and L split MI and CA. More importantly, an $R-L$ tie offset worries about an $M-L$ tie. The more likely an $R-L$ tie is relatively to an $M-L$ tie, the higher love-hate ratio is. Because every state is ex ante identical, the more likely an extreme voter will win a state, the higher ϕ^F is. When NH votes in a sequential primary, an R-L tie does not happen. If MI is R and CA is won by L in a simultaneous election, then in a sequential election, left-wing voters in CA will coordinate with moderates and thus M will win CA instead of L, and thus an $R-L$ tie in simultaneous primary turns into an $R-M$ tie in a sequential primary. If MI is won by L and CA by R, then right-wing voters in CA coordinate with moderates and ensure a victory by M in CA instead of R if M wins NH. Again an $R-L$ tie turns into an $R-M$ tie. In other words, victory by M in NH forces voters in CA to coordinate with moderates in CA and in a sense with moderates in NH so that M wins the election instead of a random draw. If this channel is important, $\phi^F > \phi^{\emptyset}$. This channel is important if the probability that an extreme candidate wins a state in a simultaneous election is low.

r voters in NH worry about $M-L$ tie. If the probability of an $M-L$ tie is smaller, then love-hate ratio is bigger. An $M-L$ tie happens in a sequential primary with half of the probability of that in a simultaneous primary. This is because if MI is M and CA is L, a victory by R in NH forces left-wing voters in

CA to coordinate with moderates, which result in a sure victory by M instead of an $M-L$ tie. This channel increases love-hate ratio under a sequential primary relative to that under a simultaneous primary.

Which channel is more important depends on whether an $R-L$ tie or an $L-R$ is more likely or a $L-M$ tie. By symmetry, $\phi^{\emptyset}$ is higher if $P^F(R) < \frac{1}{2}P^F(M)$. Because voters behave too conservatively in a strategic voting equilibrium, this is true whenever ex ante share of an extreme voter is no bigger than half of that of a moderate voter. This explains why $\eta_R^{\emptyset} < \eta_R^F$ when θ is large.

In a sequential primary, who wins NH affects voting behavior in MI and CA. In particular, it affects voting behavior in MI. In particular, victory by R in NH makes it harder for L to win MI, thus making an $L-M$ tie less likely. This effect increases NH's love-hate ratio in a sequential primary. This effect is larger when intrinsic love-hate ratio in MI is smaller, or when the ex ante share of extreme preference voters is smaller.

4.4.2 Comparing election winner between Simultaneous and Sequential Primary

I will consider parameters such that extreme voters in New Hampshire behave more aggressively in a sequential election than in a simultaneous election, i.e. $\theta < \eta_R^{\emptyset} < \eta_R^F$. In this situation, $\eta_R^h > \theta$ for any history in both election systems. Therefore, M always wins a state whenever M is the condorcet winner in that state. Thus, if M is the condorcet winner in at least two states, M will win the election regardless of primary system. I then need to discuss only cases where either an extreme candidate is the Condorcet winner in at least two states, or the Condorcet winner is different in every state.

rrr: it seems straightforward that the a better system selects candidate R more often. The Condorcet winner in state k is R if and only if $\eta^k > \theta$, but the winner in state k is R if and only if $\eta^k > \eta_R^h$ where h is either F which indicates the simultaneous system or a history in the sequential system. R wins the election if R wins at least two states. Because $h \in \{\emptyset, R, M, RR, RM, MM\}$ given that $\eta^k > \theta$ for every state k, $\eta_R^h > \eta_R^F$. Therefore, if R wins in a simultaneous system, R wins in a sequential system, and there are $(\eta^{NH}, \eta^{MI}, \eta^{CA})$ such that R wins in a sequential system but M wins in a simultaneous system.

rrm or *rmr* or *mrr*: $h \in \{\emptyset, R, M, RR, RM, MR, MM\}$.

rrl: R wins if and only if R win both $r-$states. If R wins only one r-state and M wins the $l-$state, or if R wins no r-states, then M wins the election. If R wins an r-state in simultaneous election, then R wins that state in sequential election because $\eta_R^h < \eta_R^F$ for all possible histories an r state faces in a sequential election in this case. Therefore, whenever R wins a simultaneous election, R would win a sequential election. If R wins only one r-state and L wins the l-state, then all three candidates tie in the election and the outcome is a random draw among the three. Because $\phi < 1$, all voters prefer a sure victoyr by M to a random draw among every candidate. Thus this is the worst election outcome in this case. Because $\eta_L^h = -\infty$ for $h \in \{RM, MR\}$, when California's vote matters, left-wing voters there will not vote for L and L will not win California if the first two states are both right-wing. Therefore, a sequential election never produces the worst outcome, a three-way tie, while a simultaneous election may. So sequential election produces better outcome conditional on rrl.

rlr: R wins the election if and only if R win both r-states. The set of histories that an $r-$state may face in this case is $\{\emptyset, RM, RL, MM, ML\}$. Therefore, if an R wins an r-state in simultaneous election, R would win in a sequential election as well. Therefore, R wins a simultanoues election only if R wins a sequential election. If a sequential primary results in a three-way split, then the winner of each state in the order of voting must be RLM because R will never win California in a sequential primary after history ML. The outcome under sequential is worse than that under simtulaneous election if and only if the winner in order under sequential is RLM while that under simultaneous is MLM. The election outcome would change from $\frac{1}{3}R+\frac{1}{3}M+\frac{1}{3}L$, a random draw among all, to M when the system becomes simultaneous if $\left(\eta^{NH}, \eta^{MI}, \eta^{CA}\right) \in$

$$\left(\eta_R^{\emptyset}, \eta_R^F\right) \times \left(-\infty, \eta_L^R\right) \times \left(\theta, \eta_R^* (1)\right).$$

This is the only parameter range in this case where the outcome under sequential is worse than that under simultaneous. On the other hand, the outcome would change from R to $\frac{1}{3}R + \frac{1}{3}M + \frac{1}{3}L$ when the primary becomes simultaneous if $\left(\eta^{NH}, \eta^{MI}, \eta^{CA}\right) \in$

$$\left(\eta_R^{\emptyset}, \eta_R^F\right) \times \left(-\infty, \eta_L^R\right) \times \left(\eta_R^F, \infty\right).$$

Because R is the best outcome and a random draw among all is the worst, the second effect more than cancels out the first if $G_{CA}\left(\eta_R^F, \infty\right) < G_{CA}\left(\theta, \eta_R^* (1)\right)$. Because $G_{CA}\left(-\theta, \theta\right) > \frac{1}{3}$, this is true only if the probability that R will win California in a simultaneous election is less than $\frac{1}{6}$ or only if $G_{CA}\left(\theta, \eta_R^* (1)\right) > \frac{1}{6}$. But then, $MLM \rightarrow MLR$

$$\left(\theta, \eta_R^{\emptyset}\right) \times \left(-\infty, \eta_L^M\right) \times \left(\eta_R^F, \infty\right).$$

This effect cancels out the first if $G_{CA}\left(\eta_R^{\emptyset}, \eta_R^F\right) < G_{CA}\left(\eta_R^F, \infty\right)$.

The only problem is when $G_{CA}\left(\eta_R^F, \infty\right) < \min\left\{G_{CA}\left(\theta, \eta_R^* (1)\right), G_{CA}\left(\eta_R^{\emptyset}, \eta_R^F\right)\right\}$. Then $G_{CA}\left(\theta, \eta_R^F\right) > 2G_{CA}\left(\eta_R^F, \infty\right)$.

Suppose the parameters are such that extreme voters in New Hampshire behave more aggressively in a sequential primary than a simultaneous primary. Because M always wins a state whenever M is the condorcet winner in that state, while there is always too much coordination cross camp other than in the last primary in a sequential election after the camp's favorite candidate splits with M, conditional on the median voter in every state supports the same candidate, the universally favored candidate is the winner with higher probability in a sequential election than in a simultaneous election. In addition, sequential primary facilitates coordination across camp across states and thus a three-way split between all candidates is less likely to happen under a sequentail primary. When $\phi < 1$, every voter prefers a sure victory by M than a random draw among all candidates. Thus every voter prefers a sequential primary conditional on the simultanoues primary outcome being a three-way split. In general, if the median voter in two states support the same candidate, then sequential primary is preferred unless the candidate most preferred by median voters in the last two states, MI and CA, is an extreme candidate, say R, and the median voter in the NH supports the other extreme candidate, say L. In the latter situation, the eventual winner may be M instead of R if MI is the probability of right-wing voter is not high enough. This is due to the disproportionate impact of the winner in NH in a sequential primary.

4.4.3 For NH — Sequential or Simultaneous?

If voters in NH behave more aggresively in a sequential election, then if the median voter in NH is extreme, he must prefer sequential primary to simultaneous primary. In fact, even if voting behavior is less aggressive in a sequential election, as long as the difference is small, if median voter in NH is extreme, he still prefers sequential primary. This is because the voting outcome in NH may change voting outcome in MI and/or CA toward NH's median voter's preferred candidate. For example, if $\eta_1 > \max\left\{\eta_R^F, \eta_R^\emptyset\right\}$, then the voting outcome in NH is R regardless of primary system. This makes it harder for L to win MI than if the primary system is simultaneous. If $\eta^{MI} < \eta_L^R$, then moving to a sequential primary changes the voting outcome in MI from L to M, which futher changes voting outcome in CA from L to M thus final winner from L to M if $\eta^{CA} < \eta_L^F$. Suppose η_2 is such that outcome in MI does not depend on primary system either. Then it changes CA's voting outcome from L to M and final election outcome from a random draw among all three to M if $\eta^{CA} < \eta_L^F$ and the primary system is sequential intead of simultaneous.

If in addition to the effect of changing voting outcome in MI and/or CA from L to M or from M to R, voting behavior in NH is more aggressive in a sequential primary, then moving to a sequential primary changes winner in NH from M to R if $\eta^{NH} \in \left(\eta_R^\emptyset, \eta_R^F\right)$. That r voters in NH vote for R with positive probability in equilibrium indicates that expected payoff if R wins NH is higher than that if M wins NH.

If the median voter in NH is moderate, then the median voter prefers sequential primary if and only if

$$F\left(\eta_L^F\right)^2 - F\left(\eta_L^F, -\theta\right) F\left(\eta_L^M\right) - F\left(\eta_L^F, \eta_L^M\right) F\left(\eta_L^F\right) > 0.$$

This holds if

$$F\left(\eta_L^F\right) > \frac{\left(1+\sqrt{2}\right)}{2+\sqrt{2}} F\left(-\theta\right).$$

So if voters don't behave too conservatively in a simultaneous election, then if NH's median voter is moderate, he prefers sequential election. In a sequential election, the effect of forcing forcing left-wing voters to coordinate with NH's moderates when R wins MI makes sequential election preferable to a moderate voter in NH. However, if both MI and CA are extreme on the same side, eg. both left-wing, then left-wing voters in CA are much more aggressive in a sequential election because they are now sure of an $M-L$ tie. In addition, extreme voters in MI behave more aggressively when then know that M wins NH. This increases the probability of a final victory by an extreme candidate if the primary system is sequential. Which one is better for a moderate median voter depends on which happens with higher probability.

5 Conclusion

This paper studies preference aggregation in a multi-candidate contest when the preference of the electorate is not common knowledge. In a multi-candidate contest, voters have an incentive to coordinate with supporters of their second choice to avoid a victory by the least favorite candidate. I show that the coordination incentive is stronger when love-hate ratio is weaker. I then use this model as cornerstone to compare a simultaneous election in which several states vote at the same time and a sequential election in which each state votes one by one after observing outcomes of previous states. I show that when the prior probability of extreme voters is small or when the love-hate ratio of extreme voters is small, coordination incentives are stronger for extreme voters and thus they vote more aggressively in a sequential election than in a simultaneous election. As a result, the prior probability that the winner in a state is not the first choice of the median voter is smaller in a sequential election.

6 Appendix

6.1 Proof for lemma ??.

Proof. It suffices to show that

$$\lim_{N\to\infty} N\Pr\{V_R = V_M > V_L|\hat{\eta}_i, p\} = \frac{f(\eta_R|\hat{\eta}_i)}{|p_R'(\eta_R) - p_M'(\eta_R)|}.$$

Let

$$H^u = \{(V_R, V_M, V_L)\,|V_R = V_M > V_L \text{ where } V_c \geq 0 \text{ for } c = R, M, L\}$$

Then

$$\Pr\{V_R = V_M > V_L|\hat{\eta}_i, p\} = \int_{\eta=-\infty}^{\infty} P(H^u|N, p(\eta))\, f(\eta|\hat{\eta}_i)\, d\eta.$$

Let

$$H = \{(V_R, V_M, V_L)\,|V_R = V_M \text{ where } V_c \geq 0 \text{ for } c = R, M, L\}$$

and $H^* = \{(V_R, V_M, V_L)\,|V_R = V_M \text{ where } V_c \geq 0 \text{ for } c = R, M, L\}$. Then H is a hyperplane in $(N \cup \{0\})^3$ spanned by $w_1 = (1,1,0)$ and $w_2 = (0,0,1)$.

Given η, we first show that $y_N := \left(\left[N\sqrt{p_R(\eta)\,p_M(\eta)}\right], \left[N\sqrt{p_R(\eta)\,p_M(\eta)}\right], [Np_L(\eta)]\right)$ is a near maximizer $\sum_c p_c \psi\left(\frac{x(c)}{Np_c}\right)$ over x in H^* where $\psi(\theta) = \theta(1-\log\theta) - 1$. $H^* = \{\gamma(1,1,0) + j(0,0,1)\,|\gamma \geq 0 \text{ and } j \geq 0\}$. Let

$$(\gamma^*, j^*) \in \arg\max_{\gamma\geq 0, j\geq 0} \left(p_R\psi\left(\frac{\gamma}{Np_R}\right) + p_M\psi\left(\frac{\gamma}{Np_M}\right) + p_L\psi\left(\frac{j}{Np_L}\right)\right).$$

Because the derivative is ∞ for $\gamma = 0$ or $j = 0$ and the function goes to 0 as γ or $j \to \infty$, the solution must be interior of H^*. Thus γ^*, j^* satisfy the first order condition:

$$\begin{aligned} 0 &= -\log\frac{\gamma}{Np_R} - \log\frac{\gamma}{Np_M} \\ 0 &= -\log\frac{j}{Np_L}. \end{aligned}$$

So $\gamma^* = N\sqrt{p_R p_M}$ and $j^* = Np_L$. Then y_N as defined is a near maximizer.

$$\begin{aligned} & p_R\psi\left(\frac{\gamma^*}{Np_R}\right) + p_M\psi\left(\frac{\gamma^*}{Np_M}\right) + p_L\psi\left(\frac{j^*}{Np_L}\right) \\ = \; & p_R\left(\frac{\gamma^*}{Np_R}\left(1-\log\left(\frac{\gamma^*}{Np_R}\right)\right) - 1\right) \\ & + p_M\left(\frac{\gamma^*}{Np_M}\left(1-\log\left(\frac{\gamma^*}{Np_M}\right)\right) - 1\right) \\ & + p_L\left(\frac{j^*}{Np_L}\left(1-\log\left(\frac{j^*}{Np_L}\right)\right) - 1\right) \\ = \; & -1 + \frac{\gamma^*}{N}\left(1-\log\left(\frac{\gamma^*}{Np_R}\right) + 1 - \log\left(\frac{\gamma^*}{Np_M}\right)\right) \\ & + \frac{j^*}{N}\left(1-\log\left(\frac{j^*}{Np_L}\right)\right) \\ = \; & -1 + 2\frac{\gamma^*}{N} + \frac{j^*}{N} \\ = \; & 2\sqrt{p_R p_M} - (1-p_L) \\ = \; & 2\sqrt{p_R p_M} - p_R - p_M \\ = \; & -\left(\sqrt{p_R} - \sqrt{p_M}\right)^2. \end{aligned}$$

Then using theorem 3 in Myerson (2000),

$$\lim_{N\to\infty} \frac{\Pr\{H|Np(\eta)\}}{\Pr\{y_N|Np(\eta)\}(2\pi)(\det(M(y_N)))^{-0.5}} = 1$$

where $M(y_N(\eta)) = \begin{bmatrix} \frac{2}{\left[N\sqrt{p_R(\eta)p_M(\eta)}\right]} & 0 \\ 0 & \frac{1}{[Np_L(\eta)]} \end{bmatrix}$ and $\lim_{N\to\infty} N * M(y_N) =$

$\begin{bmatrix} \frac{2}{\sqrt{p_R(\eta)p_M(\eta)}} & 0 \\ 0 & \frac{1}{p_L(\eta)} \end{bmatrix}$. By Myerson (2000),

$$\begin{aligned}
\Pr\{y_N|Np(\eta)\} &\approx \frac{e^{N*\left(p_R\psi\left(\frac{\gamma^*}{Np_R}\right)+p_M\psi\left(\frac{\gamma^*}{Np_M}\right)+p_L\psi\left(\frac{j^*}{Np_L}\right)\right)}}{\Pi_{c\in\{R,M,L\}}\sqrt{2\pi y_N(c)}} \\
&= \frac{e^{-N\left(\sqrt{p_R}-\sqrt{p_M}\right)^2}}{(2\pi)^{\frac{3}{2}}\sqrt{(\gamma^*)^2 j^*}} \\
&= \frac{e^{-N\left(\sqrt{p_R}-\sqrt{p_M}\right)^2}}{(2N\pi)^{\frac{3}{2}}\sqrt{p_R p_M p_L}}.
\end{aligned}$$

$$\begin{aligned}
(\det(M(y_N)))^{-0.5} &\approx \left(\frac{1}{N^2\sqrt{p_R p_M}p_L}\right)^{-0.5} \\
&= N\sqrt{\sqrt{p_R p_M}p_L}.
\end{aligned}$$

So

$$\begin{aligned}
\Pr\{H^*|Np(\eta)\} &\approx \Pr\{y_N|Np(\eta)\}(2\pi)(\det(M(y_N)))^{-0.5} \\
&\approx N\sqrt{\sqrt{p_R p_M}p_L}(2\pi)\frac{e^{-N\left(\sqrt{p_R}-\sqrt{p_M}\right)^2}}{(2N\pi)^{\frac{3}{2}}\sqrt{p_R p_M p_L}} \\
&= \frac{e^{-N\left(\sqrt{p_R}-\sqrt{p_M}\right)^2}}{\sqrt{2\pi N}\sqrt{\sqrt{p_R p_M}}}.
\end{aligned}$$

Given $\varepsilon > 0$, let δ be such that $|p_R(\eta) - p_M(\eta)| \geq \varepsilon$ for all η such that $|\eta - \eta_R| \geq \delta$. Define $\Lambda_\delta := \{\eta : |\eta - \eta_R| < \delta\}$. Then want to show that $\lim_{N\to\infty} \frac{\Pr\{H|Np(\eta)\}}{\Pr(H^*|Np(\eta))} = 1$ for $\eta \in \Lambda_\delta$. Then show that $\lim_{N\to\infty} N\Pr\{H^*|Np(\eta)\} = 0$ for $\eta \notin \Lambda_\delta$. Then

$$\begin{aligned}
&\lim_{N\to\infty} N\Pr\{V_R = V_M > V_L|\hat{\eta}_i, p\} \\
= &\lim_{N\to\infty} N\int_\eta \Pr\{H|Np(\eta)\} f(\eta|\hat{\eta}_i)\,d\eta \\
= &\lim_{N\to\infty}\left(N\int_{\eta\in\Lambda_\delta}\Pr\{H|Np(\eta)\} f(\eta|\hat{\eta}_i)\,d\eta + N\int_{\eta\notin\Lambda_\delta}\Pr\{H|Np(\eta)\} f(\eta|\hat{\eta}_i)\,d\eta\right) \\
= &\lim_{N\to\infty} N\int_{\eta\in\Lambda_\delta}\Pr\{H|Np(\eta)\} f(\eta|\hat{\eta}_i)\,d\eta + \lim_{N\to\infty} N\int_{\eta\notin\Lambda_\delta}\Pr\{H|Np(\eta)\} f(\eta|\hat{\eta}_i)\,d\eta.
\end{aligned}$$

$$\begin{aligned}
&\lim_{N\to\infty} N\int_{\eta\notin\Lambda_\delta}\Pr\{H|Np(\eta)\} f(\eta|\hat{\eta}_i)\,d\eta \\
\leq &\lim_{N\to\infty}\int_{\eta\notin\Lambda_\delta} N\Pr\{H|Np(\eta)\} f(\eta|\hat{\eta}_i)\,d\eta \\
= &\ 0.
\end{aligned}$$

And

$$\begin{aligned}
& N\int_{\eta\in\Lambda_\delta}\Pr\left\{H|Np\left(\eta\right)\right\}f\left(\eta|\hat{\eta}_i\right)d\eta \\
\in\quad & \left[\frac{\sqrt{2}f\left(\eta_R|\hat{\eta}_i\right)}{\sqrt{\sqrt{\frac{p_M(\eta_R)}{p_R(\eta_R)}}}p_R'\left(\eta_R\right)-\sqrt{\sqrt{\frac{p_R(\eta_R)}{p_M(\eta_R)}}}p_M'\left(\eta_R\right)}-\zeta,\frac{\sqrt{2}f\left(\eta_R|\hat{\eta}_i\right)}{\sqrt{\sqrt{\frac{p_M(\eta_R)}{p_R(\eta_R)}}}p_R'\left(\eta_R\right)-\sqrt{\sqrt{\frac{p_R(\eta_R)}{p_M(\eta_R)}}}p_M'\left(\eta_R\right)}+\zeta\right] \\
& *\int_{\eta\in\Lambda_\delta}N\left(\frac{\sqrt{2}}{\sqrt{\sqrt{\frac{p_M}{p_R}}}p_R'\left(\eta\right)-\sqrt{\sqrt{\frac{p_R}{p_M}}}p_M'\left(\eta\right)}\right)^{-1}\Pr\left\{H|Np\left(\eta\right)\right\}d\eta \\
=\quad & \left[\frac{\sqrt{2}f\left(\eta_R|\hat{\eta}_i\right)}{p_R'\left(\eta_R\right)-p_M'\left(\eta_R\right)}-\zeta,\frac{\sqrt{2}f\left(\eta_R|\hat{\eta}_i\right)}{p_R'\left(\eta_R\right)-p_M'\left(\eta_R\right)}+\zeta\right] \\
& *\int_{\eta\in\Lambda_\delta}N\left(\frac{\sqrt{2}}{\sqrt{\sqrt{\frac{p_M}{p_R}}}p_R'\left(\eta\right)-\sqrt{\sqrt{\frac{p_R}{p_M}}}p_M'\left(\eta\right)}\right)^{-1}\Pr\left\{H|Np\left(\eta\right)\right\}d\eta
\end{aligned}$$

$$\begin{aligned}
& \int_{\eta\in\Lambda_\delta}N\Pr\left\{H|Np\left(\eta\right)\right\}d\eta \\
=\quad & \int_{\eta=\eta_R-\varepsilon}^{\eta_R+\varepsilon}\frac{\sqrt{N}e^{-N\left(\sqrt{p_R(\eta)}-\sqrt{p_M(\eta)}\right)^2}}{\sqrt{2\pi}\sqrt{\sqrt{p_R\left(\eta\right)p_M\left(\eta\right)}}}d\eta.
\end{aligned}$$

Write $x=\sqrt{2N}\left(\sqrt{p_R\left(\eta\right)}-\sqrt{p_M\left(\eta\right)}\right)$. Then

$$\begin{aligned}
dx &= \sqrt{2N}\frac{\sqrt{p_M\left(\eta\right)}p_R'\left(\eta\right)-\sqrt{p_R\left(\eta\right)}p_M'\left(\eta\right)}{2\sqrt{p_Rp_M}}d\eta \\
&= \sqrt{N}\frac{\sqrt{\sqrt{\frac{p_M}{p_R}}}p_R'\left(\eta\right)-\sqrt{\sqrt{\frac{p_R}{p_M}}}p_M'\left(\eta\right)}{\sqrt{2}\sqrt{\sqrt{p_Rp_M}}}d\eta
\end{aligned}$$

$$\begin{aligned}
& \int_{\eta\in\Lambda_\delta}N\Pr\left\{H|Np\left(\eta\right)\right\}d\eta \\
=\quad & \int_{x=\sqrt{2N}\left(\sqrt{p_R(\eta_R-\varepsilon)}-\sqrt{p_M(\eta_R-\varepsilon)}\right)}^{\sqrt{2N}\left(\sqrt{p_R(\eta_R+\varepsilon)}-\sqrt{p_M(\eta_R+\varepsilon)}\right)}\frac{\sqrt{2}}{\sqrt{\sqrt{\frac{p_M}{p_R}}}p_R'\left(\eta\right)-\sqrt{\sqrt{\frac{p_R}{p_M}}}p_M'\left(\eta\right)}\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx.
\end{aligned}$$

Then

$$\begin{aligned}
&\lim_{N\to\infty}\int_{\eta\in\Lambda_\delta} N\left(\frac{\sqrt{2}}{\sqrt{\sqrt{\frac{p_M}{p_R}}p_R'(\eta)}-\sqrt{\sqrt{\frac{p_R}{p_M}}p_M'(\eta)}}\right)^{-1}\Pr\{H|Np(\eta)\}\,d\eta \\
=&\lim_{N\to\infty}\int_{x=\sqrt{2N}\left(\sqrt{p_R(\eta_R-\varepsilon)}-\sqrt{p_M(\eta_R-\varepsilon)}\right)}^{\sqrt{2N}\left(\sqrt{p_R(\eta_R+\varepsilon)}-\sqrt{p_M(\eta_R+\varepsilon)}\right)}\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx \\
=&\ 1.
\end{aligned}$$

Let $\zeta\to 0$. Then we get $\lim_{N\to\infty} N\Pr\{V_R=V_M>V_L|\hat{\eta}_i,p\}=\frac{\sqrt{2}f(\eta_R|\hat{\eta}_i)}{p_R'(\eta_R)-p_M'(\eta_R)}$. Need $\frac{\sqrt{2}f(\eta|\hat{\eta}_i)}{\sqrt{\sqrt{\frac{p_M(\eta)}{p_R(\eta)}}p_R'(\eta)}-\sqrt{\sqrt{\frac{p_R(\eta)}{p_M(\eta)}}p_M'(\eta)}}$ to be absolutely continuous. ■

6.2 Additional proofs and lemmas for Section??

6.3 Proofs for Proposition 1

Lemma 6.1 $\frac{\partial\eta_R(a;u_r,u_l)}{\partial a}\in(0,1)$ *if* $\theta>\frac{u_R+u_L}{4}$

Proof. Again, if $\eta_R>\left\{\theta,a-\frac{1}{2}\tilde{u}_R,a+\frac{\tilde{u}_L}{2}\right\}$, then

$$\begin{aligned}
\frac{\partial p_R(\eta_R;a)}{\partial a} &= -F(\eta_R-\theta)+p_R(\eta_R) \\
\frac{\partial p_L(\eta_R;a)}{\partial a} &= p_L(\eta_R).
\end{aligned}$$

So

$$\begin{aligned}
\frac{\partial\eta_R(a)}{\partial a} &= -\frac{2\frac{\partial p_R(\eta_R;a)}{\partial a}+\frac{\partial p_L(\eta_R;a)}{\partial a}}{2p_R'(\eta_R)+p_L'(\eta_R)}=\frac{2F(\eta_R-\theta)-1}{1-e^{-\eta_R+a-\frac{\tilde{u}_R}{2}}+1-e^{-(\eta_R-\theta)}} \\
&= \frac{1-e^{-(\eta_R-\theta)}}{1-e^{-\eta_R+a-\frac{\tilde{u}_R}{2}}+1-e^{-(\eta_R-\theta)}}\in(0,1) \\
&= \frac{1}{2}\frac{e^{a-\frac{u_R}{2}}-e^{\theta}+\sqrt{e^{2\theta}+e^{2a-u_R}-e^{-\theta+a+\frac{1}{2}u_L}}}{\sqrt{e^{2\theta}+e^{2a-u_R}-e^{-\theta+a+\frac{1}{2}u_L}}} \\
&= \frac{1}{2}\left(1+\frac{e^{a-\frac{u_R}{2}}-e^{\theta}}{\sqrt{e^{2\theta}+e^{2a-u_R}-e^{-\theta+a+\frac{1}{2}u_L}}}\right)
\end{aligned}$$

because $\eta_R(a)>\max\left\{a-\frac{u_R}{2},\theta\right\}$. ■

Lemma 6.2 $\frac{\partial\eta_R(a;\tilde{u}_R,\tilde{u}_L)}{\partial\tilde{u}_R}<0$ *if* $\theta>\frac{u_R+u_L}{4}$

Proof.

$$
\begin{aligned}
\frac{\partial \eta_R}{\partial u_R} &= -\frac{2\frac{\partial p_R(\eta_R;a)}{\partial u_R}+\frac{\partial p_L(\eta_R;a)}{\partial u_R}}{2p_R'(\eta_R)+p_L'(\eta_R)} \\
&= -\frac{2\left(F(\eta_R-\theta)\left(1-F\left(\eta_R-a+\frac{1}{2}\tilde{u}_R\right)\right)\right)\frac{1}{2}}{1-e^{-\eta_R+a-\frac{\tilde{u}_R}{2}}+1-e^{-(\eta_R-\theta)}} \\
&\quad \text{(this shows that it is negative)} \\
&= -\frac{\left(1-\frac{1}{2}e^{-\eta_R+\theta}\right)\frac{1}{2}e^{-\eta_R+a-\frac{u_R}{2}}}{1-e^{-\eta_R+a-\frac{\tilde{u}_R}{2}}+1-e^{-(\eta_R-\theta)}} \\
&= -\frac{\frac{1}{2}e^{a-\frac{u_R}{2}}\left(e^{a-\frac{u_R}{2}}+\sqrt{e^{2\theta}+e^{2a-u_R}-e^{-\theta+a+\frac{1}{2}u_L}}\right)}{e^{\eta_R}\sqrt{e^{2\theta}+e^{2a-u_R}-e^{-\theta+a+\frac{1}{2}u_L}}} \\
&= -\frac{e^{a-\frac{u_R}{2}}}{e^{\theta}+e^{a-\frac{u_R}{2}}+\sqrt{e^{2\theta}+e^{2a-u_R}-e^{-\theta+a+\frac{1}{2}u_L}}}\left(1+\frac{e^{a-\frac{u_R}{2}}}{\sqrt{e^{2\theta}+e^{2a-u_R}-e^{-\theta+a+\frac{1}{2}u_L}}}\right)
\end{aligned}
$$

■

Lemma 6.3 $\frac{\partial \hat{a}(a;u_R,u_L)}{\partial \tilde{u}_R}<0$ *if* $\theta>\frac{u_R+u_L}{4}$ *and* $\alpha<\min\left\{e^{-2\theta-\frac{u_R+u_L}{2}},\frac{3}{8}e^{4\theta}\right\}$.

Proof.

$$
\begin{aligned}
\frac{\partial \alpha(a;u_R,u_L,\theta)}{\partial u_R} &= -\frac{1}{2}(1+\alpha)\frac{e^{a-\frac{u_R}{2}}}{e^{\eta_R}} \\
&\quad +\frac{1}{4}\frac{e^{2a-u_R}}{\sqrt{e^{2\theta}+e^{2a-\tilde{u}_R}-e^{a-\theta+\frac{\tilde{u}_L}{2}}}}\left(\frac{1}{\sqrt{e^{2\theta}+e^{2a-\tilde{u}_R}-e^{a-\theta+\frac{\tilde{u}_L}{2}}}}-\frac{1+\alpha}{e^{\eta_R}}\right) \\
&\quad -\frac{1}{4}\frac{\frac{1}{2}e^{-\theta-a+\frac{1}{2}u_R}}{\sqrt{e^{2\theta}+e^{-2a-\tilde{u}_L}-e^{-a-\theta+\frac{\tilde{u}_R}{2}}}}\left(\frac{1}{\sqrt{e^{2\theta}+e^{-2a-\tilde{u}_L}-e^{-a-\theta+\frac{\tilde{u}_R}{2}}}}-\frac{1+\alpha}{e^{-\eta_L}}\right) \\
&= -\frac{1}{2}(1+\alpha)\frac{e^{a-\frac{u_R}{2}}}{e^{\eta_R}} \\
&\quad +\frac{1}{4}\frac{e^{a-\frac{u_R}{2}}}{e^{\eta_R}}\frac{e^{a-\frac{u_R}{2}}\left(\frac{1}{2}e^{\theta}+\frac{1}{2}e^{a-\frac{u_R}{2}}-\frac{1}{2}(1+2\alpha)\sqrt{e^{2\theta}+e^{2a-\tilde{u}_R}-e^{a-\theta+\frac{\tilde{u}_L}{2}}}\right)}{e^{2\theta}+e^{2a-\tilde{u}_R}-e^{a-\theta+\frac{\tilde{u}_L}{2}}} \\
&\quad -\frac{1}{4}\frac{\frac{1}{2}e^{-\theta-a+\frac{1}{2}u_R}}{\sqrt{e^{2\theta}+e^{-2a-\tilde{u}_L}-e^{-a-\theta+\frac{\tilde{u}_R}{2}}}}\left(\frac{1}{\sqrt{e^{2\theta}+e^{-2a-\tilde{u}_L}-e^{-a-\theta+\frac{\tilde{u}_R}{2}}}}-\frac{1+\alpha}{e^{-\eta_L}}\right).
\end{aligned}
$$

So $\frac{\partial \alpha(a;u_R,u_L,\theta)}{\partial u_R}<0$ if

$$
\begin{aligned}
& e^{a-\frac{u_R}{2}}\left(e^{\theta}+e^{a-\frac{u_R}{2}}-(1+2\alpha)\sqrt{e^{2\theta}+e^{2a-\tilde{u}_R}-e^{a-\theta+\frac{\tilde{u}_L}{2}}}\right) \\
< \;& 4\left(e^{2\theta}+e^{2a-\tilde{u}_R}-e^{a-\theta+\frac{\tilde{u}_L}{2}}\right)
\end{aligned}
$$

and

$$
\begin{aligned}
0 &< \frac{1}{\sqrt{e^{2\theta}+e^{-2a-\tilde{u}_L}-e^{-a-\theta+\frac{\tilde{u}_R}{2}}}}-\frac{1+\alpha}{e^{-\eta_L}} \\
&= e^{\eta_L}\left(\frac{e^{\theta}+e^{-a-\frac{u_L}{2}}+\sqrt{e^{2\theta}+e^{-2a-\tilde{u}_L}-e^{-a-\theta+\frac{\tilde{u}_R}{2}}}}{\sqrt{e^{2\theta}+e^{-2a-\tilde{u}_L}-e^{-a-\theta+\frac{\tilde{u}_R}{2}}}}-(1+\alpha)\right) \\
&= e^{\eta_L}\left(\frac{1}{\sqrt{1-\frac{2e^{\theta-a-\frac{u_L}{2}}+e^{-a-\theta+\frac{\tilde{u}_R}{2}}}{\left(e^{\theta}+e^{-a-\frac{u_L}{2}}\right)^2}}}-(1+\alpha)\right).
\end{aligned}
$$

The latter holds if and only if

$$
\begin{aligned}
1-\left(\frac{1}{1+\alpha}\right)^2 &< \frac{2e^{\theta-a-\frac{u_L}{2}}+e^{-a-\theta+\frac{\tilde{u}_R}{2}}}{\left(e^{\theta}+e^{-a-\frac{u_L}{2}}\right)^2} \\
&= e^{-a}\frac{2e^{\theta-\frac{u_L}{2}}+e^{-\theta+\frac{u_R}{2}}}{\left(e^{\theta}+e^{-a-\frac{u_L}{2}}\right)^2}.
\end{aligned}
$$

This holds if

$$
\begin{aligned}
e^{a} &< \frac{2e^{-\theta-\frac{u_L}{2}}+e^{-2\theta+\frac{u_R}{2}}}{1-\left(\frac{1}{1+\alpha}\right)^2} \qquad (3) \\
&= \frac{(1+\alpha)^2}{\alpha(2+\alpha)}e^{-\theta}\left(2e^{-\frac{u_L}{2}}+e^{-\theta+\frac{u_R}{2}}\right).
\end{aligned}
$$

If $\log\frac{(1+\alpha)^2}{\alpha(2+\alpha)}+\log 2-\left(\theta+\frac{u_L}{2}\right)>\theta+\frac{u_R}{2}$, i.e. if $\log\frac{(1+\alpha)^2}{\alpha(2+\alpha)}>2\theta+\frac{u_R+u_L}{2}-\log 2$, then whenever this does not not, $a-\frac{u_R}{2}>\theta$. This is true if $\alpha<e^{-2\theta-\frac{u_R+u_L}{2}}$.

If $2\theta>\frac{u_R+u_L}{2}$, then

$$
\begin{aligned}
& e^{2\theta}+e^{2a-\tilde{u}_R}-e^{a-\theta+\frac{\tilde{u}_L}{2}} \\
=\ & \left(e^{\theta}-e^{a-\frac{u_R}{2}}\right)^2+2e^{\theta+a-\frac{u_R}{2}}-e^{a-\theta+\frac{u_L}{2}} \\
\geq\ & \left(e^{\theta}-e^{a-\frac{u_R}{2}}\right)^2+e^{\theta+a-\frac{u_R}{2}}.
\end{aligned}
$$

Then

$$
\begin{aligned}
& e^{a-\frac{u_R}{2}}\left(e^{\theta}+e^{a-\frac{u_R}{2}}-(1+2\alpha)\sqrt{e^{2\theta}+e^{2a-\tilde{u}_R}-e^{a-\theta+\frac{\tilde{u}_L}{2}}}\right)\\
\leq\ & e^{a-\frac{u_R}{2}}\left(e^{\theta}+e^{a-\frac{u_R}{2}}-\left|e^{\theta}-e^{a-\frac{u_R}{2}}\right|\right)\\
=\ & e^{a-\frac{u_R}{2}}2*\min\left\{e^{\theta},e^{a-\frac{u_R}{2}}\right\}\\
\leq\ & 2e^{\theta+a-\frac{u_R}{2}}\\
\leq\ & 2\left(\left(e^{\theta}-e^{a-\frac{u_R}{2}}\right)^2+e^{\theta+a-\frac{u_R}{2}}\right)\\
\leq\ & 4\left(e^{2\theta}+e^{2a-\tilde{u}_R}-e^{a-\theta+\frac{\tilde{u}_L}{2}}\right).
\end{aligned}
$$

So $2\theta>\frac{u_R+u_L}{2}$ is sufficient for $\alpha'(a)\in(0,1)$ for all a and $\frac{\partial\alpha(a;u_R,u_L,\theta)}{\partial u_R}<0$.

In fact, because

$$
\begin{aligned}
& e^{a-\frac{u_R}{2}}\left(e^{\theta}+e^{a-\frac{u_R}{2}}-(1+2\alpha)\sqrt{e^{2\theta}+e^{2a-\tilde{u}_R}-e^{a-\theta+\frac{\tilde{u}_L}{2}}}\right)\\
<\ & 2\left(\left(e^{\theta}-e^{a-\frac{u_R}{2}}\right)^2+e^{\theta+a-\frac{u_R}{2}}\right),
\end{aligned}
$$

if (3) holds, then

$$
\begin{aligned}
& \frac{\partial\alpha(a;u_R,u_L,\theta)}{\partial u_R}\\
<\ & -\frac{1}{2}\frac{e^{a-\frac{u_R}{2}}}{e^{\eta_R}}\left(1+\alpha-\frac{e^{a-\frac{u_R}{2}}\left(e^{\theta}+e^{a-\frac{u_R}{2}}-(1+2\alpha)\sqrt{e^{2\theta}+e^{2a-\tilde{u}_R}-e^{a-\theta+\frac{\tilde{u}_L}{2}}}\right)}{4\left(e^{2\theta}+e^{2a-\tilde{u}_R}-e^{a-\theta+\frac{\tilde{u}_L}{2}}\right)}\right)\\
<\ & -\frac{1}{2}\frac{e^{a-\frac{u_R}{2}}}{e^{\eta_R}}\frac{\left(\begin{array}{c}(2+4\alpha)\left(e^{2\theta}+e^{2a-\tilde{u}_R}-e^{a-\theta+\frac{\tilde{u}_L}{2}}\right)+2\left(e^{2\theta}+e^{2a-\tilde{u}_R}-e^{a-\theta+\frac{\tilde{u}_L}{2}}\right)\\+\alpha\frac{e^{a-\frac{u_R}{2}}\left(e^{\theta}+e^{a-\frac{u_R}{2}}\right)}{2}-2\left(\left(e^{\theta}-e^{a-\frac{u_R}{2}}\right)^2+e^{\theta+a-\frac{u_R}{2}}\right)\end{array}\right)}{4\left(e^{2\theta}+e^{2a-\tilde{u}_R}-e^{a-\theta+\frac{\tilde{u}_L}{2}}\right)}\\
=\ & -\frac{1}{2}\frac{e^{a-\frac{u_R}{2}}}{e^{\eta_R}}\left(\frac{1}{2}+\alpha\left(1+\frac{1}{8}\frac{e^{a-\frac{u_R}{2}}\left(e^{\theta}+e^{a-\frac{u_R}{2}}\right)}{\left(e^{2\theta}+e^{2a-\tilde{u}_R}-e^{a-\theta+\frac{\tilde{u}_L}{2}}\right)}\right)+\frac{2\left(e^{\theta+a-\frac{u_R}{2}}-e^{a-\theta+\frac{\tilde{u}_L}{2}}\right)}{4\left(e^{2\theta}+e^{2a-\tilde{u}_R}-e^{a-\theta+\frac{\tilde{u}_L}{2}}\right)}\right)\\
<\ & -\frac{1}{4}\frac{e^{a-\frac{u_R}{2}}}{e^{\eta_R}}.
\end{aligned}
$$

And if (3) does not hold, if α is small enough such that whenever (3) does not hold, $a - \frac{u_R}{2} > \theta$, then

$$\begin{aligned}
\frac{\partial \alpha\left(a; u_R, u_L, \theta\right)}{\partial u_R} &< -\frac{1}{4}\frac{e^{a-\frac{u_R}{2}}}{e^{\eta_R}} + \frac{\alpha}{4}\frac{\frac{1}{2}e^{-\theta-a+\frac{1}{2}u_R}}{\sqrt{e^{2\theta}+e^{-2a-\tilde{u}_L}-e^{-a-\theta+\frac{\tilde{u}_R}{2}}}}e^{\eta_L} \\
&< -\frac{1}{4}\frac{e^{a-\frac{u_R}{2}}}{e^{\eta_R}} + \frac{\alpha}{4}\frac{\frac{1}{2}e^{-\theta}}{\frac{e^{\theta}+e^{-a-\frac{u_L}{2}}}{2}}\frac{e^{-a+\frac{u_R}{2}}}{\frac{3}{4}\left(e^{\theta}+e^{-a-\frac{u_L}{2}}\right)} \\
&< -\frac{1}{4}\frac{e^{a-\frac{u_R}{2}}}{e^{\eta_R}} + \frac{\alpha}{4}\frac{4}{3}e^{-4\theta} \\
&= -\frac{1}{4}\left(\frac{e^{a-\frac{u_R}{2}}}{e^{\eta_R}} - \alpha\frac{4}{3}e^{-4\theta}\right) \\
&< -\frac{1}{4}\left(\frac{1}{2} - \alpha\frac{4}{3}e^{-4\theta}\right)
\end{aligned}$$

Thus second inequality holds because $a - \frac{u_R}{2} > \eta_L$ given that r votes for R with positive probability in equilibrium. Thus $\frac{\partial a}{\partial u_R} < 0$ if α is small enough such that the negation of (3) implies that $a - \frac{u_R}{2} > \eta_R$ and such that $\alpha < \frac{3}{8}e^{4\theta}$. In addition, if $\alpha\frac{4}{3}e^{-4\theta} < \frac{1}{8}$, then

$$\begin{aligned}
\frac{\partial \alpha\left(a; u_R, u_L, \theta\right)}{\partial u_R} &< -\frac{1}{4}\left(\frac{e^{a-\frac{u_R}{2}}}{e^{\eta_R}} - \frac{1}{4}\frac{e^{a-\frac{u_R}{2}}}{e^{\eta_R}}\right) \\
&= -\frac{1}{4}\frac{3}{4}\frac{e^{a-\frac{u_R}{2}}}{e^{\eta_R}}.
\end{aligned}$$

Then either $\frac{\partial \alpha(a;u_R,u_L,\theta)}{\partial u_L} > \frac{1}{4}\frac{e^{-a-\frac{u_L}{2}}}{e^{-\eta_L}}$, or $-a - \frac{u_L}{2} > \theta$, and thus $\hat{a}'\left(a^{**}\right) > \frac{1}{4}\left(1+\alpha\right)$ and $\frac{\partial \alpha(a;u_R,u_L,\theta)}{\partial u_L} > \frac{1}{4}\frac{3}{4}\frac{e^{-a-\frac{u_L}{2}}}{e^{-\eta_L}}$. In both cases, $\frac{\partial \alpha(a;u_R,u_L,\theta)}{\partial u_L} > \frac{1}{4}\left(1-\hat{a}'\left(a^{**}\right)\right)\frac{1}{4}\frac{e^{-a-\frac{u_L}{2}}}{e^{-\eta_L}}$. ■

Lemma 6.4 *If* $\alpha < \min\left\{e^{-2\theta-\frac{u_R+u_L}{2}}, \frac{3}{32}e^{4\theta}\right\}, u_R+u_L < \log\frac{3}{2}$ *and* $2\theta - \frac{u_R+u_L}{2} > \log 3$*, then* $\frac{\partial \eta_R^{**}}{\partial u_L} > 0$.

Proof.

$$\begin{aligned}
\frac{\partial \eta_R^{**}}{\partial u_L} &= \frac{\partial \eta_R(a; u_R, u_L)}{\partial a}\frac{\partial a^{**}}{\partial u_L} + \frac{\partial \eta_R(a; u_R, u_L)}{\partial u_L} \\
&= \frac{\partial \eta_R(a; u_R, u_L)}{\partial a}\frac{\frac{\partial \hat{a}(a;u_R,u_L)}{\partial u_L}}{1-\hat{a}'(a^{**})} + \frac{\partial \eta_R(a; u_R, u_L)}{\partial u_L} \\
&> \frac{1}{4}\frac{1}{e^{\eta_R^{**}}}\frac{1}{2\sqrt{e^{2\theta}+e^{2a-\tilde{u}_R}-e^{a-\theta+\frac{\tilde{u}_L}{2}}}}\frac{1}{e^{-\eta_L}(1-\alpha'(a^{**}))} \\
&\times \left[\begin{array}{c} \left(\sqrt{e^{2\theta}+e^{2a-\tilde{u}_R}-e^{a-\theta+\frac{\tilde{u}_L}{2}}}e^{a-\frac{u_R}{2}}+e^{2a-u_R}-\frac{1}{2}e^{-\theta+a+\frac{u_L}{2}}\right)e^{-a-\frac{u_L}{2}} \\ -e^{-\eta_L}e^{-\theta+a+\frac{1}{2}u_L} \end{array}\right] \\
&\propto \left[\begin{array}{c} \left(\sqrt{e^{2\theta}+e^{2a-\tilde{u}_R}-e^{a-\theta+\frac{\tilde{u}_L}{2}}}e^{a-\frac{u_R}{2}}+e^{2a-u_R}-\frac{1}{2}e^{-\theta+a+\frac{u_L}{2}}\right)e^{-a-\frac{u_L}{2}} \\ -e^{-\eta_L}e^{-\theta+a+\frac{1}{2}u_L} \end{array}\right] \\
&> \frac{1}{2}e^{\theta-\frac{u_R+u_L}{2}} + \frac{3}{2}e^{a-u_R-\frac{u_L}{2}} - \frac{1}{2}e^{-\theta} \\
&\quad - \left(e^{\theta}+e^{-a-\frac{1}{2}u_L}\right)e^{-\theta+a+\frac{1}{2}u_L} \\
&> \frac{1}{2}e^{\theta-\frac{u_R+u_L}{2}} - \left(e^{-\theta}+e^{a+\frac{u_L}{2}}\right) + \frac{3}{2}e^{a-u_R-\frac{u_L}{2}} - \frac{1}{2}e^{-\theta} \\
&= \frac{1}{2}e^{\theta-\frac{u_R+u_L}{2}} - \frac{3}{2}e^{-\theta} + e^{a^{**}+\frac{u_L}{2}}\left(\frac{3}{2}e^{-(u_R+u_L)}-1\right).
\end{aligned}$$

Therefore, $\frac{\partial \eta_R^{**}}{\partial u_L} > 0$ if ■

1. $u_R + u_L < \log\frac{3}{2}$ and $2\theta - \frac{u_R+u_L}{2} > \log 3$, or
2. $\frac{1}{2}e^{\theta-\frac{u_R+u_L}{2}} - \frac{3}{2}e^{-\theta} + e^{\frac{u_L}{2}}\left(\frac{3}{2}e^{-(u_R+u_L)}-1\right) > 0$ and $u_R > u_L$ because in that case, $a^{**} < 0$.

Lemma 6.5 *When love-hate ratio on both sides are equal, the ex ante probability that over coordination happens decreases with θ if $\theta > u + \log\frac{3}{2}$.*

Proof. This is because

$$\begin{aligned}
\frac{\partial \eta_R^*(u,u,\theta,\alpha)}{\partial \theta} &= e^{-\eta_R^*}\frac{e^{\theta}+\frac{e^{2\theta}+\frac{1}{2}e^{-\theta+\frac{u}{2}}}{\sqrt{e^{2\theta}+e^{-\frac{u}{2}}-e^{-\theta+\frac{u}{2}}}}}{2} \\
&< e^{-\eta_R^*}\frac{e^{\theta}+e^{\frac{u_r}{2}}+\frac{e^{2\theta}+e^{-\frac{u}{2}}-e^{-\theta+\frac{u}{2}}}{\sqrt{e^{2\theta}+e^{-\frac{u}{2}}-e^{-\theta+\frac{u}{2}}}}}{2} \\
&< 1
\end{aligned}$$

because $\frac{3}{2}e^{-\theta+\frac{u}{2}+\frac{u}{2}} < 1$, and thus $\frac{3}{2}e^{-\theta+\frac{u}{2}} < e^{-\frac{u}{2}}$. Thus the derivative of the ex ante probability of over coordination w.r.t. θ is

$$\begin{aligned}\frac{\partial\left(e^{-\alpha\theta}-e^{-\alpha\eta_R^*}\right)}{\partial\theta} &= -\alpha\left(e^{-\alpha\theta}-e^{-\alpha\eta_R^*}\frac{\partial\eta_R^*(u,u,\theta,\alpha)}{\partial\theta}\right)\\ &< \alpha\left(e^{-\alpha\eta_R^*}-e^{-\alpha\theta}\right)<0.\end{aligned}$$

■

6.4 Frontloaded Primary System

Lemma 6.6 *For θ large enough and α small enough such that there exists a unique both responsive equilibrium, $\eta_R^{F^*}(u) > \eta_R^*(u)$ for all $u < \log 2$ and $\frac{\frac{1}{2}e^{-\alpha\eta_R^*(u)}}{1-e^{-\alpha\eta_R^*(u)}} < 1$ if either $\frac{1}{2}e^{-\alpha\theta} < \frac{1}{3}$ or $\frac{\alpha}{6}\frac{1-\phi^2}{\phi}\frac{\frac{1}{2}e^{-\alpha\theta}}{\left(\frac{2}{3}-\frac{1}{2}e^{-\alpha\theta}\right)^2} < 1$.*

Proof. Because $u_R = u_L = u$, $a^{**} = 0$ and thus

$$e^{\eta_R^*(u)} = \frac{e^{\theta}+e^{-\frac{u}{2}}+\sqrt{e^{2\theta}+e^{-u}-e^{-\theta+\frac{u}{2}}}}{2}.$$

$\eta_R^*(u)$ is the solution to

$$2p_R(\eta_R)+p_L(\eta_R)=1.$$

Because $\eta_R^*(u) > \max\left\{\theta, -\frac{u_R}{2}\right\}$ and this is a both responsive equilibrium, $\eta_R^*(u) > \frac{u_L}{2}$. Thus $\eta_R^*(u)$ solves

$$2\left(1-\frac{1}{2}e^{-\eta_R+\theta}\right)\left(1-\frac{1}{2}e^{-\eta_R-\frac{u}{2}}\right)+\frac{1}{2}e^{-\theta-\eta_R}\frac{1}{2}e^{-\eta_R+\frac{u}{2}}=1.$$

So

$$\left(2p_R'(\eta_R)+p_L'(\eta_R)\right)\eta_R^{*\prime}(u)+\left(1-\frac{1}{2}e^{-\eta_R+\theta}\right)\frac{1}{2}e^{-\eta_R-\frac{u}{2}}+\frac{1}{2}p_L(\eta_R)=0.$$

So

$$\eta_R^{*\prime}(u) = -\frac{\left(1-\frac{1}{2}e^{-\eta_R+\theta}\right)\frac{1}{2}e^{-\eta_R-\frac{u}{2}}+\frac{1}{2}p_L(\eta_R)}{1-e^{-\eta_R+\theta}+1-e^{-\eta_R-\frac{u}{2}}}<0$$

and

$$\begin{aligned}\eta_R^{*\prime}(u) &= -\frac{\left(1-\frac{1}{2}e^{-\eta_R+\theta}\right)\frac{1}{2}e^{-\eta_R-\frac{u}{2}}+\frac{1}{2}p_L(\eta_R)}{2\left(1-\frac{1}{2}e^{-\eta_R+\theta}\right)\frac{1}{2}e^{-\eta_R-\frac{u}{2}}+2\left(1-\frac{1}{2}e^{-\eta_R-\frac{u}{2}}\right)\frac{1}{2}e^{-\eta_R+\theta}+2p_L(\eta_R)}\\ &> -\frac{\left(1-\frac{1}{2}e^{-\eta_R+\theta}\right)\frac{1}{2}e^{-\eta_R-\frac{u}{2}}+\frac{1}{2}p_L(\eta_R)}{2\left(1-\frac{1}{2}e^{-\eta_R+\theta}\right)\frac{1}{2}e^{-\eta_R-\frac{u}{2}}+p_L(\eta_R)}\\ &> -\frac{1}{2}.\end{aligned}$$

Define $\tilde{\eta}^F$ to be the mapping in the front loaded election.

$$\begin{aligned}\tilde{\eta}^F(\eta;\phi) &= \eta_R^*\left(u_R^F(\eta,\phi)\right)\\ &= \eta_R^*\left(\log 2+\log\frac{\phi-c^F}{1-c^F\phi}\right)\\ &= \eta_R^*\left(\log 2+\log\frac{\phi-\frac{1}{2}\frac{1-\frac{P^F(R)}{P^F(M)}}{2+\frac{P^F(R)}{P^F(M)}}}{1-\frac{1}{2}\frac{1-\frac{P^F(R)}{P^F(M)}}{2+\frac{P^F(R)}{P^F(M)}}\phi}\right)\\ &= \eta_R^*\left(\log 2+\log\frac{\phi-\frac{1}{2}\frac{1-\frac{3}{2}e^{-\alpha\eta}}{2-\frac{3}{2}e^{-\alpha\eta}}}{1-\frac{1}{2}\frac{1-\frac{3}{2}e^{-\alpha\eta}}{2-\frac{3}{2}e^{-\alpha\eta}}\phi}\right).\end{aligned}$$

Then

$$\begin{aligned}\frac{\partial\tilde{\eta}^F(\eta;\phi)}{\partial\eta} &= \frac{\partial\eta_R^*\left(u^F,u^F\right)}{\partial u^F}*\frac{\partial u^F}{\partial\eta}\\ &= \frac{\partial\eta_R^*(u,u)}{\partial u}\left(\frac{-\left(1-\phi^2\right)\frac{\partial c^F}{\partial\eta}}{\left(1-c^F\phi\right)\left(\phi-c^F\right)}\right)\\ &= -\frac{\partial\eta_R^*(u,u)}{\partial u}\frac{\left(1-\phi^2\right)}{\left(1-c^F\phi\right)\left(\phi-c^F\right)}\left(3\alpha\frac{e^{-\alpha\eta}}{\left(3e^{-\alpha\eta}-4\right)^2}\right)\\ &= -\frac{\partial\eta_R^*(u,u)}{\partial u}\frac{\left(1-\phi^2\right)}{\left(1-c^F\phi\right)\left(\phi-c^F\right)}\left(\frac{1}{3}\alpha\frac{\frac{1}{2}e^{-\alpha\eta}}{\left(\frac{2}{3}-\frac{1}{2}e^{-\alpha\eta}\right)^2}\right).\end{aligned}$$

Thus for all η such that $u^F(\eta,\phi)$ is well-defined, i.e. for all η such taht $c^F<\phi$, we have $\frac{\partial\tilde{\eta}^F(\eta;\phi)}{\partial\eta}>0$ because $\frac{\partial\eta_R^*(u,u)}{\partial u}<0$ because $\phi<1$,

Because $\phi<1$ and $\frac{\frac{1}{2}e^{-\alpha\eta_R^*(u)}}{1-e^{-\alpha\eta_R^*(u)}}<1$, $u_R^F\left(\eta_R^*(u),\phi\right)<u$, and thus $\eta_R^F\left(\eta_R^*(u),u\right)>\eta_R^*(u)$. For all η such that $c^F(\eta)>0$, we have $u^F(\eta,u)<u$, and thus $\eta_R^F(\eta,u)=\eta_R^*\left(u^F(\eta,u)\right)>\eta_R^*(u)$ because $\eta_R^{*\prime}(u)<0$. Because equilibrium is both-responsive, $\eta\geq\theta$, thus $c^F(\eta,u)>0$ for all $\eta\geq\theta$ if $\frac{1}{2}e^{-\alpha\theta}<\frac{1}{3}$. Thus either the fixed point of η_R^F as a function of η is above $\eta_R^*(u)$, or it is equal to ∞, and thus in equilibrium everyone votes for M. Even if $\frac{1}{2}e^{-\alpha\theta}\geq\frac{1}{3}$, for η such that $c^F<0$,

$$\begin{aligned}\left|\frac{\partial\tilde{\eta}^F(\eta;\phi)}{\partial\eta}\right| &= \left|\frac{\partial\eta_R^*(u,u)}{\partial u}\right|\frac{\left(1-\phi^2\right)}{\left(1-c^F\phi\right)\left(\phi-c^F\right)}\left(\frac{1}{3}\alpha\frac{\frac{1}{2}e^{-\alpha\eta}}{\left(\frac{2}{3}-\frac{1}{2}e^{-\alpha\eta}\right)^2}\right)\\ &\leq \frac{1}{2}\frac{1-\phi^2}{\phi}\frac{1}{3}\frac{\frac{1}{2}e^{-\alpha\theta}}{\left(\frac{2}{3}-\frac{1}{2}e^{-\alpha\eta}\right)^2}\\ &< 1.\end{aligned}$$

Thus there exists no fixed point smaller than $\eta_R^*(u)$ given that $\eta_R^F\left(\eta_R^*(u),u\right)>\eta_R^*(u)$. ■

References

[1] Adams, W. (1987), As New Hampshire Goes. In: Orren, G., Polsby, N. (Eds.), Media and Momentum (Chatham House).

[2] Dekel, E., and M. Piccione (2000), "Sequential Voting Procedures in Symmetric Binary Elections," Journal of Political Economy, 108, 34-55.

[3] Duverger, M. (1954), Political Parties: Their Organization and Activity in the Modern State (New York: Wiley).

[4] Feddersen, T., and W. Pesendorfer (1997), "Voting Behavior and Information Aggregation in Elections with Private Information," Econometrica, 65, 1029-1058.

[5] Feddersen, T., and W. Pesendorfer (1998), "Convicting the Innocent: The Inferiority of Unanimous Jury Verdicts under Strategic Voting," The American Political Science Review, 92(1), 23-35.

[6] Klumpp, T. A., and M. K. Polborn (2006), "Primaries and the New Hampshire Effect," Journal of Public Economics, 31, 1073-1114.

[7] Lo (2007), "Language and Coordination Games," mimeo.

[8] Malbin, M., 1985. You get what you pay for, but is that what you want? In: Grassmuck, G. (Ed.), Before Nomination: Our Primary Problems (American Enterprise Institute for Public Policy Research).

[9] Morton, R. B., and K. C. Williams (1999), "Information Asymmetries and Simultaneous versus Sequential Voting," American Political Science Review, 93(1), 51-67.

[10] Myatt, D. P. (2007), "On the Theory of Strategic Voting," Review of Economic Studies, 74, 255-281.

[11] Myerson, R. (2000), "Large Poisson Games," Journal of Economic Theory, 94, 7-45.

[12] Myerson, R. (2002), "Comparison of Scoring Rules in Poisson Voting Games," 103,217-251.

[13] Myerson, R. B. and Weber, R. J. (1993), "A Theory of Voting Equilibria," American Political Science Review, 87 (1), 102-114.

[14] Riker, W. H. (1982), "The Two Party System and Duverger's Law: An Essay on the History of Political Science," American Political Science Review, 76 (4), 753-766.