Abstract

Candidates in primary elections face a trade-off between ideological proximity to the party's base and the probability of winning the general election. I describe conditions under which an incumbent politician provokes the opposition party into nominating more extreme candidates. By moving away from the center, the incumbent can exploit the divisions in the opposition party and improve her reelection prospects, despite hurting her appeal to the median voter in the general election. I identify primaries as an understudied mechanism that drives political polarization in two-party democracies. Furthermore, the analysis fits the observation that reelection-seeking incumbents sometimes move away from the center towards the end of their first term in office. I show that party gatekeeping drives this behavior: with a richer set of candidates, the effect disappears.
“Folks, we have a choice. We can go down the road that Senator Sanders and Senator Warren want to take us, which is with bad policies like Medicare for all, free everything and impossible promises that will turn off independent voters and get Trump reelected.”

— John Delaney, 2020 Democratic Party Presidential Debate, July 30, 2019

“You know, I don’t understand why anybody goes to all the trouble of running for president of the United States just to talk about what we really can’t do and shouldn’t fight for.”

— Senator Elizabeth Warren, in response to John Delaney

1 Introduction

The central feature of primary elections is the tug-of-war between a candidate’s electability and her ideological congruence with the party’s base. The former refers to a candidate’s ability to appeal to voters near the center of the ideological spectrum, and therefore her chances of winning in the general election. Electability motivations generally push towards moderation in the primaries (Owen and Grofman, 2006). The latter, ideological congruence, refers to primary voters’ desire to field a candidate who will faithfully represent their interests if elected. Because primary voters are different from general election voters, candidates who are more likely to appeal to the latter tend to be further ideologically from the party’s base (Brady et al., 2007; King et al., 2016). Primary voters thus face a trade-off between fielding a candidate who may have a greater probability of winning in the general election, but may have less similar ideology; versus someone closer who might have a harder time winning the hearts of swing voters.

I study a forward-looking incumbent politician’s ability to strategically manipulate this trade-off. I propose a model where the incumbent can improve her reelection chances by strategically pursuing more extreme policies that provoke extremist factions in the opposition party. This behavior occurs because shifting the incumbent’s position changes the
selection of candidates in the opposition, possibly benefiting the incumbent. Incumbents therefore have both the means and the motivation to influence the opposition party's primary process.

The core of this paper’s argument is as follows. As the incumbent moves away from the center, the opposition party’s payoffs change in two opposing ways. On the one hand, the value of beating the incumbent increases. Specifically, if a more extreme incumbent were reelected the policy implemented is disliked more by the members of the opposition. This strengthens the incentive to win the general election and pushes the opposition towards the center. On the other hand, a more extreme incumbent increases the probability that any given challenger would win in the general election, including extremists. This emboldens the extremists in the opposition, who see a rare window of opportunity to pursue their agenda. Thus, a more extreme incumbent may lead the opposition party to field a candidate ideologically closer to their base, pushing them away from the center. On the net, the change in the probability of reelection as the incumbent becomes more extreme is ambiguous a priori. I study a model featuring these two opposing forces to discern the conditions under which extremism begets more extremism.

I analyze primary elections under two different candidate pools. First, an open nominations model where the opposition party can choose the ideology of the challenger from the entire policy space. Second, a party elites model where they are constrained to a discrete set. In the open nominations setting, I first consider the party as a unitary actor to introduce the central trade-off between electability and ideological congruence in the simplest possible way. I describe the conditions under which behavior in this case coincides with that of a model where the challenger’s ideology is chosen by primary voters (Proposition 1). This equivalence holds even though these candidate selection mechanisms are diametrically opposite in terms of the degree of party control. I show that there is no incentive for the incumbent to provoke the opposition under either model (Proposition 2).
Then, I consider a more restricted nomination environment by analyzing a party elites model where the challenger must be chosen from a discrete set of candidates with exogenously given ideologies. Here, primary voters decide which of the given two candidates should face off against the incumbent in the general election. This approach allows me to compare challengers that emerge from an open primary field against those who are chosen from a small set of party elites. In particular, I show that the incumbent can provoke the opposition by choosing platforms more extreme than her ideal point if the opposition party must choose from a discrete set of ideologies. By intentionally hurting her appeal to the median voter, the incumbent can induce extremists to win the opposition party primary, thus improving her reelection prospects (Proposition 3).

Next, I consider a candidate-entry model. When deciding whether to run, candidates evaluate the costs of running and the impact their entry has on the eventual winner of the election. Extremists choose to stay out of races against moderate incumbents because moderate challengers are more likely to win in the general election. But they will enter themselves when their chances of beating the incumbent are high enough. Thus, an extremist incumbent induces entry by extremists in the opposition party’s primary, which results in a higher probability of reelection. This behavior occurs when the primary field is even or when it is tilted in favor of extremists, but not when there is a primary advantage for moderates (Proposition 4). Additionally, I show that it may be preferable for the incumbent to weaken her appeal to the median voter in this setting too (Proposition 5).

This paper also makes several theoretical contributions. Existing models often make explicit functional form assumptions about payoffs and probabilities of winning. Instead, this model provides sharp predictions that hold under a large class of functional forms, including all widely used loss and distribution functions. This generality allows the model presented here to be widely applicable for future researchers who want to study primary elections. Furthermore, it generalizes results in previous papers and shows that they extend
to a larger class of functions. Finally, it describes conditions on voters’ risk preferences that ensure median voters’ decisiveness in the primary context, a result that had hitherto eluded researchers.\footnote{In most papers that study primaries, median primary voter’s decisiveness is assumed either explicitly (Owen and Grofman, 2006; Serra, 2011; Snyder and Ting, 2011) or implicitly (Grofman et al., 2019). Adams and Merrill (2008) assume primary voters ignore the electability of candidates and vote as if they were voting in the general election. Mirhosseini (2015) shows the median primary voter is decisive when losses are quadratic and uncertainty is normally distributed. I generalize this result to much broader classes of loss functions and distributions.}

2 RELATED LITERATURE

After the foundational papers studying this trade-off between electability and congruence in the 1970s (Coleman, 1971; Aronson and Ordeshook, 1972), the formal literature on primaries lay mostly dormant until it recently saw a revival (Owen and Grofman, 2006). These theoretical models find that the two opposing forces highlighted in the introduction lead primary voters to support candidates with platforms between their ideal points and the population median, conceding slightly on policy for greater electability. When aggregated, this results in candidates who are between the median primary voter and the median general election voter (Jackson et al., 2007). In other words, the winner of the primary election is an interior solution of this trade-off between electability and ideology; she is not quite so extreme as to be unelectable, and not quite so moderate she is indistinguishable from the other party’s candidate. This prediction has broad empirical support: according to both observational (Abramowitz, 1989; Abramson et al., 1992; Stone et al., 1992) and experimental studies (Rickershauser and Aldrich, 2007; Woon, 2018), primary voters seem to have both of these considerations in mind while casting ballots, and candidates position themselves accordingly (Brady et al., 2007).

Recent empirical evidence lends support to this idea that there is pressure from both sides. On the one hand, voters with extreme ideologies from either party are both more
likely to vote and donate in primary elections than moderates (Hill and Huber, 2017; Barber et al., 2017). On the other hand, extremists who win primaries are more likely to lose general elections. For example, Hall (2015) finds that when an extremist candidate running for US Congress barely wins a primary, this leads to a smaller vote share and a lower probability of winning in the general election for their party and that downstream roll-call voting shifts towards the other party. Thus a party benefits from their opponents electing an extremist; this enables them to retain office and pursue their agenda.

How do parties deal with this trade-off? As others have recognized, we need a better theoretical understanding of this trade-off and the underlying dynamics that drive extremists to win primaries (Tausanovitch and Warshaw, 2018). Many factors play into primary voters’ voting decisions, like the ideological distribution of voters (Meirowitz, 2005; Serra, 2011), the uncertainty about candidates’ abilities to appeal to swing voters (Adams and Merrill, 2008; Snyder and Ting, 2011), concerns about candidates flip-flopping (Hummel, 2010), and the main focus of this paper: the incumbent’s policies (Mirhosseini, 2015). Previous formal literature has identified interesting comparative statics about how the incumbent’s position affects the opposition party’s primaries, but it has done so taking the incumbent’s position as given. Instead, the present paper focuses on a reelection-seeking incumbent’s ability to strategically manipulate the opposition party’s primary voters and shows how extremists can emerge victorious from primaries endogenously in equilibrium as a result of strategic provocation by the incumbent.

This paper also relates to the literature on candidate entry. Banks and Kiewiet (1989) and Buisseret and Van Weelden (Forthcoming) explore candidate entry to primary races. Thomsen (2014, 2017) finds that more liberal Republicans think they are less likely to win their party’s primary election, and they value winning elections less, relative to more conservative Republicans. This leads to fewer moderate Republicans running for office. Hall (2019) argues that it is the high costs of running, and low benefits for office, that drive moderates
out of running. When moderates stay out, extremists have no choice but to run themselves
to try to keep the opposition party from pursuing their agenda. His argument therefore
applies to settings where moderates prefer to stay out, and extremists’ entry decreases the
probability opposition party will win the election. The general model of primaries I pro-
vide allows these observations to be recovered and expands on them by considering races
where extremists join moderates in running for office, despite their understanding that this
increases the probability that the opposition party will win the election.

3 PRIMARY ELECTIONS

I start by presenting the primitives of the model. There are two periods and a unidimen-
sional policy space. There is a measure one of citizens whose ideologies are summarized
by their ideal points $x_i \in \mathbb{R}$. Each citizen’s ideology payoff from the implemented policy $x$ is
captured by $-\ell(|x - x_i|)$, where the loss function $\ell$ is increasing and convex in the absolute
distance between the implemented policy and voter $i$’s ideal point.$^2$ Citizens also receive a
global utility shock $\psi$ for the incumbent, drawn from distribution function $F$ that has an as-
sociated log-concave probability density function, $f$. Accordingly, their utility in the second
period is the ideological loss from the incumbent’s platform plus the shock if the incumbent
is reelected, and the ideological loss from the challenger’s platform otherwise.$^3$

Citizens are affiliated with at most one of two parties: $L$ and $R$. Without loss of gener-
ality I assume that the ideal point of the median member of $L$ is less than the ideal point
of the median member of $R$, and fix the ideal point of the overall median to zero, so that
$x_{m_L} < 0 < x_{m_R}$. Candidates in the model are also citizens, in that they have ideal points,
and receive ideology payoffs given by the loss function $\ell$. The incumbent cares about the

$^2$Formally, I assume $\ell: \mathbb{R}_+ \rightarrow \mathbb{R}_+$ is thrice continuously right differentiable, with $\ell(0) = 0$, $\ell' > 0$, and $\ell'' \geq 0$. This class includes all widely used loss functions, including linear, quadratic, and exponential loss.

$^3$This is known as the stochastic partisanship model and is described in great detail in Duggan (2006).
policy implemented in the second period only if she is reelected. Both the incumbent and
the opposition party candidates enjoy office benefits of $B \geq 0$ if they win the general election.
There is no discounting. The timing of the game is as follows:

1. Incumbent chooses her platform for the second period, $x_I \geq 0$.
2. The $L$ party primary election is held.
3. Global utility shock against the incumbent is revealed.
4. The general election is held between the incumbent and the challenger.
5. The winner of the general election implements her platform in the second period.

I focus on Subgame Perfect Nash Equilibria in undominated strategies where voters
vote for the candidate with the lower platform when they are indifferent between multiple
candidates. Ties are resolved by coin flips.

I proceed by backward induction. Analysis of the general election is straightforward. In
the second period, the winner of the first-period general election implements her platform.
Thus, voters vote for the candidate who gives them a higher payoff in the second period.
Formally, given a challenger $J$ with ideal point $x_J$, and the realization of the utility shock
$\psi$, voter $x_i$ votes for the challenger if $-\ell(|x_J - x_i|) \geq -\ell(|x_I - x_i|) + \psi$. Thus, voter $i$ votes
for the challenger with probability $F(\ell(|x_I - x_i|) - \ell(|x_J - x_i|))$. In particular, the median
voter votes for the challenger with probability $F(\ell(x_I) - \ell(x_J))$. Because this is an election
between two candidates on a single policy dimension with single-peaked preferences, the
median voter is decisive. Thus, this expression equals the probability that a challenger with
platform $x_J$ wins. This gives the probability that any potential challenger would beat the
incumbent. Having described these probabilities, we are now ready to explore the primary

---

$^4$This simplifies the exposition by stripping the model of motivations for the incumbent to manipulate the
ideology of the challenger other than improving her reelection prospects. Results can be recovered if the incum-
bent’s payoff function includes a continuation value that takes into account her successor’s ideology. Although
not without loss of generality, this assumption is standard in the literature. An exception is Bernhardt et al.
(2009) who show such concerns push incumbents towards moderation.
election stage. I start the analysis of primary elections with a model where members of $L$ can freely choose the ideology of the challenger from the entire policy space.

3.1 Open Nominations

Suppose first that control of the challenger selection process is concentrated in the hands of party insiders. Specifically, suppose that they can act as a unitary actor and decide unilaterally who the challenger will be. I denote the party insiders by $d$, and their ideal point by $x_d < 0$. When choosing the ideology of the challenger $x \in \mathbb{R}$, party insiders care not only about the positions of candidates but also their probability of beating the incumbent. This is because their expected payoff is given by the probability each politician wins in the general election multiplied by the policy payoffs she provides, summed over the challenger and the incumbent. Formally, insiders solve:

$$\max_{x \in \mathbb{R}} EU_d(x, x_I) = \max_{x \in \mathbb{R}} -\ell(|x_d - x|)F(\ell(x_I) - \ell(|x|)) - \ell(x_I - x_d)(1 - F(\ell(x_I) - \ell(|x|))).$$

(1)

The solution to the above problem must be between party insiders’ ideal point and that of the median voter. To see why, notice first that a candidate with platform $x_d$ is preferred to anybody more extreme than she because she is both more likely to be elected, and results in a higher payoff for party insiders if she does. Similarly, a candidate whose platform is the ideal point of the median voter, 0, is both more likely to beat the incumbent and also yields a higher payoff to $d$ than anybody to the right of her. These imply that the optimal challenger’s ideology must be in $[x_d, 0]$. Thus, we can write the derivative of the insiders’ problem as:

$$-\ell'(x - x_d)F(\ell(x_I) - \ell(-x)) + \ell'(-x)f(\ell(x_I) - \ell(-x))(\ell(x_I - x_d) - \ell(x - x_d)).$$

(2)
The first term in this expression is the marginal loss of policy payoff as the challenger moves away from \(d\)'s ideal point. The second term is the marginal gain of beating the incumbent with a higher probability. If the \(\ell'(x - x_d)F(\ell(x_I) - \ell(-x))\) term is less than the second term, this expression is always positive. This means that when the probability of defeating the incumbent is very low for all \(x\), insiders always find it worthwhile to trade-off ideological congruence for more electability, and so the optimal challenger has maximal electability at 0. In contrast, if this expression is always negative, the party insiders do not want to give up on policy at all. This happens when \(L\) has a high chance of beating the incumbent even with a challenger whose platform is the same as insiders’ ideal point. Finally, if expression 2 is positive for some platforms and negative for others, there must be a unique interior optimum. I summarize these observations in the following lemma:

**Lemma 1.** When party insiders choose the ideology of the challenger, the challenger has an ideal point between those of the party insiders and the median voter.

*Proof.* All proofs in Appendix A.

Next, I examine the opposite case of extremely diffuse control over the selection process of challengers. All observations above about party insiders carry through when we consider primary voters, except the selection method of the challenger. Specifically, because there no longer is a unitary actor picking the ideology of the challenger, we must consider a preference aggregation rule for primary voters. As is standard, I consider the Condorcet method, which means that the winner of the primary must win a majority of votes in a pairwise race against any other candidate.

Notice that despite single-peaked preferences over a single policy dimension, the decisiveness of the median primary voter does not readily obtain in this setting. Instead, centrists and extremists in primaries may vote together against the median primary voter’s favored candidate. There are two ways this can happen. First, centrist primary voters may
side with extremists to vote for an extremist challenger. This may happen because centrist primary voters do not necessarily have single-peaked preferences over challenger platforms. They may rather vote for an extremist challenger over a moderate one in the primary election to improve the incumbent’s chance of reelection, even though they would prefer the moderate challenger over the extremist to be in office. This is because although the extremist challenger would give them a very low payoff if elected, her probability of being elected is also very low. In contrast, a moderate challenger can both have a high probability of beating the incumbent and result in a much lower payoff than her. Such voters, if they are allowed to vote in the primary, would vote for an extremist candidate to decrease the probability the party whose primary they are voting in wins in the general election. This is known as “raiding” or “strategic crossover voting,” and may lead to Condorcet cycles as centrist voters side with extremist voters to vote for extremist candidates (Chen and Yang, 2002; Oak, 2006).

Second, extremist primary voters may side with the centrists to vote for a candidate more moderate than the median primary voter’s preferred candidate. This is because voters’ induced rankings over candidates may be non-monotonic in their ideal points. This happens when a primary voter who is slightly more extreme than another receives so much greater disutility from the incumbent’s reelection that he is willing to support a more moderate candidate who has a higher chance of beating the incumbent, even if her platform is only slightly closer than that of the incumbent. This can lead to extremist voters joining centrists in supporting a candidate more moderate than the median primary voter’s optimal candidate. Then, the Condorcet winner may be different than the median $L$ voter’s optimal candidate.\footnote{For example, suppose there are three primary voters whose ideal points are $-6$, $-2$, and $-1$, and whose preferences are described by the following loss function

$$\ell(x-x_i) = \begin{cases} |x-x_i| & \text{if } |x-x_i| \leq 5 \\ (x-x_i)^2 - 9(x-x_i) + 25 & \text{otherwise.} \end{cases}$$

This function is increasing, convex, and right differentiable. Here, the median primary voter’s optimal candidate, located at $-1.8$, loses the primary election against a candidate at $-1$, who is the optimal candidate of both the centrist and the extremist primary voter.}
A sufficient condition for the existence of a Condorcet winner and its coincidence with the median primary voter’s optimal candidate is that the third derivative of the loss function is not too high, and so the rate of increase of the marginal loss function does not increase very fast.\footnote{Formally, I require that the first derivative of the loss function be log-concave, or $\ell''' \ell' \leq (\ell'')^2$.} This assumption is satisfied by all widely used loss functions including linear, quadratic, and exponential; and precludes the above-mentioned pathological cases by ensuring primary voters’ payoff functions satisfy an “increasing ratios” property in their ideal points and candidates’ platforms.\footnote{The interested reader can compare this property with increasing differences (Ashworth and Bueno de Mesquita, 2006), log-supermodularity (Milgrom and Roberts, 1990), and ratio dominance (Kartik et al., 2019).} This ensures the median primary voter’s optimal candidate, $x^*_{m_L}$, is the Condorcet winning ideology. Henceforth, I maintain this assumption.

**Proposition 1.** The median primary voter’s optimal candidate is the Condorcet winner.\footnote{This result is a generalization of Proposition 1 in Mirhosseini (2015) who shows that the median primary voter is decisive when losses are quadratic: $\ell(|x-x_i|) = (x-x_i)^2$ and uncertainty about voter preferences are captured by the normal distribution.}  

Proposition 1 shows that under a broad range of loss functions, the median primary voter is decisive, and effectively replaces the party insiders in choosing the ideology of the challenger.\footnote{A substantive implication of this is that if voters’ policy preferences satisfy the increases ratios property, raiders cannot influence outcomes of primaries. This is in line with the empirical literature, which finds openness of primaries has little effect on the ideology of the elected legislators (McGhee et al., 2014), or the platforms they run on (Rogowski and Langella, 2015).} Put differently, when $L$ holds primaries to choose the ideology of the challenger instead of insiders choosing unilaterally, the only change is that the challenger’s ideology coincides with that of the median primary voter’s optimal candidate $x^*_{m_L}$, instead of that of the party insiders $x^*_{d}$.  

The exact location of the Condorcet winner depends on the functional forms. For example, if losses are linear or exponential, $\ell(|x-x_i|) = |x-x_i|$, or $\ell(|x-x_i|) = e^{-|x-x_i|}$, there exists a platform $x^*$ who is the optimal candidate of everyone to her left. Then, it follows from Proposition 1 that the Condorcet winner must be either $x^*$, or the median primary voter, whichever is more moderate:
Figure 1: Optimal candidates of $L$ voters when losses are linear or exponential, and $x_{mL} > x^*$. 

**Lemma 2.** When losses are linear or exponential, there exists a unique candidate platform $x^*$ such that all voters with ideologies to the left of $x^*$ prefer a candidate with this platform to face off the incumbent in the general election over any other candidate. Then, the Condorcet winner is the more moderate of $x_{mL}$ and $x^*$.

Thus, for the special cases of linear and exponential losses, there exists a platform $x^*$ such that all primary voters with more extreme ideal points agree that a candidate with that ideology trades off electability and ideology perfectly. That is to say, for voters with ideologies $x \leq x^*$, the gain in the probability of winning with someone slightly more moderate than $x^*$ would be too little to justify the ideological loss, and someone slightly more extreme would be too unlikely to win to make the ideology gain worthwhile.

More generally, we do not know where the ideal point of the challenger chosen by party insiders or primary voters is located. However, we can still study how it changes with the incumbent’s platform. Recall the trade-off faced by $L$ voters described in the introduction: against a more extreme incumbent, every challenger has a higher chance of winning, and this strengthens the incentives $L$ voters have to push forward candidates whose platforms they like more. This is because the $F(\ell(x_I) - \ell(-x))$ term in $L$ voters’ payoff increases. On the other hand, a more extreme incumbent increases the value of defeating the incumbent by increasing the $\ell(x_I - x_i) - \ell(x - x_i)$ component of the payoff function for $i \in \{l, mL\}$, which pushes $L$ voters to favor more electable candidates.

---

10This result is a generalization of Theorem 1 of Owen and Grofman (2006), which shows such an ideology exists for exponential loss and normal distribution.

12
Shapes of $\ell$ and $F$ determine whether the challenger chosen by $L$ becomes more or less extreme as the incumbent moves away from the median voter's ideal point. If a slightly more extreme incumbent causes a much larger policy loss for $L$ voters without causing a big change in the probability of reelection, they would rather choose a more moderate candidate who has a better chance of beating a more extreme incumbent. In contrast, if a slightly more extreme incumbent leads to a substantially lower chance of reelection without causing much additional disutility for $L$ voters, they would rather choose somebody more congruent, even though such a candidate would be less able to appeal to the median general election voter. Whether the ideology of $L$'s optimal challenger becomes more or less extreme as the incumbent moves away from the center depends on the functional forms of $\ell$ and $F$.\textsuperscript{11} Nevertheless, we can prove that it never becomes so extreme as to improve the incumbent's probability of reelection. This is because $L$ always responds to an incumbent moving away from the center by fielding a candidate who will defeat her with a higher probability. Formally, let $x^*$ and $x^{**}$ denote challengers $L$ chooses against $x_I$ and $x'_I$ respectively, and let $x_I < x'_I$. Then, $F(\ell(x'_I) - \ell(-x^{**})) \geq F(\ell(x_I) - \ell(-x^*))$.

**Proposition 2.** When the opposition party can choose the ideology of the challenger from the entire policy space, against a more extreme incumbent they choose a challenger who has a higher probability of winning.

The intuition behind this proposition is straightforward; when the opposition party can choose the ideology of the challenger, their best response against a more extreme incumbent improves both the probability component $F(\ell(x_I) - \ell(-x))$, and the policy gain component, $\ell(x_I - x_i) - \ell(x - x_i)$, $i \in \{d, m_L\}$. This means that even if an increase in the incumbent's platform leads to a decrease in the challenger's platform, this decrease cannot be so large to lead to an overall lower probability of winning in the general election.

\textsuperscript{11}For example, when losses are linear, the optimal challenger becomes more extreme as the incumbent moves away from the center if and only if $f'(x_I + x^*) < 0$. 

13
Under open nominations, then, the incumbent does not have an incentive to provoke the opposition by pursuing policies more extreme than her ideal point. Next, I study a model where the party cannot freely choose the ideology of the challenger and instead must choose from an exogenously given set of party elites.

3.2 Party Elites

The previous section established that when the opposition party can choose the challenger’s ideology from the entire policy space, they respond to a more extreme incumbent by increasing the probability they win in the general election. Therefore, there is no incentive for the incumbent to provoke the opposition. Fielding candidates, however, is rarely an unconstrained optimization problem. When deciding who should run in the general election, a constraint parties face is the set of politicians that they have. There is much evidence showing that the set of politicians is a particular, and not necessarily representative, subset of the population who chose to become politicians (Mattozzi and Merlo, 2008; Dal Bó et al., 2017), and were not screened out by interest groups and party insiders (Cohen et al., 2009; Masket, 2011; Bawn et al., 2012). Furthermore, these politicians are themselves constrained in their policy platforms by their previous records. A more realistic model thus would have the opposition party choose from a set of candidates with exogenously given ideologies. This is the approach I take in this section.

Suppose that at the start of the game there is a pair of candidates, Extremist and Moderate, whose platforms are given exogenously. Let their platforms be $x_E < x_M \leq 0$. Here, primary voter $i$’s problem is to vote for the candidate that gives him a higher expected payoff. An implication of Proposition 1 is that the median primary voter is decisive in the primary between $E$ and $M$. The winner and therefore the challenger against the incumbent
is then \( E \) if and only if she provides a higher expected payoff to \( m_L \) than \( M \):

\[
F(\ell(x_I) - \ell(-x_E))\left(\ell(x_I - x_{m_L} - \ell(|x_{m_L} - x_E|)) \geq F(\ell(x_I) - \ell(-x_M))\left(\ell(x_I - x_{m_L}) - \ell(|x_{m_L} - x_M|)\right).
\]

In this setting, a more extreme incumbent can be reelected with a higher probability. This is in contrast to Proposition 2 which establishes that this is impossible when the opposition party can choose directly the ideology of the challenger. When primary voters must choose from a discrete set of exogenously given ideologies instead, there are discontinuities in how they can respond to changes in the incumbent’s platform. In other words, here the only tool the median primary voter has to respond to changes in the incumbent’s platform is picking one candidate with a given platform over another. This allows for configurations such that the median primary voter chooses a challenger who wins with a lower probability against a more extreme incumbent.

To analyze the opposition party primary when their choice is restricted to \( E \) and \( M \), let us briefly consider the case when the opposition party always fields the same challenger against all incumbents. If the median primary voter always prefers one candidate over another regardless of the incumbent’s platform, the incumbent’s probability of reelection is monotone decreasing in her platform. It follows that there cannot be an incentive to provoke to opposition by moving away from the center. Therefore, I restrict attention to the case where the median primary voter’s choice of \( E \) or \( M \) is responsive to the incumbent’s position. In other words, \( x_E \) and \( x_M \) are such that the median primary voter with ideal point \( x_{m_L} \) prefers \( E \) to become the challenger against some incumbents, and \( M \) against others. In particular, I assume that the median primary voter is ideologically closer to \( E \), but \( M \)’s probability of beating a very moderate incumbent is sufficiently higher:
**Assumption 1A.**

\[ x_{mL} < \frac{x_M + x_E}{2} \quad \text{and} \quad \frac{F(-\ell(x_M))}{F(-\ell(x_E))} > \frac{\ell(-x_{mL}) - \ell(|x_{mL} - x_E|)}{\ell(-x_{mL}) - \ell(x_{mL})}. \]

Specifically, under Assumption 1A very moderate incumbents induce the median primary voter to vote for \( M \), and very extreme incumbents induce him to vote for \( E \). Because \( f \) and \( \ell \) are both continuous in \( x_I \in \mathbb{R}_+ \), we know the median primary voter’s payoff must be continuous as well. Then, there must exist a platform for the incumbent such that the median primary voter is indifferent between \( E \) and \( M \). Let \( \tilde{x}_I \) denote this platform. To simplify the exposition, I assume this platform is unique.\(^{12}\) Assumption 2A is a sufficient condition for single-crossing of the median voter’s net payoff and is formally stated in the Appendix. Define \( x_I \) as the incumbent platform that leads to the incumbent’s reelection against the moderate challenger \( M \) with the same probability as \( \tilde{x}_I \) beats the extremist challenger \( E \), or \( F(\ell(x_I) - \ell(x_M)) = F(\ell(\tilde{x}_I) - \ell(x_E)) \), if such a platform exists. Otherwise let \( x_I := 0 \). Incumbents with platforms in the interval \((x_I, \tilde{x}_I)\) face the moderate opponent in the general election and are reelected with a lower probability than incumbents with the more extreme platform \( \tilde{x}_I \) who face the extremist.

**Lemma 3.** When the median primary voter’s ideal point satisfies Assumptions 1A and 2A, there exists an interval of incumbent platforms that result in her facing the moderate opponent and being reelected with a lower probability than if she chose the more extreme platform \( \tilde{x}_I \) and faced the extremist.

Lemma 3 finds that an incumbent with a more extreme platform can be reelected with a higher probability when the challenger is chosen from a discrete set. To see that this can indeed cause the incumbent to pursue platforms more extreme than her ideal point, observe

\(^{12}\) Allowing for multiplicity complicates the analysis but does not lead to additional substantively meaningful insights.
Figure 2: Incumbent’s probability of reelection is plotted as a function of her platform. \( \psi \) is drawn from a standard normal distribution, and losses are linear on the left plot and quadratic on the right. Parameter values are fixed at \( x_E = -2, x_M = -0.5 \), and \( x_{mL} = -4 \).

that her expected payoff is

\[
\max_{x_I \in \mathbb{R}^+} \text{EU}_I(x_I) = \max_{x_I \in \mathbb{R}^+} (1 - F(\ell(x_I) - \ell(-x_J)))(B - \ell(|x_I - t|)).
\] (3)

where \( J \) is the challenger chosen by \( L \) in equilibrium. Solving the incumbent’s problem reveals the conditions under which she provokes the opposition. Specifically, for incumbents whose ideal points lie in the interval \( t \in \left( \frac{x_I + \tilde{x}_I}{2}, \tilde{x}_I \right) \), there exist some office rents \( B \) such that the incumbent’s optimal platform induces the extremist \( E \) to win the opposition party primary, \( \tilde{x}_I > t \).

**Proposition 3.** When the incumbent’s ideal point \( t \) is more moderate than but sufficiently close to the threshold that induces the extremist opposition candidate to win the primary, there exists an interval of office rents \( B \) such that the incumbent provokes the opposition by choosing the platform \( \tilde{x}_I > t \) for her reelection bid in equilibrium.

Therefore, when \( L \) chooses the challenger from a set of party elites whose ideologies are given exogenously, there are parameter values such that the incumbent pursues policies more extreme than her ideal point to improve her reelection chance. She does this solely to weaken her appeal to the median voter. This emboldens the median primary voter in
the opposition party, and induces him to vote for the extremist $E$ in the primary election whose platform he prefers to that of the moderate $M$. $E$ gathers the votes of all voters to the left of the median primary voter, and beats $M$ to face the incumbent in the general election. However, $E$’s ideology is further from the general election median voter than $M$, so her winning the primary causes a boon to the incumbent’s reelection prospects, surpassing the harm caused by the incumbent’s move away from the center.

4 Endogenous Entry

The previous section supposed that opposition candidates always run in primary races. In this section, I study the strategic considerations of candidates with a model of costly entry. Candidates care about both getting elected and the policies implemented. Thus, they take into account their probabilities of winning the primary and the general election, as well as the effect their entry has on the equilibrium outcome. To lay bare the effect of the incumbent’s position on the candidates’ considerations alone, I assume the winner of the primary is decided by the flip of a (possibly biased) coin.

As before, there are two candidates whose platforms are $x_E < x_M \leq 0$. Candidates announce their running decisions sequentially. If only one candidate runs, she faces the incumbent in the general election. If both candidates run, $E$ wins the primary with probability $p$. If neither candidate runs, the incumbent is retained. The only decision candidates make is whether to run, and they run when they are indifferent between running and staying out. The cost of running for office is given by $c \geq 0$. Throughout I assume that this cost is low enough so that each candidate prefers to run when in equilibrium the other is not running.\(^1\) The game is otherwise identical to the one described in Section 3. The timing of the endogenous entry game is as follows:

\(^1\)Formally, I assume that $c \leq F(\ell(x_I) - \ell(-x_J))(B + \ell(x_I - x_J))$ for $J \in \{E, M\}$.  

18
1. Incumbent chooses her platform for the second period, \( x_I \geq 0 \).

2. \( E \) and \( M \) announce their running decisions according to a predetermined sequence.

3. Global utility shock against the incumbent is revealed.

4. If there is a challenger, the general election is held between her and the incumbent.

5. The winner of the general election implements her platform in the second period.

I again proceed by backward induction. The second period and general election play are identical to the previous model, so I skip to candidates’ entry decisions. Because it is simpler, I start with the limiting case of no cost of running, \( c = 0 \) before analyzing positive costs.

In deciding whether to enter the primary race when the other candidate is running, each candidate evaluates the benefits of running —choosing the policy and obtaining office rents if they win— and the impact of their entry on both the primary and the general elections. In particular, candidates must consider the effect their entry has on the probability that the incumbent is reelected. If by entering the race a candidate increases the probability that the incumbent is reelected, this may induce them to stay out. This is never the case for the moderate candidate. Because the moderate has a higher probability of beating the incumbent than the extremist, \( M \) can always decrease the probability that the incumbent is retained by entering the primary race. It follows that \( M \) always enters, even when \( E \) is running. Formally, this is because when \( E \) is running, \( M \)'s net expected payoff of entering the race versus staying out, \((1 - p)\Delta_M(x_I)\) is positive, where:

\[
\Delta_M(x_I) = F(\ell(x_I) - \ell(-x_M))(B + \ell(x_I - x_M)) - F(\ell(x_I) - \ell(-x_E))(\ell(x_I - x_M) - \ell(x_M - x_E)).
\]

The first term in this expression is the direct effect: by entering, \( M \) increases the probability he wins. The second term is the indirect effect \( M \)'s entry has on the decreased probability \( E \) faces the incumbent in the general election. Because the probability \( M \) beats the incumbent
is higher, the first term dominates the second, and $\Delta_M(x_I)$ must be positive.

The same does not hold for $E$; when the moderate is running, extremist’s entry increases the probability the incumbent is reelected. Thus, she may prefer to stay out of the race if $M$ has a sufficiently better chance of beating the incumbent in the general election. In other words, similar to primary voters who vote for the moderate candidate despite liking the policies of the extremist more; $E$ can concede on her policy goals and office rent to help her party win the general election by letting the more electable $M$ face the incumbent. Formally, when $M$ is running, the net expected utility of $E$ of running versus staying out is $p\Delta_E(x_I)$, where the latter term is:

$$\Delta_E(x_I) = F(\ell(x_I) - \ell(-x_E))(B + \ell(x_I - x_E)) - F(\ell(x_I) - \ell(-x_M))(\ell(x_I - x_E) - \ell(x_M - x_E)).$$

As argued above, this expression can be positive or negative. Notice that it must be positive for sufficiently high values of $x_I$: as $F(\ell(x_I) - \ell(-x_M)) - F(\ell(x_I) - \ell(-x_E))$ goes to zero, entering becomes strictly preferred for $E$. Formally, $\lim_{x_I \to \infty} \Delta_E(x_I) > 0$. For $E$ to prefer to stay out when the incumbent is very moderate, $\Delta_E(x_I)$ must be negative against some $x_I$. I make the following modification to Assumption 1A which ensures $M$’s probability of beating a moderate incumbent is sufficiently higher than $E$:

**Assumption 1B.**

$$\frac{F(-\ell(-x_M))}{F(-\ell(-x_E))} > \frac{\ell(-x_E) + B}{\ell(-x_E) - \ell(x_M - x_E)}.$$  

Under Assumption 1B, we have $\Delta_E(0) < 0$. Then, there exists an incumbent platform that leaves the extremist indifferent between entering and staying out. Let $\tilde{x}'_I$ denote this platform.\(^{14}\) Also define $x'_I(p)$ as the platform that leads to the same probability of incumbent’s reelection as $\tilde{x}'_I$, if such a point exists. If such a point does not exist, let $x'_I(p) = 0$. I

\(^{14}\)As before, I assume for simplicity that this point is unique. The formal statement of Assumption 2B which ensures single-crossing of $\Delta_E$ can be found in the Appendix.
can then prove that despite being more moderate, incumbents with platforms in the interval $(\bar{x}'_I(p), \tilde{x}'_I)$ are reelected with a lower probability than incumbents with the platform $\tilde{x}'_I$. Notice that this result is very similar to Lemma 3 in Section 3.

**Lemma 4.** When the ideal point of $E$ satisfies Assumptions 1B and 2B, there exists an interval of incumbent platforms that result in the incumbent facing the moderate opponent with probability one, and being reelected with a lower probability than the more extreme platform $\tilde{x}'_I$ which results in the incumbent facing the extremist opponent with probability $p$.\[^{15}\]

Lemma 4 extends to positive costs of running for intermediate values of $p$. When $c > 0$ is small and $p$ close to 1/2, there exists an incumbent platform that induces the extremist opponent’s entry to the race and hence leads to a higher probability of reelection than platforms slightly more moderate than it. Thus, assuming that candidates face a small but positive cost to enter the race does not give us substantively different insights for intermediate values of $p$.

Next, consider a primary field that is slanted in favor of moderates: suppose that $p$ is close to zero, meaning that the moderate is very likely to win a competitive primary. In

\[^{15}\text{Because } M \text{ plays a dominant strategy in equilibrium, this is true regardless of the sequence of announcements.}\]
Figure 4: Orange (SW) and blue (NE) regions respectively correspond to parameter values where only $M$ and $E$ run in equilibrium. In green (E) both candidates run. $\psi$ is drawn from a standard normal distribution, and losses are linear on the left plot and quadratic on the right. Parameter values are fixed at $x_E = -1.2$, $x_M = -0.5$, $c = 0.5$, and $B = 2$. In both plots, $M$ announces first. Plots for equilibria when $E$ announces first are presented in Appendix C.

In this case too, the moderate always runs. The extremist, in contrast, prefers to stay out of the race, even when the incumbent’s platform is extreme. This is because even if $E$ knew she would likely beat the incumbent in the general election, it is unlikely she can get there, so she decides to stay out of the race. This is true regardless of the sequence of candidates’ announcements. It follows that when the primary field is slanted towards moderates, $M$ can drive $E$ out of running, and the general election is held between the moderate challenger and the incumbent, regardless of the platform of the latter.

Finally, suppose there is an extremist advantage in the primary, or $p$ close to one. Here, the moderate only enters the race if the extremist does not. This is because the probability $M$ makes it through to the general election from a competitive primary field is very low, despite having a higher chance of beating the incumbent if he did. Thus, it is possible for the extremist to be the only candidate in equilibrium. If the incumbent is sufficiently extreme, a strong primary advantage induces the extremist to enter the race even if the moderate were running, driving the moderate out. In contrast, when $M$ has a much better chance of beating
the incumbent, $E$ prefers to stay out and let $M$ face the incumbent in the general election. For intermediate incumbent platforms, who runs in equilibrium depends on the order in which candidates announce their entry decisions. If $M$ announces first, he runs if and only if $E$ prefers to stay out when he is running, $p\Delta_E(x_I) < c$. If $E$ announces first instead, she runs whenever she prefers to face the incumbent herself, $\Delta_E(x_I) \geq c$, knowing that $M$ stays out if she enters. It follows that when there is a primary advantage for extremists, for either sequence of announcements, there exist thresholds such that only the moderate challenger runs against incumbents whose platforms are more moderate than this threshold, and only the extremist challenger runs against those more extreme. These are visualized in Figure 4 and summarized in the following Proposition:

**Proposition 4.** Suppose the cost of running is small but strictly positive. Then,

1. When neither the extremist nor the moderate has an advantage in the primary ($p$ close to $1/2$), there exists a platform for the incumbent that leads to her facing $E$ in the general election with a positive probability, and thus being reelected with a higher probability than if she had a more moderate platform and faced $M$ with probability one;

2. When there is a primary advantage for moderates ($p$ close to $0$), the incumbent faces $M$ in the general election regardless of her platform;

3. When there is a primary advantage for extremists ($p$ close to $1$), for either sequence of announcements, there exists a platform for the incumbent that leads to her facing $E$ in the general election, and thus being reelected with a higher probability than if she had a more moderate platform and faced $M$.

Proposition 4 finds that when running is costly, unless the primary election is slanted in favor of moderates, an incumbent with a more extreme platform can be reelected with a higher probability. In particular, if the primary election is slanted in favor of extremists, there cannot be a competitive primary: when $E$ runs, $M$ does not want to pay the cost of
running, knowing he has a very low probability of making it through to the general election. If the incumbent is sufficiently extreme, $E$ would rather be the challenger herself to implement her ideal point and obtain office rents with some probability, even though this probability is still lower than the probability $M$ would win in the general election. This drives $M$ out, and the equilibrium challenger is $E$. If the incumbent is sufficiently moderate, $E$ recognizes that she has a low chance of beating her, and therefore stays out of the race, and the equilibrium challenger is $M$.

When instead the primary election is balanced and both sides have roughly equal chances of winning a competitive primary, $M$ always runs, regardless of the incumbent’s platform and whether $E$ is also running or not. This is in part because his entry always decreases the probability the incumbent is reelected, and in part because his probability of winning both the primary and the general elections is high enough to cover the cost of entry. In contrast, against a sufficiently moderate incumbent, $E$ stays out. This is in part because the probability of winning both the primary and the general election is too low for $E$ to justify paying the cost of running, and in part because $E$ gets a sufficiently large disutility from the incumbent’s reelection that she prefers $M$ to be the challenger and beat the incumbent with a higher probability. When instead the incumbent’s platform is sufficiently extreme and $E$ can plausibly beat her in the general election, there is a competitive primary. Here, with probability of about one half, the incumbent faces $M$, and otherwise faces $E$ in the general election. Thus, when there is a primary advantage for extremists or no advantage for either side, a more extreme incumbent induces $E$’s entry, which increases the overall probability the incumbent is reelected.\footnote{Evidence presented in Hall and Snyder Jr (2015) suggests that these two cases are more relevant than a primary advantage for moderates; they find that extremist candidates tend to have an advantage in primary elections as measured by vote share and probability of winning.}

It follows that an incumbent with an ideal point more moderate than the platform that induces the extremist’s entry may thus find it preferable to pursue that platform instead of
Incumbent’s probability of reelection is plotted as a function of her platform. $\psi$ is drawn from a standard normal distribution, and losses are linear on the left plots and quadratic on the right. Parameter values are fixed at $x_E = -1.2$, $x_M = -0.5$, $B = 2$, and $c = 0.5$. There is a primary advantage for moderates in the top plots ($p = 0.05$), and for extremists in the bottom ($p = 0.95$).

her ideal point. Despite hurting her policy-wise and electorally against any given opponent, going more extreme increases the probability she faces a weaker challenger in the general election. Specifically, by choosing the threshold platform, the incumbent can increase the probability she faces $E$ in the general election from zero to one if there is a primary advantage for extremists, and to $p$ if there is no primary advantage for either extremists or moderates. For an incumbent with an ideal point sufficiently close to this platform, this leads to a strictly higher expected payoff for appropriate levels of office rents. This result is presented below, which is an analog of Proposition 3:

**Proposition 5.** Given $x_M$, $x_E$, and $c$, there exist intervals of office rents $B$ and incumbent ideal points $t$ such that when there is a primary advantage for extremists or when there is no primary advantage for either extremists or moderates, the incumbent provokes the opposition
by choosing platforms more extreme than her ideal point. When there is a primary advantage for moderates, provoking the opposition cannot occur.

Thus, like the model of primary elections, provoking the opposition is possible in a model of costly entry. Specifically, moderate incumbents can benefit from hurting themselves electorally by decreasing their appeal. This increases the probability the extremist opposition party candidate wins in the general election, and hence induces her to run in her party’s primary. If she emerges as the challenger to face the incumbent in the general election, the incumbent is reelected with a higher probability. In particular, for incumbents with ideal points close to the threshold that induces extremist’s entry, provoking the opposition leads to a large enough boost to make up for the decreased appeal caused by going more extreme.\textsuperscript{17}

5 CONCLUSION

This paper studies a model where an incumbent politician may find it optimal to pursue extreme policies to improve her reelection prospects. Such policies provoke extremist candidates in the opposition party to run for office, and extremist opposition primary voters to support them, potentially shifting the ideology of the challenger. This leads to a weaker challenger replacing a strong one, and thus improves the incumbent’s chance of reelection. This result holds both under a primary elections model and a candidate entry model, but only when the opposition party is constrained to choose from a discrete set of candidates. If parties are unconstrained in their choice, this mechanism cannot work.

The model is based on the idea that an incumbent with a more extreme platform makes it both more likely and more important to defeat her. The first follows from the fact that an

\textsuperscript{17}A natural extension is combining the two models presented in this paper to see how the two forces that lead to provoking the opposition interact. In Appendix B, I present simulations that demonstrate my results are robust to when both primary elections and entry decisions by candidates are endogenous and show how the effects on primary voters and candidates can complement each other.
extreme platform is further from the ideal point of the median voter, which increases the probability that any given challenger can beat the incumbent. This emboldens the extremist factions within the opposition party who see an opportunity to pursue their agenda. On the other hand, an incumbent with an extreme platform also increases the payoff gain of defeating her. This is because the policy that would be implemented if the incumbent were reelected is disliked more by the members of the opposition. This pushes the opposition towards moderation to increase their appeal to the median voter, and therefore their probability of beating the incumbent. Whether the mechanism identified in this paper occurs depends on how these two forces play out, and the set of candidates parties have access to. I show that when the opposition party can choose the ideology of their candidate, they always improve the probability of winning against a more extreme incumbent. However, when candidate ideologies are given exogenously, it is possible for an extremist incumbent to be reelected with a higher probability than a moderate one. This highlights a novel implication of parties’ gatekeeping of candidates: the inability of the opposition party to choose challengers from a rich enough set of candidates enables incumbents to provoke the opposition.

Primaries selecting extremists as a driver of polarization has been argued and tested before. This paper identifies another mechanism by which primaries may result in polarization. It demonstrates how a strategic incumbent may move away from the center to improve her reelection prospects by provoking the opposition to select extremists. Thus, primaries may also drive polarization by leading politicians to pursue policies more extreme than their ideal points after they are elected. Here, centrist candidates lose primaries or decide not to run at all, and elected moderates are pressured to move away from the center. This decreases the welfare of moderate voters, who must choose between extremists on both sides in the general election.
REFERENCES


Bawn, Kathleen, Martin Cohen, David Karol, Seth Masket, Hans Noel, and John Zaller, “A Theory of Political Parties: Groups, Policy Demands and Nominations in Amer-

Brady, David W., Hahrie Han, and Jeremy C. Pope, “Primary Elections and Candidate Ideology: Out of Step with the Primary Electorate?,” *Legislative Studies Quarterly*, 2007, 32 (1), 79–105.


Hill, Seth J. and Gregory A. Huber, “Representativeness and Motivations of the Con-


Owen, Guillermo and Bernard Grofman, “Two-stage electoral competition in two-party


APPENDIX A  PROOFS

Proof of Lemma 1. Recall Equation 1

\[ EU_d(x, x_I) = -\ell(|x_d - x|)F(\ell(x_I) - \ell(|x|)) - \ell(x_I - x_d)(1 - F(\ell(x_I) - \ell(|x|))). \]

The derivative of this is

\[ \frac{d EU_d}{dx} = \text{sgn}(x_d - x)\ell'(|x_d - x|)F(\ell(x_I) - \ell(|x|)) - \text{sgn}(x)\ell'(|x|)f(\ell(x_I) - \ell(|x|))(\ell(x_I - x_d) - \ell(|x_d - x|)). \]

The challenger cannot be someone whose platform the insiders like less than that of the incumbent, because a candidate whose platform is the same as \( d \)'s ideal point always yields a higher expected payoff than such a candidate. Restricting attention to platforms \( d \) prefers to that of the incumbent: \(|x_d - x_I| > |x_d - x|\), notice that \( d \)'s payoff is increasing in \( x \) for \( x < x_d \) because \( \frac{d EU_d}{dx} > 0 \) in that region. This implies that \( EU_d(x_d, x_I) > EU_d(x, x_I) \) for all \( x < x_d \). Also notice that \( d \)'s payoff is decreasing in \( x \) for \( x > 0 \) because \( \frac{d EU_d}{dx} < 0 \) for that region, which implies \( EU_d(0, x_I) > EU_d(x, x_I) \) for all \( x > 0 \). Thus, as argued in the text we can restrict attention to \( x \in [x_d, 0] \).

Next, notice that the first order condition of the insiders’ problem is:

\[ \frac{F(\ell(x_I) - \ell(-x^*))}{f(\ell(x_I) - \ell(-x^*))} \ell'(x^* - x_d) = \ell'(-x^*)(\ell(x_I - x_d) - \ell(x^* - x_d)). \]  \hspace{1cm} (4)

Both sides of Equation 4 are positive for \( x \in [x_d, 0] \). Log-concavity of \( f \) implies that the left-hand side is increasing in \( x \). The right-hand side is decreasing in \( x \) because \( \ell \) is increasing and convex. Thus, there can be at most one solution to Equation 4 in \( x_d \leq x \leq 0 \).
Notice also

$$\frac{d^2 \text{EU}_d}{dx^2} = -\ell''(x - x_d)F(\ell(x_I) - \ell(x)) - 2\ell'(x - x_d)\ell'(-x)f(\ell(x_I) - \ell(-x)) +$$

$$(\ell(x_I - x_d) - \ell(x - x_d))(f'(\ell(x_I) - \ell(-x))\ell'(-x)^2 - f(\ell(x_I) - \ell(-x))\ell''(-x)).$$

Evaluating this expression at the solution of Equation 4 yields:

$$-\ell''(x^* - x_d)F(\ell(x_I) - \ell(x^*)) - F(\ell(x_I) - \ell(-x^*))\ell'(x^* - x_d)\ell''(-x^*)$$

$$- \ell'(x^* - x_d)\ell'(-x^*)^2 \left(2f(\ell(x_I) - \ell(-x^*)) - F(\ell(x_I) - \ell(-x^*)) \frac{f'(\ell(x_I) - \ell(-x^*))}{f(\ell(x_I) - \ell(-x^*))} \right) < 0.$$ 

Therefore, if $x$ solves Equation 4, it is a maximum. \hfill \square

**Proof of Proposition 1.** Start by noting that applying Lemma 1 to the median primary voter implies that his optimal candidate must have a platform between his ideal point and the median general election voter: $x^*_{mL} \in [x_{mL}, 0]$. To prove that the median primary voter is pivotal, we need to show that every primary voter to his left (right) prefers a candidate with ideology $x^*_{mL}$ to any candidate with a more moderate (extreme) ideology. Formally, a sufficient condition for median primary voter’s pivotality is that we have $\text{EU}_I(x^*_{mL}, x_I) \geq \text{EU}_I(x, x_I)$ for both all $x_i \leq x_{mL}$ and $x \geq x^*_{mL}$, and all $x_i \geq x_{mL}$ and $x \leq x^*_{mL}$. Notice that these imply

$$\frac{\ell(x_I - x_i) - \ell(|x^*_{mL} - x_i|)}{\ell(x_I - x_i) - \ell(|x - x_i|)} \geq \frac{F(\ell(x_I) - \ell(|x|))}{F(\ell(x_I) - \ell(-x^*_{mL}))},$$

for all $x_i \leq x_{mL}$, $x \geq x^*_{mL}$ as well as all $x_i \geq x_{mL}$, $x \leq x^*_{mL}$. By definition of $x^*_{mL}$, we know that for all $x$:

$$\frac{\ell(x_I - x_{mL}) - \ell(x^*_{mL} - x_{mL})}{\ell(x_I - x_{mL}) - \ell(|x - x_{mL}|)} \geq \frac{F(\ell(x_I) - \ell(|x|))}{F(\ell(x_I) - \ell(-x^*_{mL}))}.$$
So a sufficient condition is that for \((x_i - x_{m_L})(x - x_{m_L}^*) \leq 0\) we have

\[
\frac{\ell(x - x_i) - \ell(|x_{m_L} - x_i|)}{\ell(x - x_i) - \ell(|x - x_i|)} \geq \frac{\ell(x - x_{m_L}) - \ell(x_{m_L}^* - x_{m_L})}{\ell(x - x_{m_L}) - \ell(|x - x_{m_L}|)}.
\]

To show that log-concavity of \(\ell'\) satisfies this, we first need to prove the following lemma:

**Lemma 5.** Suppose that \(\ell'\) is log-concave. Then, for any \(x_2 > x_1 > x_0 > x_i\):

\[
\frac{d}{dx_i} \frac{\ell(x_2 - x_i) - \ell(x_0 - x_i)}{\ell(x_2 - x_i) - \ell(x_1 - x_i)} \leq 0.
\]

**Proof of Lemma 5.** Suppose that \(\ell'\) is log-concave. By definition of log-concavity, this means that for all \(x\), we have \(\ell'(x)\ell''(x) \leq (\ell''(x))^2\), where \(\ell', \ell'', \text{ and } \ell'''\) refer to first, second, and third derivatives of \(\ell\) respectively. This in turn implies that the cross-partial of logarithm of \(\ell'\) with respect to any \(x\) and \(x_i\) is positive because

\[
\frac{\partial^2 \ln \ell'(|x - x_i|)}{\partial x \partial x_i} = \frac{\partial \text{sgn}(x - x_i) \ell''(|x - x_i|)}{\partial x_i} = -\ell'''(|x - x_i|)\ell'(|x - x_i|) + (\ell''(|x - x_i|))^2 \geq 0,
\]

where the last inequality follows from log-concavity of \(\ell'\). This in turn implies that for any \(x_1 > x_0\):

\[
\frac{\partial \ln(\ell'(x_1 - x_i))}{\partial x_i} - \frac{\partial \ln(\ell'(x_0 - x_i))}{\partial x_i} \geq 0 \iff \frac{\partial \ln(\ell'(x_1 - x_i))}{\partial x_i} \geq 0 \iff \frac{\partial \ell'(x_1 - x_i)}{\partial x_i} \geq 0.
\]

Now let \(x_j < x_i\), and define \(s(x) = \frac{\ell'(x - x_j)}{\ell'(x - x_i)}\) for \(x > x_i\). The previous condition implies that \(\frac{\partial s(x)}{\partial x} \leq 0\). Let \(x_2 > x_1 > x_0\), and notice we can write

\[
\frac{\ell(x_2 - x_i) - \ell(x_1 - x_i)}{\ell(x_1 - x_i) - \ell(x_0 - x_i)} \geq \frac{\int_{x_1}^{x_2} s(x) \ell'(x - x_i) dx}{\int_{x_0}^{x_1} s(x) \ell'(x - x_i) dx} = \frac{\int_{x_1}^{x_2} \ell'(x - x_i) dx}{\int_{x_0}^{x_1} \ell'(x - x_i) dx} = \frac{\ell(x_2 - x_i) - \ell(x_1 - x_i)}{\ell(x_1 - x_i) - \ell(x_0 - x_j)},
\]

where the inequality follows from the fact that \(s(x)\) is lower everywhere it's evaluated in the
numerator than everywhere in the denominator. Thus
\[
\frac{\partial \ell(x_2-x_i) - \ell(x_1-x_i)}{\partial x_i} \geq 0 \iff \frac{\partial \ell(x_2-x_i) - \ell(x_1-x_i)}{\partial x_i} \leq 0.
\]

Finally, add 1 to the above expression. Because this is a constant, the derivative does not change, and we get
\[
\frac{\partial \ell(x_1-x_i) - \ell(x_0-x_i)}{\partial x_i} = \frac{\partial \ell(x_2-x_i) - \ell(x_0-x_i)}{\partial x_i} \geq 0.
\]

We can now use Lemma 5 to show that for \( x \geq x_{m_L}^* \), we have
\[
\frac{d}{dx_i} \frac{\ell(x_1-x_i) - \ell(x_{m_L}^*-x_i)}{\ell(x_1-x_i) - \ell(x-x_i)} \leq 0.
\]

Similarly, for \( x \leq x_{m_L}^* \) we have
\[
\frac{d}{dx_i} \frac{\ell(x_1-x_i) - \ell(x_{m_L}^*-x_i)}{\ell(x_1-x_i) - \ell(x-x_i)} \geq 0.
\]

It follows that
\[
\frac{\ell(x_1-x_i) - \ell(|x_{m_L}^*-x_i|)}{\ell(x_1-x_i) - \ell(x-x_i)} \geq \frac{\ell(x_1-x_{m_L}) - \ell(x_{m_L}^*-x_{m_L})}{\ell(x_1-x_{m_L}) - \ell(|x-x_{m_L}|)}
\]
for both \( x_i \leq x_{m_L}, x \geq x_{m_L}^* \) and \( x_i \geq x_{m_L}, x \leq x_{m_L}^* \). This means that if the median primary voter prefers a more moderate candidate to a more extreme one, all primary voters to his right also prefer that more moderate candidate. Similarly, if the median primary voter prefers a more extreme candidate to a more moderate one, all primary voters to his left also prefer that more extreme candidate. It must then be that the median primary voter's optimal candidate is the Condorcet winner.

\[\square\]

**Proof of Lemma 2.** Replacing \( \ell \) with either absolute or exponential loss in Equation 4
gives

\[ \frac{F(\ell(x_I) - \ell(-x^*))}{f(\ell(x_I) - \ell(-x^*))} = \ell'(-2x^*)(\ell(x_I) - \ell(x^*)). \]  \tag{5} \]

This must have a solution with \( x < 0 \) by log-concavity of \( f \), the fact that \( \ell \) is minimized at 0, and a simple application of the intermediate value theorem. The uniqueness follows from the facts that the left-hand side is strictly increasing because of the log-concavity of \( f \), and the right-hand side is strictly decreasing in \( x^* \). Notice that the \( x^* \) in Equation 5 does not depend on \( x_i \).

Next, recall from Proposition 1 that the optimal candidate of the median \( L \) voter is the Condorcet winner. If \( x^* \) is to the right of the median \( L \) voter, then she is \( m_L \)'s optimal candidate and therefore the Condorcet winner. If \( x^* \) is to the left of the median \( L \) voter, \( m_L \)'s optimal candidate must have the same ideology as him. This is because we know \( \frac{dEU_{m_L}}{dx} < 0 \) for \( x \geq x_{m_L} \). It follows that the Condorcet winner must be the either \( x_{m_L} \) or \( x^* \), whichever is greater.

\[ \square \]

**Proof of Proposition 2.** Let \( x'_I > x_I \) be the platforms of two incumbents, and \( x^{**} \) and \( x^* \) challengers chosen by \( L \) against \( x'_I \) and \( x_I \) respectively. Recall from Lemma 1 that the optimal candidate must be in \([x_d, 0]\). Notice first that if either \( x^* = x_d \) or \( x^{**} = 0 \), it must be that \( x^* \leq x^{**} \), and the proposition follows immediately. Thus we only need to prove the proposition for \( x^* \in (x_d, 0] \) and \( x^{**} \in [x_d, 0) \). These require:

\[ - \ell'(x^*-x_d)F(\ell(x_I) - \ell(-x^*)) + \ell'(-x^*)f(\ell(x_I) - \ell(-x^*)) (\ell(x_I - x_d) - \ell(x^* - x_d)) \geq 0 \] and

\[ - \ell'(x^{**}-x_d)F(\ell(x'_I) - \ell(-x^{**})) + \ell'(-x^{**})f(\ell(x'_I) - \ell(-x^{**})) (\ell(x'_I - x_d) - \ell(x^{**} - x_d)) \leq 0. \]

These must hold with equality if the challenger is in the interior \((x_d, 0)\), and with strict inequality if she is on a corner. I need to show \( F(\ell(x'_I) - \ell(-x^{**})) \geq F(\ell(x_I) - \ell(-x^*)) \). Suppose
for a contradiction that

\[ F(\ell(x_I^*) - \ell(-x^{**})) < F(\ell(x_I) - \ell(-x^*)). \]

By log-concavity of \( f \), it follows that

\[ \ell'(x^{**} - x_d)\ell'(-x^*)(\ell(x_I - x_d) - \ell(x^* - x_d)) > \ell'(x^* - x_d)\ell'(x^{**} - x_d)(\ell(x_I^* - x_d) - \ell(x^{**} - x_d)). \]

This requires either \( \ell(x^{**} - x_d) > \ell(x^* - x_d) \), or \( \ell'(x^{**} - x_d)\ell'(-x^*) > \ell'(x^* - x_d)\ell'(-x^{**}) \). Both of these imply \( x^{**} > x^* \); former because \( \ell \) is increasing, and latter because \( \ell \) is convex. But, \( x^{**} > x^* \) leads to a contradiction with the premises \( x'_I > x_I \) and \( F(\ell(x'_I) - \ell(-x^{**})) < F(\ell(x_I) - \ell(-x^*)) \). Thus, it must be that \( F(\ell(x'_I) - \ell(-x^{**})) \geq F(\ell(x_I) - \ell(-x^*)) \).

Proof of Lemma 3. Let us start by restating the sufficient conditions for an incumbent platform \( \tilde{x}_I \) that leaves the median primary voter indifferent to exist. By the differentiability of the loss function and log-concavity of \( f \), we know that the median primary voter’s payoff must be continuous in the incumbent’s platform. Thus, if there is an incumbent against which \( m_L \) prefers \( E \) over \( M \), and another who induces a preference for \( M \) over \( E \), there must then exist \( \tilde{x}_I \) such that he is indifferent between the two. To recover the conditions under which the above premise holds, notice we can write the median primary voter’s net expected utility of \( E \) over \( M \) as

\[
\left(\ell(x_I - x_{m_L}) - \ell(|x_{m_L} - x_E|)\right) F(\ell(x_I) - \ell(-x_E)) - \left(\ell(x_I - x_{m_L}) - \ell(x_M - x_{m_L})\right) F(\ell(x_I) - \ell(-x_M)) \]

\[ \geq 0. \tag{6} \]

Notice that for all \( x_I, x_E, \) and \( x_M, F(\ell(x_I) - \ell(-x_M)) \geq F(\ell(x_I) - \ell(-x_E)). \) Moreover, as \( x_I \to \infty \), Chebyshev’s Inequality tells us that the \( \ell(x_I - x_{m_L})F(\ell(x_I) - \ell(-x_M)) - F(\ell(x_I) - \ell(-x_E)) \) term goes to zero, meaning that \( \lim_{x_I \to \infty} EU_{m_L}(x_E, x_I) - EU_{m_L}(x_M, x_I) = \ell(x_M - x_{m_L}) - \ell(|x_{m_L} - x_E|). \)

If this term is negative, this means that the median primary voter always prefers \( M \) to \( E \),
and $E$ never wins the primary election. The first part of Assumption 1A in the main text ensures that against sufficiently extreme incumbents the median primary voter votes for $E$.

To rule out the other case where the median primary voter always votes for $E$, notice that when the incumbent’s platform coincides with the ideal point of the median general election voter $0$, expression 6 becomes

$$
\left(\ell(-x_{mL}) - \ell(|x_{mL} - x_E|)\right)F(-\ell(-x_M)) - \left(\ell(-x_{mL}) - \ell(x_M - x_{mL})\right)F(-\ell(-x_M)).
$$

It follows that against a very moderate incumbent the median primary voter votes for $M$ whenever the second part of Assumption 1A holds. Therefore when both parts of Assumption 1A hold and so the median primary voter votes for $M$ against some incumbents and for $E$ against others, it follows by continuity that there must exist at least one incumbent platform $\tilde{x}_I$ that leaves him indifferent.

Next, I assume a condition that ensures there cannot be multiple such platforms. This is not a critical assumption, and most arguments made below apply to the case when there are multiple incumbent platforms that leave $m_L$ indifferent between $E$ and $M$. Uniqueness of $\tilde{x}_I$ however greatly simplifies exposition, and so I assume the following condition that ensures single-crossing.

**Assumption 2A.**

$$(\ell(x_I - x_{mL}) - \ell(x_I - x_M))\frac{\ell'(x_I)}{\ell'(x_I - x_{mL})} > \frac{(F(x_I, x_M) - F(x_I, x_E))F(x_I, x_E)}{F(x_I, x_M)f(x_I, x_E) - F(x_I, x_E)f(x_I, x_M)},$$

where $f(x_I, x_J) := f(\ell(x_I) - \ell(-x_J))$ and $F(x_I, x_J) := F(\ell(x_I) - \ell(-x_J))$.

This condition states that the derivative of the expected payoff of the median $L$ voter with respect to $x_I$ must be positive when evaluated in the region where he prefers $E$ to $M$. In other words, as long as $m_L$ prefers $M$, his net payoff from electing $E$ may increase or
decrease as the incumbent becomes more extreme. But once \( m_L \) has a weak preference for \( E \), he never goes back to preferring \( M \) as the incumbent goes even more extreme.

When Assumptions 1A and 2A hold, there is a unique incumbent platform \( \tilde{x}_I \) that leaves the median primary voter indifferent between \( E \) and \( M \). This platform satisfies:

\[
F(\ell(\tilde{x}_I) - \ell(-x_E))\left(\ell(\tilde{x}_I - x_{m_L}) - \ell(|x_{m_L} - x_E|)\right) = F(\ell(\tilde{x}_I) - \ell(-x_M))\left(\ell(\tilde{x}_I - x_{m_L}) - \ell(x_M - x_{m_L})\right).
\]

Because indifferent voters vote for the more extreme candidate, when the incumbent’s platform is \( \tilde{x}_I \) the median primary voter votes for \( E \). We know from the proof of Proposition 1 that if the median primary voter votes for \( E \) (\( M \)), then so must all primary voters to his left (right). It follows then when the incumbent’s platform is \( x_I < \tilde{x}_I \), the primary winner is \( M \); and otherwise it is \( E \).

When primary voters choose \( E \) as the challenger to face off against the incumbent in the general election, the incumbent’s probability of reelection is \( 1 - F(\ell(x_I) - \ell(-x_E)) \). Because \( x_E < x_M \), we know that for all \( x_I \), \( F(\ell(x_I) - \ell(-x_E)) < F(\ell(x_I) - \ell(-x_M)) \). Furthermore, continuity of \( f \) ensures the existence of an interval of platforms where the incumbent is reelected with a lower probability than \( \tilde{x}_I \) because she faces \( M \) instead of \( E \). The upper bound of this interval is \( \tilde{x}_I \), exclusive, because the probability of reelection is monotonic and continuous in the incumbent’s platform holding fixed the identity of the opponent. To prove that the lower bound of this interval is \( x_I \), formally define it as \( x_I := \ell^{-1}(\max(0, \ell(\tilde{x}_I) + \ell(-x_M) - \ell(-x_E))) \), and notice that for any platform in the interval \( x_I \in (x_I, \tilde{x}_I) \), the challenger is \( M \). The net expected utility of the median general election voter from the moderate challenger over an incumbent in the interval \((x_I, \tilde{x}_I)\) is greater than his net expected utility from the extremist challenger over \( I' \). This is because the former is strictly greater than

\[
\max(0, \ell(\tilde{x}_I) + \ell(-x_M) - \ell(-x_E)) - \ell(-x_M) \geq \ell(\tilde{x}_I) - \ell(-x_E),
\]
whereas the latter is exactly equal to \( \ell(\tilde{x}_I) - \ell(-x_E) \). It follows that \( I' \) is reelected with a strictly higher probability than an incumbent with a platform \( x_I \in (x_I, \tilde{x}_I) \).

**Proof of Proposition 3.** Consider an incumbent whose ideal point is \( t \in \left( x_I + \frac{\tilde{x}_I}{2}, \tilde{x}_I \right) \). Taking the derivative of incumbent’s payoff in Equation 3 yields

\[
- \ell'(x_I)f(\ell(x_I) - \ell(-x_J))(B - \ell(|x_I - t|)) - \text{sgn}(x_I - t)\ell'(|x_I - t|)(1 - F(\ell(x_I) - \ell(-x_J))),
\]

such that \( J = S \) if and only if \( x_I \geq \tilde{x}_I \). To eliminate potential regions, and therefore narrow the set of possible solutions, let us study this derivative separately in the following three regions: \( x_I \leq t, x_I \in (t, \tilde{x}_I), \) and \( x_I \geq \tilde{x}_I \).

1. First, let us explore \( x_I \in (t, \tilde{x}_I) \). Notice that for this region the expression in 7 becomes

\[
- \ell'(x_I)f(\ell(x_I) - \ell(-x_M))(B - \ell(x_I - t)) - \ell'(x_I - t)(1 - F(\ell(x_I) - \ell(-x_M))),
\]

which is negative, meaning that the optimal platform cannot be in this region.

2. Next, for \( x_I \geq \tilde{x}_I \), this is also negative:

\[
- \ell'(x_I)f(\ell(x_I) - \ell(-x_E))(B - \ell(x_I - t)) - \ell'(x_I - t)(1 - F(\ell(x_I) - \ell(-x_E))).
\]

This means that \( t = \tilde{x}_I \) is strictly preferred to every platform strictly greater than it, and the optimal platform cannot be strictly greater than \( \tilde{x}_I \).

3. Finally, for \( x_I \leq t \), we can rewrite the expression in 7 as

\[
- \ell'(x_I)f(\ell(x_I) - \ell(-x_M))(B - \ell(t - x_I)) + \ell'(t - x_I)(1 - F(\ell(x_I) - \ell(-x_M))).
\]

For sufficiently low \( B \), the above expression may be positive for all \( x_I \leq t \), which means
that the incumbent’s optimal platform in this region is her own ideal point. The intuition is that when office rents are low and the incumbent is very likely to be reelected already at her ideal point, she does not find it worthwhile to moderate her platform to improve her reelection chance. In contrast, for sufficiently high $B$, the expression in 8 may be negative for all $x_I \leq t$. This implies that the optimal platform for the incumbent is 0 in this interval, which corresponds to the case when the incumbent always finds it preferable to moderate her platform to improve her reelection chance and obtain large office rents. For intermediate values of $B$, there is an interior optimum that solves

$$\ell'(x_I^{\text{int}})(B - \ell(t - x_I^{\text{int}})) = \ell'(t - x_I^{\text{int}}) \frac{1 - F(\ell(x_I^{\text{int}}) - \ell(-x_M))}{f(\ell(x_I^{\text{int}}) - \ell(-x_M))}.$$  

(9)

There can be at most one interior solution in this interval. This is because the left-hand side is increasing and the right-hand side is decreasing in $x_I$ as the inverse hazard function on the right inherits log-concavity from $f$ (Bagnoli and Bergstrom, 2005). Let us denote the platform in this region that gives the highest expected utility by $x_I^* \in \{0, x_I^{\text{int}}, t\}$.

Therefore, there are four possible optimal platforms for an incumbent with an ideal point in $\left(\frac{x_I + \tilde{x}_I}{2}, \tilde{x}_I\right)$: $x_I = 0$ that maximizes the probability of beating $M$, $x_I = t$ that minimizes policy loss, $x_I = x_I^{\text{int}}$ that satisfies Equation 9, and $x_I = \tilde{x}_I$, the most moderate platform that induces $E$ to win the opposition party primary.

Notice that $x_I^*$ is monotone decreasing in $B$. This is intuitive; as office rents increase the incumbent responds by increasing her reelection probability. Also notice that the cross-partial of the incumbent’s payoff with respect to $x_I$ and $B$ is $-\ell'(x_I)f(\ell(x_I) - \ell(-x_M))$. It must then be that $x_I^*$ is continuously decreasing in $B$. We can then write $b(x_I): (0, t) \rightarrow \mathbb{R}_+$ as a surjection that maps incumbent platforms to the office rents $B$ that make them optimal for the incumbent, subject to the constraint $x_I \leq t$. 

41
Take \( b(x_I) \). By construction, this means that \( \text{EU}_I(x_I) \geq \text{EU}_I(x_I) \) for all \( x_I \leq t \). But we know by the definition of \( x_I \) that \( \tilde{x}_I \) results in weakly higher probability of reelection for the incumbent. Furthermore, our restriction of \( t > \frac{x_I + \tilde{x}_I}{2} \) ensures that \( \tilde{x}_I \) is closer to the incumbent’s ideal point than \( x_I \), \( \ell(\tilde{x}_I - t) < \ell(t - x_I) \). Therefore, by running on \( \tilde{x}_I \) instead, the incumbent can be reelected with as high a probability as \( x_I \) and get a policy she strictly prefers. Thus, it follows that for \( b(x_I) \), the incumbent’s expected utility is maximized at \( x_I = \tilde{x}_I \).

Next, take \( B = b(\hat{x}_I) \), where \( \hat{x}_I := 2t - \tilde{x}_I \). Again, by construction of \( b \) we have \( \text{EU}_I(\hat{x}_I) \geq \text{EU}_I(x_I) \) for all \( x_I \leq t \). Notice that \( \hat{x}_I > x_I \), which implies both that \( b(\hat{x}_I) \leq b(x_I) \) because \( b \) is decreasing, and that \( \hat{x}_I \) leads to a strictly lower probability of reelection for the incumbent than \( x_I \). Because \( x_I \) leads to the same probability of reelection as \( \tilde{x}_I \), it follows that \( \hat{x}_I \) leads to a lower probability of reelection than \( \tilde{x}_I \) and results in the same policy payoff conditional on election. Thus, when \( B = b(\hat{x}_I) \), the incumbent can improve her expected payoff by running on \( \hat{x}_I \) instead. Because we know \( \hat{x}_I \) is the constrained optimum, \( \tilde{x}_I \) must be the unconstrained optimum.

So we know that for both \( b(x_I) \) and \( b(\hat{x}_I) \) the incumbent’s optimal strategy is to provoke the opposition by playing \( \tilde{x}_I \). Recall next that the incumbent’s payoff function is

\[
(1 - F(\ell(x_I) - \ell(-x_J)))(B - \ell(|x_I - t|)).
\]

The derivative of this with respect to \( B \) is \( 1 - F(\ell(x_I) - \ell(-x_J)) \geq 0 \).

Because the value function is monotone increasing in \( B \) in the interval \([0, t]\), it follows by the Envelope Theorem that \( \tilde{x}_I \) is optimal for all \( B \in [b(\hat{x}_I), b(x_I)] \).

**Proof of Lemma 4.** Under Assumption 1B, all arguments from the proof of Lemma 3 involving the existence of an incumbent platform that leaves the extremist indifferent between
entering and staying out carry through. For this platform to be unique, I make the following 
assumption, which replaces $m_L$ with $E$ but is otherwise identical to Assumption 2A:

**Assumption 2B.**

$$
\frac{\ell'(x_I) - \ell'(x_M)}{\ell'(x_I - x_E)} > \frac{(F(x_I, x_M) - F(x_I, x_E))F(x_I, x_E)}{F(x_I, x_M)f(x_I, x_E) - F(x_I, x_E)f(x_I, x_M)}.
$$

This assumption ensures that while $\Delta_E(x_I)$ can be non-monotonic when $E$ prefers to stay 
out of the race, once $\Delta_E(x_I)$ is positive it is monotone increasing. When Assumptions 1B and 
2B hold, there is a unique platform $\tilde{x}'_I$ which solves $\Delta_E(\tilde{x}'_I) = 0$ and that leaves $E$ indifferent 
between running and staying out. Candidates run when they are indifferent, meaning that 
there is a competitive opposition party primary if and only if the incumbent’s platform is 
at least $\tilde{x}_I$. With probability $p$, the extremist wins a competitive primary and faces the 
incumbent in the general election. It follows then an incumbent with platform $x_I \geq \tilde{x}'_I$ faces 
$E$ with probability $p$, and $M$ with probability $1 - p$. Her probability of being reelected is 
the sum of the probabilities she faces each candidate times she beats them in the general 
election, so $p(F(\ell(x_I) - \ell(-x_E)) + (1 - p)F(\ell(x_I) - \ell(-x_M))$, if $x_I \geq \tilde{x}'_I$, and $F(\ell(x_I) - \ell(-x_M))$ 
otherwise. In particular, when $x_I = \tilde{x}'_I$, the incumbent is reelected with probability

$$
p(F(\ell(\tilde{x}'_I) - \ell(-x_E)) + (1 - p)F(\ell(\tilde{x}'_I) - \ell(-x_M)).
$$

Because $F(\ell(\tilde{x}'_I) - \ell(-x_M)) > F(\ell(\tilde{x}'_I) - \ell(-x_E))$ for all $x_I$, it follows by continuity of $\ell$ and $f$ that 
there exists a some interval to the left of $\tilde{x}'_I$ that lead to a lower probability of incumbent’s 
reelection.

Next, take $F(-\ell(-x_M))$. If this is less than the probability in expression 10, then $\tilde{x}'_I$ leads 
to the highest possible reelection probability. If it is larger, then by continuity there must
exist a platform \( x'_I(p) \) such that

\[
F(\ell(x'_I(p)) - \ell(-x_M)) = p(F(\ell(\tilde{x}'_I(p)) - \ell(-x_E)) + (1-p)F(\ell(\tilde{x}'_I(p)) - \ell(-x_M)).
\]

It follows that every incumbent with a platform \( x_I \in (x'_I(p), \tilde{x}'_I(p)) \) is reelected with a lower probability than \( \tilde{x}'_I(p) \).

**Proof of Proposition 4.** I prove each part of this proposition in the order they are presented. Throughout, I use \( \Delta_E := \lim_{x_I \to \infty} \Delta_E(x_I), \Delta_M := \sup(\Delta_M(x_I)) \), and \( \Delta' := \min(\Delta_M(x_I)). \)

Let \( 2c < \min(\Delta_M, \Delta_E) \). We know that \( 0 < \Delta_M, \Delta_E < \infty \), so this is well-defined. In the first two parts, the order of announcements does not matter because \( M \) plays a dominant strategy.

1. To see how Lemma 4 extends to positive costs of running, notice that because \( \Delta_M \) is bounded away from zero, we can find some \( p \) that satisfies \( p < 1 - \frac{c}{\Delta_M} \). This means that for such values of \( p \), \( M \) always runs. Also notice that from Assumption 1B, \( \Delta_E(0) < \frac{c}{p} \) immediately follows for any \( c, p > 0 \). Finally, because \( f \) has finite variance and \( x_E < x_M \), we can find some \( p > \frac{c}{\Delta_E} \). Then, by continuity and Assumption 2B there exists a unique incumbent platform \( \tilde{x}'_I \) that leaves \( E \) indifferent; she enters for \( x_I \geq \tilde{x}'_I \) and stays out otherwise. Thus, for \( p \in \left( \frac{c}{\Delta_E}, 1 - \frac{c}{\Delta_M} \right) \), we have our result.

2. Let \( p = \frac{c}{\Delta_E} \). Then, we have \( \Delta_E(x_I) < \frac{c}{p} \) for all \( x_I \geq 0 \) and for any \( p \in (0, p) \). This means that \( E \) never enters the race when \( M \) runs. Furthermore, for any \( p \in (0, \bar{p}) \) we have \( \frac{c}{1-p} < \Delta_M \leq \Delta_M(x_I) \) for all \( x_I \). It follows that \( M \) always prefers to run, driving \( E \) out. It must then be that \( M \) is the only candidate.

3. Let \( \bar{p} = 1 - \frac{c}{\Delta_M} \), and suppose first that \( M \) announces his decision to run, followed by \( E \). Then, for all \( p \in (\bar{p}, 1) \) and \( x_I \), we have \( c < \Delta_M < \Delta_M < \frac{c}{1-p} \). This means that \( M \) enters if and only if \( E \) will not join him, and so \( E \) can drive \( M \) out of running. Whether she chooses to depends on whether \( \Delta_E(x_I) \) is greater than \( \frac{c}{p} \) or not. Notice that we have
\[ \Delta_E(x_I) > \frac{c}{p} \text{ for } p \in (\bar{p}, 1) \] sufficiently high \( x_I \) because \( 1 - \frac{c}{\Delta_M} > 1 - \frac{c}{\Delta'} > \frac{c}{\Delta_E} \). Furthermore, from Assumption 1B it follows that \( \Delta_E(0) < 0 < \frac{c}{\bar{p}} \). By continuity it must be then for some intermediate values of \( \tilde{x}'_I \) such that \( \Delta_E(\tilde{x}') = \frac{c}{\bar{p}} \). When the incumbent’s platform is \( \tilde{x}'_I \) \( E \) is indifferent between entering and staying out, and enters. It follows that for \( p \in (\bar{p}, 1) \), against an incumbent with a platform \( x_I < \tilde{x}'_I \) only \( M \) runs; against an incumbent with a platform \( x_I \geq \tilde{x}'_I \) only \( E \) runs.

Suppose now \( E \) announces first, and \( M \) second. Here, \( E \) runs if and only if she prefers facing the incumbent herself rather than \( M \). As before, for all \( p \in (\bar{p}, 1) \) we have \( c < \Delta_M < \Delta_E < \frac{c}{1 - \bar{p}} \), and so \( M \) enters if and only if \( E \) has stayed out. \( E \) enters when \( \Delta_E(x_I) \geq c \). For sufficiently high values of \( x_I \) this must hold because \( 2c < \Delta_E \). Again, from Assumption 1B it follows that \( \Delta_E(0) < 0 < c \). Then, by Assumption 2B there exists a unique incumbent platform \( \tilde{x}'_I \) such that when \( p \in (\bar{p}, 1) \), against an incumbent with a platform \( x_I < \tilde{x}'_I \) only \( M \) runs and against an incumbent with a platform \( x_I \geq \tilde{x}'_I \) only \( E \) runs.

\[ \square \]

**Proof of Proposition 5.** When there is a primary advantage for extremists, we know from Lemma 4 that depending on the order of announcements, there exists a unique incumbent platform \( \tilde{x}'_I \) such that for \( x_I \geq \tilde{x}'_I \) the challenger is \( E \), and for \( x_I < \tilde{x}'_I \) the challenger is \( M \). Define as before \( x_I := \ell^{-1}(\max(0, \ell(\tilde{x}'_I) + \ell(-x_M) - \ell(-x_E))) \), and take \( t \in \left( \frac{x_I + \tilde{x}'_I}{2}, \tilde{x}'_I \right) \). The proof of Proposition 3 carries through with \( \tilde{x}'_I \) replacing \( \tilde{x}_I \).

Suppose now there is no primary advantage for either side. Then, we know from Lemma 4 that \( M \) always runs regardless of \( x_I \), and that there exists a unique incumbent platform \( \tilde{x}'_I \) such that \( E \) enters the race alongside \( M \) if and only if \( x_I \geq \tilde{x}'_I \). When \( E \) enters, she wins the primary and becomes the challenger with probability \( p \). Thus, the probability of reelection for the incumbent is \( F(\ell(x_I) - \ell(-x_M)) \) for \( x_I < \tilde{x}'_I \), and \( p(F(\ell(x_I) - \ell(-x_E)) + (1 - p)F(\ell(x_I) - \ell(-x_M)) \) for \( x_I \geq \tilde{x}'_I \).
ell(-x_M)) for \( x_I \geq \tilde{x}_I' \). Take an incumbent with ideal point \( t \in \left( \frac{\tilde{x}_I'(p) + \tilde{x}_I'}{2}, \tilde{x}_I' \right) \), where \( \tilde{x}_I'(p) \) is defined as in the text. Once again, the proof of Proposition 3 carries through with \( \tilde{x}_I' \) replacing \( \tilde{x}_I \), and \( x_I'(p) \) replacing \( x_I \).

That provoking the opposition cannot occur when there is a primary advantage for moderates follows from the fact that \( E \) never runs, and \( M \) always runs when \( p \) is sufficiently low.

**Appendix B  Combining the Two Models**

Here, I present results from simulations of a model that has both primary voters, and endogenous entry decisions by candidates. To ensure probabilities of winning the primary are on the interior, I assume primary voters receive a utility shock before the primary election. This is similar to the shock general election voters receive against the incumbent before the general election. I denote this shock, drawn from a log-concave distribution \( f_D \) by \( \eta \). For simplicity, I assume these two shocks are independent, \( \eta \perp \psi \), and each shock only lasts for one election: for any \( \eta \in \mathbb{R} \), \( F(ell(x_I) - ell(|x|) | \eta) = F(ell(x_I) - ell(|x|)) \). In other words, given the platforms of the incumbent and the challenger, realization of the primary election shock does not inform or change the probability the incumbent is reelected.

Simulations reveal that as the incumbent moves away from the median voter's ideal point, the probability median primary voter votes for the extremist candidate first fall, and then increase until it flattens out. The intuition is that both \( M \) and \( E \) have low chances of beating a very moderate incumbent, so the median primary voter only has a slight preference for \( M \). Against incumbents with moderate platforms, \( M \) has a much better chance of winning, which pushes the primary voters towards supporting him more. When the incumbent is more extreme, the probabilities \( M \) and \( E \) would beat her converge again, and so do
Figure 6: Orange (SW) and blue (NE) regions respectively correspond to parameter values where only $M$ and $E$ run in equilibrium. In green (E) both candidates run. $\psi$ is drawn from a standard normal distribution, and losses are linear on the left plot and quadratic on the right. Parameter values are fixed at $x_E = -1.2$, $x_M = -0.5$, $c = 0.5$, and $B = 2$. Lines refer to the probabilities $E$ wins when primary voters face a utility shock drawn from a standard normal distribution, and the median primary voter’s ideal point is $-2$, $-0.8$, and $-0.4$, depicted by the red, blue, and orange lines respectively.

The payoffs of primary voters from the two candidates. Finally, when the incumbent is very extreme, the probability either candidate beats her approaches one, and voters tend to vote for the candidate whose ideology is closer to their ideal point.

The net payoff for $E$ over $M$ of primary voters with ideal points close to $E$ is very similar to $\Delta E(x_I)$. This is intuitive, the extreme opposition primary voters and candidates face similar problems and have similar payoffs, with the exception that the candidate also cares about office rents and costs of running for office. As such, extremist primary voters support the extremist candidate in similar conditions as when she wants to enter the race. It follows that if the median primary voter has an ideal point close to $x_E$, incumbent moving away from the center increases both the probability $E$ wins in the general election, and the probability she wins in the primary election. Thus, the effect of the incumbent’s platform on the primary voters’ calculus reinforces the extremist candidate’s entry decision.
Here, I reproduce Figure 4 for when $E$ moves first instead to show that the order of announcements does not lead to significant changes in who runs in equilibrium. Notice that unless $p$ is close to one, running is a dominant strategy for $M$, and thus he runs regardless of the sequence of announcements. The order only matters when both candidates prefer to be the challenger themselves, but not so much to induce a competitive primary where they might lose. This can only happen when $p$ is high and so $(1 − p)\Delta_M(x_I) < c$, meaning that $M$ wants to stay out when $E$ enters. The condition for $E$ to be only candidate running in equilibrium when $M$ moves first is $p\Delta_E(x_I) > c$, that is, $E$ prefers to run even when $M$ is running, and so she only wins the primary with probability $p$. The same condition when $E$ moves first is $\Delta_E(x_I) > c$, because she knows her entry will deter $M$. Thus, the only case where the identity of the challenger depends on the order is $p\Delta_E(x_I) < c < \Delta_E(x_I)$ and $(1 − p)\Delta_M(x_I) < c$. The latter condition requires $p$ to be close to one, which means the region where the identity of the challenger is sensitive to the order of announcements must be narrow.

**Figure 7:** Orange (SW) and blue (NE) regions respectively correspond to parameter values where only $M$ and $E$ run in equilibrium. In green (E) both candidates run. $\psi$ is drawn from a standard normal distribution, and losses are linear on the left plot and quadratic on the right. Parameter values are fixed at $x_E = −1.2$, $x_M = −0.5$, $c = 0.5$, and $B = 2$. $E$ moves first.