Primaries, Strategic Voters and Heterogenous Valences*

Shino Takayama† Yuki Tamura‡ Terence Yeo§

May 24, 2019

Abstract. We propose a two-party model of policy promises and valence for office-seeking candidates under a two-stage electoral process with strategic voters. There are two equilibrium regimes depending on whether a good quality candidate of one party can win elections at both stages with certainty. We then provide the conditions for the existence of each equilibrium regime. We further analyze the case where only one party holds a primary, and conduct comparative statics analyses including how the change of public opinion affects equilibrium outcomes. Using a modified model including the decision of each candidate on entering the primary, we also show that if a low quality candidate places less importance on policy outcomes, and the good quality candidate in the opponent party has sufficiently low valence relative to the good quality candidate in her own party, she does not enter. This is because her entry will potentially damage the probability of her party winning the general election.

Key Words: primary election, median voter, uncertainty, valence.
JEL Classification Code: D72.

*We would like to thank Youngsub Chun, Jinwoo Kim, Andrew McLennan, Masahiro Okuno-Fujiwara, and Haishan Yuan for discussion and their many suggestions for improving the paper. We are also grateful for participants in the seminars at Australian National University, Hitotsubashi University, Tohoku University, and Keio University, as well as participants at the Econometric Society Australasian Meeting held in Tasmania, and the Kyoto-Aomori International Workshop. This research was partly funded with a grant from the Zengin Foundation. All errors remaining are our own.

†Takayama; School of Economics, University of Queensland, St Lucia, QLD 4072, Australia; email: s.takayama@economics.uq.edu.au; tel: +61-7-3346-7379; fax: +61- 7-3365-7299.

‡Tamura; Department of Economics, University of Rochester

§Yeo (corresponding author); School of Economics, University of Queensland; email: t.yeo@uq.edu.au; tel: +61-7-3365-6570; fax: +61- 7-3365-7299.

1 Introduction

Across the world, primaries have become commonplace as a means to select a party's representative to compete in an upcoming general election. This applies to many countries other than the US, including some European and Latin American countries (see Morton, 2006, Carey and Polga-Hecimovich, 2006). Similarly, in 2009, the Conservative Party in the UK experimented with the use of open primaries to select two of its parliamentary candidates. Preselection where party members vote for candidates to run for office at particular elections is also common in Australia, New Zealand, and Canada.

In much of the literature, and in this paper, a candidate is modeled as a combination of a policy promise and a one-dimensional measure of personal characteristics (or valence). There is a large body of empirical work that studies how these two factors affect electoral outcomes. Studies of the US, and evidence from recent cross-country research of Western Europe suggest that changes to candidates' valence have important electoral consequences and show that there is some relationship between valence, voters' choices, and the ideological distance between the candidates' policy promises (see Abney et al., 2013, Clark and Leiter, 2014, Green and Hobolt, 2008, Adams et al., 2012, Buttice and Stone, 2012). Many studies have also analyzed how the valence of party elites affects voters' electoral support for parties (Mondak, 1995, McCurley and Mondak, 1995, Stone and Simas, 2010), although the relationship between valence, policy decisions, and electability remains uncertain.

There is a view that having primaries results in more extreme polarization because candidates need to appeal to partisan primary voters, whose preferences are typically more extreme than those of voters in the general election, and so candidates tend to take more extreme positions. Polarization, which refers to the process by which public opinion divides and moves to the extremes, is a significant contemporary concern (Benoit and Laver, 2006, McCarty et al., 2006, 2013, Woon, 2018, Hill and Tausanovitch, 2018). Several works have considered the view that primaries also induce a high level of polarization, although the empirical results are mixed. Overall, while primaries may cause polarization in limited circumstances (Bullock and Clinton, 2011), they could also serve to reduce polarization because the general elections will continue to exert some moderating pressure on candidates (Hirano et al., 2010, Hall, 2015).

This paper contributes to this literature by providing a theoretical framework for a two-stage electoral process where voters consider the policy promises, valence, and electability of candidates when voting in primaries. In the theoretical literature, there are several attempts to model a primary election. Among these, the closest to our analysis is Hummel (2013). Hummel (2013) examines a scenario that we term the *symmetric game*. Our model is an extended version of the Hummel (2013) model.

Our model works in the following way. There are two parties and each party has two candidates, one of which has higher valence than the other, and this candidate is called the party's *advantaged* candidate. Before the primaries, the four candidates simultaneously select their policy promises and both parties simultaneously hold primary elections. In the second stage, the winning candidate from each primary election competes in the general election. However, there is uncertainty about the realization of the median voter's bliss point in the general election. Voters judge candidates on two aspects: policy promises and valence. Voters in each party's primary choose strategically between the two candidates of their party. In the model, we make the *no Etch a Sketch* assumption[1] such that politicians are committed to the one policy promise for both the primary and general elections. We define a political equilibrium as a pure strategy Nash equilibrium.

In equilibrium, there are two possible regimes. Without loss of generality, we assume that the left party's advantaged candidate's valence is at least as high as the valence of the right party's advantaged candidate. In regime 1, the left party's advantaged candidate has much higher valence than all the other candidates, such that there is some policy promise for the left party's advantaged candidate that guarantees her victories in both the primary and general elections. In regime 2, the differences in valence between the candidates are not as significant as in regime 1, and there is a policy promise for the right party's advantaged candidate that allows her to win the general election with some strictly positive probability. In regime 2, our result demonstrates that in equilibrium, the advantaged candidate in each party chooses that policy which maximizes the probability of winning the general election subject to just beating the disadvantaged candidate in the primary election.

In reality, we have observed both competitive and uncompetitive primaries.[2] For example, we could well imagine that the presence of a very strong candidate could result in an uncompetitive primary. In our analysis, a regime 1 equilibrium is the situation where the best advantaged candidate has a policy promise that guarantees victories for her, and in this sense, she is less constrained by her party's primary, which can be interpreted as relatively uncompetitive. In contrast, a regime 2 equilibrium is the situation where advantaged candidates are more constrained by primaries, which can be interpreted as more competitive. Our first set of theorems provides sufficient conditions under which an equilibrium exists and belongs to each regime.

Further, as a special case of our general model, we consider the symmetric game where both

[1]This is a common assumption in the literature, with Hummel (2013) and Grofman et al. (2019) making the same. Its origin lies in the 2012 Republican presidential primaries. Asked whether candidate Mitt Romney had moved too far to the right for the upcoming general election, his top aide, Eric Fehrnstrom, said that they would hit a reset button for the fall campaign, "almost like an Etch A Sketch." Two rival candidates then held up an Etch A Sketch toy as a visual aid to their audiences and criticized the remark. For further details, see: `http://en.wikipedia.org/wiki/Mitt_Romney_presidential_campaign,_2012`, viewed on May 4, 2019.

[2]For instance, see `https://edition.cnn.com/election/2018/primaries` for the results of the US presidential primaries in 2018, viewed on May 4, 2019.

parties' advantaged candidates have the same level of valence, while both parties' disadvantaged candidates also share the same but lower level of valence. In addition, the distribution of the overall median voter's bliss point is also symmetric. We show that the equilibrium in the symmetric game always exists, and prove that if the difference in valences between the advantaged and disadvantaged candidates is small relative to the distance between the bliss points of the parties' median voters and the median of the overall median voter's bliss point, which we call the *center*, then the equilibrium in Hummel (2013), which is a regime 2 equilibrium, exists. Otherwise, both advantaged candidates choose the center.

Our model also covers another special case where one party's sole candidate is an incumbent already in office committed to a predetermined policy promise, while the other party holds a primary between its two candidates. In this scenario, we characterize the equilibrium. Further, we conduct comparative static analyses. The first result of these examines the impact of a change in the left party's advantaged candidate's valence. We show that when the left party's advantaged candidate's valence increases, most of her rivals' policy promises become more extreme, in the sense that they shift their policy promises closer to their own party's median voters and away from the center.

Another set of comparative statics analyses several scenarios where the distribution of public opinion changes. First, we compare an equilibrium in two societies, such that in one society, public opinion is clustered closer to the center than the other society. Second, we consider two scenarios: under the first scenario, there are more left extremists in one society than the other; under the second scenario, the centrists' opinion shifts toward the left. Our general model allows us to consider how these changes in public opinion affect the equilibrium outcomes and provide results that can be empirically testable.

Finally, we consider a modified version of the game where the candidates strategically choose to enter a primary. We show that if the disadvantaged candidate cares more about policy outcomes, the disadvantaged candidate enters the primary. Conversely, if that candidate places less importance on policy outcomes, and the advantaged candidate in the opponent party has a sufficiently low valence level relative to the advantaged candidate in her own party, she does not enter. This is because her entry will constrain the policy promise of the advantaged candidate in her own party, potentially damaging the probability of her party winning the general election.

In the symmetric game in the Hummel (2013) model, Hummel (2013) provides the set of sufficient conditions for the existence of a regime 2 equilibrium. We extend this analysis to the setting in which the four candidates' valence values are heterogeneous, and the distribution F is not necessarily symmetric. One of our results obtains the result in Hummel (2013) as a special case. We further show that equilibrium in the symmetric game always exists, and prove that if the difference in valences between the advantaged and disadvantaged candidates is small relative to the distance between the bliss points of the parties' median voters and the center, then the equilibrium in Hummel (2013) arises, otherwise both advantaged candidates choose the center. However, in contrast to Hummel (2013), our results also apply to games that are not symmetric and to candidates with heterogeneous valence levels. Hummel (2013) also considers strategic entry into party primaries in the symmetric game. As the valence level of the advantaged candidates is the same for the two parties in Hummel's analysis, the impact of the opponent party's advantaged candidate's valence on the disadvantaged candidate's decision to enter the primary election cannot be captured. Our result here adds yet another interesting insight into the decisions by politicians whether to enter the race.

Further, in the literature on pure strategy Nash equilibria in political competition models, Aragones and Palfrey (2002) employ the standard Downsian model of two-candidate elections of two politicians, such that one is more advantaged in terms of valence than the other, and show that pure strategy Nash equilibria often do not exist.[3] Intuitively, a pure strategy Nash equilibrium may fail to arise because an unending "policy chase" arises in which the advantaged candidate continually pursues the disadvantaged candidate's policy promise over the entire policy space. In our model, a Nash equilibrium could also fail to exist. However, the existence of the disadvantaged candidate in the primary could limit the extent of the chase by the advantaged candidate, and in this circumstance, an unending policy chase does not occur. Then, a regime 2 equilibrium could exist, for which one of our main theorems specifies the condition. Interestingly, the existence of primaries could eliminate the possibility of an unending policy race.

Despite many real examples of two-stage electoral processes around the world, theoretical studies involving party primaries are still stylized relative to their importance. We provide a vehicle for responding to questions about the effects of primary elections, valence, and changes in public opinion on policy promises. Our findings are also testable and generally consistent with the empirical research on primaries and the roles of valence in elections. We provide an equilibrium analysis that is analytically tractable and deals with the multidimensional issues of policy promises and valence.

Further, our model is robust under various extensions, and particularly useful for conducting comparative statics. In our analysis, we also provide testable predictions for electoral outcomes as public opinion changes. As Harmon (2018) describes, an influx of immigrants could affect public opinion and the world has experienced profound demographic changes, particularly this century (for example, see Frey, 2015). The framework presented in this paper can also be a vehicle to examine how changes in public opinion affect electoral outcomes.

Recently, Hirano and Snyder (2014) have argued that the literature underestimates the value

[3] See also Saporiti (2008), which uses the concept of the *better-reply security* proposed by Reny (1999) to consider the existence of pure strategy Nash equilibria, and Takayama et al. (2019) which uses the recently developed concept in the literature of discontinuous games by McLennan et al. (2011) to prove the existence of pure strategy Nash equilibria, in games where the median voter's bliss point is stochastic.

of primaries. While there is an increasing body of empirical research that has considered the effects of primaries on electoral prospects using more recent and international data sets (for example, see Carey and Polga-Hecimovich, 2006, Stone et al., 1992), theoretical research continues to lag behind. In fact, until very recently, most models of electoral competition in the theoretical literature assume that candidates representing their political parties compete in single-stage elections. Only of late has there been a greater attempt at providing a theoretical framework for primary elections.

Owen and Grofman (2006) develop a model of two-stage electoral competition based on the framework proposed by Coleman (1971), and examine the candidates' policy positions. Furthermore, Adams and Merrill (2008) propose a model where there are two candidates in each party, each of whom attempts to maximize the joint winning probability in both the general and primary elections. In their model, voters vote sincerely and they prove the existence of a Nash equilibrium when the voter's utility is concave and peaks at the voter's ideal point. Furthermore, they show that when the candidates in the same party have the same valence, they choose the same policy position in equilibrium. They also show that as the other party's candidates' valence increases, candidates have an incentive to shift unilaterally toward the general election median voter's bliss point. Our results show that this tendency does not hold in our framework when voters vote strategically. When one advantaged candidate's valence increases in our setting, this pushes the other candidates' equilibrium policies away from the center and toward their own party's median voter's bliss point.

Another similar analysis to ours is Grofman et al. (2019). Grofman et al. (2019) assumes that voters are sincere in the sense that they always vote for the candidate with the policy closest to their own bliss points without considering the viability of the candidate in the general election. Using this, Grofman et al. (2019) shows that an equilibrium always exists with sincere voters. In our model, voters are strategic in casting their votes in the primary elections in the sense that they also take into consideration the chances of each candidate in the general election. As discussed, when an unending policy chase arises, an equilibrium may fail to exist because the best response is not well defined for at least one of the two advantaged candidates. We show that in the symmetric game, even when this happens, at the center, the probability of winning the general election is continuous, and hence the best responses are well defined for the two advantaged candidates. Thus, there is always an equilibrium in the symmetric game.

There is yet another stream of the literature that considers how primaries affect policy outcomes. For example, Serra (2011) proposes a model of party elites in which those who have a different bliss point from the party's median voter decide whether or not to hold a primary election. Snyder and Ting (2011) consider a model with primaries in which the candidate can differ in quality, but they do not include endogenous candidate policy promises in their model. Woon (2018) uses behavioral game theory and experiments to examine how primaries affect policy promises. Our model can also be used to examine how having a primary affects policy outcomes. Using the modified version of the model, we also analyze the scenario where candidates endogenously decide whether to enter the primary, and show that if a low quality candidate places less importance on policy outcomes, and the good quality candidate in the opponent party has sufficiently low valence relative to the good quality candidate in her own party, she does not enter, because her entry will potentially damage the probability of her party winning the general election.

In the literature (e.g., see Adams and Merrill, 2008), it is widely acknowledged that there are two opposing forces determining the candidates' policy promises: namely, how do office-seeking candidates balance the *centrifugal* incentive to appeal to their primary voters against the *centripetal* motivation to appeal to voters in the general election if they have to commit to the same policy promises at both stages of the electoral competition? Realistically, at the later stage, politicians are constrained by policy promises made at an earlier stage of the electoral competition. One difficulty is to characterize an optimal policy promise for each candidate who faces these opposing incentives, especially when voters in the primaries are strategic. The literature so far has assumed that primary voters vote sincerely and do not consider the candidates' prospects in the general election when casting their primary votes. Put differently, the electoral competition game is symmetric (see McGann, 2002, Adams and Merrill, 2008, Owen and Grofman, 2006), even though empirical studies suggest that some primary voters are strategic and do consider the candidates' prospects in the general election (Abramowitz, 1989).

The remainder of the paper is organized as follows. Section 2 describes the model and Section 3 provides the equilibrium analysis. Section 4 demonstrates the equilibrium under the scenario of the one primary, while Section 5 conducts the comparative statics. Section 6 discusses some extensions and extends the model with endogenous entry. Section 7 concludes.

2 The Model

In the model, there are two parties, which we refer to as party R and party L. The policy space $X = [-1, 1]$ is unidimensional. Party R and party L have their respective core constituencies toward the right (closer to 1) and the left (closer to -1) in X.

We assume that the electoral process involves two elections: a primary election and a general election. First, both parties hold a primary election. Voters in each party's primary choose strategically between the two candidates in their primary. Second, in the general election, the winning candidate in the primary election competes with the other party's winner of their primary election.

Each party has two candidates. The candidates are assumed to be motivated solely by winning office. Each candidate is characterized by two aspects: strategy and characteristics. Each candidate chooses a *policy promise* and their characteristics are measured by a variable called *valence*. A candidate's valence variable is fixed exogenously, and benefits all voters independently of their political opinions if the candidate is elected. We assume that in each party, there is one advantaged candidate with *high valence* denoted A, and one disadvantaged candidate with *low valence* denoted D. We can then identify each candidate using their valence. Party R has two candidates R_A and R_D, and their valence satisfies $R_A > R_D$, while party L also has two candidates L_A and L_D, and their valence satisfies $L_A > L_D$. Without loss of generality, we assume $L_A \geq R_A$.

There is a continuum of voters. Each voter has a bliss point $b \in X$, and we define the utility of voters with bliss point $b \in X$ when the winning candidate has valence P_j for $P = L, R$ and $j = A, D$, and policy x by:

$$v_b(x, P_j) = P_j - |x - b|. \tag{1}$$

We assume that the set of voters in the primary election of party k is a subset of the overall voting population, and has a measure denoted by μ_k for each $k = L, R$. Assume that μ_L and μ_R are public information.

Let M_L and M_R be party L's and party R's median voter's bliss point, respectively. We assume $M_R > M_L$. We call the median voter in the general election *the overall median voter*. We denote the bliss point of the overall median voter using M, and assume this is randomly distributed in $[M_L, M_R]$, with a cumulative distribution function F. We denote the probability density function of F by f. Then, $F(x) = 0$ for $x \leq M_L$ and $F(x) = 1$ for $x \geq M_R$. We assume that F is continuously differentiable in its support $[M_L, M_R]$, and F is common knowledge.

Assumption 1. *The hazard rate $\frac{f(x)}{1-F(x)}$ is increasing in x for all $x \in [M_L, M_R]$ and $\frac{f(x)}{F(x)}$ is decreasing in x for all $x \in [M_L, M_R]$.*

Assumption 1 is a standard assumption in the literature known as the *monotone hazard rate*. See, for example, Roemer (1997), Saporiti (2008), Hummel (2013), and Drouvelis et al. (2014). This assumption guarantees the unique existence of maximizers in the expected payoffs and is mainly for tractability, although we use the hazard rate to conduct comparative statics later in Section 5.

The sequence of the game is as follows.

Stage 0 At the beginning of the game, each candidate in both parties simultaneously and independently chooses their policy promises. We assume that a politician chooses a policy promise that she commits to in both the primary and general elections if she wins the candidacy in the primary election.

Stage 1 The primary election is held for each party.

Stage 2 The general election is held.

Suppose that party R's candidate with policy r and valence R, and party L's candidate with policy l and valence L, are respectively the winners of their parties' primaries. Then party R candidate's probability of winning in the general election is

$$\pi_R(r, R, l, L) = \Pr[v_M(r, R) \geq v_M(l, L)]. \tag{2}$$

When $|r - l| > |R - L|$, if $r < l$,

$$\pi_R(r, R, l, L) = \begin{cases} 0 & \text{if } \frac{R-L+l+r}{2} \leq M_L \\ F(\frac{R-L+l+r}{2}) & \text{if } \frac{R-L+l+r}{2} \in (M_L, M_R) \\ 1 & \text{if } \frac{R-L+l+r}{2} \geq M_R. \end{cases} \tag{3}$$

and if $r > l$,

$$\pi_R(r, R, l, L) = \begin{cases} 1 & \text{if } \frac{L-R+l+r}{2} \leq M_L \\ 1 - F(\frac{L-R+l+r}{2}) & \text{if } \frac{L-R+l+r}{2} \in (M_L, M_R) \\ 0 & \text{if } \frac{L-R+l+r}{2} \geq M_R. \end{cases} \tag{4}$$

Note that when $R = L$, (3) and (4) hold, but we also need to specify the probability of winning the general election for each candidate when $l = r$. This situation only arises in Theorem 5 when the game is symmetric (see Definition 2). Define a tiebreaking rule in the general election when $L_A = R_A$, $p : X \to [0, 1]$, such that for every $x \in X$,

$$\pi_R(x, R_A, x, L_A) = p(x). \tag{5}$$

For each $i, n = A, D$, define the expected payoff V_b of voter b when candidate L_i is the winner in party L's primary and candidate R_n is the winner in party R by

$$V_b(l_i, L_i, r_n, R_n) \;=\; \pi_R(r_n, R_n, l_i, L_i)v_b(r_n, R_n) + (1 - \pi_R(r_n, R_n, l_i, L_i))v_b(l_i, L_i).$$

Let $w^k = (w^k_A, w^k_D)$ for each $k = L, R$, where each w^k_i denotes the probability that type $i = A, D$ wins the primary of party $k = L, R$. Then for each $i = A, D$, the expected payoff of voter b in party L when candidate L_i with policy l_i wins party L's primary is given by

$$\mathrm{E}V_b(l_i, L_i, r_A, r_D) \;=\; w^R_A V_b(l_i, L_i, r_A, R_A) + (1 - w^R_A)V_b(l_i, L_i, r_D, R_D),$$

and the expected payoff of voter b in party R when candidate R_i with policy r_i wins party R's primary is given by

$$\mathrm{E}V_b(r_i, R_i, l_A, l_D) \;=\; w^L_A V_b(l_A, L_A, r_i, R_i) + (1 - w^L_A)V_b(l_D, L_D, r_i, R_i).$$

In words, V_b is the interim expected payoff after knowing the winners of both parties' primaries, while $\mathrm{E}V_b$ is the ex ante expected payoff before the primaries.

The following definition expresses the notion of a pure strategy Nash equilibrium for the model.

Definition 1 (pure strategy Nash equilibrium)**.** *An equilibrium in a game of two-stage electoral competition is a collection of policy promises,* (l_A^*, l_D^*) *and* (r_A^*, r_D^*)*, and a collection of winning probabilities in the primary elections* $(w^{*k})_{k\in\{L,R\}} = (w_A^{*k}, w_D^{*k})_{k\in\{L,R\}}$ *such that:*

I. $w_A^{*k} + w_D^{*k} = 1$ *for every* $k = L, R$*;*

II. for $i, j = A, D$*, in response to* (l_A^*, l_D^*, r_j^*)*, each* r_i^* *is a best response so that the following holds:*

(A) If $\mathrm{EV}_{M_R}(r_i, R_i, l_A^*, l_D^*) > \mathrm{EV}_{M_R}(r_j, R_j, l_A^*, l_D^*)$*, then* $w_i^{*R} = 1$*;*

(B) Whenever $w_n^{*L} > 0$ *for any* $n = A, D$*, there is no policy* r_i *such that candidate* R_i *can* profitably deviate *in the sense that*

$$\mathrm{EV}_{M_R}(r_i, R_i, l_A^*, l_D^*) > \mathrm{EV}_{M_R}(r_j^*, R_j, l_A^*, l_D^*),$$

and $\pi_R(r_i, R_i, l_n^*, L_n) > w_i^{*R} \cdot \pi_R(r_i^*, R_i, l_n^*, L_n)$*;*

III. for $i, j = A, D$*, in response to* (r_A^*, r_D^*, l_j^*)*, each* l_i^* *is a best response so that the following holds:*

(A) If $\mathrm{EV}_{M_L}(l_i^*, L_i, r_A^*, r_D^*) > \mathrm{EV}_{M_L}(l_j^*, L_j, r_A^*, r_D^*)$*, then* $w_i^{*L} = 1$*;*

(B) Whenever $w_n^{*R} > 0$ *for any* $n = A, D$*, there is no policy* l_i *such that candidate* L_i *can* profitably deviate *in the sense that*

$$\mathrm{EV}_{M_L}(l_i, L_i, r_A, r_D) > \mathrm{EV}_{M_L}(l_j^*, L_j, r_A^*, r_D^*),$$

and $(1 - \pi_R(r_n^*, R_n, l_i, L_i)) > w_i^{*L} \cdot (1 - \pi_R(r_n^*, R_n, l_i^*, L_i))$.

The equilibrium condition (A) states that a candidate who obtains the majority wins the primary with certainty. The equilibrium condition (B) requires that there are no other policy promises for any of the four candidates that would win them the primary election with certainty and yield a higher probability of winning the general election. Intuitively, the motivation of each candidate is to maximize the probability of winning office. In the definition of equilibrium, we assume that the median voter theorem holds in primary elections. This is for maintaining analytical tractability and keeping the analysis simple. An earlier version of this paper, Takayama (2014), shows that when there is a reasonable concentration of primary voters around M_R or M_L compared with those around the tails of μ_R or μ_L, there is a Condorcet winner, and the median voter theorem holds in primary elections.

In Hummel (2013), the assumption is that the number of primary voters, say m, is odd, and that $\frac{m+3}{2}$ voters have bliss points smaller than M_L in party L's primary, while $\frac{m+3}{2}$ voters have bliss points larger than M_R in party R's primary. Thus, the median voter theorem also holds in primary elections in Hummel (2013).[4]

In the next section, we present the main theorems. In doing so, it is useful to keep in mind that there are two possible regimes in equilibrium.

Regime 1. There is a $l_A^* \in [M_L, M_R]$ that guarantees candidate L_A victories in both the primary and general elections.

Regime 2. Advantaged candidates choose more centrist policy promises that are just close enough to their parties' median bliss points to guarantee the advantaged candidates victory in their primaries. Candidate L_A does not win the general election with certainty.

The definition of a pure strategy Nash equilibrium requires an endogenous tiebreaking rule, whose nature and motivation may not be immediately intuitive.[5] A similar issue arises in a first price auction with the item where bidders' valuations are publicly known to bidders. The natural equilibrium concept has the agent with the highest valuation bidding the second highest valuation and winning with certainty.

In a regime 2 equilibrium of our analysis, the disadvantaged candidates choose the policy promises that maximize the expected payoff of their own party's median voters. Then, we can show that a set of policy promises that guarantee the advantaged candidates victories in the primaries is nonempty. Thus, the advantaged candidates choose the policy promises that maximize their probabilities of winning the general election in the set of policy promises that yield a victory to the advantaged candidates. When the advantaged candidate chooses the policy promise that yields the same level of expected payoff to their own party's median voter with the level obtained from the disadvantaged candidate and her policy promise, if the advantaged candidate's winning probability is smaller than one, the advantaged candidate would obtain a certain victory in the primary by deviating from the policy promise to a policy promise that is slightly smaller or larger than the policy promise. Our endogenous tiebreaking rule allows an equilibrium in which the advantaged candidate can choose the policy promise that yields the same level of expected payoff for their own party's median voter as the policy promise chosen by the disadvantaged candidate and still win the primary with certainty.

[4]Although the focus is very different from ours, Bouton (2013) considers a runoff election, and demonstrates that runoff systems with a threshold below 50 percent produce an Ortega effect such that the Condorcet loser of the election may win. Further, in contrast with Grofman et al. (2019), Bouton (2013) concludes that a sincere voting equilibrium does not always exist.

[5]The notion of an *endogenous sharing rule* with Nash equilibrium is proposed in Simon and Zame (1990) and used in several studies to obtain the existence of pure strategy Nash equilibrium. For example, Ausubel and Milgrom (2002) and Day and Milgrom (2008) use a similar definition of Nash equilibrium in the context of auction theory. See also Dell'Ariccia and Marquez (2006) in the banking literature and Prat and Rustichini (2003) in the common agency theory literature. Myerson (1993) uses it in the context of political competition.

In other words, whenever a tie arises in the primary, there is a slight deviation such that the median voter of each party strictly prefers the advantaged candidate over the disadvantaged candidate. For this slight deviation, we would imagine that the advantaged candidate wins the primary and an endogenous tiebreaking rule is consistent with this expectation. However, an endogenous tiebreaking rule in the primaries is not relevant for a regime 1 equilibrium because candidate L_A wins party L's primary with certainty, and an equilibrium strategy is unrestricted for party R's candidates (as we discuss further in the following section).

3 Equilibrium Analysis

3.1 The Main Theorems

There are three possible cases that result in regime 1. In the first case, candidate L_A's valence advantage over the other candidates is sufficiently high that candidate L_A wins the primary and general elections with certainty regardless of the policy promises of all other candidates. Theorem 1 provides the conditions under which this case arises. In Theorems 2 and 3, we consider the situations in which candidate L_A can win both the primary and general elections with certainty by choosing M_L (Theorem 2) or M_R (Theorem 3). Thus, these two theorems provide the conditions under which candidate L_A chooses M_L or M_R and regime 1 arises.

Obviously, regime 1 arises when candidate L_A is strong in terms of valence. Consider two cases: case 1, $M_R - M_L \leq L_A - R_A$, and case 2, $M_R - M_L > L_A - R_A$. In case 1, candidate L_A is so strong that candidate L_A wins the general election with certainty regardless of the policy promises of candidate R_A or R_D. Furthermore, for any primary voter in party R, the expected payoff resulting from either party's R candidate with any policy promise is simply the payoff of having candidate L_A elected, so the winning probabilities w_D^{*R} and w_A^{*R} are unrestricted. Theorem 2 considers this situation. In case 2, both regimes 1 and 2 can arise. Theorem 3 provides the conditions under which regime 1 arises, even in case 2.

Alternatively, Theorem 4 provides the conditions under which regime 2 arises. In Theorem 4, we see that regime 2 occurs when each advantaged candidate chooses the maximizer of the winning probability in the general election subject to winning the primary election. In this way, the advantaged candidate wins the primary of each party with certainty.

Finally, Theorem 5 considers the case where candidates L_A and R_A have equal valence. Indeed, Theorem 5 is a special case of Theorem 4 in the context of a symmetric game, such that the advantaged candidates have the same valence across both parties as is also the case for the disadvantaged candidates, and the distribution of the median voter's bliss point is symmetric (see Definition 2).

Now we state our theorems.

Theorem 1. *Suppose that* $\frac{M_R - M_L}{2} < \min\{L_A - L_D, L_A - R_A\}$. *Then any strategy profile such that* $l_A^* = \frac{M_R + M_L}{2}$ *and* (l_D^*, r_A^*, r_D^*), *and any collection of primary winning probabilities* $(w_A^{*R}, w_D^{*R}) \in [0,1]^2$ *such that* $w_A^{*R} + w_D^{*R} = 1$, *and* $(w_A^{*L}, w_D^{*L}) = (1, 0)$ *constitutes an equilibrium. Finally,* $\pi_R(l_A^*, L_A, r_n^*, R_n) = 0$ *for every* $n = A, D$.

Theorem 2. *Suppose that* $L_A - R_A \geq M_R - M_L$. *Then any strategy profile such that* $l_A^* = M_L$ *and* (l_D^*, r_A^*, r_D^*), *and any collection of primary winning probabilities* $(w_A^{*R}, w_D^{*R}) \in [0,1]^2$ *such that* $w_A^{*R} + w_D^{*R} = 1$, *and* $(w_A^{*L}, w_D^{*L}) = (1, 0)$ *constitutes an equilibrium. Finally,* $\pi_R(l_A^*, L_A, r_n^*, R_n) = 0$ *for every* $n = A, D$.

Theorem 3. *Suppose* $M_R - M_L > L_A - R_A > 0$.

1. *If* $M_R - M_L \leq R_A - L_D$, *then any strategy profile such that* $l_A^* = M_R$ *and* (l_D^*, r_A^*, r_D^*) *such that* $r_A^*, r_D^* \in [M_R - (L_A - R_A), M_R]$, *and any collection of primary winning probabilities* $(w_A^{*R}, w_D^{*R}) \in [0,1]^2$ *such that* $w_A^{*R} + w_D^{*R} = 1$, *and* $(w_A^{*L}, w_D^{*L}) = (1, 0)$ *constitute an equilibrium. Finally,* $\pi_R(l_A^*, L_A, r_n^*, R_n) = 0$ *for every* $n = A, D$.

2. *If* $M_R - M_L > R_A - L_D$ *and for each* $n = A, D$, *there is* $\bar{d}_n > 0$ *that satisfies*

$$\bar{d}_n = (L_A - R_n) - \frac{\left(F(M_R + \frac{L_D - R_n}{2})\right)^2}{\frac{1}{2} f(M_R + \frac{L_D - R_n}{2})}, \tag{6}$$

then any strategy profile such that $l_A^* = M_R$ *and* (l_D^*, r_A^*, r_D^*) *such that* $r_A^* \geq M_R - \bar{d}_A$, $r_D^* \geq M_R - \bar{d}_D$ *and any collection of primary winning probabilities* $(w_A^{*R}, w_D^{*R}) \in [0,1]^2$ *such that* $w_A^{*R} + w_D^{*R} = 1$, *and* $(w_A^{*L}, w_D^{*L}) = (1, 0)$ *constitute an equilibrium. Finally,* $\pi_R(l_A^*, L_A, r_n^*, R_n) = 0$ *for every* $n = A, D$.

Overall, the first three theorems yield the conditions under which, in equilibrium, regime 1 arises and candidate L_A chooses $\frac{M_R + M_L}{2}$ (Theorem 1), M_L (Theorem 2), and M_R (Theorem 3).[6]

Theorem 1 shows how regime 1 occurs when candidate L_A is strong relative to both candidates L_D and R_A. Theorem 2 considers how regime 1 occurs when candidate L_A is strong relative to candidate R_A. Theorem 3 considers how regime 1 occurs when candidate L_A is strong relative to candidate L_D.

Theorem 1 provides minimal values for candidate L_A's valence advantage over both L_D and R_A that guarantee victories in both the primary and general elections, irrespective of all

[6] Hummel (2013) assumes that if a candidate loses in the primary election, then the candidate obtains a utility of 0, if a candidate wins party k's primary, then she obtains a utility of $u_k \geq 0$ for each $k = L, R$, and the winner of the general election receives an additional utility of $u \geq 0$. We do not impose any particular utility and simply assume that each candidate attempts to maximize the probability of winning office. If we make a similar assumption to that of Hummel (2013), a regime 1 equilibrium still arises under the conditions of Theorems 1, 2, and 3, although candidates R_A and R_D would attempt to maximize the chance of winning party R's primary.

other candidates' choices. The cutoff value is obtained by examining the scenario in which $l_A^* = \frac{M_L+M_R}{2} \in \alpha(r_n, R_n)$ for all $r_n \in X$ and $n = A, D$.

In Theorem 2, candidate L_A's valence advantage over R_A is sufficiently high such that candidate L_A wins the general election against R_A with certainty by choosing M_L, regardless of R_A's policy promise. No matter how small the difference in valence between L_A and L_D, the median voter M_L in party L's primary strictly prefers candidate L_A. By the definition of equilibrium, candidate L_A wins the primary of party L.

In Theorem 3, candidate L_A's valence advantage over L_D is sufficiently high such that even if candidate L_A chooses M_R, the median voter M_L in party L's primary strictly prefers candidate L_A regardless of candidate L_D's policy promise. As candidate L_A's valence is greater than that of both R_A and R_D, the median voter M_R in party R's primary strictly prefers candidate L_A to be voted into office in the general election. In order to maximize their probability of winning party R's primary, each candidate in party R's primary chooses a policy promise that maximizes the median voter M_R's expected payoff, which is equivalent to choosing a policy promise that guarantees candidate L_A a victory in the general election.

Theorem 3 requires that both party R candidates choose policy promises that satisfy $r_n^* \geq M_R - \bar{d}_n$ for each $n = A, D$. If $r_n^* < M_R - \bar{d}_n$, candidate L_D can choose a policy promise that party L's median voter M_L strictly prefers to candidate L_A's choice of M_R. Candidate L_A could then lose the primary by choosing M_R. Thus, $r_n^* \geq M_R - \bar{d}_n$ for each $n = A, D$ is required.

It is easy to see that when L_A increases, $\bar{d}_n > 0$ is likely to hold for each $n = A, D$. When L_D increases, by Assumption 1, $\bar{d}_n > 0$ is less likely to hold. Being more advantaged with a higher valence makes it easier for candidate L_A to win both elections by choosing M_R, whereas facing a stronger competitor L_D in the primary makes it more difficult to win office by choosing M_R.

Now we state the fourth theorem, which comprises regime 2. As we show in Propositions 1 and 2, in regime 2, advantaged candidates win the primaries for both parties and neither of the advantaged candidates are guaranteed to win the general election in equilibrium. Given $w_A^{*L} = w_A^{*R} = 1$ in the equilibrium of regime 2, when examining the primary results, we need to only consider

$$\begin{aligned} V_b(l_i, L_i, r_A, R_A) &= \mathrm{E}V_b(l_i, L_i, r_A, r_D); \\ V_b(l_A, L_A, r_i, R_i) &= \mathrm{E}V_b(r_i, R_i, l_A, l_D). \end{aligned}$$

To explain our results, we now introduce sets of possible equilibrium strategies. Given a policy promise r_n of party R's primary winner with valence R_n, let $\alpha(r_n, R_n)$ be the set of policy promises that guarantees candidate L_A a victory in the general election, such that

$$\alpha(r_n, R_n) = \left[r_n - (L_A - R_n), r_n + (L_A - R_n)\right].$$

Next, define correspondences β_L and β_R as follows: for each $r_A, l_A \in X$,

$$\beta_L(r_A, R_A) = \{l : V_{M_L}(l, L_A, r_A, R_A) \geq \max_{l_D \in X} V_{M_L}(l_D, L_D, r_A, R_A)\};$$
$$\beta_R(l_A, L_A) = \{r : V_{M_R}(l_A, L_A, r, R_A) \geq \max_{r_D \in X} V_{M_R}(l_A, L_A, r_D, R_D)\}.$$

Then β_L and β_R are the sets of policy promises that guarantee candidates L_A and R_A victory in their respective primaries when the opponent party's primary winner is known to be the advantaged candidate. When β_L and β_R are nonempty, it is optimal for the advantaged candidates to select the policy promises in these sets.

For each $r_A, l_A \in X$, define $\hat{l}_D(r_A)$, $\hat{r}_D(l_A)$, $\hat{l}_A(r_A)$, and $\hat{r}_A(l_A)$ by

$$\hat{l}_D(r_A) = \arg\max_{l_D \in X} V_{M_L}(l_D, L_D, r_A, R_A); \tag{7}$$
$$\hat{r}_D(l_A) = \arg\max_{r_D \in X} V_{M_R}(l_A, L_A, r_D, R_D); \tag{8}$$
$$\hat{l}_A(r_A) = \max\{l_A | V_{M_L}(l_A, L_A, r_A, R_A) \geq V_{M_L}(\bar{l}_D, L_D, r_A, R_A) \text{ for } \bar{l}_D \in \hat{l}_D(r_A)\}; \tag{9}$$
$$\hat{r}_A(l_A) = \min\{r_A | V_{M_R}(l_A, L_A, r_A, R_A) \geq V_{M_R}(l_A, L_A, \bar{r}_D, R_D) \text{ for } \bar{r}_D \in \hat{r}_D(l_A)\}. \tag{10}$$

Notice that $\hat{l}_D(r_A)$ and $\hat{r}_D(l_A)$ are correspondences. In Lemma A-1, we show that if candidate L_D's probability of winning the general election against candidate R_A is strictly positive when $l_D < r_A$, then $\hat{l}_D(r_A)$ is singleton, and the same holds for $\hat{r}_D(l_A)$ when $r_D > l_A$ and candidate R_D's probability of winning the general election against candidate L_A is strictly positive. Otherwise, if $\pi_R(r_A, R_A, l_D, L_D) = 1$ for all $l_D \leq r_A$, then $V_{M_L}(l_D, L_D, r_A, R_A) = R_A - r_A + M_L$ for $l_D \leq r_A$, and $\hat{l}_D(r_A) \supseteq [M_L, r_A]$, because for every $l_D > r_A$ with $\pi_R(r_A, R_A, l_D, L_D) < 1$, we must have $R_A - r_A + M_L > L_D - l_D + M_L$ and thus, $V_{M_L}(l_D, L_D, r_A, R_A) < R_A - r_A + M_L$. Therefore, $\hat{l}_D(r_A)$ and $\hat{r}_D(l_A)$ are nonempty.

Further, note that $\hat{l}_A(r_A) = \max \beta_L(r_A, R_A)$ and $\hat{r}_A(l_A) = \min \beta_R(l_A, L_A)$. As we show in Lemma A-4, for every $l_A, r_A \in X$, $\beta_L(r_A, R_A)$ and $\beta_R(l_A, L_A)$ are nonempty. When the conditions for Theorem 4 hold, there are $\bar{x}$ and $\bar{y}$ such that for every $l_A \in [M_L, \bar{x}]$ and $r_A \in [\bar{y}, M_R]$, $\beta_L(r_A, R_A)$ and $\beta_R(l_A, L_A)$ are compact and disjoint. Therefore, for such l_A and r_A, $\hat{l}_A(r_A)$ and $\hat{r}_A(l_A)$ are well defined.

Theorem 4. *Suppose $M_R - M_L > L_A - R_A \geq 0$, and one of the following two conditions holds:*

- *$L_A = R_A$, and the following $\bar{x}$ and $\bar{y}$ with $\bar{x} \leq \bar{y}$ exist:*

$$\{\hat{l}_A(r_A) : r_A \in [\bar{y}, M_R]\} \cap \{\hat{r}_A(l_A) : l_A \in [M_L, \bar{x}]\} = \varnothing;$$

- *$L_A > R_A$, and the following $\bar{x}$ and $\bar{y}$ with $\bar{x} \leq \bar{y}$ exist:*

$$\frac{\left(1 - F\left(\frac{M_R+\bar{x}+L_A-R_D}{2}\right)\right)^2}{\frac{1}{2}f\left(\frac{M_R+\bar{x}+L_A-R_D}{2}\right)} = L_A - R_A; \tag{11}$$

$$\frac{\left(F\left(\frac{\bar{y}+M_L+L_D-R_A}{2}\right)\right)^2}{\frac{1}{2}f\left(\frac{\bar{y}+M_L+L_D-R_A}{2}\right)} = 2(L_A - R_A). \tag{12}$$

Then an equilibrium exists, and $(w_A^{*L}, w_D^{*L}) = (w_A^{*R}, w_D^{*R}) = (1, 0)$. Further, $\hat{l}_D(r_A^*)$ and $\hat{r}_D(l_A^*)$ are singletons, and the equilibrium satisfies $l_D^* \in \hat{l}_D(r_A^*)$, $r_D^* \in \hat{r}_D(l_A^*)$, $l_A^* = \hat{l}_A(r_A^*)$ and $r_A^* = \hat{r}_A(r_A^*)$.

In contrast to Theorems 1, 2, and 3, Theorem 4 deals with the case where candidate L_A is not so strong. In this case, candidates L_D and R_D choose policy promises that maximize the expected payoffs of their respective party's median voter, while candidates L_A and R_A choose policy promises to maximize their probability of winning the general election subject to winning their party's primary. Equations (11) and (12) ensure that in equilibrium, the sets $\alpha(r_A^*, R_A)$ and $\beta_L(r_A^*, R_A)$ do not overlap, because otherwise, candidate L_A can choose a policy promise in $\alpha(r_A^*, R_A) \cap \beta_L(r_A^*, R_A)$ that guarantees victory in both elections in equilibrium. However, this equilibrium is no longer in regime 2, whereas Theorem 4 is focused solely on the equilibrium in regime 2.

Further, as we show in the proof of Theorem 4, $\beta_L(r_A^*, R_A)$ and $\beta_R(l_A^*, L_A)$ do not overlap. Then, each candidate chooses the policy promises that are in their respective β set in order to win their primary. In order to maximize the probability of winning the general election, each candidate chooses the policy at the boundary of the β set that is closest to their opponent's policy. In this way, these conditions are sufficient for the existence of equilibrium in regime 2.

Finally, we state Theorem 5. Theorem 5 is a special case of Theorem 4 in the context of a symmetric game as described in Definition 2 below. The median of the distribution of the overall median voter's bliss point is important particularly in this case. Formally, we define $\bar{m} = F^{-1}(\frac{1}{2})$ as the median of the distribution of the overall median voter's bliss point, while f is assumed to be symmetric with respect to $\bar{m} = 0$. We call this setting *symmetric* and formally define it as follows.

Definition 2. *Let $\bar{m} = 0$. Suppose that f is symmetric with respect to 0 with its support $[-a, a]$, and $R_A = L_A$ and $R_D = L_D$. We say that the game is* symmetric.

The next theorem states that when the game is symmetric, an equilibrium always exists. An example of a symmetric game is Hummel (2013), which shows that in a symmetric game, if there is a solution to (7), (8), (9), and (10), then the game has an equilibrium. Theorem 5 shows that an equilibrium always exists in a symmetric game, regardless of whether the solution to the set of equations exists, and gives a condition under which the equilibrium that Hummel (2013) has shown arises and belongs to regime 2. In this sense, Theorem 5 generalizes Proposition 2 in Hummel (2013).

The following result holds.

Theorem 5. *Suppose that the game is symmetric. Then either (I) or (II) holds:*

(I) If $a \leq L_A - L_D$ and $p(\bar{m}) = \frac{1}{2}$, then $(w_A^{*L}, w_D^{*L}) = (w_A^{*R}, w_D^{*R}) = (1, 0)$, $(l_A^*, r_A^*) = (\bar{m}, \bar{m})$ and $l_D^*, r_D^* \in X$.

(II) Otherwise, if $a > L_A - L_D$, then $(w_A^{*L}, w_D^{*L}) = (w_A^{*R}, w_D^{*R}) = (1, 0)$, and $(l_A^*, l_D^*, r_A^*, r_D^*)$ solve (7), (8), (9), *and* (10). *The equilibrium satisfies $r_A^* = -l_A^*$, and $r_D^* = -l_D^*$ where r_A^* and r_D^* respectively satisfy*

$$r_A^* + r_D^* - (L_A - L_D) = \frac{F(\frac{r_A^* - r_D^* - (L_A - L_D)}{2})}{\frac{1}{2} f(\frac{r_A^* - r_D^* - (L_A - L_D)}{2})} \tag{13}$$

$$r_A^* = (r_A^* + r_D^* - (L_A - L_D)) \cdot F(\frac{r_A^* - r_D^* - (L_A - L_D)}{2}). \tag{14}$$

An intuition behind Theorem 5 is as follows. As we formally state later, Proposition 4 in Hummel (2013) shows that when $(L_A - L_D)$ increases, the β sets expand. Consider the situation where $(L_A - L_D)$ is greater than a so that $\beta_R(\bar{m}, L_A)$ and $\beta_L(\bar{m}, R_A)$ overlap (as shown in Lemma A-8). Notice that in a symmetric game, $\bar{m} \in \beta_R(\bar{m}, L_A)$ if and only if $\bar{m} \in \beta_L(\bar{m}, R_A)$. Thus, when $(L_A - L_D)$ is sufficiently large, $\bar{m} \in \beta_R(\bar{m}, L_A) \cap \beta_L(\bar{m}, R_A)$ holds. In this situation, as in Hotelling's model, both advantaged candidates choose $\bar{m} = 0$ in equilibrium. Conversely, when $(L_A - L_D)$ is low, both advantaged candidates are more constrained by competition in their primaries. Therefore, the equilibrium in regime 2 that Hummel (2013) studied arises.

3.2 Propositions of Characterization and the Proofs of the Theorems

In this section, we provide the proofs of the main theorems. The proofs of Theorems 1, 2, and 3 are straightforward.

Proof of Theorem 1. We claim that $l_A^* = \frac{M_R + M_L}{2}$ is a policy promise for candidate L_A that guarantees the candidate a victory in both party L's primary as well as in the general election, regardless of the policy promises of any of the other candidates.

First, we show that if candidate L_A wins party L's primary, then L_A wins the general election with certainty, regardless of the identity or policy promise of party R's primary winner.

Given $\frac{M_R - M_L}{2} < L_A - R_A$, for any $r_A \in X$ we must have $\pi_R(r_A, R_A, l_A^*, L_A) = 0$ given $v_b(l_A^*, L_A) > v_b(r_A, R_A)$ for all voters with bliss points $b \in [M_L, M_R]$. The same holds for any $r_D \in X$, i.e. $\pi_R(r_D, R_D, l_A^*, L_A) = 0$ given $R_D < R_A$. Hence, $\pi_R(r_i, R_i, l_A, L_A) = 0$ for $i = A, D$ and any $r_i \in X$.

Second, we show that candidate L_A wins party L's primary with certainty in equilibrium, and $w_A^{*L} = 1$. By assumption, $\frac{M_R - M_L}{2} < \min\{L_A - R_A, L_A - L_D\}$ and $R_A > R_D$, hence $v_{M_L}(l_A^*, L_A) > \max\{L_D, R_A, R_D\}$. As $\pi_R(r_i, R_i, l_A, L_A) = 0$ for any $i \in A, D$, therefore $\mathrm{E}V_{M_L}(l_A^*, L_A, r_A, r_D) = v_{M_L}(l_A^*, L_A)$ for any $r_A, r_D \in X$. Then because $V_{M_L}(l_D, L_D, r_i, R_i) \leq \max\{L_D, R_A, R_D\}$ for all $l_D \in X$, we obtain

$$\mathrm{E}V_{M_L}(l_A^*, L_A, r_A, r_D) > V_{M_L}(l_D, L_D, r_i, R_i),$$

for all $l_D \in X$, and $i \in \{A, D\}$.

Thus, $\mathrm{E}V_{M_L}(l_A^*, L_A, r_A, r_D) > \mathrm{E}V_{M_L}(l_D, L_D, r_A, r_D)$, and the median voter in party L, M_L, strictly prefers candidate L_A in the primary election regardless of candidate L_D's policy promise or the outcome of party R's primary. Then by the definition of equilibrium, candidate L_A is guaranteed victory in party L's primary in equilibrium. □

Proof of Theorem 2. Let $l_A^* = M_L$. Because $L_A - R_A \geq M_R - M_L$, $\pi_R(r_n^*, R_n, l_A^*, L_A) = 0$ for any $r_n^* \in X$ and thus $V_{M_R}(r_n^*, R_n, l_A^*, L_A) = L_A - M_R + M_L$. Further, obviously we also have $V_{M_L}(l_A^*, L_A, r_n^*, R_n) = L_A > V_{M_L}(l_D^*, L_D, r_n^*, R_n)$ for any $l_D^* \in X$ and each $n = A, D$. Therefore,

$$\mathrm{E}V_{M_L}(l_A^*, L_A, r_A^*, r_D^*) = L_A > \mathrm{E}V_{M_L}(l_D^*, L_D, r_A^*, r_D^*),$$

which completes the proof. □

Proof of Theorem 3. Suppose $M_R - M_L > L_A - R_A > 0$. Here, we consider the two cases: $M_R - M_L \leq R_A - L_D$ and $M_R - M_L > R_A - L_D$.

First, suppose $M_R - M_L \leq R_A - L_D$. We show that by choosing M_R, candidate L_A wins the primary and $w_A^{*L} = 1$. For all $r_A \in [M_L, M_R]$ and $l_D \in X$,

$$\pi_R(r_A, R_A, l_D, L_D) = 1;$$
$$V_{M_L}(l_D, L_D, r_A, R_A) = R_A - r_A + M_L.$$

If $r_A, r_D \geq M_R - (L_A - R_A)$, for each $n = A, D$,

$$V_{M_L}(M_R, L_A, r_n, R_n) = L_A - M_R + M_L.$$

Thus, $V_{M_L}(l_D, L_D, r_A, R_A) \leq V_{M_L}(M_R, L_A, r_A, R_A)$ for $r_A \geq M_R - (L_A - R_A)$. Next, notice that $M_R - M_L \leq R_A - L_D \leq L_A - L_D$. When $r_D \geq M_R - (L_A - R_A)$, regardless of $\pi_R(r_D, R_D, l_D, L_D)$, because $L_A - M_R \geq L_D - M_L$ and $R_D - r_D < L_A - M_R$,

$$\begin{aligned}
&V_{M_L}(l_D, L_D, r_D, R_D) = \\
&\pi_R(r_D, R_D, l_D, L_D)(L_D - l_D) + (1 - \pi_R(r_D, R_D, l_D, L_D))(R_D - r_D) + M_L \\
&\leq V_{M_L}(M_R, L_A, r_D, R_D).
\end{aligned}$$

Therefore, $V_{M_L}(l_D, L_D, r_n, R_n) \leq V_{M_L}(M_R, L_A, r_n, R_n)$ for each $n = A, D$. Then

$$\mathrm{E}V_{M_L}(l_D, L_D, r_A, r_D) < \mathrm{E}V_{M_L}(M_R, L_A, r_A, r_D).$$

Thus, we obtain $w_A^{*L} = 1$ in this case. Further, note that $r_A^*, r_D^* \in [M_R - (L_A - R_A), M_R]$, $l_D^* \in X$, and for each $n = A, D$, the following holds:

$$\pi_R(r_n^*, R_n, l_A^*, L_A) = 0$$
$$\max_{r_n} \mathrm{E}V_{M_R}(r_n, R_n, l_A^*, l_D^*) = \mathrm{E}V_{M_R}(r_n^*, R_n, l_A^*, l_D^*).$$

Notice that there is no l_D that candidate L_D can profitably deviate to by (15) and because for each $n = A, D$, $\pi_R(r_n^*, R_n, l_A^*, L_A) = 0$. Thus, l_D^* is unrestricted and we obtain the desired result.

Second, suppose $M_R - M_L > R_A - L_D$. Suppose that for each $n = A, D$, there is $\bar{d}_n > 0$ such that $\bar{d}_n = (L_A - R_n) - \frac{\left(F(M_R + \frac{L_D - R_n}{2})\right)^2}{\frac{1}{2}f(M_R + \frac{L_D - R_n}{2})}$. Fix $n = A, D$. When $\pi_R(r_n^*, R_n, l_A^*, L_A) = 0$ and $l_A^* = M_R$,

$$V_{M_L}(M_R, L_A, r_n^*, R_n) = L_A - M_R + M_L.$$

As shown in Lemma A-1, when candidate L_D maximizes $V_{M_L}(l_D, L_D, r_n^*, R_n)$, we have

$$V_{M_L}(\hat{l}_D(r_n^*), L_D, r_n^*, R_n) = R_n - r_n^* + \frac{\left(F(\frac{r_n^* + \hat{l}_D(r_n^*) + L_D - R_n}{2})\right)^2}{\frac{1}{2}f(\frac{r_n^* + \hat{l}_D(r_n^*) + L_D - R_n}{2})} + M_L.$$

Now, fix $n = A, D$. First, suppose $r_n^* \leq M_R$. Because $\bar{d}_n = (L_A - R_n) - \frac{\left(F(M_R + \frac{L_D - R_n}{2})\right)^2}{\frac{1}{2}f(M_R + \frac{L_D - R_n}{2})}$, we have

$$R_n - \left(M_R - \bar{d}_n\right) + \frac{\left(F(\frac{r_n^* + \hat{l}_D(r_n^*) + L_D - R_n}{2})\right)^2}{\frac{1}{2}f(\frac{r_n^* + \hat{l}_D(r_n^*) + L_D - R_n}{2})} = L_A - M_R.$$

By Lemma A-1, $M_L \leq \hat{l}_D(r_n^*) < r_n^* \leq M_R$ and thus $r_n^* + \hat{l}_D(r_n^*) < 2M_R$. By Assumption 1, we have

$$\max \frac{\left(F(\frac{r_n^* + \hat{l}_D(r_n^*) + L_D - R_n}{2})\right)^2}{\frac{1}{2}f(\frac{r_n^* + \hat{l}_D(r_n^*) + L_D - R_n}{2})} < F(M_R + \frac{L_D - R_n}{2}) \cdot \frac{F(M_R + \frac{L_D - R_n}{2})}{\frac{1}{2}f(M_R + \frac{L_D - R_n}{2})}.$$

Thus,

$$R_n - r_n^* + \frac{\left(F(\frac{r_n^* + \hat{l}_D(r_n^*) + L_D - R_n}{2})\right)^2}{\frac{1}{2}f(\frac{r_n^* + \hat{l}_D(r_n^*) + L_D - R_n}{2})} < L_A - M_R.$$

Thus $V_{M_L}(\hat{l}_D(r_n^*), L_D, r_n^*, R_n) < L_A - M_R + M_L$.

Second, $r_n^* > M_R$. Notice that $L_D \geq R_A - M_R + M_L$. Thus, for every $r_n \in [M_L, M_R]$,

$$\max_{l_D} V_{M_L}(\hat{l}_D(r_n), L_D, r_n, R_n) \geq L_D \geq R_n - M_R + M_L.$$

Then $V_{M_L}(\hat{l}_D(r_n), L_D, r_n, R_n) < L_A - M_R + M_L$ for $r_n \in [M_L, M_R]$ by the previous step of this proof.

Because n is arbitrary, for each $n = A, D$, $r_n^* \in [M_R - \bar{d}_n, M_R]$, $l_D^* \in X$, and the following holds:

$$\begin{aligned}
\pi_R(r_n^*, R_n, l_A^*, L_A) &= 0 \\
\max_{r_n} \mathrm{E}V_{M_R}(r_n, R_n, l_A^*, l_D^*) &= \mathrm{E}V_{M_R}(r_n^*, R_n, l_A^*, l_D^*) \\
\max_{l_D} \mathrm{E}V_{M_L}(l_D, L_D, r_A^*, r_D^*) &< \mathrm{E}V_{M_L}(M_R, L_A, r_A^*, r_D^*). \qquad (15)
\end{aligned}$$

Thus, we obtain $w_A^{*L} = 1$. This completes the proof. □

Now we sketch the proofs for Theorems 4 and 5. The detailed proofs are in the Appendix together with the other lemmas. We begin by stating the propositions that help characterize an equilibrium. Our strategy is that we first provide the properties that a set of policy promises must satisfy as an equilibrium in regime 2. Then we give the conditions under which such a set of equilibrium policy promises exists.

Suppose that $(l_A^*, l_D^*, r_A^*, r_D^*)$ and (w^{*L}, w^{*R}) constitute an equilibrium. Proposition 1 shows that in equilibrium, candidate L_A wins the primary of party L if $r_A^* > M_L$.

Proposition 1. *Suppose that the conditions for Theorem 4 hold. If $r_A^* > M_L$, then $w_A^{*L} = 1$ in equilibrium.*

Next, Proposition 2 provides the condition under which candidate R_A wins party R's primary. Proposition 2 is relevant in an equilibrium where candidate L_A's valence advantage is not sufficiently high to guarantee that candidate L_A always beats candidate R_A in the general election.

To formally state Proposition 2, for every $l_A \in X$, we define the set of candidate R_A's strategies that the median voter M_R strictly prefers to any policy promise that can be made by candidate R_D, by $\beta_R^\circ(l_A, L_A) \subseteq \beta_R(l_A, L_A)$ as follows:

$$\beta_R^\circ(l_A, L_A) \equiv \{r : V_{M_R}(l_A, L_A, r, R_A) > \max_{r_D \in X} V_{M_R}(l_A, L_A, r_D, R_D)\}.$$

If $\beta_R^\circ(l_A^*, L_A)$ is nonempty, Proposition 2 states that candidate R_A can win the primary with a probability of one by adopting a policy promise in $\beta_R^\circ(l_A^*, L_A)$. In contrast, $\beta_R^\circ(l_A^*, L_A)$ is empty only if candidate L_A is so strong that candidate R_A's winning probability of the general election is zero. This case belongs to regime 1.

Proposition 2. *Suppose that the conditions for Theorem 4 hold. If candidate L_A wins party L's primary and $\beta_R^\circ(l_A^*, L_A)$ is nonempty, $w_A^{*R} = 1$ in equilibrium.*

Proposition 3 provides the set of conditions that the equilibrium policy promises in regime 2 must satisfy.

Proposition 3. *Suppose $L_A - R_A < M_R - M_L$ and the conditions for Theorem 4 hold. Then an equilibrium satisfies*

$$L_D + r_A^* - R_A - l_D^* = \frac{F(\frac{r_A^*+l_D^*+L_D-R_A}{2})}{\frac{1}{2}f(\frac{r_A^*+l_D^*+L_D-R_A}{2})} \tag{16}$$

$$F(\frac{l_A^* + r_A^* + L_A - R_A}{2})(L_A - l_A^* - R_A + r_A^*) = (L_D + r_A^* - R_A - l_D^*) \cdot F(\frac{l_D^* + r_A^* + L_D - R_A}{2}) \tag{17}$$

$$R_D + r_D^* - L_A - l_A^* = \frac{(1 - F(\frac{r_D^*+l_A^*+L_A-R_D}{2}))}{\frac{1}{2}f(\frac{r_D^*+l_A^*+L_A-R_D}{2})} \tag{18}$$

$$(1 - F(\frac{l_A^* + r_A^* + L_A - R_A}{2}))(R_A + r_A^* - L_A - l_A^*) = (R_D + r_D^* - L_A - l_A^*) \cdot (1 - F(\frac{l_A^* + r_D^* + L_A - R_D}{2})). \tag{19}$$

An outline of the proof for Theorem 4 is as follows. Define a correspondence $BR = (BR_A^L, BR_D^L, BR_A^R, BR_D^R)$ with $BR : X^4 \to X^4$ as follows: for every $l_A, l_D, r_A, r_D \in X$, $BR_D^L(r_A) = \hat{l}_D(r_A), BR_D^R(l_A) = \hat{r}_D(l_A)$ and

$$BR_A^L(l_D) = \max\{l_A | V_{M_L}(l_A, L_A, r_A, R_A) = V_{M_L}(l_D, L_D, r_A, R_A)\};$$
$$BR_A^R(r_D) = \min\{r_A | V_{M_R}(l_A, L_A, r_A, R_A) = V_{M_R}(l_A, L_A, \bar{r}_D, R_D)\}.$$

Under two conditions (11) and (12), we show that $\beta_R(l_A, L_A) \cap \beta_L(r_A, R_A) = \varnothing$ when $r_A \geq \bar{y}$ and $l_A \leq \bar{x}$. After demonstrating that $\beta_R(l_A, L_A)$ and $\beta_L(r_A, R_A)$ are compact, by Kakutani's fixed point theorem, we show that there is a fixed point $(l_A^*, l_D^*, r_A^*, r_D^*)$ for a correspondence $BR = (BR_A^L, BR_D^L, BR_A^R, BR_D^R)$ in $[M_L, \bar{x}]^2 \times [\bar{y}, M_R]^2$. Then we explain that this fixed point is an equilibrium. To do so, we show that under conditions (11) and (12), $r_A^* = \hat{r}_A(l_A^*) > M_L$, and that $\beta_R^\circ(l_A^*, L_A)$ is nonempty. Then, by Proposition 1, we know that candidate L_A wins party L's primary with certainty. By Proposition 2, we also know that candidate R_A wins party R's primary with certainty. By showing that for each candidate, there is no profitable deviation from this fixed point, we obtain Theorem 4.

To provide an outline of the proof for Theorem 5, because a game is symmetric, we have either $\beta_R(l_A, L_A) \cap \beta_L(r_A, R_A) = \{\bar{m}\}$ (case (I)) or $\beta_R(l_A, L_A) \cap \beta_L(r_A, R_A) = \varnothing$ (case (II)). As described following the statement of Theorem 5, we show $\beta_R(l_A, L_A) \cap \beta_L(r_A, R_A) = \varnothing$ in case (II) and then, it follows that case (II) is a special case of Theorem 4. Otherwise, in case (I), $\bar{m} \in \beta_R(l_A, L_A) \cap \beta_L(r_A, R_A)$, and because the probability of winning the general election is maximized by choosing $\bar{m}$ for both candidates L_A and R_A when $p(\bar{m}) = \frac{1}{2}$, $\bar{m}$ is the equilibrium policy promises for the two advantaged candidates. If $p(\bar{m}) \neq \frac{1}{2}$, then the probability of winning the general election becomes discontinuous at $\bar{m}$, and $\bar{m}$ fails to be the equilibrium policy promises for the two advantaged candidates.

4 The One-Primary Scenario

The content of this section is developed from an earlier version of this paper, Takayama (2014). The following two propositions characterize the equilibrium in the case where only one party holds a primary. In the scenario we consider in this section, we assume that an incumbent in office is committed to a predetermined policy. Therefore, a policy chase between the two advantaged candidates as in the general scenario does not occur.

In Proposition 4, we consider the scenario where only party R holds a primary and candidate L_A is the incumbent in the office, committed to a policy promise $\bar{l}_A$. In Proposition 5, we consider the scenario where candidate R_A is the incumbent in the office, committed to a policy promise $\bar{r}_A$. In these propositions, we assume $\bar{r}_A > M_L$ and $\bar{l}_A < M_R$. It is highly unlikely that any incumbent would be committed to a policy promise that reflects the opponent's party ideology.

Proposition 4. *Let $\bar{l}_A < M_R$. The following holds in an equilibrium of the scenario where only party R holds a primary.*

- *(a) Suppose that $L_A - R_A \geq M_R - \bar{l}_A$. Then any (r_D^*, r_A^*) and any (w_D^{*R}, w_A^{*R}) such that $w_D^{*R} + w_A^{*R} = 1$ is an equilibrium.*
- *(b) Suppose that $L_A - R_A < M_R - \bar{l}_A$ and $M_R - \bar{l}_A > L_A - R_D$. If $\bar{l}_A + L_A - R_A \notin \beta_R(\bar{l}_A, L_A)$, then $\hat{r}_D(\bar{l}_A)$ is singleton and in the unique equilibrium, $r_D^* \in \hat{r}_D(\bar{l}_A)$, $r_A^* = \hat{r}_A(\bar{l}_A)$ and $(w_D^{*R}, w_A^{*R}) = (0, 1)$.*

Proposition 5. *Let $\bar{r}_A > M_L$. An equilibrium always exists in the scenario where only party L holds a primary. Further, the following holds in this scenario.*

- *(a) Suppose that $\min\{L_A - R_A, R_A - L_D\} \geq \bar{r}_A - M_L$. Then any (l_D^*, l_A^*) and $(w_D^{*L}, w_A^{*L}) = (0, 1)$ is an equilibrium.*
- *(b) Suppose that $R_A - L_D < \bar{r}_A - M_L$. Then $\hat{l}_D(\bar{r}_A)$ is singleton, and $l_D^* \in \hat{l}_D(\bar{r})$ and $(w_D^{*L}, w_A^{*L}) = (0, 1)$ constitute an equilibrium. Further, in equilibrium,*
  - *(1) If $\alpha(\bar{r}_A, R_A) \cap \beta_L(\bar{r}_A, R_A) \neq \emptyset$, then $l_A^* \in \alpha(\bar{r}_A, R_A) \cap \beta_L(\bar{r}_A, R_A)$,*
  - *(2) If $\alpha(\bar{r}_A, R_A) \cap \beta_L(\bar{r}_A, R_A) = \emptyset$, then $l_A^* = \hat{l}_A(\bar{r}_A)$.*

(c) *Suppose that* $R_A - L_D \geq \bar{r}_A - M_L > L_A - R_A$. *Then in equilibrium,* $l_A^* \in [\bar{r}_A - L_A + R_A, \bar{r}_A]$, $l_D^* \in X$, *and* $(w_D^{*L}, w_A^{*L}) = (0, 1)$.

Because $L_A \geq R_A$ by assumption, cases in Propositions 4 and 5 are not exactly symmetric. In Proposition 4, party R is the challenging party, while in Proposition 5, party L is the challenging party.

In both propositions, case (a) is the situation where candidate L_A is strong in terms of valence relative to both candidates R_A and the disadvantaged candidate. Case (b) is the situation where candidate L_A's valence advantage over candidate R_D is not sufficiently high to guarantee certain victory in the general election against candidate R_D. Case (c) of Proposition 5 is the situation where the disadvantaged candidate cannot win the general election with a strictly positive probability, while candidates L_A and R_A do not have significant difference in valence levels.

In Propositions 4 and 5, case (a) belongs to regime 1, and case (b) in Proposition 4 and case (b-2) in Proposition 5 belong to regime 2. Cases (b-1) and (c) in Proposition 5 are special cases of regime 1. In both cases, $\alpha_L(R_A, \bar{r}_A) \cap \beta_L(\bar{r}_A, R_A)$ is nonempty; however, a policy chase does not occur when candidate L_A chooses inside $\alpha_L(R_A, \bar{r}_A)$ because $\bar{r}_A$ is fixed.

5 Comparative Statics

In this section, we examine how changing model parameters such as candidate valence and voters' bliss points affects the equilibrium outcome. Our analyses yield testable hypotheses as propositions. First, we consider how a unilateral change of valence for candidate L_A affects the equilibrium. Second, we also evaluate the impact of a decrease in the variance of the distribution of the overall median voter's bliss point. We can interpret this as public opinion becoming more centrally concentrated. Third, we appraise the impact of a leftward shift in the distribution of the overall median voter's bliss point. We interpret this as public opinion shifting toward party L's ideology.

The extensive literature on candidate characteristics makes clear that voters care about the valence of candidates. Among many, Buttice and Stone (2012) use US congressional elections data and find that candidates' valence differences affect voting choice. Assessing the interplay of policy promises and valence is a complicated task. Theorem 6 is a theoretical attempt in this direction.

In Theorems 7 and 8, we evaluate how changes in public opinion affect electoral outcomes, by comparing two equilibria under two different distribution functions F_S and F_N. Although we compare equilibria between societies S and N, the comparison can also be interpreted as how electoral outcomes change when public opinions in a society change over time. For example, demographic changes could affect public opinions. As an example, Harmon (2018) uses data from Denmark and shows that an influx of immigrants could shift public opinions to the political right. Theorems 7 and 8 provide a better understanding of how electoral outcomes are affected by the change in the distribution of the overall median voter's bliss point.

Changing Valence

We first consider how a unilateral change in valence affects the equilibrium. The literature provides examples of the impact of changes to the valence of opposing parties' candidates in symmetric games. Adams and Merrill (2008) examine the model where each party has two candidates with identical valence. Then they show that as one of the two opposing party's candidates valence increases, the primary candidates of the other party locate their policy closer to the center to become more competitive in the general election. Intuitively, in the analysis of Adams and Merrill (2008), the candidates with relatively lower valence have to moderate their policy promises toward the center.

In contrast, Hummel (2013) considers the symmetric game defined in Definition 2 and shows that as the valence difference between the advantaged and disadvantaged candidates ($R_A - R_D = L_A - L_D$) increases, a candidate with higher valence is able to choose a more moderate policy promise in order to appeal to the general electorate while still winning the primary election. The following is the formal statement of this result (Proposition 4 from Hummel (2013)). Because the equilibrium in Hummel (2013) is equivalent to the one in the symmetric game when $a > L_A - L_D$ in our analysis, and is also characterized by (13) and (14) of Theorem 5, the following result also holds in the symmetric game of our analysis.

Proposition (Hummel, 2013). *Consider the symmetric game studied in Theorem 5. Then when $R_A - R_D = L_A - L_D$ decreases, r_D^* decreases while r_A^* increases.*

In contrast, the next theorem assesses a scenario where only candidate L_A's valence increases, while every other candidates' valence remains the same. The next theorem shows that when L_A increases, r_A^* and l_D^* move away from the center. The theorem also shows that r_D^* moves in the same direction with l_A^*. This is because candidate R_D chooses r_D^* to maximize the median voter M_R's expected payoff for voting for candidate R_D, given that candidate R_D will be facing candidate L_A with a policy promise l_A^* in the general election. More interestingly, the change in the direction of l_A^* depends on how strong in terms of valence candidate L_D is relative to candidate R_A. The theorem also states that there is an upper bound on candidate L_D's probability of winning the general election against candidate R_A, and if candidate L_D is strong enough relative to candidate R_A, then l_A^* moves toward M_L.

Theorem 6. *Suppose that the game is symmetric and $a > R_A - R_D = L_A - L_D$ holds.*[7] *Holding everything else except L_A constant, when L_A increases,*

[7]Recall from Definition 2 that f has a support $[-a, a]$.

- *r_A^* and $1-\pi_R(r_A^*, R_A, l_A^*, L_A)$ increase.*
- *l_D^* decreases.*
- *Define K to be a solution in $(0, \frac{1}{2})$ for $8K^2 - 6K + (\frac{1}{2} + f(0)r_A^*) = 0$. Then K exists uniquely. Further, the following holds.*
  - *If $1-\pi_R(r_A^*, R_A, l_D^*, L_D) < K$, l_A^*, and r_D^* increase.*
  - *Otherwise, l_A^* and r_D^* decrease.*

Increasing Diversity of Bliss Points

Here, we examine the situation where the center is the same in both societies, but the variance differs. For simplicity, we consider the symmetric game defined in Definition 2. Formally, we make the following assumption. We denote the two societies S and N.

Assumption 2. *Let F_S and F_N be cumulative distribution functions with symmetric probability density function f_S and f_N and a common support $[-a, a]$. Suppose that f_S and f_N satisfy the following: (1) $f_S(-a) < f_N(-a)$, (2) $f_S(0) > f_N(0)$, and (3) for all $x, x' \in [-a, 0]$ with $x < x'$, $\frac{f_S(x)}{f_S(x')} < \frac{f_N(x)}{f_N(x')} < 1$ and for all $x, x' \in [0, a]$ with $x < x'$, $\frac{f_S(x)}{f_S(x')} > \frac{f_N(x)}{f_N(x')} > 1$.*

Assumption 2 implies the following properties.

Property 1: For every $x \in [-a, 0]$, $\frac{F_S(x)}{f_S(x)} < \frac{F_N(x)}{f_N(x)}$. For every $x \in [0, a]$, $\frac{1-F_S(x)}{f_S(x)} < \frac{1-F_N(x)}{f_N(x)}$.

Property 2: For all $x \in [-a, 0]$, $F_S(x) < F_N(x)$ and for all $x \in [0, a]$, $F_S(x) > F_N(x)$.

Property 1 is verified as follows. For all $x, x' \in [-a, 0]$ with $x < x'$, we have

$$\int_{-\infty}^{x} \frac{f_S(t)}{f_S(x)} \mathrm{d}t = \frac{F_S(x)}{f_S(x)} < \frac{F_N(x)}{f_N(x)} = \int_{-\infty}^{x} \frac{f_N(t)}{f_N(x)} \mathrm{d}t.$$

By symmetry, $\frac{1-F_S(-x)}{f_S(-x)} < \frac{1-F_N(-x)}{f_N(-x)}$ for all $x \in [-a, 0]$, and thus we can rewrite it as follows: for all $x \in [0, a]$,

$$\frac{1-F_S(x)}{f_S(x)} < \frac{1-F_N(x)}{f_N(x)}.$$

Property 2 is obtained by comparing the regions below $f_S(x)$ or $f_N(x)$ and the line $x = 0$. By Assumption 2 and because $F_S(0) = F_N(0)$, at one point $x' \in [-a, 0]$, we have $f_S(x') = f_N(x')$. For $x \in [-a, x']$, $f_S(x) < f_N(x)$, which gives us $F_S(x) < F_N(x)$. Because $F_S(0) = F_N(0)$, we also have $F_S(x) < F_N(x)$ for every $x \in [x', 0]$. By symmetry, for all $x \in [-a, 0]$,

$$1 - F_S(-x) = F_S(x) < 1 - F_N(-x) = F_N(x),$$

which we can rewrite that for all $x \in [0, a]$, $F_S(x) > F_N(x)$.

We can interpret the two societies under Assumption 2 as where in society S, public opinion is clustered closer to the center, whereas in society N, public opinion is more diverse. Proposition 7 shows that when public opinion is more spread in society N, the distance between policy promises made by the advantaged and disadvantaged candidates of each party becomes larger. In other words, the policy promises of the two candidates in each party are also spread. An interesting point is that the advantaged candidate chooses a policy promise closer to the center, while the disadvantaged candidate chooses a policy promise closer to their party's median voter.

In what follows, y^s denotes the equilibrium variable y in society $s = N, S$.

Theorem 7. *Suppose that there are two societies N and S as defined in Assumption 2. Suppose that both societies satisfy the conditions for the symmetric game as defined in Definition 2, and a regime 2 equilibrium arises. Then* $r_A^S < r_A^N$ *and* $l_A^S > l_A^N$.

In equilibrium, the advantaged candidates R_A and L_A attempt to appeal to the voters of their own party because holding policy promises at the level of society S, the disadvantaged candidates' probability of winning the general election against the opponent party's nominee increases as the distribution shifts from F_S to F_N. This increases the attractiveness of the disadvantaged candidates to primary voters. The advantaged candidates have to shift their policy promises towards their own parties' bliss points to defend their prospects in the primaries. Then, the distance between the two policy promises increases. Proposition 7 describes this scenario.

Proposition 5 of Hummel (2013) analyses the scenario where $F_N(x) = F_S(\frac{x}{\sigma})$ and shows that when σ increases, the advantaged candidates choose policy promises that are closer to their parties' bliss points. Theorem 7 is consistent with this result and shows that for a general F_N that satisfies Assumption 2, this tendency holds.

Shifting Median Voter's Bliss Point

We consider two different societies N and S with differing distributions of the overall median voter's bliss points. In society $s = N, S$, the cumulative distribution function is given by F_s with a well-defined probability density function f_s. In what follows, $(y)_s$ denotes the equilibrium variable y in society s.

Suppose that for all $x, x' \in [-a, a]$, $\frac{f_S(x)}{f_S(x')} \leq \frac{f_N(x)}{f_N(x')}$ if $x < x'$. Let $\bar{m}_N$ satisfy $F_S(0) = F_N(\bar{m}_N) = \frac{1}{2}$.

Under this assumption, the distribution of the overall median voter is more skewed to the left in society N compared with society S. Further, the following three properties hold.

Property 3: For every $x \in [-a, a]$, $\frac{f_N(x)}{1-F_N(x)} \geq \frac{f_S(x)}{1-F_S(x)}$.

Property 4: For every $x \in [-a, a]$, $\frac{f_N(x)}{F_N(x)} \leq \frac{f_S(x)}{F_S(x)}$.

Property 5: For every $x \in [-a, a]$, $F_S(x) \le F_N(x)$ and hence $\bar{m}_S \ge \bar{m}_N$.

Property 3 can be verified by

$$\frac{1-F_N(x)}{f_N(x)} = \int_x^\infty \frac{f_N(t)}{f_N(x)}\mathrm{d}t \le \int_x^\infty \frac{f_S(t)}{f_S(x)}\mathrm{d}t = \frac{1-F_S(x)}{f_S(x)},$$

and similarly property 4 can be verified by

$$\frac{F_N(x)}{f_N(x)} = \int_0^x \frac{f_N(t)}{f_N(x)}\mathrm{d}t \ge \int_0^x \frac{f_S(t)}{f_S(x)}\mathrm{d}t = \frac{F_S(x)}{f_S(x)}.$$

Property 5 can be obtained by either property 3 or 4, because

$$F_S(x) = \exp\left(-\int_x^\infty \frac{f_S(t)}{F_S(t)}\mathrm{d}t\right) \le \exp\left(-\int_x^\infty \frac{f_N(t)}{F_N(t)}\mathrm{d}t\right) = F_N(x). \tag{20}$$

Property 5 is often called *first-order stochastic dominance* when the inequality is strict, and we can say that F_N first-order stochastically dominates F_S. This implies that the overall median voter tends to be closer to $-a$ with a higher probability in society N than in society S. Properties 1 and 2 imply the same.

We consider how the equilibrium policy promises move, when society S shifts to society N. Although the formula of the changes that we obtain in the proof can be applied to any size of change, to obtain a simple intuition, we consider an infinitesimal change in the distribution function from F_S to F_N:

- Under the scenario of *extremists' change*, for $x \in [-a, -a+\bar{\epsilon}]$, $f_N(x) = (1+d(x))f_S(x)$, while for $x \in [a-\bar{\epsilon}, a]$, $f_N(x) = (1-d(x))f_S(x)$ where $d(x)$ is strictly smaller than ϵ for every x.

- Under the scenario of *centrists' change*, for $x \in [-\bar{\epsilon}, 0]$, $f_N(x) = (1+d(x))f_S(x)$, while for $x \in [0, \bar{\epsilon}]$, $f_N(x) = (1-d(x))f_S(x)$ where $d(x)$ is strictly smaller than ϵ for every x.

Under the scenario of centrists' change, the shift of public opinion implies that there is a large number of voters shifting their bliss points closer towards candidate L_A's policy promise from the center. As the shift is from the center, these voters' bliss points remain further away from candidate L_D's policy promise. The shift in public opinion towards party L's bliss point also benefits candidate L_A much more than candidate L_D in the primary election, because these new voters are attracted by both candidate L_A's policy promise and valence. So candidate L_A faces less constraint by candidate L_D's primary challenge, and hence is able to shift her policy towards the center to capture a greater share of votes in the general election.

Under the scenario of extremists' change, the shift in public opinion implies a large increase in voters with bliss points between party L's bliss point and candidate L_D's policy promise. These new voters strictly prefer candidate L_A in the general election, forcing candidate L_A to shift her policy promise towards the party's bliss point to protect her chances of winning the primary election.

Theorem 8. *Suppose that society S satisfies the conditions for the symmetric game. Suppose that in both societies, a regime 2 equilibrium arises. Then when the change between F_S and F_N is sufficiently small:*

- *Under the scenario of centrists' change, $l_A^N > l_A^S$.*
- *Under the scenario of extremists' change, $l_A^N < l_A^S$.*

6 Extensions Including Endogenous Entry

In this section, we discuss the robustness of our results under various extensions and evaluate a modified version of our model including the endogenous entry of a candidate.

Hummel (2013) explains that in the symmetric game, a regime 2 equilibrium is robust under the extensions such that there are three candidates in the primaries, candidates can take different policy promises in two stages, parties hold open primaries, and politicians also care about the policy outcomes. We do not repeat the same argument about a regime 2 equilibrium because the arguments in Hummel (2013) also apply to our results in the general setting. However, a couple of points require noting.

First, a regime 1 equilibrium is also robust under these extensions, because in this regime, candidate L_A is so strong in terms of valence and her status does not change even under these extensions. As mentioned in footnote 6, Hummel (2013) assumes that if a candidate loses the primary election, then the candidate obtains a utility of 0, if a candidate wins party k's primary, then she obtains a utility of $u_k \geq 0$ for each $k = L, R$, and the winner of the general election receives an additional utility of $u \geq 0$. We do not impose any particular utility and simply assume that each candidate attempts to maximize the probability of winning office. If we make a similar assumption to that of Hummel (2013), a regime 1 equilibrium still arises under the conditions of Theorems 1, 2, and 3, although candidates R_A and R_D would attempt to maximize the chance of winning party R's primary.

Second, in a regime 2 equilibrium, when candidates also care about the policy outcomes, because the disadvantaged candidates cannot win the primaries with a strictly positive probability, they still maximize the expected payoffs of their own parties' median voter. The electoral competition with mixed motivation where politicians care about both winning office and policy outcomes is studied in Takayama et al. (2019). In the Takayama et al. (2019) model, there are two candidates, who we can consider as our advantaged candidates. In a similar way as for our current proof for Theorem 4 or Takayama et al. (2019) which uses the recently developed concept in the literature of discontinuous games by McLennan et al. (2011), we can show that when the β sets do not overlap, there is an equilibrium. However, depending on the positions of the advantaged candidates' bliss points, advantaged candidates may take policy promises inside of the β sets instead of the boundaries of the sets.

Finally, we consider a modified version of our model such that a candidate has a choice of entry to the elections. We can model this situation by using the concept of *subgame-perfect equilibria*. By using Propositions 4 and 5 together with Theorems 1, 2, 3, 4, and 5, we can consider various scenarios where an advantaged or disadvantaged candidate decides to enter the electoral race or not or the opponent party does not or does hold a primary.

Here, we consider a scenario where candidate R_A is committed to a policy promise $\bar{r}_A$ and there is no primary in party R. We evaluate the endogenous decision of entry by candidate L_D. Suppose that candidate L_D has a bliss point $\bar{l}_D$ and the following utility: for $P_j = L_A, R_A$,

$$u_{\bar{l}_D}(x, P_j) = P_j - \kappa|x - \bar{l}_b|. \tag{21}$$

Compared with a voter's preference defined in (1), (21) has an additional element κ for the disutility caused by the policy difference with $\bar{l}_D$. One can imagine that as a politician, candidate L_D values valence more, or policy promises more. This is captured by κ. To consider a general case, we do not impose any assumption on κ, but obviously when $\kappa = 1$, candidate L_D's utility coincides with a voter's utility.

Proposition 6. *Suppose $L_A - R_A > 0$ holds, and candidate R_A is committed to policy $\bar{r}_A$, and candidate L_D's bliss point $\bar{l}_D$ satisfies $\bar{l}_D \leq M_L$. Suppose that if candidate L_D does not enter, candidate L_A chooses $l_A^0 = \bar{r}_A - d$, while if candidate L_D enters, candidate L_A chooses $l_A^1 = \bar{l}_A(\bar{r}_A, R_A)$ as in a regime 2 equilibrium. Then $d \leq L_A - R_A$ holds. Let $\bar{\kappa} = \frac{(1-F(\frac{\bar{r}_A + l_A^1 + L_A - R_A}{2}))(L_A - R_A)}{F(\frac{\bar{r}_A + l_A^1 + L_A - R_A}{2})(\bar{r}_A - l_A^1) - d}$. If $\kappa < \bar{\kappa}$, candidate L_D does not enter. Otherwise, if $\kappa > \bar{\kappa}$, candidate L_D enters.*

Proof of Proposition 6. If candidate L_D does not run for the primary, candidate L_A chooses a policy promise in $\alpha_L(\bar{r}_A, R_A)$ as described in Proposition 5. Suppose candidate L_A chooses $l_A^0 = \bar{r}_A - d \in \alpha_L(\bar{r}_A, R_A)$, and thus we have $d \leq L_A - R_A$. Then the payoff that candidate L_D receives is $L_A - \kappa(\bar{r}_A - d) + \kappa\bar{l}_D$. If candidate L_D runs and candidate L_A chooses l_A^1, because $l_A^1 \geq M_L \leq \bar{l}_D$, the expected payoff of candidate L_D is

$$(1 - F(\frac{\bar{r}_A + l_A^1 + L_A - R_A}{2}))(R_A - \kappa\bar{r}_A + \kappa\bar{l}_D) + F(\frac{\bar{r}_A + l_A^1 + L_A - R_A}{2})(L_A - \kappa l_A^1 + \kappa\bar{l}_D).$$

Thus, if $(1 - F(\frac{\bar{r}_A + l_A^1 + L_A - R_A}{2}))(L_A - R_A) < F(\frac{\bar{r}_A + l_A^1 + L_A - R_A}{2})\kappa(\bar{r}_A - l_A^1) - \kappa d$, candidate L_D enters. Notice that because $\pi_R(\bar{r}_A, R_A, l_A^1, L_A) > 0$ in a regime 2 equilibrium, $\bar{r}_A > l_A^1 + L_A - R_A$. Then, when $\kappa = 0$, $(1 - F(\frac{\bar{r}_A + l_A^1 + L_A - R_A}{2}))(L_A - R_A) > 0$. On the other hand, suppose $F(\frac{\bar{r}_A + l_A^1 + L_A - R_A}{2})(\bar{r}_A - l_A^1) > d$. In this case, when κ is sufficiently large, the above inequality holds. Thus, we obtain the desired result. □

Proposition 6 points out several interesting features of the revised game with endogenous entry. First, if κ is large and the difference in valences between the two advantaged candidates is less of a concern, candidate L_D enters and influences the policy outcome. However, if κ is small and candidate L_D is less concerned about the policy outcome, she values more the difference in valences between the two advantaged candidates and does not enter. In other words, if candidate L_D places less importance on policy outcomes, and the advantaged candidate in the opponent party has a sufficiently low valence level relative to the advantaged candidate in her own party, she does not enter, because her entry will constrain the policy promise of the advantaged candidate in her own party, potentially damaging the probability of her party winning the general election.

Second, as $l_A^0 = \bar{r}_A - d$ is candidate L_A's choice, candidate L_A can influence candidate L_D's decision of entry. By choosing d such that $F(\frac{\bar{r}_A + l_A^1 + L_A - R_A}{2})(\bar{r}_A - l_A^1) - d$ is greater than 0 but sufficiently small, candidate L_A can make the situation of $\kappa < \bar{\kappa}$ happen. In a sense, by choosing a d that is sufficiently close to $F(\frac{\bar{r}_A + l_A^1 + L_A - R_A}{2})(\bar{r}_A - l_A^1)$ and compensating for the disutility coming from the difference of policy outcomes for not entering to candidate L_D, candidate L_A can induce the situation where candidate L_D does not enter.

Hummel (2013) also considers the endogenous entry decision into primaries in the symmetric game. Given that the valence levels of the advantaged candidates are identical for the two parties in Hummel (2013)'s analysis, the impact of the opponent party's advantaged candidate's valence on the disadvantaged candidate's decision to enter the primary election cannot be captured. Proposition 6 adds another interesting insight into the decisions of candidates about whether or not to enter the race.

7 Concluding Remarks

In this paper, we proposed a simple model of a two-stage electoral process with primaries followed by a general election. By allowing voters to value both policy promises and valence, we showed that there are possibly two equilibrium regimes, and provided the conditions under which each regime arises. Our model is simple and tractable. At the same time, it is sufficiently general to cover special scenarios such as a symmetric game or the case of a one-party primary, and may also lend itself to further extensions such as uncertainty over candidates' valence. Our model is also rich as we provide the sufficient conditions for each equilibrium regime, and can obtain the equilibrium in Hummel (2013) as a special case of our equilibrium analysis.

Aragones and Palfrey (2002) use the standard Downsian model of two-candidate elections of two politicians such that one is more advantaged in terms of valence than the other, and show that pure strategy Nash equilibria often fail to exist, because the more advantaged candidate's optimal choice is to choose a policy promise that is as close as possible to the less advantaged candidate's policy promise, but the disadvantaged candidate has an incentive to move away from the advantaged candidate's policy promise. Intuitively, a Nash equilibrium fails to exist because an unending policy chase arises in which the advantaged candidate continually pursues the disadvantaged candidate's policy promise over the entire policy space. In our model, a

Nash equilibrium could also fail to exist. However, when β limits the extent of the chase by the advantaged candidate, a regime 2 equilibrium exists. For example, if party R's general election candidate chooses a policy promise outside of β_L, party L's general election candidate cannot follow without losing the primary election.

Because our framework is simple but general, there are many ways to extend the simple framework studied here. In the real world, primaries sequentially aggregate a number of different individual choices as well as party leaders and decide over time which candidate to endorse. Next, we could attempt to relax the assumption about the same policy across the two stages of the elections by allowing politicians to revise a policy promise within a certain constraint in the general election. As there is valence in the model, and valence also captures the honesty of candidates, it may be interesting to see how the equilibrium policy promises vary, if candidates can choose a different policy in the general election within a certain distance, which is a function of valence, from their policy promise in the primary.

Ali and Kartik (2012) consider a model such that a sequence of agents choose between two actions in a fixed sequential order, observing the choices of all prior agents, and allowing an agent's payoff to depend on the profile of actions of any subset of all agents. Callander (2007) develops a model of sequential voting to shed light on when and why momentum is observed and bandwagons begin in order to understand how electoral dynamics affect the selection of candidates. Dekel and Piccione (2000) compare the sets of equilibria in sequential voting and simultaneous voting. Deltas et al. (2016) propose the simplest framework in which the issues of learning and coordination interact with each other, estimate the structural parameters of the theoretical model, and conduct some policy experiments.

In this paper, we assumed that valence and the policy promise are independent and that voters' utility is a function of these two variables. Also, in our paper, we assumed that the valence values are exogenously given, and are public information. It may also be interesting to relax these assumptions and analyze how the equilibrium changes.

Klumpp and Polborn (2006) and Schwabe (2010) develop a model of campaign finance in primary elections in which campaigns supply hard information about candidates' electability and evaluate the role of campaign advertising in primaries. While our settings are quite different, Anh and Oliveros (2012) examine elections that simultaneously decide multiple issues, where voters have independent private values over bundles of issues, while Lizzeri and Persico (2001) consider public goods provision under various electoral incentives. Adams and Merrill (2009) have developed a model of policy-seeking parties in a parliamentary democracy, and in their model, party elites are uncertain about voters' evaluations of the parties' valence. Finally, Adams et al. (2013) extend their model to situations where voters hold coalitions of parties collectively responsible for their valence-related performances. It would then be interesting to specify uncertainty about valence and consider aspects of the collective responsibility for this, as in Adams et al. (2013).

Recently, Hirano and Snyder (2012) examined the impact of scandals on primary and general election competitions using data for US House incumbents running for reelection over the period 1978 to 2008. They find that incumbents involved in relatively serious scandals are very likely to be challenged in the primary and lose. In their analysis, they point out that primary elections can improve accountability by removing incumbents who face reelection in districts dominated by the one political party.

Examining how different institutional arrangements affect the equilibrium outcomes in our model would also be interesting. One example is crossover voting. Chen and Yang (2002) and Cho and Kang (2014) consider an open primary election with two competing candidates and the possibility of crossover voting. In our current model, crossover voting is eliminated by assumption and our setting is then more akin to closed primaries.

Another interesting question is whether the open primary system leads to more moderate outcomes than our current model. Galderisi and Ezra (2001) state that in general, the direct primary where voters directly vote for a primary candidate was thought to be a method of controlling the corrupting influences of the urban political machine. Ware (2002) reviews the political history of the US, and asserts that the direct primary was the result of an attempt by politicians to subject previously informal procedures to formal rules, which started in the late 1880s. Gerber and Morton (1998) empirically examine how different primary election laws affect the types of candidates elected in nonpresidential American elections. It would be useful to evaluate whether these theoretical predictions match the empirical observations in Gerber and Morton (1998). Casas (2018) proposes a model and shows that open primaries result in candidates that are expected to be ideologically more extreme than those in closed primaries. Our model is stable enough to lend itself to these various extensions.

Appendix: Proofs

Derivation of Equations (3) *and* (4). Let $\sigma(l, L, r, R)$ satisfy

$$v_{\sigma(l,L,r,R)}(l, L) = v_{\sigma(l,L,r,R)}(r, R).$$

Then, $\sigma(l, L, r, R)$ does not exist, if $|r - l| < |R - L|$. If $|r - l| > |R - L|$, let

$$\sigma(l, L, r, R) = \begin{cases} \frac{(L-R)+(l+r)}{2} & \text{if } l < r \\ \frac{(R-L)+(l+r)}{2} & \text{if } r < l, \end{cases} \tag{A-1}$$

and the probability that candidate R wins the general election is

$$\pi_R(r, R, l, L) = \begin{cases} \Pr(M > \sigma(l, L, r, R)) = 1 - F(\sigma(l, L, r, R)) & \text{if } l < r \\ \Pr(M < \sigma(l, L, r, R)) = F(\sigma(l, L, r, R)) & \text{if } r < l. \end{cases}$$

□

Lemma A-1.

1. *Suppose that a winner in party L's primary has a policy promise $\bar{l}$ and a valence $\bar{L}$. Then, for each $i = A, D$, $V_{M_R}(\bar{l}, \bar{L}, x, R_i)$ is a continuous function of x on $[\bar{l}, M_R]$, and when $\pi_R(x, R_i, \bar{l}, \bar{L}) > 0$,*
   - *$V_{M_R}(\bar{l}, \bar{L}, x, R_i)$ is*
     - *single-peaked if $\frac{\partial}{\partial x}\left(V_{M_R}(\bar{l}, \bar{L}, x, R_j)\right)\big|_{x=M_R} \leq 0$;*
     - *increasing otherwise.*
   - *When $V_{M_R}(\bar{l}, \bar{L}, x, R_i)$ is singled-peaked, the maximal value is*

     $$\bar{l} + \bar{L} - M_R + \frac{\left(1 - F(\frac{x^* + \bar{l} + \bar{L} - R_i}{2})\right)^2}{\frac{1}{2} f(\frac{x^* + \bar{l} + \bar{L} - R_i}{2})},$$

     where x^* satisfies the first-order condition

     $$R_i + x^* - \bar{L} - \bar{l} = \frac{(1 - F(\frac{x^* + \bar{l} + \bar{L} - R_i}{2}))}{\frac{1}{2} f(\frac{x^* + \bar{l} + \bar{L} - R_i}{2})}.$$
   - *The maximizer of $V_{M_R}(\bar{l}, \bar{L}, x, R_i)$ exists uniquely on $[\bar{l}, M_R]$.*

2. *Suppose that a winner in party R's primary has a policy promise $\bar{r}$ and a valence $\bar{R}$. For every $i = A, D$, $V_{M_L}(x, L_i, \bar{r}, \bar{R})$ is a continuous function of x on $[M_L, \bar{r}]$, and when $\pi_R(\bar{r}, \bar{R}, x, L_i) < 1$,*

- *$V_{M_L}(x, L_i, \bar{r}, \bar{R})$ is*
  - *single-peaked if $\frac{\partial}{\partial x}\left(V_{M_L}(x, L_i, \bar{r}, \bar{R})\right)\big|_{x=M_L} \geq 0$;*
  - *decreasing otherwise.*
- *When $V_{M_L}(x, L_i, \bar{r}, \bar{R})$ is singled-peaked, the maximal value is*

$$\bar{R} - \bar{r} + \frac{\left(F(\frac{\bar{r}+x^*+L_i-\bar{R}}{2})\right)^2}{\frac{1}{2}f(\frac{\bar{r}+x^*+L_i-\bar{R}}{2})} + M_L,$$

  where x^* satisfies the first-order condition:

$$L_i + \bar{r} - \bar{R} - x^* = \frac{F(\frac{\bar{r}+x^*+L_i-\bar{R}}{2})}{\frac{1}{2}f(\frac{\bar{r}+x^*+L_i-\bar{R}}{2})}.$$

- *The maximizer of $V_{M_L}(x, L_i, \bar{r}, \bar{R})$ always exists uniquely on $[M_L, \bar{r}]$.*

Proof of Lemma A-1. By symmetry, we only prove the first case. Taking the first derivatives, then

$$\frac{\partial}{\partial x}\left(V_{M_R}(\bar{l}, \bar{L}, x, R_i)\right) = \pi_R'(x, R_i, \bar{l}, \bar{L})(R_i + x - (\bar{l} + \bar{L})) + \pi_R(x, R_i, \bar{l}, \bar{L}),$$

where $\pi_R'(x, R_i, \bar{l}, \bar{L})$ denotes the first derivative of $\pi_R(x, R_i, \bar{l}, \bar{L})$ with respect to x.

When $x = \bar{l}$, $R_i + x - (\bar{l} + \bar{L}) \leq 0$. Further, $\pi_R'(x, R_i, \bar{l}, \bar{L}) < 0$ when $x > \bar{l}$. Thus $\frac{\partial}{\partial x}\left(V_{M_R}(\bar{l}, \bar{L}, x, R_i)\right) > 0$. If

$$\frac{\partial}{\partial x}\left(V_{M_R}(\bar{l}, \bar{L}, x, R_i)\right)\Big|_{x=M_R} \leq 0,$$

then there is an $\bar{x}$ to satisfy

$$\pi_R'(\bar{x}, R_i, \bar{l}, \bar{L})(R_i + \bar{x} - (\bar{l} + \bar{L})) + \pi_R(\bar{x}, R_i, \bar{l}, \bar{L}) = 0.$$

Then, note that when $\bar{l} > x$ and $\pi_R(x, R_i, \bar{l}, \bar{L}) > 0$,

$$-\frac{\pi_R'(\bar{x}, R_i, \bar{l}, \bar{L})}{\pi_R(\bar{x}, R_i, \bar{l}, \bar{L})} = \frac{\frac{1}{2}f(\frac{\bar{x}+\bar{l}+\bar{L}-R_i}{2})}{1 - F(\frac{\bar{x}+\bar{l}+\bar{L}-R_i}{2})}. \tag{A-2}$$

Also,

$$-\frac{\pi_R'(\bar{x}, R_i, \bar{l}, \bar{L})}{\pi_R(\bar{x}, R_i, \bar{l}, \bar{L})} = \frac{1}{R_i + \bar{x} - (\bar{l} + \bar{L})}. \tag{A-3}$$

By Assumption 1, for every $x > \bar{x}$, we have

$$\frac{\frac{1}{2}f(\frac{\bar{x}+\bar{l}+\bar{L}-R_i}{2})}{1 - F(\frac{\bar{x}+\bar{l}+\bar{L}-R_i}{2})} > \frac{1}{R_i + x - (\bar{l} + \bar{L})}$$

and for every $x < \bar{x}$, we have

$$\frac{\frac{1}{2}f(\frac{\bar{x}+\bar{l}+\bar{L}-R_i}{2})}{1-F(\frac{\bar{x}+\bar{l}+\bar{L}-R_i}{2})} < \frac{1}{R_i + x - (\bar{l}+\bar{L})}.$$

When there is no such $\bar{x}$, the expected payoff is increasing. Thus we obtain the first and second statements.

By (A-3), we obtain the first-order condition, and the maximal value is obtained by

$$\begin{aligned}
&V_{M_R}(\bar{l},\bar{L},x^*,R_i) + M_R \\
&= F(\frac{x^*+\bar{l}+\bar{L}-R_i}{2})\left(\bar{L}+\bar{l}\right) + (1-F(\frac{x^*+\bar{l}+\bar{L}-R_i}{2}))\left(R_i + x^*\right) \\
&= F(\frac{x^*+\bar{l}+\bar{L}-R_i}{2})\left(\bar{L}+\bar{l}\right) \\
&\qquad + (1-F(\frac{x^*+\bar{l}+\bar{L}-R_i}{2}))\left(\bar{l}+\bar{L}+\frac{(1-F(\frac{x^*+\bar{l}+\bar{L}-R_i}{2}))}{\frac{1}{2}f(\frac{x^*+\bar{l}+\bar{L}-R_i}{2})}\right).
\end{aligned}$$

To prove the last item, notice that when $\pi_R(x,R_i,\bar{l},\bar{L}) > 0$,

$$x - \bar{l} > \max\{R_i - \bar{L}, \bar{L} - R_i\}.$$

Then, we obtain $v_{M_R}(x,R_i) > v_{M_R}(\bar{l},\bar{L})$. Hence, if there is some $r \geq \bar{l}$ with $\pi_R(r,R_i,\bar{l},\bar{L}) = 0$, we have

$$V_{M_R}(\bar{l},\bar{L},r,R_i) = \bar{L}+\bar{l}-M_R < \max V_{M_R}(\bar{l},\bar{L},x,R_i).$$

Thus, from the first item of this lemma, the maximizer is either x^* that satisfies the first-order condition, or M_R. This completes the proof. $\square$

Lemma A-2.

1. *Suppose that a winner in party L's primary has a policy promise $\bar{l}$ and a valence $\bar{L}$. Then, for each $i = A, D$, $V_{M_R}(\bar{l},\bar{L},x,R_i)$ is a continuous function of x on $[\bar{l}, M_R]$, and when $\pi_R(x,R_i,\bar{l},\bar{L}) > 0$, $V_{M_R}(\bar{l},\bar{L},x,R_i)$ is increasing for $x \leq \bar{l}$.*

2. *Suppose that a winner in party R's primary has a policy promise $\bar{r}$ and a valence $\bar{R}$. For every $i = A, D$, $V_{M_L}(x,L_i,\bar{r},\bar{R})$ is a continuous function of x on $[M_L, \bar{r}]$, and when $\pi_R(\bar{r},\bar{R},x,L_i) < 1$, $V_{M_L}(x,L_i,\bar{r},\bar{R})$ is decreasing for $x \geq \bar{r}$.*

Proof of Lemma A-2. By symmetry, we only prove the second statement. Suppose that for some $x \geq \bar{r}$, $\pi_R(\bar{r},\bar{R},x,L_i) < 1$, and note that the first derivative of $V_{M_L}(x,L_i,\bar{r},\bar{R})$ with respect to x when $x > \bar{r}$ is

$$f(\frac{\bar{R}-L_i+x+\bar{r}}{2})((L_i - x) - (\bar{R}-\bar{r})) - F(\frac{\bar{R}-L_i+x+\bar{r}}{2}).$$

When $\pi_R(\bar{r},\bar{R},x,L_i) < 1$ and $\bar{r} < x$, $\bar{r} < x - (L_i - \bar{R})$. Thus, the first derivative of $V_{M_L}(x,L_i,\bar{r},\bar{R})$ is negative. This completes the proof. $\square$

Lemma A-3.

1. *Suppose that a winner in party L's primary has a policy promise $\bar{l}$ and a valence $\bar{L}$. Then, for each $i = A, D$, the maximizer of $V_{M_R}(\bar{l}, \bar{L}, x, R_i)$ for $x > \bar{l}$ maximizes $V_{M_R}(\bar{l}, \bar{L}, x, R_i)$ for $x \in X$.*

2. *Suppose that a winner in party R's primary has a policy promise $\bar{r}$ and a valence $\bar{R}$. For every $i = A, D$, the maximizer of $V_{M_L}(x, L_i, \bar{r}, \bar{R})$ maximizes $V_{M_L}(x, L_i, \bar{r}, \bar{R})$ for $x \in X$.*

Proof of Lemma A-3. By Lemmas A-1 and A-2, the maximizer of $V_{M_L}(x, L_i, \bar{r}, \bar{R})$ for $x < \bar{r}$ maximizes $V_{M_L}(x, L_i, \bar{r}, \bar{R})$ for $x \in X$ with $\pi_R(\bar{r}, \bar{R}, x, L_i) < 1$. When $\pi_R(\bar{r}, \bar{R}, x, L_i) = 1$, $L_i < \bar{R}$ and $x \in [\bar{r} - \bar{R} + L_i, \bar{r} + \bar{R} - L_i]$. Thus, $L_i - x \leq \bar{R} - \bar{r}$. Thus,

$$V_{M_L}(x, L_i, \bar{r}, \bar{R}) = \bar{R} - \bar{r} + M_L.$$

If there is $x < \bar{r}$ with $\pi_R(\bar{r}, \bar{R}, x, L_i) < 1$, the maximizer of $V_{M_L}(x, L_i, \bar{r}, \bar{R})$ for $x < \bar{r}$ maximizes $V_{M_L}(x, L_i, \bar{r}, \bar{R})$ for $x \in X$. If there is no such x, for every $x < \bar{r}$,

$$V_{M_L}(x, L_i, \bar{r}, \bar{R}) = \bar{R} - \bar{r} + M_L.$$

Thus, even in this case, every $x < \bar{r}$ is a maximizer of $V_{M_L}(x, L_i, \bar{r}, \bar{R})$. This completes the proof. □

Lemma A-4. *For every $l_A, r_A \in X$,*

- *If $\pi_R(r, R_A, l_A, L_A) > 0$ for some $r \in X$, $\beta^\circ_R(l_A, L_A)$ is nonempty.*
- *If $\pi_R(r_A, R_A, l, L_A) < 1$ for some $l \in X$, $\beta^\circ_L(r_A, R_A)$ is nonempty.*
- *$\beta_L(r_A, R_A)$ and $\beta_R(l_A, L_A)$ are nonempty.*

Proof of Lemma A-4. As the proof is symmetric, we only prove the claims for $\beta_R(l_A, L_A)$. We show that

$$\max_r V_{M_R}(l_A, L_A, r, R_A) \geq \max_r V_{M_R}(l_A, L_A, r, R_D) \tag{A-4}$$

and if there is an r with $\pi_R(r, R_A, l_A, L_A) > 0$, $\beta^\circ_R(l_A, L_A)$ is nonempty.

Fix l_A and r in X. If $\pi_R(r, R_A, l_A, L_A) > 0$, we have $v_{M_R}(r, R_A) > v_{M_R}(l_A, L_A)$. Because $v_{M_R}(r, R_A) > v_{M_R}(r, R_D)$ and $\pi_R(r, R_A, l_A, L_A) > 0$ implies $\pi_R(r, R_A, l_A, L_A) > \pi_R(r, R_D, l_A, L_A)$,

$$V_{M_R}(l_A, L_A, r, R_A) > V_{M_R}(l_A, L_A, r, R_D). \tag{A-5}$$

Thus, indeed $\beta_R^\circ(l_A, L_A)$ is nonempty, which gives us the first statement. Alternatively, if $\pi_R(r, R_A, l_A, L_A) = 0$, then

$$V_{M_R}(l_A, L_A, r, R_A) = V_{M_R}(l_A, L_A, r, R_D) = L_A - M_R + l_A. \tag{A-6}$$

Because r and l_A are arbitrary and either (A-5) or (A-6) holds, we conclude that for every $r \in X$, (A-4) holds. $\square$

For every $l_A, r_A \in X$, recall that we have defined $\hat{l}_D(r_A)$ and $\hat{r}_D(l_A)$ by

$$\hat{l}_D(r_A) = \arg\max_{l_D \in X} V_{M_L}(l_D, L_D, r_A, R_A);$$
$$\hat{r}_D(l_A) = \arg\max_{r_D \in X} V_{M_R}(l_A, L_A, r_D, R_D),$$

and $\hat{l}_A(r_A)$ and $\hat{r}_A(l_A)$ by

$$\hat{l}_A(r_A) = \max\{l_A | V_{M_L}(l_A, L_A, r_A, R_A) \geq V_{M_L}(\bar{l}_D, L_D, r_A, R_A) \text{ for } \bar{l}_D \in \hat{l}_D(r_A)\};$$
$$\hat{r}_A(l_A) = \min\{r_A | V_{M_R}(l_A, L_A, r_A, R_A) \geq V_{M_R}(l_A, L_A, \bar{r}_D, R_D) \text{ for } \bar{r}_D \in \hat{r}_D(l_A)\}.$$

Lemma A-5. *Suppose $M_R - M_L > L_A - R_A$. If $L_A = R_A$, or $L_A > R_A$ and there is $\bar{x}$ such that*

$$\frac{\frac{1}{2}f(\frac{M_R+\bar{x}+L_A-R_D}{2})}{\left(1-F(\frac{M_R+\bar{x}+L_A-R_D}{2})\right)^2} = \frac{1}{L_A - R_A},$$

then for every $x \leq \bar{x}$, $x + L_A - R_A$ does not belong to $\beta_R(x, L_A)$.

Proof of Lemma A-5. Fix $l_A = x$. Let $\hat{x}_R \in \bar{r}_D(x)$. Then $x + L_A - R_A$ does not belong to $\beta_R(x, L_A)$ if and only if

$$V_{M_R}(x, L_A, \hat{x}_R, R_D) > \lim_{y^+ \to x+L_A-R_A} V_{M_R}(x, L_A, y, R_A). \tag{A-7}$$

By the first-order condition in Lemma A-1,

$$R_D + \hat{x}_R = L_A + x + \frac{(1 - F(\frac{\hat{x}_R+x+L_A-R_D}{2}))}{\frac{1}{2}f(\frac{\hat{x}_R+x+L_A-R_D}{2})}. \tag{A-8}$$

The left-hand side (LHS) of (A-7) is

$$\begin{aligned}&(1-F(\frac{\hat{x}_R+x+L_A-R_D}{2}))(R_D+\hat{x}_R) + F(\frac{\hat{x}_R+x+L_A-R_D}{2})(L_A+x) - M_R\\ =&L_A + x + \frac{\left(1-F(\frac{\hat{x}_R+x+L_A-R_D}{2})\right)^2}{\frac{1}{2}f(\frac{\hat{x}_R+x+L_A-R_D}{2})} - M_R.\end{aligned}$$

The right-hand side (RHS) of (A-7) is

$$\begin{aligned}&(1-F(x+L_A-R_A))(L_A-M_R+x+L_A-R_A) + F(x+L_A-R_A)(L_A-M_R+x)\\ &= L_A - M_R + x + (1-F(x+L_A-R_A))(L_A-R_A).\end{aligned}$$

Thus, (A-7) holds when

$$\frac{\frac{1}{2}f(\frac{\hat{x}_R+x+L_A-R_D}{2})(1-F(x+L_A-R_A))}{\big(1-F(\frac{\hat{x}_R+x+L_A-R_D}{2})\big)^2} \leq \frac{1}{(L_A-R_A)}. \tag{A-9}$$

Notice that when $L_A = R_A$, (A-7) holds. Thus, suppose $L_A > R_A$. Then, if

$$\max \frac{\frac{1}{2}f(\frac{\hat{x}_R+x+L_A-R_D}{2})(1-F(x+L_A-R_A))}{\big(1-F(\frac{\hat{x}_R+x+L_A-R_D}{2})\big)^2} \leq \frac{1}{(L_A-R_A)}, \tag{A-10}$$

then $x+L_A-R_A$ does not belong to $\beta_R(x, L_A)$. Then if there is $\bar{x}$ such that

$$\frac{\frac{1}{2}f(\frac{M_R+\bar{x}+L_A-R_D}{2})}{\big(1-F(\frac{M_R+\bar{x}+L_A-R_D}{2})\big)^2} = \frac{1}{(L_A-R_A)}, \tag{A-11}$$

then for every $x \leq \bar{x}$, $x+L_A-R_A$ does not belong to $\beta_R(x, L_A)$. $\square$

Lemma A-6. *Suppose* $\frac{M_R-M_L}{2} > L_A - R_A$*. If* $L_A = R_A$*, or* $L_A > R_A$ *and there is a* $\bar{y}$ *to satisfy*

$$\frac{\frac{1}{2}f(\frac{\bar{y}+M_L+L_D-R_A}{2})}{\big(F(\frac{\bar{y}+M_L+L_D-R_A}{2})\big)^2} = \frac{1}{2(L_A-R_A)},$$

then for every $y \geq \bar{y}$*,* $y-L_A+R_A$ *does not belong to* $\beta_L(y, R_A)$*.*

Proof of Lemma A-6. Fix $r_A = y$. Let $\hat{y}_L \in \hat{l}_D(y)$. Let $p = 0$ if $L_A > R_A$, and $p = F(y)$ if $L_A = R_A$. Then $y - L_A + R_A$ does not belong to $\beta_L(y, R_A)$ if and only if

$$V_{M_L}(\hat{y}_L, L_D, y, R_A) > \lim_{x^- \to y-L_A+R_A} V_{M_L}(x, R_A, y, L_A), \tag{A-12}$$

which implies

$$\begin{aligned}&(1-F(\tfrac{y+\hat{y}_L+L_D-R_A}{2}))(R_A-y)+F(\tfrac{y+\hat{y}_L+L_D-R_A}{2})(L_D-\hat{y}_L)+M_L\\&>(1-p)(L_A-y+L_A-R_A+M_L)+p(L_A-y+M_L).\end{aligned} \tag{A-13}$$

By Lemma A-1, the LHS of (A-13) is

$$R_A - y + \frac{\big(F(\frac{y+\hat{y}_L+L_D-R_A}{2})\big)^2}{\frac{1}{2}f(\frac{y+\hat{y}_L+L_D-R_A}{2})} + M_L.$$

Thus, (A-13) is rewritten as

$$R_A - y + \frac{\big(F(\frac{y+\hat{y}_L+L_D-R_A}{2})\big)^2}{\frac{1}{2}f(\frac{y+\hat{y}_L+L_D-R_A}{2})} + M_L > L_A - y + M_L + (1-p)(L_A-R_A). \tag{A-14}$$

If $L_A = R_A$, (A-14) holds. Thus, suppose that $L_A > R_A$, and then (A-14) holds if and only if

$$\frac{\frac{1}{2}f(\frac{y+\hat{y}_L+L_D-R_A}{2})}{\big(F(\frac{y+\hat{y}_L+L_D-R_A}{2})\big)^2} < \frac{1}{2(L_A-R_A)}. \tag{A-15}$$

Thus, if

$$\max \frac{\frac{1}{2} f\left(\frac{y+\hat{y}_L+L_D-R_A}{2}\right)}{\left(F\left(\frac{y+\hat{y}_L+L_D-R_A}{2}\right)\right)^2} < \frac{1}{2(L_A - R_A)}, \tag{A-16}$$

then $y + L_A - R_A$ does not belong to $\beta_L(y, R_A)$. Thus, if there is a $\bar{y}$ to satisfy

$$\frac{\frac{1}{2} f\left(\frac{\bar{y}+M_L+L_D-R_A}{2}\right)}{\left(F\left(\frac{\bar{y}+M_L+L_D-R_A}{2}\right)\right)^2} = \frac{1}{2(L_A - R_A)}, \tag{A-17}$$

then for every $y \geq \bar{y}$, $y + L_A - R_A$ does not belong to $\beta_L(y, R_A)$. □

Lemma A-7. *Suppose that the conditions of Theorem 4 hold. Then there is a fixed point* $(l_A^*, l_D^*, r_A^*, r_D^*)$ *such that* $l_A^* = \hat{l}_A(r_A^*), l_D^* \in \hat{l}_D(r_A^*), r_A^* = \hat{r}_A(l_A^*), r_D^* \in \hat{r}_D(l_A^*)$.

Proof of Lemma A-7. Notice that $\{\hat{l}_A(r_A) : r_A \in [\bar{y}, M_R]\} = \{\max \beta_L(r_A, R_A) : r_A \in [\bar{y}, M_R]\}$ and $\{\hat{r}_A(l_A) : l_A \in [M_L, \bar{x}]\} = \{\min \beta_R(l_A, L_A) : l_A \in [M_L, \bar{x}]\}$.

When $L_A = R_A$, by assumption, there is $\bar{x}$ and $\bar{y}$ such that $\{\max \beta_L(r_A, R_A) : r_A \in [\bar{y}, M_R]\} \cap \{\min \beta_R(l_A, L_A) : l_A \in [M_L, \bar{x}]\} = \varnothing$ hold. We show that this also holds when $L_A > R_A$. Suppose $L_A > R_A$. Let $r_A \geq \bar{y}$ and $l_A = \max \beta_L(r_A, R_A)$. Then $l_A < l_A + L_A - R_A \notin \beta_R(l_A, L_A)$ by Lemma A-5. Similarly, let $l_A \leq \bar{x}$ and $r_A = \min \beta_R(l_A, L_A)$. Then $r_A > r_A - L_A + R_A \notin \beta_L(r_A, R_A)$ by Lemma A-6. Thus, we conclude that $\beta_R(l_A, L_A) \cap \beta_L(r_A, R_A) = \varnothing$ when $r_A \geq \bar{y}$ and $l_A \leq \bar{x}$. Therefore, $\{\max \beta_L(r_A, R_A) : r_A \in [\bar{y}, M_R]\} \cap \{\min \beta_R(l_A, L_A) : l_A \in [M_L, \bar{x}]\} = \varnothing$ holds when $L_A \geq R_A$.

Now, $[M_L, \bar{x}]^2 \times [\bar{y}, M_R]^2$ is a compact and convex subset of a metric space. By the maximum theorem, $BR_D^L(r_A) = \hat{l}_D(r_A)$ and $BR_D^R(l_A) = \hat{r}_D(l_A)$ are nonempty, upper semicontinuous, and compact-valued. Notice that for every $r_A \in [\bar{y}, M_R]$ and $l_A \in [M_L, \bar{x}]$, there are r_D and l_D that satisfy $\pi_R(r_D, R_D, l_A, L_A) > 0$ and $\pi_R(r_A, R_A, l_D, L_D) < 1$. Then Lemma A-1 is applicable, and the maximizers of the expected payoffs uniquely exist for every $r_A \in [\bar{y}, M_R]$ and $l_A \in [M_L, \bar{x}]$. Therefore, $BR_D^L(r_A) = \hat{l}_D(r_A)$ and $BR_D^R(l_A) = \hat{r}_D(l_A)$ are convex-valued for every $r_A \in [\bar{y}, M_R]$ and $l_A \in [M_L, \bar{x}]$.

Further, because $\beta_R(l_A, L_A) \cap \beta_L(r_A, R_A) = \varnothing$ when $r_A \geq \bar{y}$ and $l_A \leq \bar{x}$, for every $r \in \beta_R(l_A, L_A)$, $\pi_R(r, R_A, l_A, L_A) > 0$ and for every $l \in \beta_L(r_A, R_A)$, $\pi_R(r_A, R_A, l, L_A) < 1$. By Lemma A-4, $\beta_R^\circ(l_A, L_A)$ and $\beta_L^\circ(r_A, R_A)$ are nonempty. Then also applying Lemma A-1, $\beta_R(l_A, L_A)$ and $\beta_L(r_A, R_A)$ are nonempty and compact intervals. Therefore, $BR_A^L(r_D)$ and $BR_A^R(l_D)$ are nonempty values for every $r_D \geq \bar{y}$ and $l_D \leq \bar{x}$, and indeed singletons. Because the expected payoffs are continuous with respect to each candidate's policy promise, $BR_A^L(r_D)$ and $BR_A^R(l_D)$ are also continuous and convex because they are singletons.

Thus, by Kakutani's fixed point theorem, there is a fixed point $(l_A^*, l_D^*, r_A^*, r_D^*)$ in $[M_L, \bar{x}]^2 \times [\bar{y}, M_R]^2$ such that

$$BR(l_A^*, l_D^*, r_A^*, r_D^*) = (BR_A^L(l_D^*), BR_D^L(r_A^*), BR_A^R(r_D^*), BR_D^R(l_A^*)) = (l_A^*, l_D^*, r_A^*, r_D^*).$$

□

Proof of Proposition 1. Suppose $L_A - R_A < M_R - M_L$ and suppose $r_A^* > M_L$. We will show that there is always an l_A such that

$$\mathrm{E}V_{M_L}(l_A, L_A, r_A^*, r_D^*) > \mathrm{E}V_{M_L}(l_D^*, L_D, r_A^*, r_D^*) \tag{A-18}$$
$$1 - \pi_R(r_A^*, R_A, l_A, L_A) > 0. \tag{A-19}$$

Then by the definition of equilibrium, we obtain $w_A^{*L} = 1$.

Because $r_A^* > M_L$ and $L_A - R_A < M_R - M_L$, there is always an $l_A \in [M_L, r_A^*]$ such that $1 - \pi_R(r_A^*, R_A, l_A, L_A) > 0$. Notice that when $1 - \pi_R(r_A^*, R_A, l_A, L_A) > 0$, we have $1 - \pi_R(r_D^*, R_D, l_A, L_A) > 0$ (for any r_D^*) and

$$\mathrm{E}V_{M_L}(l_A, L_A, r_A^*, r_D^*) > \mathrm{E}V_{M_L}(l_A, L_D, r_A^*, r_D^*),$$

because

$$V_{M_L}(l_A, L_D, r_A^*, R_A) > V_{M_L}(l_A, L_D, r_A^*, R_A);$$
$$V_{M_L}(l_A, L_D, r_D^*, R_D) > V_{M_L}(l_A, L_D, r_D^*, R_D).$$

Then we have $\max_{l_A} \mathrm{E}V_{M_L}(l_A, L_A, r_A^*, r_D^*) > \max_{l_D} \mathrm{E}V_{M_L}(l_D, L_D, r_A^*, r_D^*)$. By Lemma A-3, the maximizer of $\mathrm{E}V_{M_L}(l_A, L_A, r_A^*, r_D^*)$, denoted by l_A^{**}, belongs to the interval $[M_L, r_A^*]$, and because $r_A^* > M_L$ and $L_A - R_A < M_R - M_L$, $1 - \pi_R(r_A^*, R_A, l_A^{**}, L_A) > 0$. Therefore, l_A^{**} always satisfies (A-18) and (A-19). This completes the proof. $\square$

Proof of Proposition 2. Suppose $L_A - R_A < M_R - M_L$. Suppose that candidate L_A wins party L's primary, and $w_A^{*L} = 1$. Suppose that $\beta_R^\circ(l_A^*, L_A)$ is nonempty. Now, on the contrary, suppose $w_A^{*R} \neq 1$. Then by the definition of equilibrium,

$$V_{M_R}(l_A^*, L_A, r_A^*, R_A) \leq V_{M_R}(l_A^*, L_A, r_D^*, R_D). \tag{A-20}$$

Case 1. (A-20) *holds with strict inequality.* Then, $w_D^{*R} = 1$ and $w_A^{*R} = 0$ by the definition of equilibrium. Then immediately we obtain a contradiction, and because $\beta_R^\circ(l_A^*, L_A)$ is nonempty, there is an $r_A \in \beta_R^\circ(l_A^*, L_A)$ such that

$$V_{M_R}(l_A^*, L_A, r_A, R_A) > V_{M_R}(l_A^*, L_A, r_D^*, R_D);$$
$$\pi_R(r_A, R_A, l_A^*, L_A) > w_A^{*R} \cdot \pi_R(r_A^*, R_A, l_A^*, L_A) = 0.$$

Case 2. (A-20) *holds with equality.* Then because $\beta_R^\circ(l_A^*, L_A)$ is nonempty and $w_A^{*R} < 1$, by Lemma A-1, we can take $r_A = r_A^* + \epsilon$ for ϵ sufficiently small such that

$$V_{M_R}(l_A^*, L_A, r_A, R_A) > V_{M_R}(l_A^*, L_A, r_D^*, R_D);$$
$$\pi_R(r_A, R_A, l_A^*, L_A) > w_A^{*R} \cdot \pi_R(r_A^*, R_A, l_A^*, L_A).$$

By the definition of equilibrium, this contradicts the assumption that w_A^{*R} and r_A^* constitute an equilibrium. $\square$

Proof of Theorem 4. By Lemma A-7, we know that there is a fixed point $(l_A^*, l_D^*, r_A^*, r_D^*)$ such that $l_A^* = \hat{l}_A(r_A^*), l_D^* \in \hat{l}_D(r_A^*), r_A^* = \hat{r}_A(l_A^*), r_D^* \in \hat{r}_D(l_A^*))$. We show that this fixed point is an equilibrium. Notice that when $r_A^* = \hat{r}_A(l_A^*)$, $r_A^* > M_L$. Thus by Proposition 1, $w_A^{*L} = 1$. Further, when $l_A^* + L_A - R_A \notin \beta_R(l_A^*, L_A)$ by Lemma A-5, $\pi_R(r_A, R_A, l_A^*, L_A) > 0$ for every $r_A \in \beta_R(l_A^*, L_A)$ and thus $\beta_R^\circ(l_A^*, L_A)$ is nonempty by the first statement of Lemma A-4. Thus by Proposition 2, $w_A^{*R} = 1$.
Party R. First, we will show that there is no incentive for candidate R_A to deviate from $\hat{r}_A(l_A^*)$. Notice that $\hat{r}_A(l_A^*)$ maximizes the probability of winning the general election among policy promises in $\beta_R(l_A^*, L_A)$. Because candidate R_A's probability of winning the general election is a strictly decreasing continuous function of r_A, and otherwise candidate R_A could improve her winning chances by moving further to the right while still winning the primary election, that is, there is an r_A such that

$$\begin{aligned} V_{M_R}(l_A^*, L_A, r_A, R_A) &> V_{M_R}(l_A^*, L_A, r_A^*, R_A); \text{ and} \quad &\text{(A-21)} \\ \pi_R(r_A, R_A, l_A^*, L_A) &> \pi_R(r_A^*, R_A, l_A^*, L_A). &\text{(A-22)} \end{aligned}$$

This contradicts an equilibrium condition. Thus, candidate R_A does not have other strategies to deviate from $\hat{r}_A(l_A^*)$. Meanwhile, if $r_D^* \notin \hat{r}_D(l_A^*)$, then there is an r_A for which (A-21) and (A-22) simultaneously hold. This is a contradiction with the definition of equilibrium.
Party L. The proof for party L's candidates can be done in the same fashion as the first part. Intuitively, this is because candidate L_A's probability of winning the general election is a strictly decreasing continuous function of l_A, and otherwise candidate L_A could improve her winning chances by moving further to the left while still winning the primary election, which contradicts an equilibrium condition. Then candidate L_A can win the primary with probability one by adopting a policy promise in the interior of $\beta_L(r_A^*, R_A)$, and if she adopts a position outside this interval, candidate L_D could defeat her by adopting a policy promise in $\hat{l}_D(r_A^*)$. In addition, her probability of winning the general election is a strictly decreasing continuous function of l_A, so in any equilibrium we must have $l_A^* = \hat{l}_A(r_A^*)$. □

Proof of Proposition 3. By the first-order condition in Lemma A-1, we obtain (16) and (18). Similarly, by Lemma A-1, we obtain the results. □

Lemma A-8. *In a symmetric game,*

- *when regime 2 arises, $l_A^* = -r_A^*$, and $l_D^* = -r_D^*$ where r_A^* and r_D^* respectively solve* (13) *and* (14).
- $\bar{m} \notin \beta_R(\bar{m}, L_A) \cap \beta_L(\bar{m}, R_A)$ *if and only if* $a > L_A - L_D$.

Proof of Lemma A-8. The first item can be shown by the symmetry of the equilibrium and the four equations in Theorem 4. Now, we illustrate the second item. Let $r*_A = 0(= \bar{m})$.

Then by (8) of Theorem 4, we obtain

$$r_D^* - (L_A - L_D) = \frac{1 - F(\frac{r_D^* + L_A - L_D}{2})}{\frac{1}{2} f(\frac{r_D^* + L_A - L_D}{2})}.$$

When $r_A^* = l_A^* = 0$, we obtain

$$V_{M_R}(l_A^*, L_A, r_A^*, R_A) = L_A - M_R$$

and by Proposition 3,

$$V_{M_R}(l_A^*, L_A, r_D^*, R_D) = L_A - M_R + (r_D^* - (L_A - L_D))(1 - F(\frac{r_D^* + L_A - L_D}{2})).$$

Then, $V_{M_R}(l_A^*, L_A, r_A^*, R_A) < V_{M_R}(l_A^*, L_A, r_D^*, R_D)$, which implies $\bar{m} \notin \beta_R(\bar{m}, L_A)$, if and only if $a > L_A - L_D$. $\square$

Proof of Theorem 5. First, as in case (I), when $(l_A^*, r_A^*) = (\bar{m}, \bar{m})$, because a tiebreaking rule $p(\bar{m})$ specifies $\pi_R(\bar{m}, R_A, \bar{m}, L_A) = \frac{1}{2}$, by Lemma A-1, $\beta_R^\circ(\bar{m}, L_A)$ and $\beta_L^\circ(\bar{m}, R_A)$ are nonempty. By Propositions 1 and 2, $(w_A^{*L}, w_D^{*L}) = (w_A^{*R}, w_D^{*R}) = (1, 0)$. Case (II) is a special case of Theorem 4. When the set of equilibrium policy promises is as in case (II), again by Lemma A-1 and Propositions 1 and 2, we obtain $(w_A^{*L}, w_D^{*L}) = (w_A^{*R}, w_D^{*R}) = (1, 0)$.

Now, we show that if $\bar{m} \in \beta_R(\bar{m}, L_A) \cap \beta_L(\bar{m}, R_A)$ and $p(\bar{m}) = \frac{1}{2}$, case (I) arises, and otherwise, case (II) arises. Then the rest follows by Lemma A-8. Case (II) can be proved in the same way with Theorem 4. Thus, we focus on the proof of case (I). First, we prove that $r_A^* = l_A^* = \bar{m}$ is an equilibrium. Suppose that $r_A^* = l_A^* = \bar{m}$ and then $F(\frac{r_A^* + l_A^*}{2}) = 1 - F(\frac{r_A^* + l_A^*}{2}) = p(\bar{m}) = \frac{1}{2}$. Given $l_A^* = \bar{m}$, for every $r < l_A^*$, candidate R_A's probability of winning the general election is increasing and for every $r > l_A^*$, it is decreasing. Symmetrically, given $r_A^* = \bar{m}$, this is true for candidate L_A. Thus, $r_A^* = l_A^* = \bar{m}$ is an equilibrium.

Second, consider the disadvantaged candidates. Because choosing $\bar{m}$ gives a primary win with certainty for both advantaged candidates, there is no strategy for each disadvantaged candidate to improve their chance of winning in the primary. Thus any strategy in X can be an equilibrium strategy for each disadvantaged candidate. $\square$

Proof of Proposition 4. The first case (a) is proved in a similar manner with Theorem 1. Now, we consider case (b). Because $M_R - \bar{l}_A > L_A - R_D$, for some r_D, $\pi_R(r_D, R_D, \bar{l}_A, L_A) > 0$. By Lemma A-1, $\hat{r}_D(\bar{l}_A)$ is singleton. Further, $\beta_R^\circ(\bar{l}_A, L_A)$ is nonempty. Thus, $(w_D^{*R}, w_A^{*R}) = (0, 1)$ by Proposition 2. The rest follows from the proof for party R in Theorem 4. $\square$

Proof of Proposition 5. The proof of case (a) is the same with case (a) in Proposition 4. First, $(w_D^{*L}, w_A^{*L}) = (0, 1)$ is obtained by Proposition 1 because $\bar{r}_A > M_L$. In case (b), $l_D^* \in \hat{l}_D(\bar{r}_A)$ is proved in the same way as for the proof for party L in Theorem 4. Thus it suffices to show $l_A^* \in \alpha(\bar{r}_A, R_A) \cap \beta_L(\bar{r}_A, R_A)$ in case (b-1), and otherwise, in case (b-2), $l_A^* = \hat{l}_A(\bar{r}_A)$. In case (b-1), candidate L_A can win the primary with probability one and still win the general election with a probability arbitrarily close to one by moving slightly to the right of the rightmost such policy, and in equilibrium she must take a position in $\alpha(\bar{r}_A) \cap \beta_L(\bar{r}_A, R_A)$. No matter what position candidate L_D adopts, neither candidate L_D nor candidate R_A have any way to improve. The proof of case (c)'s $l^*_A \in [\bar{r}_A - L_A + R_A, \bar{r}_A + L_A - R_A]$ is immediate because in this region, candidate L_A can secure a victory in the general election. Because $\bar{r}_A - M_L \leq L_A - R_D$, candidate L_D cannot win the general election against candidate R_A and thus the equilibrium policy promise is unrestricted for candidate L_D.

Finally, if $L_A - R_A \geq R_A - L_D$, then cases (a) and (c) are the only ones possible. In contrast, if $L_A - R_A < R_A - L_D$, then all of cases (a), (b), and (c) cover all those possible. As a result, in either situation, there is always an equilibrium. This completes the proof. □

Proof of Theorem 6. In a symmetric game, if $\bar{m} \notin \beta_R(\bar{m}, L_A) \cap \beta_L(\bar{m}, R_A)$, by Theorem 5, regime 2 arises. In the proof, the size of $\left|\frac{\mathrm{d}V_{M_L}(l^*_A, L_A, r^*_A, R_A)}{\mathrm{d}l_A}\right|$ is important. Thus, first we note the following. When the game is symmetric, by (13) and (14) of Theorem 5,

$$\begin{aligned} L_A - l^*_A - R_A + r^*_A &= 2r^*_A & \text{(A-23)} \\ L_A + l^*_A - R_A + r^*_A &= 0. & \text{(A-24)} \end{aligned}$$

Thus, because the expected payoff for candidate L_A is decreasing at l^*_A,

$$\begin{aligned} \frac{\mathrm{d}V_{M_L}(l^*_A, L_A, r^*_A, R_A)}{\mathrm{d}l_A} &= (1-\pi_R)'(L_A - l^*_A - R_A + r^*_A) - (1 - \pi_R(r^*_A, R_A, l^*_A, L_A)) \\ = -F(\frac{l^*_A + r^*_A + L_A - R_A}{2}) &+ f(\frac{l^*_A + r^*_A + L_A - R_A}{2})\frac{(L_A - l^*_A - R_A + r^*_A)}{2} < 0 \\ &= -\frac{1}{2} + f(0) r^*_A < 0. \qquad \text{(A-25)} \end{aligned}$$

Step I: Computing the formula for $\frac{\mathrm{d}r^*_A}{\mathrm{d}L_A}$, and $\frac{\mathrm{d}l^*_A}{\mathrm{d}L_A}$.

By taking the derivative of (17), we have

$$\begin{aligned} &\frac{\frac{\mathrm{d}l^*_A}{\mathrm{d}L_A} + \frac{\mathrm{d}r^*_A}{\mathrm{d}L_A} + 1}{2} f(\frac{l^*_A + r^*_A}{2})(r^*_A - l^*_A) + F(\frac{l^*_A + r^*_A}{2})(1 - \frac{\mathrm{d}l^*_A}{\mathrm{d}L_A} + \frac{\mathrm{d}r^*_A}{\mathrm{d}L_A}) \\ &= \left(\frac{\mathrm{d}r^*_A}{\mathrm{d}L_A} - \frac{\mathrm{d}l^*_D}{\mathrm{d}L_A}\right) F(\frac{l^*_D + r^*_A + L_D - R_A}{2}) \\ &\quad + (L_D + r^*_A - R_A - l^*_D) f(\frac{l^*_D + r^*_A + L_D - R_A}{2}) \frac{(\frac{\mathrm{d}r^*_A}{\mathrm{d}L_A} + \frac{\mathrm{d}l^*_D}{\mathrm{d}L_A})}{2}. \end{aligned}$$

Then by (16),

$$\left(\frac{\mathrm{d}l^*_A}{\mathrm{d}L_A} + \frac{\mathrm{d}r^*_A}{\mathrm{d}L_A} + 1\right) f(0) \cdot r^*_A + \frac{1}{2}(1 - \frac{\mathrm{d}l^*_A}{\mathrm{d}L_A} + \frac{\mathrm{d}r^*_A}{\mathrm{d}L_A}) = 2\frac{\mathrm{d}r^*_A}{\mathrm{d}L_A} F(\frac{l^*_D + r^*_A + L_D - R_A}{2}). \quad \text{(A-26)}$$

Conversely, by taking the derivative of (19),

$$\begin{aligned}
&-\frac{\frac{\mathrm{d}r_A^*}{\mathrm{d}L_A}+\frac{\mathrm{d}l_A^*}{\mathrm{d}L_A}+1}{2} f(\frac{l_A^*+r_A^*+L_A-R_A}{2})(R_A+r_A^*-L_A-l_A^*) \\
&\quad+(1-F(\frac{l_A^*+r_A^*+L_A-R_A}{2}))(\frac{\mathrm{d}r_A^*}{\mathrm{d}L_A}-\frac{\mathrm{d}l_A^*}{\mathrm{d}L_A}-1) \\
&\qquad=(1-F(\frac{l_A^*+r_D^*+L_A-R_D}{2}))(\frac{\mathrm{d}r_D^*}{\mathrm{d}L_A}-\frac{\mathrm{d}l_A^*}{\mathrm{d}L_A}-1) \\
&\qquad\quad-(R_D+r_D^*-L_A-l_A^*) f(\frac{l_A^*+r_D^*+L_A-R_D}{2}) \frac{\frac{\mathrm{d}r_D^*}{\mathrm{d}L_A}+\frac{\mathrm{d}l_A^*}{\mathrm{d}L_A}+1}{2}.
\end{aligned}$$

By applying (17) and (A-24),

$$\begin{aligned}
&-\left(\frac{\mathrm{d}r_A^*}{\mathrm{d}L_A}+\frac{\mathrm{d}l_A^*}{\mathrm{d}L_A}+1\right) f(0) r_A^*+\frac{1}{2}(\frac{\mathrm{d}r_A^*}{\mathrm{d}L_A}-\frac{\mathrm{d}l_A^*}{\mathrm{d}L_A}-1) \\
&\qquad=-2(1-F(\frac{l_A^*+r_D^*+L_A-R_D}{2}))\left(\frac{\mathrm{d}l_A^*}{\mathrm{d}L_A}+1\right). \qquad \text{(A-27)}
\end{aligned}$$

By adding (A-26) and (A-27), because $F_D \equiv F(\frac{l_D^*+r_A^*+L_D-R_A}{2})=1-F(\frac{l_A^*+r_D^*+L_A-R_D}{2})$, we obtain

$$\frac{\mathrm{d}r_A^*}{\mathrm{d}L_A}\left(\frac{1}{2}-2F_D\right)=\left(\frac{1}{2}+2F_D\right)\left(\frac{\mathrm{d}l_A^*}{\mathrm{d}L_A}+1\right)-\frac{1}{2}. \qquad \text{(A-28)}$$

Further, by (A-27),

$$(\frac{1}{2}-f(0) r_A^*) \frac{\mathrm{d}r_A^*}{\mathrm{d}L_A}=\left(\frac{\mathrm{d}l_A^*}{\mathrm{d}L_A}+1\right) \times a, \qquad \text{(A-29)}$$

where

$$a=(\frac{1}{2}+f(0) r_A^*)-2(1-F(\frac{l_A^*+r_D^*+L_A-R_D}{2}))=\frac{1}{2}+f(0) r_A^*-2F_D.$$

By (A-28) and (A-29), we obtain

$$\frac{\mathrm{d}r_A^*}{\mathrm{d}L_A}\left(\frac{1}{2}-2F_D-\frac{\frac{1}{2}-f(0) r_A^*}{a}\left(\frac{1}{2}+2F_D\right)\right)=-\frac{1}{2}. \qquad \text{(A-30)}$$

Thus,

$$\frac{\mathrm{d}r_A^*}{\mathrm{d}L_A}=\frac{\frac{1}{2}}{2F_D+\frac{\frac{1}{2}-f(0) r_A^*}{a}\left(\frac{1}{2}+2F_D\right)-\frac{1}{2}}; \qquad \text{(A-31)}$$

$$\frac{\mathrm{d}l_A^*}{\mathrm{d}L_A}=\frac{\frac{1}{2}-f(0) r_A^*}{a} \cdot \frac{\frac{1}{2}}{2F_D+\frac{\frac{1}{2}-f(0) r_A^*}{a}\left(\frac{1}{2}+2F_D\right)-\frac{1}{2}}-1; \qquad \text{(A-32)}$$

$$\frac{\mathrm{d}F(\frac{l_A^*+r_A^*+L_A-R_A}{2})}{\mathrm{d}L_A}=\frac{\frac{\mathrm{d}l_A^*}{\mathrm{d}L_A}+\frac{\mathrm{d}r_A^*}{\mathrm{d}L_A}+1}{2} \cdot f(\frac{l_D^*+r_A^*+L_D-R_A}{2}). \qquad \text{(A-33)}$$

Step II: Proving the signs of $\frac{\mathrm{d}r_A^*}{\mathrm{d}L_A}$ *and* $\frac{\mathrm{d}l_A^*}{\mathrm{d}L_A}$.

Step II-1: Showing $\frac{1}{2}-f(0)r_A^*>a>0$. Notice that by (17) and (A-24),

$$0<L_A-l_A^*-R_A+r_A^*=2r_A^*<L_D-l_D^*-R_A+r_A^*. \tag{A-34}$$

Further, by (A-23)

$$0=L_A+l_A^*-R_A+r_A^*>L_D+l_D^*-R_A+r_A^*. \tag{A-35}$$

Let $F_A\equiv F(\frac{l_A^*+r_A^*+L_A-R_A}{2})$, $f_D\equiv f(\frac{l_D^*+r_A^*+L_D-R_A}{2})$ and $f_A\equiv f(\frac{l_A^*+r_A^*+L_A-R_A}{2})$. Then by (A-35), $F_A>F_D$ and thus we can set $F_A=kF_D$ for some $k>1$. Also, given the game is symmetric with respect to 0, by (A-35), we can set $f_A=mf_D$ for some $m>1$. Further by Assumption 1 and (A-35), $\frac{F_A}{f_A}>\frac{F_D}{f_D}$ implies $k>m>1$. By (17),

$$L_A-l_A^*-R_A+r_A^*:L_D-l_D^*-R_A+r_A^*=F_D:F_A=1:k. \tag{A-36}$$

Then, by applying (A-36) and (17),

$$\begin{aligned}
a&=(F_A-F_D)+\left(f_Ar_A^*-f_D\frac{L_D-l_D^*-R_A+r_A^*}{2}\right)\\
&=(F_A-F_D)+\left(mr_A^*-\frac{L_D-l_D^*-R_A+r_A^*}{2}\right)f_D\\
&=(k-1)\,F_D+(m-k)\,r_A^*f_D=\left((k-1)\frac{L_D-l_D^*-R_A+r_A^*}{2}+(m-k)r_A^*\right)f_D\\
&=((k-1)k+(m-k))\,f_Dr_A^*=(k^2-2k+m)f_Dr_A^*>(k-1)^2f_Dr_A^*>0.
\end{aligned}$$

The first term in a is strictly positive by (A-34), and the second term is strictly negative. Thus,

$$\begin{aligned}
a&<F(\frac{l_A^*+r_A^*+L_A-R_A}{2})-F(\frac{l_D^*+r_A^*+L_D-R_A}{2})\\
&<F(\frac{l_A^*+r_A^*+L_A-R_A}{2})-f(\frac{l_A^*+r_A^*+L_A-R_A}{2})\frac{(L_A-l_A^*-R_A+r_A^*)}{2}\\
&=\frac{1}{2}-f(0)r_A^*.
\end{aligned}$$

Notice that

$$2F_D+\frac{\frac{1}{2}-f(0)r_A^*}{a}\left(\frac{1}{2}+2F_D\right)-\frac{1}{2}>2F_D+\frac{1}{2}+2F_D-\frac{1}{2}=4F_D>0. \tag{A-37}$$

Thus,

$$\frac{\mathrm{d}r_A^*}{\mathrm{d}L_A}=\frac{\frac{1}{2}}{2F_D+\frac{\frac{1}{2}-f(0)r_A^*}{a}\left(\frac{1}{2}+2F_D\right)-\frac{1}{2}}>0.$$

Also, because

$$\frac{\mathrm{d}l_A^*}{\mathrm{d}L_A}+1=\frac{\frac{1}{2}-f(0)r_A^*}{a}\cdot\frac{\frac{1}{2}}{2F_D+\frac{\frac{1}{2}-f(0)r_A^*}{a}\left(\frac{1}{2}+2F_D\right)-\frac{1}{2}}>0,$$

we obtain

$$\frac{\mathrm{d}F(\frac{l_A^*+r_A^*+L_A-R_A}{2})}{\mathrm{d}L_A}=\frac{\frac{\mathrm{d}l_A^*}{\mathrm{d}L_A}+\frac{\mathrm{d}r_A^*}{\mathrm{d}L_A}+1}{2}\cdot f(\frac{l_D^*+r_A^*+L_D-R_A}{2})>0.$$

Step III: Proving the signs of $\frac{\mathrm{d}r_D^*}{\mathrm{d}L_A}$ and $\frac{\mathrm{d}l_D^*}{\mathrm{d}L_A}$ and Concluding.

Step III-1. Showing $\frac{\mathrm{d}l_D^*}{\mathrm{d}L_A}<0$.

By taking the derivative of (16),

$$\begin{aligned}&\left(\frac{\mathrm{d}r_A^*}{\mathrm{d}L_A}-\frac{\mathrm{d}l_D^*}{\mathrm{d}L_A}\right)\frac{f(\frac{r_A^*+l_D^*+L_D-R_A}{2})}{2}\\ &=-\left((L_D+r_A^*-R_A-l_D^*)\frac{f'(\frac{r_A^*+l_D^*+L_D-R_A}{2})}{2}+f(\frac{r_A^*+l_D^*+L_D-R_A}{2})\right)\frac{\left(\frac{\mathrm{d}r_A^*}{\mathrm{d}L_A}+\frac{\mathrm{d}l_D^*}{\mathrm{d}L_A}\right)}{2}.\end{aligned}$$

Then

$$\begin{aligned}&\frac{\mathrm{d}r_A^*}{\mathrm{d}L_A}\left(f(\frac{r_A^*+l_D^*+L_D-R_A}{2})-\frac{(L_D+r_A^*-R_A-l_D^*)}{2}f'(\frac{r_A^*+l_D^*+L_D-R_A}{2})\right)\\ &=-\frac{(L_D+r_A^*-R_A-l_D^*)}{2}f'(\frac{r_A^*+l_D^*+L_D-R_A}{2})\frac{\mathrm{d}l_D^*}{\mathrm{d}L_A}.\end{aligned}$$

Because $\frac{\mathrm{d}r_A^*}{\mathrm{d}L_A}>0$ and $f'(\frac{r_A^*+l_D^*+L_D-R_A}{2})>0$, $\frac{\mathrm{d}l_D^*}{\mathrm{d}L_A}<0$.

Step III-2. Showing that there is a desired K and completing the proof.

By taking the derivative of (18),

$$\begin{aligned}&\left(\frac{\mathrm{d}r_D^*}{\mathrm{d}L_A}-\frac{\mathrm{d}l_A^*}{\mathrm{d}L_A}-1\right)\frac{f(\frac{r_D^*+l_A^*+L_A-R_D}{2})}{2}\\ &=-\left((R_D+r_D^*-L_A-l_A^*)\frac{f'(\frac{r_D^*+l_A^*+L_A-R_D}{2})}{2}+f(\frac{r_D^*+l_A^*+L_A-R_D}{2})\right)\frac{\left(\frac{\mathrm{d}r_D^*}{\mathrm{d}L_A}+\frac{\mathrm{d}l_A^*}{\mathrm{d}L_A}+1\right)}{2}.\end{aligned}$$

Let $\Gamma=f(\frac{r_D^*+l_A^*+L_A-R_D}{2})+\frac{\left(R_D+r_D^*-L_A-l_A^*\right)}{2}f'(\frac{r_D^*+l_A^*+L_A-R_D}{2})$. Thus, we obtain

$$\frac{\mathrm{d}r_D^*}{\mathrm{d}L_A}=\frac{f(\frac{r_D^*+l_A^*+L_A-R_D}{2})-\Gamma}{f(\frac{r_D^*+l_A^*+L_A-R_D}{2})+\Gamma}\left(\frac{\mathrm{d}l_A^*}{\mathrm{d}L_A}\right).\tag{A-38}$$

Notice that by Assumption 1, $\frac{\mathrm{d}\left(\frac{f(x)}{1-F(x)}\right)}{\mathrm{d}x}\geq 0$. Thus $\frac{f(x)}{1-F(x)}+\frac{f'(x)}{f(x)}>0$. Applying this, because

$$\Gamma = f(\frac{r_D^* + l_A^* + L_A - R_D}{2}) + \frac{(R_D + r_D^* - L_A - l_A^*)}{2} f'(\frac{r_D^* + l_A^* + L_A - R_D}{2})$$
$$= \left(1 - F(\frac{r_D^* + l_A^* + L_A - R_D}{2})\right) \times$$
$$\left(\frac{f(\frac{r_D^*+l_A^*+L_A-R_D}{2})}{1 - F(\frac{r_D^*+l_A^*+L_A-R_D}{2})} + \frac{\frac{(R_D+r_D^*-L_A-l_A^*)}{2}}{1 - F(\frac{r_D^*+l_A^*+L_A-R_D}{2})} f'(\frac{r_D^* + l_A^* + L_A - R_D}{2}) \right)$$
$$= \left(1 - F(\frac{r_D^* + l_A^* + L_A - R_D}{2})\right) \times \left(\frac{f(\frac{r_D^*+l_A^*+L_A-R_D}{2})}{1 - F(\frac{r_D^*+l_A^*+L_A-R_D}{2})} + \frac{f'(\frac{r_D^*+l_A^*+L_A-R_D}{2})}{f(\frac{r_D^*+l_A^*+L_A-R_D}{2})} \right)$$
$$> 0, \tag{A-39}$$

together with the fact that $\frac{r_D^*+l_A^*+L_A-R_D}{2} > 0$ implies $f'(\frac{r_D^*+l_A^*+L_A-R_D}{2}) < 0$, it is held that

$$1 > \frac{f(\frac{r_D^*+l_A^*+L_A-R_D}{2}) - \Gamma}{f(\frac{r_D^*+l_A^*+L_A-R_D}{2}) + \Gamma} > 0. \tag{A-40}$$

Therefore, by (A-38), r_D^* moves in the same direction as l_A^*.

Remember that

$$\frac{\mathrm{d}l_A^*}{\mathrm{d}L_A} = \frac{\frac{\frac{1}{2}-f(0)r_A^*}{a} \cdot \frac{1}{2}}{2F_D + \frac{\frac{1}{2}-f(0)r_A^*}{a}\left(\frac{1}{2} + 2F_D\right) - \frac{1}{2}} - 1,$$

which implies that $\frac{\mathrm{d}l_A^*}{\mathrm{d}L_A} > 0$ if and only if

$$2F_D\left(1 + \frac{\frac{1}{2} - f(0)r_A^*}{a}\right) = 2F_D\left(1 + \frac{\frac{1}{2} - f(0)r_A^*}{\frac{1}{2} + f(0)r_A^* - 2F_D}\right) < \frac{1}{2}.$$

Define K to satisfy $2K\left(1 + \frac{\frac{1}{2}-f(0)r_A^*}{\frac{1}{2}+f(0)r_A^*-2K}\right) = \frac{1}{2}$. Then, $8K^2 - 6K + (\frac{1}{2} + f(0)r_A^*) = 0$. Notice that $8(0) - 6(0) + (\frac{1}{2} + f(0)r_A^*) > 0$, and $8(\frac{1}{2})^2 - 6(\frac{1}{2}) + (\frac{1}{2} + f(0)r_A^*) = f(0)r_A^* - \frac{1}{2} < 0$. Thus, K exists in $(0, \frac{1}{2})$. Further, because the quadratic equation has two solutions and $8(1)^2 - 6(1) + (\frac{1}{2} + f(0)r_A^*) > 0$, another solution exists in $(\frac{1}{2}, 1)$. Thus, a solution that is between 0 and $\frac{1}{2}$ is unique. Now, it is clear that $2F_D\left(1 + \frac{\frac{1}{2}-f(0)r_A^*}{a}\right) < \frac{1}{2}$ if and only if $F_D > K$. □

Proof of Theorem 7. Let $R_A - R_D = \delta$. In what follows, $(g(x))'$ denotes the first derivative of $g(x)$. We denote $r_A^S + l_D^S + L_D - R_A = x^S$ and $r_A^N + l_D^N + L_D - R_A = x^N$. We denote $r_D^S + l_A^S + L_A - R_D = y^S$ and $r_D^N + l_A^N + L_A - R_D = y^N$.

Suppose that $|f_N(x) - f_S(x)| < \bar{\epsilon}$ for some arbitrarily small $\bar{\epsilon} > 0$. We will show $r_A^S < r_A^N$ and $l_A^S > l_A^N$. Then, because we can recursively repeat the same argument, we will obtain the desired result. By the equilibrium condition for regime 2, we must have

$$V_{M_R}^S(l_A^S, L_A, r_A^S, R_A) + M_R = V_{M_R}^S(l_A^S, L_A, r_D^S, R_D) + M_R. \tag{A-41}$$

Also, by symmetry,

$$V_{M_R}^S(l_A^S, L_A, r_A^S, R_A) = V_{M_R}^N(l_A^N, L_A, r_A^N, R_A) = L_A - M_R. \tag{A-42}$$

By Lemma A-1, in regime 2, for each society $s = N, S$, we must have

$$V_{M_R}^s(l_A^s, L_A, r_D^s, R_D) + M_R = l_A^s + L_A + (1 - F_s(\frac{y^s}{2})) \cdot (r_D^s - l_A^s - \delta). \tag{A-43}$$

By comparing (A-42) and (A-43) for each society $s = N, S$, we must have

$$l_A^s + \frac{(1 - F_s(\frac{y^s}{2}))^2}{\frac{1}{2} f_s(\frac{y^s}{2})} = l_A^s + (1 - F_s(\frac{y^s}{2})) \cdot (r_D^s - l_A^s - \delta) = 0. \tag{A-44}$$

By comparing (A-44) for societies S and N, we obtain

$$(F_S(\frac{y^S}{2}) - F_N(\frac{y^N}{2}))\delta + F_S(\frac{y^S}{2})l_A^S - F_N(\frac{y^N}{2})l_A^N = (1 - F_N(\frac{y^N}{2}))r_D^N - (1 - F_S(\frac{y^S}{2}))r_D^S. \tag{A-45}$$

As $y^S > 0$, property 2 implies $F_S(\frac{y^S}{2}) > F_N(\frac{y^S}{2})$ and thus for some small $\epsilon_y > 0$, $F_S(\frac{y^S}{2}) - F_N(\frac{y^S}{2}) = \epsilon_y$ and we have

$$\begin{aligned} F_S(\frac{y^S}{2}) - F_N(\frac{y^N}{2}) &= F_S(\frac{y^S}{2}) - F_N(\frac{y^S}{2}) - F_N(\frac{y^N}{2}) + F_N(\frac{y^S}{2}) \\ &= \epsilon_y - f_N(\frac{y^N}{2}) \cdot \frac{(r_D^N - r_D^S) + (l_A^N - l_A^S)}{2}. \end{aligned} \tag{A-46}$$

Consider how $F_N(\frac{r+l+\delta}{2}) \cdot l$ changes when r and l change infinitesimally, denoted by $\mathrm{d}F_N(\frac{r+l+\delta}{2}) \cdot l$, which equals

$$\left(\frac{f_N(\frac{r+l+\delta}{2})l}{2}\right)\mathrm{d}r + \left(\frac{f_N(\frac{r+l+\delta}{2})}{2} + F_N(\frac{r+l+\delta}{2})\right)\mathrm{d}l.$$

Thus, $F_S(\frac{y^S}{2})l_A^S - F_N(\frac{y^N}{2})l_A^N$ is given by $F_S(\frac{y^S}{2})l_A^S - F_N(\frac{y^S}{2})l_A^S = \epsilon_y l_A^S$ plus $F_N(\frac{y^S}{2})l_A^S - F_N(\frac{y^N}{2})l_A^N$, which is

$$\epsilon_y l_A^S + \left(\frac{f_N(\frac{y^N}{2})l_A^N}{2}\right)(r_D^S - r_D^N) + \left(\frac{f_N(\frac{y^N}{2})l_A^N}{2} + F_N(\frac{y^N}{2})\right)(l_A^S - l_A^N). \tag{A-47}$$

Similarly, consider how $(1 - F_N(\frac{r+l+\delta}{2})) \cdot r$ changes when r and l change infinitesimally, denoted by $\mathrm{d}(1 - F_N(\frac{r+l+\delta}{2})) \cdot r$, which equals

$$\left(\frac{-f_N(\frac{r+l+\delta}{2})r}{2} + (1 - F_N(\frac{r+l+\delta}{2}))\right)\mathrm{d}r - \frac{f_N(\frac{r+l+\delta}{2})r}{2}\mathrm{d}l.$$

Thus, $(1-F_N(\frac{y^N}{2}))r_D^N-(1-F_S(\frac{y^S}{2}))r_D^S$ is given by $(1-F_N(\frac{y^S}{2}))r_D^S-(1-F_S(\frac{y^S}{2}))r_D^S=\epsilon_y r_D^S$ plus $(1-F_N(\frac{y^N}{2}))r_D^N-(1-F_N(\frac{y^S}{2}))r_D^S$, which is

$$\epsilon_y r_D^S+\left(\frac{-f_N(\frac{y^N}{2})r_D^N}{2}+(1-F_N(\frac{y^N}{2}))\right)(r_D^N-r_D^S)-\frac{f_N(\frac{y^N}{2})r_D^N}{2}(l_A^N-l_A^S). \quad \text{(A-48)}$$

Combining the three terms (A-46), (A-47), and (A-48) into (A-45), we obtain

$$\begin{aligned}&\left(\epsilon_y-f_N(\frac{y^N}{2})\cdot\frac{(r_D^N-r_D^S)+(l_A^N-l_A^S)}{2}\right)\delta\\&+\epsilon_y l_A^S+\left(\frac{f_N(\frac{y^N}{2})l_A^N}{2}\right)(r_D^S-r_D^N)+\left(\frac{f_N(\frac{y^N}{2})l_A^N}{2}+F_N(\frac{y^N}{2})\right)(l_A^S-l_A^N)\\&=\epsilon_y r_D^S+\left(\frac{-f_N(\frac{y^N}{2})r_D^N}{2}+(1-F_N(\frac{y^N}{2}))\right)(r_D^N-r_D^S)-\frac{f_N(\frac{y^N}{2})r_D^N}{2}(l_A^N-l_A^S).\end{aligned}$$

By organizing terms, we obtain

$$\begin{aligned}\epsilon_y(\delta+l_A^S-r_D^S)&=\left(\frac{f_N(\frac{y^N}{2})}{2}\delta+\frac{f_N(\frac{y^N}{2})l_A^N}{2}-\frac{f_N(\frac{y^N}{2})r_D^N}{2}+(1-F_N(\frac{y^N}{2}))\right)(r_D^N-r_D^S)\\&+\left(\frac{f_N(\frac{y^N}{2})}{2}(L_A-R_D)+\frac{f_N(\frac{y^N}{2})l_A^N}{2}+F_N(\frac{y^N}{2})-\frac{f_N(\frac{y^N}{2})r_D^N}{2}\right)(l_A^N-l_A^S). \quad \text{(A-49)}\end{aligned}$$

We can simplify the RHS because by (18), which implies $1-F_N(\frac{y^N}{2})=\frac{f_N(\frac{y^N}{2})}{2}(r_D^N-\delta-l_A^N)$, the first term becomes

$$\frac{f_N(\frac{y^N}{2})}{2}\delta+\frac{f_N(\frac{y^N}{2})l_A^N}{2}-\frac{f_N(\frac{y^N}{2})r_D^N}{2}+(1-F_N(\frac{y^N}{2}))=0,$$

and the second term becomes

$$\frac{f_N(\frac{y^N}{2})}{2}\delta+\frac{f_N(\frac{y^N}{2})l_A^N}{2}+F_N(\frac{y^N}{2})-\frac{f_N(\frac{y^N}{2})r_D^N}{2}=1-f_N(\frac{y^N}{2})(r_D^N-\delta-l_A^N).$$

Because $1-F_N(\frac{y^N}{2})=\frac{f_N(\frac{y^N}{2})}{2}(r_D^N-\delta-l_A^N)<\frac{1}{2}$, $f_N(\frac{y^N}{2})\cdot(r_D^N-\delta-l_A^N)<1$. As a result, by (A-49), we obtain

$$l_A^N-l_A^S=\frac{-\epsilon_y(r_D^S-l_A^S-\delta)}{1-f_N(\frac{y^N}{2})(r_D^N-\delta-l_A^N)}<0.$$

Thus, we obtain $r_A^S<r_A^N$. □

Lemma A-9. *In the symmetric game, an equilibrium policy for party R's candidate satisfies $r_A<r_D-(R_A-R_D)$.*

Proof of Lemma A-9. See Proposition 3 in Hummel (2013). □

Proof of Theorem 8. We continue to use the notation for x^S, x^N, y^S, y^N as defined in the proof of Theorem 7. Note that under both scenarios, $0 \leq F_N(x) - F_S(x) < \epsilon \cdot \bar{\epsilon}$ for every $x \in [-a, a]$ and further,

(A) Under the scenario of extremists' change, for every $x_0, x_1 \in [-a + \bar{\epsilon}, a]$, $F_N(x_0) - F_S(x_0) = F_N(x_1) - F_S(x_1)$;

(B) Under the scenario of centrists' change, for every $x \in [-a, a] \setminus [-\bar{\epsilon}, \bar{\epsilon}]$, $F_N(x) = F_S(x)$.

Because $V^S_{M_R}(l^S_A, L_A, r^S_A, R_A) = F_S(0)(L_A + l^S_A) + (1 - F_S(0))(R_A + r^S_A) - M_R$ and $l^S_A = -r^S_A$, we obtain

$$\begin{aligned} &V^N_{M_L}(l^N_A, L_A, r^N_A, R_A) - V^S_{M_L}(l^S_A, L_A, r^S_A, R_A) \\ &= V^S_{M_R}(l^S_A, L_A, r^S_A, R_A) - V^N_{M_R}(l^N_A, L_A, r^N_A, R_A) \\ &= F_N(\frac{r^N_A + l^N_A}{2})l^N_A + (1 - F_N(\frac{r^N_A + l^N_A}{2}))r^N_A. \end{aligned} \tag{A-50}$$

In regime 2 equilibria, the following holds:

$$\begin{aligned} &V^N_{M_L}(l^N_A, L_A, r^N_A, R_A) - V^S_{M_L}(l^S_A, L_A, r^S_A, R_A) \\ &= V^N_{M_L}(l^N_D, L_D, r^N_A, R_A) - V^S_{M_L}(l^S_D, L_D, r^S_A, R_A) \\ &= V^S_{M_R}(l^S_A, L_A, r^S_D, R_D) - V^N_{M_R}(l^N_A, L_A, r^N_D, R_D). \end{aligned} \tag{A-51}$$

<u>*Step I.*</u> *Calculating* (A-51).

Consider $\mathrm{d}(r - \delta - l)(1 - F(\frac{r+l+\delta}{2}))$ when r and l change. It is given by

$$\begin{aligned} &\left((1 - F(\frac{r + l + \delta}{2})) - \frac{r - \delta - l}{2} f(\frac{r + l + \delta}{2})\right) \mathrm{d}r \\ &- \left((1 - F(\frac{r + l + \delta}{2})) + \frac{r - \delta - l}{2} f(\frac{r + l + \delta}{2})\right) \mathrm{d}l. \end{aligned}$$

Similarly to (A-47) and (A-48) in the proof of Theorem 7, by applying the first-order condition (18), we can compute that $(r^S_D - \delta - l^S_A)(1 - F_N(\frac{y^S}{2})) - (r^N_D - \delta - l^N_A)(1 - F_N(\frac{y^N}{2}))$ equals

$$-\left((1 - F_N(\frac{y^N}{2})) + \frac{r^N_D - \delta - l^N_A}{2} f_N(\frac{y^N}{2})\right)(l^S_A - l^N_A).$$

Let $F_N(\frac{y^S}{2}) - F_S(\frac{y^S}{2}) = \epsilon_1$. By property 5, $\epsilon_1 > 0$. The RHS of (A-50) is given by

$$\begin{aligned} &V^S_{M_R}(l^S_A, L_A, r^S_D, R_D) - V^N_{M_R}(l^N_A, L_A, r^N_D, R_D) \\ &= l^S_A + (r^S_D - \delta - l^S_A)(1 - F_S(\frac{y^S}{2})) - l^N_A - (r^N_D - \delta - l^N_A)(1 - F_N(\frac{y^N}{2})) \\ &- (r^S_D - \delta - l^S_A)(1 - F_N(\frac{y^S}{2})) + (r^S_D - \delta - l^S_A)(1 - F_N(\frac{y^S}{2})) \\ &= l^S_A - l^N_A + (r^S_D - \delta - l^S_A)\epsilon_1 - \left((1 - F_N(\frac{y^N}{2})) + \frac{r^N_D - \delta - l^N_A}{2} f_N(\frac{y^N}{2})\right)(l^S_A - l^N_A) \\ &= (r^S_D - \delta - l^S_A)\epsilon_1 + \left(F_N(\frac{y^N}{2}) - \frac{r^N_D - \delta - l^N_A}{2} f_N(\frac{y^N}{2})\right)(l^S_A - l^N_A). \end{aligned} \tag{A-52}$$

Similarly, consider $\mathrm{d}(r-\delta-l)F(\frac{r+l-\delta}{2})$ when r and l change. It is given by

$$\left(F(\frac{r+l-\delta}{2})+\frac{r-\delta-l}{2}f(\frac{r+l-\delta}{2})\right)\mathrm{d}r$$
$$-\left(F(\frac{r+l-\delta}{2})-\frac{r-\delta-l}{2}f(\frac{r+l-\delta}{2})\right)\mathrm{d}l.$$

Then, by applying the first-order condition (16), we can compute that $(r_A^S-\delta-l_D^S)F_N(\frac{x^S}{2})-(r_A^N-\delta-l_D^N)F_N(\frac{x^N}{2})$ equals

$$\left(F_N(\frac{x^N}{2})+\frac{r_A^N-\delta-l_D^N}{2}f_N(\frac{x^N}{2})\right)(r_A^S-r_A^N).$$

Let $F_N(\frac{x^S}{2})-F_S(\frac{x^S}{2})=\epsilon_0$. By property 5, $\epsilon_0>0$. The RHS of (A-51) is given by

$$\begin{aligned}
&V_{M_L}^S(l_D^S,L_D,r_A^S,R_A)-V_{M_L}^N(l_D^N,L_D,r_A^N,R_A)\\
&=-r_A^S+(r_A^S-\delta-l_D^S)F_S(\frac{x^S}{2})+r_A^N-(r_A^N-\delta-l_D^N)F_N(\frac{x^N}{2})\\
&-(r_A^S-\delta-l_D^S)F_N(\frac{x^S}{2})+(r_A^S-\delta-l_D^S)F_N(\frac{x^S}{2})\\
&=r_A^N-r_A^S-(r_A^S-\delta-l_D^S)\epsilon_0+\left(F_N(\frac{x^N}{2})+\frac{r_A^N-\delta-l_D^N}{2}f_N(\frac{x^N}{2})\right)(r_A^S-r_A^N). \qquad \text{(A-53)}
\end{aligned}$$

Let

$$\chi_1=1-F_N(\frac{x^N}{2})-\frac{r_A^N-\delta-l_D^N}{2}f_N(\frac{x^N}{2}) \text{ and } \chi_2=F_N(\frac{y^N}{2})-\frac{r_D^N-\delta-l_A^N}{2}f_N(\frac{y^N}{2}).$$

By the results of Step I into (A-51), equating $V_{M_L}^N(l_D^N,L_D,r_A^N,R_A)-V_{M_L}^S(l_D^S,L_D,r_A^S,R_A)$ and $V_{M_R}^S(l_A^S,L_A,r_D^S,R_D)-V_{M_R}^N(l_A^N,L_A,r_D^N,R_D)$ yields

$$(r_A^S-\delta-l_D^S)\epsilon_0-\chi_1(r_A^S-r_A^N)=(r_D^S-\delta-l_A^S)\epsilon_1+\chi_2(l_A^S-l_A^N). \qquad \text{(A-54)}$$

Therefore,

$$r_A^S-r_A^N=\frac{\chi_2}{\chi_1}(l_A^S-l_A^N)+\frac{(r_D^S-\delta-l_A^S)\epsilon_1-(r_A^S-\delta-l_D^S)\epsilon_0}{\chi_1}.$$

Finally, we obtain

$$r_A^N-r_A^S=\frac{\chi_2}{\chi_1}(l_A^N-l_A^S)-\frac{(r_D^S-\delta-l_A^S)(\epsilon_1-\epsilon_0)}{\chi_1}. \qquad \text{(A-55)}$$

Let $F_N(0)-F_S(0)=\epsilon_A$. Then $V_{M_L}^N(l_A^N,L_A,r_A^N,R_A)-V_{M_L}^S(l_A^S,L_A,r_A^S,R_A)$ is rewritten as

$$\begin{aligned}
&F_S(0)l_A^S+(1-F_S(0))r_A^S-F_N(\frac{r_A^N+l_A^N}{2})l_A^N-(1-F_N(\frac{r_A^N+l_A^N}{2}))r_A^N\\
&=F_S(0)l_A^S+(1-F_S(0))r_A^S-F_N(\frac{r_A^N+l_A^N}{2})l_A^N-(1-F_N(\frac{r_A^N+l_A^N}{2}))r_A^N\\
&-F_N(0)l_A^S-(1-F_N(0))r_A^S+F_N(0)l_A^S+(1-F_N(0))r_A^S\\
&=2\epsilon_A r_A^S+\left(F_N(0)l_A^S-F_N(\frac{r_A^N+l_A^N}{2})l_A^N\right)+\left((1-F_N(0))r_A^S-(1-F_N(\frac{r_A^N+l_A^N}{2}))r_A^N\right). \qquad \text{(A-56)}
\end{aligned}$$

Consider how $G(\frac{x+y}{2})\cdot x$ when x and y change where we denote $g(x)=\frac{\mathrm{d}G(x)}{\mathrm{d}x}$. Then

$$\mathrm{d}(G(\frac{x+y}{2})\cdot x)=(\frac{g(\frac{x+y}{2})\cdot x}{2}+G(\frac{x+y}{2}))\mathrm{d}x+\frac{g(\frac{x+y}{2})\cdot x}{2}\mathrm{d}y.$$

Consider how $(1-G(\frac{x+y}{2}))\cdot y$ when x and y change. Then

$$\mathrm{d}(1-G(\frac{x+y}{2}))\cdot y=(\frac{-g(\frac{x+y}{2})\cdot y}{2}+(1-G(\frac{x+y}{2})))\mathrm{d}y-\frac{g(\frac{x+y}{2})\cdot y}{2}\mathrm{d}x.$$

Then the second term of (A-56) is

$$\begin{aligned}&F_N(\frac{r_A^S+l_A^S}{2})l_A^S-F_N(\frac{r_A^N+l_A^N}{2})l_A^N\\&=\left[\frac{f_N(\frac{r_A^S+l_A^S}{2})\cdot l_A^S}{2}+F_N(\frac{r_A^S+l_A^S}{2})\right]\cdot(l_A^S-l_A^N)+\frac{f_N(\frac{r_A^S+l_A^S}{2})\cdot l_A^S}{2}\cdot(r_A^S-r_A^N).\end{aligned}$$

Similarly, the third term of (A-56) is

$$\begin{aligned}&(1-F_N(\frac{r_A^S+l_A^S}{2}))r_A^S-(1-F_N(\frac{r_A^N+l_A^N}{2}))r_A^N\\&=\left[(1-F_N(\frac{r_A^S+l_A^S}{2}))-\frac{f_N(\frac{r_A^S+l_A^S}{2})\cdot r_A^S}{2}\right]\cdot(r_A^S-r_A^N)-\frac{f_N(\frac{r_A^S+l_A^S}{2})\cdot r_A^S}{2}\cdot(l_A^S-l_A^N).\end{aligned}$$

Substituting the two terms into (A-56), we obtain

$$\begin{aligned}&V_{M_L}^N(l_A^N,L_A,r_A^N,R_A)-V_{M_L}^S(l_A^S,L_A,r_A^S,R_A)\\&=2\epsilon_A r_A^S+\left[\frac{f_N(0)\cdot r_A^S}{2}-F_N(0)\right]\cdot(l_A^N-l_A^S)+\frac{f_N(0)\cdot r_A^S}{2}\cdot(r_A^N-r_A^S)\\&+\left[\frac{f_N(0)\cdot r_A^S}{2}-(1-F_N(0))\right]\cdot(r_A^N-r_A^S)+\frac{f_N(0)\cdot r_A^S}{2}\cdot(l_A^N-l_A^S)\\&=2\epsilon_A r_A^S+(f_N(0)r_A^S-F_N(0))\cdot(l_A^N-l_A^S)+(f_N(0)r_A^S-(1-F_N(0)))(r_A^N-r_A^S).\end{aligned}\tag{A-57}$$

Thus, by (A-51),

$$\begin{aligned}&2\epsilon_A r_A^S+(f_N(0)r_A^S-F_N(0))\cdot(l_A^N-l_A^S)+(f_N(0)r_A^S-(1-F_N(0)))(r_A^N-r_A^S)\\&=(r_D^S-\delta-l_A^S)\epsilon_1-\chi_2(l_A^N-l_A^S).\end{aligned}$$

Then,

$$\begin{aligned}&\chi_2\left(1+\tfrac{f_N(0)r_A^S-F_N(0)}{\chi_2}+\tfrac{f_N(0)r_A^S-(1-F_N(0))}{\chi_1}\right)(l_A^N-l_A^S)\\&=(r_D^S-\delta-l_A^S)\epsilon_1-2\epsilon_A r_A^S-\tfrac{f_N(0)r_A^S-(1-F_N(0))}{\chi_1}(r_D^S-\delta-l_A^S)(\epsilon_1-\epsilon_0).\end{aligned}\tag{A-58}$$

Step II. Completing the proof.

By applying (16) and (18),

$$\chi_1 = 1 - 2F_N(\frac{x^N}{2}) > 0 \text{ and } \chi_2 = 2F_N(\frac{y^N}{2}) - 1 > 0.$$

By the slopes of expected payoffs, $f_N(0)r_A^S - F_N(0) < 0$ and $f_N(0)r_A^S - (1 - F_N(0)) < 0$. Therefore, $f_N(0)r_A^S < \frac{1}{2}$. Therefore, for some γ, we have

$$1 + \frac{f_N(0)r_A^S - F_N(0)}{\chi_2} + \frac{f_N(0)r_A^S - (1 - F_N(0))}{\chi_1} = 2 \cdot \frac{f_N(0)r_A^S - F_N(\frac{x^N}{2})}{\chi_1} + \gamma.$$

When $\epsilon_0 = \epsilon_1 = \epsilon_A$ by (A) under the scenario of extremists' change, the RHS of (A-58) is

$$\begin{aligned}
&(r_D^S - \delta - l_A^S)\epsilon_1 - 2\epsilon_A r_A^S + \frac{(1 - F_N(0)) - f_N(0)r_A^S}{\chi_1}(r_D^S - \delta - l_A^S)(\epsilon_1 - \epsilon_0) \\
&= \epsilon_1(r_D^S - \delta - l_A^S - 2r_A^S) = \epsilon_1(r_D^S - \delta - r_A^S) > 0,
\end{aligned}$$

where the last inequality is due to Lemma A-9, which shows $r_D^S - \delta - r_A^S > 0$.

When $0 = \epsilon_0 = \epsilon_1 < \epsilon_A$ by (B) under the scenario of centrists' change, the RHS of (A-58) is

$$\begin{aligned}
&(r_D^S - \delta - l_A^S)\epsilon_1 - 2\epsilon_A r_A^S + \frac{(1 - F_N(0)) - f_N(0)r_A^S}{\chi_1}(r_D^S - \delta - l_A^S)(\epsilon_1 - \epsilon_0) \\
&= -2\epsilon_A r_A^S < 0.
\end{aligned}$$

Consider the LHS of (A-58). By (16) and (17), $r_A^S = \frac{(F_S(\frac{x^S}{2}))^2}{\frac{1}{2}f_S(\frac{x^S}{2})}$. Thus, the LHS is positive if and only if $\frac{(F_S(\frac{x^S}{2}))^2}{\frac{1}{2}f_S(\frac{x^S}{2})} < \frac{F_N(\frac{x^N}{2})}{f_N(0)}$. By Assumption 1, we have $\frac{F_S(\frac{x^S}{2})}{f_S(\frac{x^S}{2})} < \frac{F_S(0)}{f_S(0)} = \frac{\frac{1}{2}}{f_S(0)}$. By property 5, we have $F_S(\frac{x^S}{2}) < F_N(\frac{x^S}{2})$. Thus, we have $\frac{(F_S(\frac{x^S}{2}))^2}{f_S(\frac{x^S}{2})} < \frac{\frac{1}{2}F_N(\frac{x^S}{2})}{f_S(0)}$. Because x^S and x^N can be arbitrarily close, we have $r_A^S = \frac{(F_S(\frac{x^S}{2}))^2}{f_S(\frac{x^S}{2})} < \frac{\frac{1}{2}F_N(\frac{x^N}{2})}{f_S(0)}$. Given $f_N(0)r_A^S < F_N(\frac{x^N}{2})$, the LHS of (A-58) is negative. Combining these observations, we obtain the desired result. □

References

Abney, R., Adams, J., Clark, M., Easton, M., Ezrow, L., Kosmidis, S. and Neundorf, A. (2013), 'When does valence matter? heightened valence effects for governing parties during election campaigns', *Party Politics* **19**(1), 61–82.

Abramowitz, A. I. (1989), 'Viability, electability, and candidate choice in a presidential primary election: A test of competing models', *The Journal of Politics* **51**(4), 977–992.

Adams, J., Ezrow, L., Merrill, S. I. and Somer-Topcu, Z. (2013), 'Does collective responsibility for performance alter party strategies? policy-seeking parties in proportional systems', *British Journal of Political Science* **43**(10), 1–23.

Adams, J., Green, J. and Milazzo, C. (2012), 'Has the british public depolarized along with political elites? an american perspective on british public opinion', *Comparative Political Studies* **45**(4), 507–530.

Adams, J. and Merrill, S. I. (2008), 'Candidate and party strategies in two-stage elections beginning with a primary', *American Journal of Political Science* **52**(2), 344–359.

Adams, J. and Merrill, S. I. (2009), 'Policy-seeking parties in a parliamentary democracy with proportional representation: A valence-uncertainty model', *British Journal of Political Science* **39**(3), 539–558.

Ali, S. N. and Kartik, N. (2012), 'Herding with collective preferences', *Economic Theory* **51**(3), 601–626.

Anh, D. S. and Oliveros, S. (2012), 'Combinatorial voting', *Econometrica* **80**(1), 89–141.

Aragones, E. and Palfrey, T. R. (2002), 'Mixed equilibrium in a downsian model with a favored candidate', *Journal of Economic Theory* **103**(1), 131–161.

Ausubel, L. M. and Milgrom, P. (2002), 'Ascending auctions with package bidding', *Frontiers of Theoretical Economics* **1**, 1–42.

Benoit, K. and Laver, M. (2006), *Party policy in modern democracies*, Routledge.

Bouton, L. (2013), 'A theory of strategic voting in runoff elections', *The American Economic Review* **103**(4), 1248–1288.

Bullock, W. and Clinton, J. D. (2011), 'More a molehill than a mountain: The effects of the blanket primary on elected officials behavior from california', *The Journal of Politics* **73**(3), 915–930.

Buttice, M. K. and Stone, W. J. (2012), 'Candidates matter: Policy and quality differences in congressional elections', *The Journal of Politics* **74**(12), 870–887.

Callander, S. (2007), 'Bandwagons and momentum in sequential voting', *The Review of Economic Studies* **74**(3), 653–684.

Carey, J. M. and Polga-Hecimovich, J. (2006), 'Primary elections and candidate strength in latin america', *The Journal of Politics* **68**(3), 530–543.

Casas, A. (2018), 'Ideological extremism and primaries', *Working Paper* .

Chen, K.-P. and Yang, S.-Z. (2002), 'Strategic voting in open primaries', *Public Choice* **112**(1/2), 1–30.

Cho, S.-j. and Kang, I. (2014), 'Open primaries and crossover voting', *Journal of Theoretical Politics* . DOI: 10.1177/0951629814531671.

Clark, M. and Leiter, D. (2014), 'Does the ideological dispersion of parties mediate the electoral impact of valence? a cross-National study of party support in nine western european democracies', *Comparative Political Studies* **47**(2), 171–202.

Coleman, J. S. (1971), 'Internal processes governing party positions in elections', *Public Choice* **11**(1), 35–60.

Day, R. and Milgrom, P. (2008), 'Core-selecting package auctions', *International Journal of Game Theory* **36**, 393–407.

Dekel, E. and Piccione, M. (2000), 'Sequential voting procedures in symmetric binary elections', *Journal of political Economy* **108**(1), 34–55.

Dell'Ariccia, G. and Marquez, R. (2006), 'Lending booms and lending standards', *The Journal of Finance* **61**(5), 2511–2546.
URL: *http://www.jstor.org/stable/3874718*

Deltas, G., Herrera, H. and Polborn, M. (2016), 'Learning and coordination in the presidential primary system', *Review of Economic Studies* **83**(4), 1544–1578.

Drouvelis, M., Saporiti, A. and Vriend, N. J. (2014), 'Political motivations and electoral competition: Equilibrium analysis and experimental evidence', *Games and Economic Behavior* **83**, 86–115.

Frey, W. (2015), *Diversity Explosion: How New Racial Demographics are Remaking America*, Brookings Institution Press, Washington.

Galderisi, P. and Ezra, M. (2001), Congressional primaries in historical and theoretical context, *in* P. Galderisi, M. Ezra and M. Lyons, eds, 'Congressional Primaries and the Politics of Representation', Rowman & Littlefield Publishers, INC, Oxford.

Gerber, E. and Morton, R. (1998), 'Primary election systems and representation', *The Journal of Law, Economics, & Organization* **14**(2), 304–324.

Green, J. and Hobolt, S. B. (2008), 'Owning the issue agenda: Party strategies and vote choices in british elections', *Electoral Studies* **27**(3), 460–476.

Grofman, B., Troumpounis, O. and Xefteris, D. (2019), 'Electoral competition with primaries and quality asymmetries', *Journal of Politics* **81**(1), 260–273.

Hall, A. B. (2015), 'What happens when extremists win primaries?', *American Political Science Review* **109**(1), 1842.

Harmon, N. A. (2018), 'Immigration, ethnic diversity, and political outcomes: Evidence from Denmark', *Scandinavian Journal of Economics* **120**, 1043–1074.

Hill, S. J. and Tausanovitch, C. (2018), 'Southern realignment, party sorting, and the polarization of american primary electorates, 1958–2012', *Public Choice* **176**(1), 107–132.
URL: *https://doi.org/10.1007/s11127-017-0478-0*

Hirano, S. and Snyder, Jr., J. M. (2012), 'What happens to incumbents in scandals?', *Quarterly Journal of Political Science* **7**(4), 447–456.

Hirano, S. and Snyder, Jr., J. M. (2014), Primary elections and the quality of elected officials. Working Paper.

Hirano, S., Snyder, Jr., J. M., Ansolabehere, S. D. and Hansen, J. M. (2010), 'Primary elections and partisan polarization in the U.S. congress', *Quarterly Journal of Political Science* **5**(2), 169–191.

Hummel, P. (2013), 'Candidate strategies in primaries and general elections with candidates of heterogeneous quality', *Games and Economic Behavior* **78**, 85–102.

Klumpp, T. and Polborn, M. K. (2006), 'Primaries and the new hampshire effect', *Journal of Public Economics* **90**(6), 1073–1114.

Lizzeri, A. and Persico, N. (2001), 'The provision of public goods under alternative electoral incentives', *The American Economic Review* **91**(1), 225–239.

McCarty, N., Poole, K. T. and Rosenthal, H. (2006), *Polarized America: The Dance of Ideology and Unequal Riches*, The MIT University Press, Cambridge, MA.

McCarty, N., Poole, K. T. and Rosenthal, H. (2013), *Political Bubbles: Financial Crises and the Failure of American Democracy*, Princeton University Press, Princeton, NJ.

McCurley, C. and Mondak, J. J. (1995), 'Inspected by #1184063113: The influence of incumbents' competence and integrity in U.S. house elections', *American Journal of Political Science* **39**(4), 864–885.

McGann, A. J. (2002), 'The advantages of ideological cohesion: a model of constituency representation and electoral competition in multi-party democracies', *Journal of Theoretical Politics* **14**(1), 37–70.

McLennan, A., Monteiro, P. K. and Tourky, R. (2011), 'Games with discontinuous payoffs: a strengthening of Reny's existence theorem', *Econometrica* **79**(5), 1643–1664.

Mondak, J. J. (1995), 'Competence, integrity, and the electoral success of congressional incumbents', *The Journal of Politics* **57**(4), 1043–1069.

Morton, R. (2006), *Analyzing Elections: New Institutionalism in American Politics*, W. W. Norton & Company, New York, NY.

Myerson, R. (1993), 'Incentives to cultivate favored minorities under alternative electoral systems', *The American Political Science Review* **87**(4), 856–869.

Owen, G. and Grofman, B. (2006), 'Two-stage electoral competition in two-party contest: Persistent divergence of party positions', *Social Choice and Welfare* **26**(3), 547–569.

Prat, A. and Rustichini, A. (2003), 'Games played through agents', *Econometrica* **71**(4), 989–1026.
URL: *http://www.jstor.org/stable/1555488*

Reny, P. J. (1999), 'On the existence of pure and mixed strategy Nash equilibria in discontinuous games', *Econometrica* **67**(5), 1029–1056.

Roemer, J. E. (1997), 'Political-economic equilibrium when parties represent constituents: The unidimensional case', *Social Choice and Welfare* **14**(4), 479–502.

Saporiti, A. (2008), 'Existence and uniqueness of Nash equilibrium in electoral competition games: The hybrid case', *Journal of Public Economic Theory* **10**(5), 827–857.

Schwabe, R. (2010), Super tuesday: campaign finance and the dynamics of sequential elections. Working Paper.

Serra, G. (2011), 'Why primaries? the party's tradeoff between policy and valence', *Journal of Theoretical Politics* **23**(21), 21–55.

Simon, L. K. and Zame, W. R. (1990), 'Discontinuous games and endogenous sharing rules', *Econometrica* **58**(4), 861–872.
URL: *http://www.jstor.org/stable/2938353*

Snyder, J. M. J. and Ting, M. M. (2011), 'Electoral selection with parties and primaries', *American Journal of Political Science* **55**, 782–796.

Stone, W. J., Rapoport, R. B. and Abramowitz, A. I. (1992), 'Candidate support in presidential nomination campaigns: The case of iowa in 1984', *The Journal of Politics* **54**(4), 1074–1097.

Stone, W. J. and Simas, E. N. (2010), 'Candidate valence and ideological placement in U. S. house elections', *American Journal of Political Science* **54**(2), 371–388.

Takayama, S. (2014), A model of two-stage electoral competition with strategic voters. School of Economics Discussion Paper Series No. 525, University of Queensland.

Takayama, S., Tamura, Y. and Yeo, T. (2019), 'Nash equilibrium and party polarization in an electoral competition model', *Journal of Public Economic Theory* **21**(2), 219–240.

Ware, A. (2002), *The American Direct Primary: Party Institutionalization and Transformation in the North*, Cambridge University Press, UK.

Woon, J. (2018), 'Primaries and candidate polarization: Behavioral theory and experimental evidence', *American Political Science Review* **112**(4), 826843.