Provoking the Opposition

Korhan Kocak*

First draft: August 30, 2019

This draft: November 11, 2025

Abstract

Parties in primary elections often choose between candidates who appeal to their base and those who appeal to the broader electorate. I present a model of primaries where incumbents can exploit this trade-off the opposition faces between ideological congruence and electability. In the model, incumbents reduce their appeal to the median voter by moving away from the center. In doing so, they provoke the opposition into nominating extremists, improving their reelection prospects. This mechanism generates elite polarization as politicians leapfrog voters — not despite electoral concerns, but because of them. The analysis fits the observation that incumbents sometimes move away from the center near the end of their term. Provoking the opposition relies on two conditions: divergence of primary and general electorates and a limited set of potential nominees. I argue that partisan sorting and changes in nomination procedures over the last decades made this strategy viable.

JEL classification: D7, P5.
Keywords: Primary elections, Calvert-Wittman, polarization.

*Assistant Professor, SPEGA, IE University, IE Tower — Paseo de la Castellana 259E, Madrid, Spain. E-mail: kkocak@faculty.ie.edu. Web: www.korhankocak.com.

We're not running to make a statement. We're not running to pressure the incumbent to the left. We're running to win.[1]

— Alexandria Ocasio-Cortez

Don't let extremists give Trump four more years.[2]

— John W. Hickenlooper

1 Introduction

The distinguishing feature of primary elections is the tug-of-war between a candidate's electability and her ideological compatibility with the party's base. Electability motivations push candidates towards moderation (Owen and Grofman, 2006). Because of partisan sorting, however, candidates who can better appeal to the median voter tend to be further ideologically from the party's base (Hall and Snyder Jr, 2015; Levendusky, 2009; Gerber and Morton, 1998). Primary voters thus often face a trade-off between nominating an incongruent centrist who has a greater probability of winning in the general election versus someone closer to them, but may have a harder time winning the hearts of swing voters.

In this paper, I study when and how incumbents can manipulate this trade-off in a standard unidimensional spatial voting framework. Incumbents have both the *means* and the *motivation* to influence the opposition party's primary process. I argue they can improve their reelection chances by strategically pursuing extreme platforms that embolden the extremist factions in the opposition party. This move affects the selection of candidates on the other side, benefiting the incumbent despite hurting most voters. The mechanism described here demonstrates how elite polarization can spiral in a feedback loop where extremism on one side begets more extremism on the other.

[1] https://www.instagram.com/p/BePOZY1lxCZ/

[2] https://www.facebook.com/JohnHickenlooper/videos/2545066648889450/

The core of this paper's argument is as follows. As the incumbent moves away from the center, the opposition's payoffs change in two conflicting ways. On the one hand, the value of defeating the incumbent increases. The reelection of a more extreme incumbent would result in worse policies for the members of the opposition, which strengthens their incentive to win the general election — pushing them toward the center. On the other hand, a more extreme incumbent increases the probability that any given challenger would win in the general election. In this case, the opposition extremists see a rare window of opportunity to pursue their agenda. When this latter effect is stronger — as recent evidence suggests (Lockhart and Hill, 2023) — a more extreme incumbent leads the opposition party to nominate a candidate ideologically closer to their base. Because she faces a weaker challenger, the incumbent's reelection probability increases — even though her appeal to the median voter declines.

The mechanism described in this paper explains several recent empirical findings in the literature on polarization. Although evidence of a polarizing electorate is mixed (Fiorina, Abrams, and Pope, 2008; Bafumi and Herron, 2010; Barberá, 2015), scholars agree Congress is becoming more polarized. This is happening via two main mechanisms. First, elected officials often move towards more extreme positions over the course of their tenure (McCarty, Poole, and Rosenthal, 2016). Second, moderates are replaced by extremists at every level of government (Thomsen, 2014, 2017; Hall, 2019). These observations are surprising given the ample evidence suggesting extremists face significant penalties at the ballot box (Hall and Snyder Jr, 2015; Canes-Wrone, Brady, and Cogan, 2002; Canes-Wrone and Kistner, 2022). A third related empirical observation consistent with the mechanism proposed here is that the electoral penalty for extremism has been on the decline — but only for incumbents (Bonica and Cox, 2018; Tausanovitch and Warshaw, 2018). The model in this paper provides a way to reconcile these seemingly contradictory findings. Incumbents move towards more extreme positions — the first mechanism that increases polarization — to provoke the opposition into nominating extremists. Extremist challengers, despite having a lower chance of winning in the general election, nevertheless sometimes do win, leapfrogging the median voter — the second mechanism by which Congress is becoming more polarized (Utych, 2020). Finally, that this penalty is declining for incumbents is in line with the logic that they sometimes adopt extreme positions strategically, anticipating their effect on the opposition's candidate selection. Incumbents' lower appeal to the median voter is compensated when they succeed in provoking the opposition.

I show that two conditions are necessary for provoking the opposition to be a viable strategy — conditions I argue became more pertinent over the last few decades due to institutional changes in the US. The first condition requires the primary and general electorates to be sufficiently divergent. If the primary electorate was similar to the general electorate, there would be little tension, as any candidate who appeals to one would also appeal to the other. Abundant evidence shows that the distance between primary and general election voters has been growing since at least the 1950s (Brady, Han, and Pope, 2007; Hill and Tausanovitch, 2018; King, Orlando, and Sparks, 2016) and that this is true regardless of the openness of primaries (Hill, 2015; Sides et al., 2020).

The second condition is that the opposition must be constrained in their choice of candidates. That is, they must be forced to choose between a limited set of primary candidates who are either *too* centrist or *too* extreme, given the incumbent's platform and the primary electorate's ideologies. If they could freely choose any ideology, the opposition would produce a goldilocks challenger who improves on both electability and congruence whenever the incumbent moves away from the center. It is well-documented that challenger-party primaries feature few candidates: there are on average 1.5 candidates in primary races against out-party incumbents (Thomsen, 2023). Furthermore, parties have less control over the nomination process than they did in the past. Party delegates used to be able to nominate dark horse candidates who could better respond to changes in the incumbent's platform. But the McGovern–Fraser reforms in the 1970s, by broadening the selectorate, made it harder for parties to coordinate on dark horse candidates (Shafer, 2014; Steger, 2000). Since then, party elites became increasingly reluctant to endorse candidates who do not have considerable support or are polling poorly (Polsby, 1983; McCarty and Schickler, 2018). Although until the early 2000s party elites maintained some degree of control over the nomination process through the invisible primary (Cohen et al., 2009), this control has since dwindled because of enhanced intra-party divisions, the rise of new media, and the grass-roots campaigning and the fundraising it allowed (Cohen et al., 2016), thus opening the way for the mechanism described in this paper.

I build my theory by first proving an analogue for the median voter theorem for primary elections. This is necessary, because single-peaked policy preferences do not ensure that primary voters have single-peaked preferences over primary candidates — each primary candidate produces a lottery between her and the incumbent. Previous studies typically assume that the median primary voter is decisive either explicitly (Owen and Grofman, 2006; Serra, 2011; Snyder and Ting, 2011) or implicitly (Grofman, Troumpounis, and Xefteris, 2019).[3] Others, notably Hummel (2013), Duggan (2014), and Mirhosseini (2015) prove the median primary voter's decisiveness for specific loss functions. This paper advances the literature by showing that log-concavity of the marginal loss function and the distribution of the median voter's ideal point are sufficient for primary voters' preferences over candidates to be single-peaked. In particular, Proposition 1 shows that under these conditions, the winner of the primary is the median primary voter's preferred candidate. This allows the subsequent analyses to restrict attention to the median primary voter.

I then analyze primaries under an open nominations model where the opposition is unconstrained: they can nominate a challenger with any ideology. Proposition 2 shows that under open nominations, the incumbent cannot provoke the opposition by adopting more extreme platforms. In this case, the opposition can always respond to a more extreme incumbent by nominating a challenger that improves both their probability of winning and policy gain.

When primary voters are limited to a discrete set of ideologies to choose from, the incumbent can provoke the opposition into selecting more extreme candidates by moving away from the median voter's position. In particular, I show the existence of a threshold platform for the

[3]Adams and Merrill (2008) assume primary voters ignore the electability of candidates and vote as if they were voting in the general election.

incumbent that induces the median primary voter to switch to supporting a more extreme candidate. By adopting a more radical position, the incumbent can induce an extremist challenger to win the opposition party primary and thus improve her reelection prospects. Proposition 3 presents the conditions under which the incumbent adopts a platform more extreme than her ideal point to provoke the opposition. This finding highlights how the centripetal force that is at the heart of much of the spatial voting literature — that politicians move towards the median voter's ideal point — may be reversed when primaries are taken into consideration, even when the primary electorate is forward-looking.

In the Appendix, I analyze how provoking the opposition can work on the supply side by considering 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 stay out to let moderate challengers face moderate incumbents. But they enter the race against sufficiently extremist — and therefore weak — incumbents. Thus, an extremist incumbent induces entry by extremists in the opposition party's primary, which results in a discrete increase in the probability of reelection. Proposition A.1 shows that this can occur when the primary field is even or when it is tilted in favor of extremists, but not when it is tilted in favor of moderates. Lastly, Proposition A.2 presents the conditions under which the incumbent prefers to weaken her appeal to the median voter to provoke opposition extremists to enter the race.

2 Related Literature

The formal literature studying the trade-off between electability and congruence finds that it leads primary voters to support candidates with platforms between their ideal points and that of the population median, conceding slightly on policy for greater electability (Owen and Grofman, 2006; Coleman, 1971). Candidates who win primaries should thus be between the median primary voter and the median general election voter (Gerber and Morton, 1998; Jackson, Mathevet, and Mattes, 2007). In other words, the winner of the primary 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 that 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; Hirano and Snyder Jr, 2019) and experimental studies (Woon, 2018), primary voters seem to have both of these considerations in mind while casting ballots

Recent empirical evidence lends support to the idea that primaries feature both centrifugal and centripetal forces. On the one hand, voters with extreme ideologies from either party are more likely to both vote and donate in primaries than moderates (Hill and Huber, 2017; Barber, Canes-Wrone, and Thrower, 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 this effect extends to downstream votes as well. Thus, parties benefit from their opponents electing an extremist.

How do parties deal with this trade-off? As others have recognized, we need a better theoretical understanding of the underlying dynamics that drive extremists to win primaries (Tausanovitch and Warshaw, 2018). Many factors play into primary voters' decisions, like the ideological distribution of voters (Serra, 2011; Meirowitz, 2005), uncertainty about candidates' appeal to swing voters (Snyder and Ting, 2011; Adams and Merrill, 2008; Ascencio, 2023), concerns about candidates flip-flopping (Hummel, 2010; Agranov, 2016), 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 and shows how extremists can emerge victorious from primaries in equilibrium, as a result of strategic provocation by the incumbent.

3 Model

I start by presenting the primitives of the model. In the main text, there are two players: an incumbent and a unit mass of citizens, a subset of whom are also primary voters.[4] Citizens' ideologies are summarized by their ideal points on a unidimensional policy space $x_i \in \mathbb{R}$. Each citizen's payoff from the implemented policy x is captured by $-\ell(|x - x_i|)$, where the loss function ℓ is increasing and convex in the absolute distance between the implemented policy and voter i's ideal point. Facing an incumbent platform x_I and a challenger platform x_J, voter i's payoff is given by:

$$-\mathbb{1}\{\text{Incumbent is reelected}\}\ell(|x_I - x_i|) - \mathbb{1}\{\text{Challenger is elected}\}\ell(|x_J - x_i|). \tag{1}$$

I assume $\ell(0) = 0$, and that the first derivative of the loss function satisfies log-concavity:

Assumption 1. *The marginal loss function, l', is log-concave: $\ell'''\ell' \leq (\ell'')^2$, where ℓ', ℓ'', and ℓ''' refer to the first, second, and third derivatives of the loss function, respectively.*

Two parties, L and R, compete for office. At the start of the game there is a party R incumbent in power. The opposition party, L, must hold a primary to nominate a challenger to run against the incumbent in the general election. I denote by x_{m_L} and x_m the ideal points of the median primary voter in L and the median general election voter, respectively. For simplicity, I assume that the ideal points of L primary voters — and x_{m_L} in particular — are negative and common knowledge. The uncertainty about the location of x_m is captured by distribution F.

Assumption 2. *The distribution of the median voter's ideology, F, is log-concave: $f'F \leq f^2$, where f and f' refer to the first and second derivatives of F, respectively.*

Together, Assumptions 1 and 2 ensure that the primary voters' preferences are ratio-ordered and thus the median primary voter is decisive.

[4] I present an extension that models opposition party elites' entry decisions in the Appendix.

The incumbent is both policy and office-motivated. She obtains office rents $B > 0$ if she wins the general election. I denote the incumbent's ideal point by $t > 0$ and I restrict her platform choice x_I to be positive. This restriction simplifies the exposition by ensuring no L primary voters have ideal points greater than the incumbent's platform, but does not affect substantive conclusions. The incumbent's objective function is:

$$\mathbb{1}\{\text{Incumbent is reelected}\}(B - \ell(|x_I - t|)). \tag{2}$$

The timing of the game is as follows:

1. The incumbent commits to a second period platform, x_I.
2. The L party primary voters observe x_I and simultaneously vote for a primary candidate from the set of available candidates, identified below. The candidate that obtains a majority of votes becomes the challenger.
3. The general election is held between the incumbent and the challenger.
4. The winner of the general election implements her platform in the second period.

I focus on Subgame Perfect Nash Equilibria in undominated strategies. Voters vote for one of the candidates with the leftmost platform when they are indifferent between multiple candidates. Ties are resolved by coin flips.

Before turning to the analysis, it is useful to discuss some assumptions. An important feature of the model is the incumbent's ability to commit to a second term platform. She cannot deceive the opposition into thinking she is an extremist, only to back-pedal to a more moderate platform after the opposition's primary. If this were possible, the incumbent could reap benefits from provoking the opposition — facing a weak opponent — without suffering an electoral penalty for extremism in the general election. But primary voters would anticipate such a "flip-flop" and ignore the incumbent's temporary move away from the center.

Thus, the incumbent must be able to commit to a platform before the opposition party primary for the model's strategic logic to obtain. I argue that incumbents have greater ability to commit, by virtue of being able to enact policies or nominate cabinet members that diverge from their ideal points. Opposition candidates, however, cannot credibly commit to a platform. In contrast to incumbents who can establish track records of pursuing platforms different from their policy preferences, the only signals challengers have are cheap-talk campaign promises. Put differently, candidates for office cannot convince voters that they would stick to their campaign platforms once elected in the same way incumbents can, through observable actions that credibly signal divergence from their ideal points.[5]

Second, I assume that players know the position of the median primary voter, but not that of the median general election voter. This assumption is motivated by the observation that turnout in primaries is significantly lower than general elections and consists almost entirely of strong partisans (Gerber, Huber, et al., 2017). The outcomes of general elections, however, are often decided by swing voters, whose choices are harder to predict. General elections are also one shot — in contrast to primaries which take place over the course of months, each stage producing more information about primary voters' preferences.

4 Analysis

To solve the model, I proceed by backward induction. Analysis of the general election is straightforward. In the second period, the winner of the general election implements her platform. Knowing this, general election voters vote for the candidate who would give them a higher payoff in the second period. Formally, given a challenger J with platform x_J, voter x_i votes for the challenger if and only if $-\ell(|x_J - x_i|) \geq -\ell(|x_I - x_i|)$. This means that all voters whose ideal points are to the left of the midpoint between the challenger and the incumbent, $\frac{x_I + x_J}{2}$, vote for the chal-

[5]One micro-foundation for this asymmetry is as follows. Politicians do not have ideal points *per se*; they have platforms voters attribute to them based on observables. Voters expect that politicians will implement their platforms if elected. The only way a politician can change voters' expectations about her is by enacting policies different from voters' attributed platforms.

lenger, and those to the right vote for the incumbent. In particular, the median voter votes for the challenger with probability $F\left(\frac{x_I+x_J}{2}\right)$. 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 defeats an incumbent seeking reelection on platform x_I. Having described these probabilities, we are now ready to explore the primary stage.

4.1 Decisiveness in Primaries

Before analyzing the incumbent's influence on the choice of challengers in the opposition party, I first present a technical result regarding preference aggregation in primaries. This analogue of the median voter theorem for primary elections enables us to restrict attention to the median primary voter when we consider the full game. It is also non-trivial, because primary voters face a fundamentally different problem than general election voters. While the latter choose from a set of platforms, primary voters choose from a set of *lotteries* over platforms — a challenger platform x and an incumbent platform x_I. For a primary voter with ideal point x_i, the expected utility of voting for an opposition candidate with platform x against an incumbent with platform x_I is:

$$\mathrm{EU}_i(x, x_I) = -\ell(|x_i - x|)F\left(\frac{x_I + x}{2}\right) - \ell(x_I - x_i)\left(1 - F\left(\frac{x_I + x}{2}\right)\right). \tag{3}$$

Equation (3) reveals the central tension in the model: the Calvert-Wittman trade-off between ideological congruence and electability. On the one hand, extreme challengers produce risky lotteries — a lower platform produces a lower probability that the party wins the general election and a higher probability the incumbent is reelected. On the other hand, if an extremist challenger wins, she may also generate better policies for some primary voters. It follows that, although voters have single-peaked preferences over policies, their preferences over primary candidates are not necessarily single-peaked. Without more structure on the payoff functions, centrists and extremists may vote together against the median primary voter's preferred candidate. Thus, the median primary voter may not be decisive.

In principle, there are two ways centrist and extremist primary voters may support a candidate against the one favored by the median primary voter. First, centrist primary voters may join extremists to vote for an extremist challenger to improve the incumbent's chances of reelection. This is known as *raiding* or *strategic crossover voting* (Chen and Yang, 2002). Second, extremist primary voters may side with centrists to vote for a candidate more moderate than the median primary voter's preferred candidate. When a primary voter who is slightly more extreme than another receives disproportionately greater disutility from the incumbent's reelection, he may be willing to support a more moderate candidate who has a higher chance of beating the incumbent, even if that candidate's platform is only slightly closer than that of the incumbent.[6]

My first result shows that Assumptions 1 and 2 are sufficient for the decisiveness of the median primary voter. These assumptions — satisfied by virtually all loss functions and distributions used in the spatial voting literature, respectively — ensure that primary voters' preferences over lotteries between the incumbent and potential challengers are *ratio ordered* and their payoffs satisfy single-crossing differences (Kartik, Lee, and Rappoport, 2023). Thus, primary voters who are more extreme (moderate) than the median prefer her optimal candidate to any candidate who is

[6]For example, suppose there are three primary voters whose ideal points are -6, -1, and -0.5 with preferences described by the following loss function

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

This function is increasing, convex, and right differentiable. Let $x_I = 3$ and the ideal point of the general election median voter be drawn from the standard normal distribution. Here, the median primary voter's optimal candidate, located at -1, loses the primary against a candidate near -0.5, who is preferred by both the centrist and the extremist primary voter.

more moderate (extreme). In other words, under Assumptions 1 and 2, the Condorcet winner platform is the median primary voter's optimal candidate, x_{m_L}.

Proposition 1. *Suppose Assumptions (1) and (2) hold. Then, the median primary voter's preferred candidate is the Condorcet winner.*

Proof. All proofs are in the Appendix. □

Proposition 1 shows that log-concavity of F and l' are sufficient for the median primary voter to be decisive. This result is a generalization of Proposition 1 in Mirhosseini (2015) who builds on Banks and Duggan (2006) to show that the median primary voter is decisive when losses are quadratic: $\ell(|x - x_i|) = (x - x_i)^2$. These results rely on the observation that under quadratic loss, the disutility from uncertainty is independent of voters' ideal points. Proposition 1 generalizes this result to a broader class of functions.[7] A substantive implication is that 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 or the platforms they run on (Hill, 2015; McGhee et al., 2014).

The exact location of the Condorcet winner depends on the functional forms. An interesting case that previous literature focuses on obtains when losses are linear or exponential, $\ell(|x-x_i|) = |x - x_i|$, or $\ell(|x - x_i|) = e^{-|x-x_i|}$. Then, there exists a platform x^* who is the optimal candidate of everyone to her left. For everyone to the right of x^*, each voter's optimal candidate has the same ideology as him. Proposition 1 implies that the Condorcet winner must be the more moderate of x^* and the median primary voter's ideology, x_{m_L}:

Corollary 1. *Suppose Assumption (2) holds, and losses are linear or exponential. Then, there exists a unique x^* such that it is the optimal candidate platform for all voters whose ideal points more*

[7] In a related result, Duggan (2014) shows that the median voter is decisive over lotteries when the difference in the expected utilities of lotteries is monotonic in voter types. This condition does not hold in the present context.

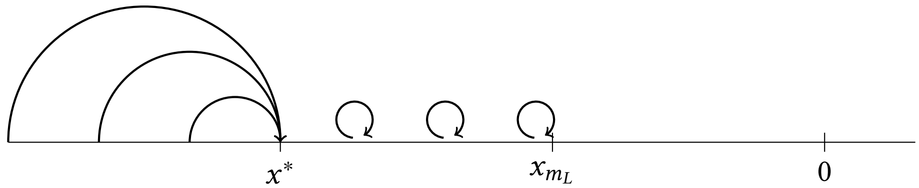


Figure 1: Optimal candidates of L voters when losses are linear or exponential, and $x_{m_L} > x^*$.

extreme; for all voters whose ideal points are more moderate than x^*, the optimal challenger shares their ideology. The Condorcet winner is the more moderate of x_{m_L} and x^* .

Thus, for the special cases of linear and exponential losses, there exists a platform x^* such that all primary voters to the left of that platform agree that a candidate with ideology x^* optimally trades off electability for ideology. 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.[8] This is visualized in Figure 1.

Having established that the median primary voter is decisive under Assumptions 1 and 2, I now turn to the case of open nominations, where the opposition party can freely choose a challenger from the ideological spectrum.

4.2 Open Nominations

When choosing the ideology of the challenger $x \in \mathbb{R}$, primary voters care not only about the positions of candidates but also their probability of beating the incumbent. Formally, primary voters choose $x \in \mathbb{R}$ to maximize Equation (3).

The solution to this problem must be between primary voters' ideal points and the incumbent's platform, $x_i \leq x \leq x_I$. To see why, notice first that a candidate with platform x_i is preferred to anybody more extreme than her. This is because she is both more likely to win the election and results in a higher payoff for primary voters if she does. Furthermore, nominating

[8]This result is a generalization of Theorem 1 of Owen and Grofman (2006), which shows this for the case of exponential loss and normal f.

a candidate whose platform is further right than that of the incumbent is strictly dominated for the primary voters than nominating a candidate to the incumbent's left.[9] These imply that the optimal challenger's ideology must be in $[x_i, x_I]$. In this interval, we can write the derivative of the primary voters' problem as:

$$\frac{d\mathrm{EU}_i}{dx} = -\ell'(x - x_i)F\left(\frac{x_I + x}{2}\right) + \frac{1}{2}f\left(\frac{x_I + x}{2}\right)\left(\ell(x_I - x_i) - \ell(x - x_i)\right). \tag{4}$$

This expression captures the familiar trade-off in probabilistic voting models. The first term is the marginal policy loss as the challenger moves away from i's ideal point: a more moderate challenger results in a lower payoff for i if she wins. The second term is the marginal gain from winning with a higher probability. Notice that when $x = x_I$, the second term is zero, which means that this expression is negative. It follows that the optimal challenger platform must be strictly lower than the incumbent's platform, $x < x_I$. If the first term is larger in absolute value than the second term, Equation (4) is always negative. In this case, the incumbent is sufficiently weak, and primary voters always find it worthwhile to trade off electability for greater ideological congruence. Thus, the optimal challenger for each voter i has maximal congruence at x_i. If Equation (4) is positive for some platforms and negative for others, there must be a unique interior optimum.

We know from Proposition 1 that the median primary voter's optimal candidate wins the primary. Thus, we can restrict focus to the median primary voter, m_L, and investigate how the optimal challenger ideology changes as a function of the incumbent's platform. Recall the trade-off faced by primary voters described in the introduction: against a more extreme incumbent, every challenger has a higher chance of winning, because the $F\left(\frac{x_I+x}{2}\right)$ term increases in x_I. This strengthens the incentives m_L has to push forward candidates whose platforms he likes more. On the other hand, a more extreme incumbent increases the value of defeating the incumbent as the $\ell(x_I - x_{m_L}) - \ell(x - x_{m_L})$ component of the payoff function goes up, which pushes m_L to favor

[9]The arguments sketched here are formally presented and proved in the Appendix.

more electable candidates.

Shapes of ℓ and F determine whether the optimal challenger 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 greater policy loss for the median primary voter without causing a big change in the probability of reelection (*i.e.* changes in ℓ dominate changes in F), he 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 the median primary voter (*i.e.* changes in F dominate changes in ℓ), he would rather choose somebody more congruent. Depending on the functional forms, the optimal challenger ideology may become more or less extreme as the incumbent moves away from the center. Nevertheless, we can prove that it never becomes *so extreme* as to improve the incumbent's probability of reelection. This is because the opposition party always responds to an incumbent moving away from the center by nominating a candidate who will defeat her with a higher probability.

Formally, let x^* and x^{**} denote the optimal challenger platforms for the median primary voter against x_I and x_I' respectively, and let $x_I < x_I'$. Then, $F\left(\frac{x_I'+x^{**}}{2}\right) \geq F\left(\frac{x_I+x^*}{2}\right)$.

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 result is as follows: when the opposition party can choose the ideology of the challenger, their best response against a more extreme incumbent improves both the probability component $F\left(\frac{x_I+x}{2}\right)$, and the policy gain component, $\ell(x_I - x_i) - \ell(x - x_{m_L})$. This means that even if a more extremist incumbent platform leads the optimal challenger platform to also be more extreme, this shift cannot be so large to lead to an overall lower probability of winning in the general election. 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; instead, they must choose from an exogenously given set of party elites.

4.3 Party Elites

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. Evidence shows that politicians who self-select into the profession (Dal Bó et al., 2017) and are not screened out by interest groups and party insiders (Cohen et al., 2009; Broockman et al., 2021) are not representative of the larger population. Furthermore, politicians themselves are 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. This is true whenever the net expected utility from nominating E over M, $\Delta_{m_L}(x_I)$, is positive. Formally:

$$\Delta_{m_L}(x_I) \coloneqq F\left(\frac{x_I + x_E}{2}\right)\left(\ell(x_I - x_{m_L}) - \ell(|x_{m_L} - x_E|)\right) - F\left(\frac{x_I + x_M}{2}\right)\left(\ell(x_I - x_{m_L}) - \ell(|x_{m_L} - x_M|)\right).$$

In this setting, a more extreme incumbent can be reelected with a higher probability. When primary voters must choose from a discrete set of exogenously given ideologies, their response to the incumbent's platform are discontinuous. In other words, the only tool the median primary voter has here 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.

Notice that 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. Thus, let us restrict attention to the case where the median primary voter's choice of E or M is responsive to the incumbent's position: he prefers E to become the challenger against some incumbents and M against others. Specifically, I suppose that the median primary voter's ideology is closer to E's,[10] but M's probability of beating a very moderate incumbent is sufficiently higher:

Assumption 3. $x_{m_L} < \frac{x_M+x_E}{2}$*, and* $\Delta_{m_L}(0) < 0$.[11]

Under Assumption 3, incumbents close to the center induce the median primary voter to vote for M and those who are far induce him to vote for E. Because f and ℓ are both continuous in $x_I \in \mathbb{R}_+$, the median primary voter's payoff is also continuous. Then, there exists a platform for the incumbent such that the median primary voter is indifferent between E and M. Let $\tilde{x}_I$ denote this platform so that when the incumbent's platform is $x_I < \tilde{x}_I$ the opposition nominates M and otherwise nominates E. Allowing for multiple incumbent platforms that leave the median primary voter indifferent complicates the analysis, but does not lead to additional substantively meaningful insights. To simplify exposition, I assume this platform is unique:

[10] Patty and Penn (2019) shows that forward-looking voters may exhibit a preference for an extremist even if their ideal points are closer to a moderate. This taste-for-extremism results from institutional or chance factors that preclude representatives from implementing their ideal points, biasing the implemented policy towards the center.

[11] To get a sense of what the second part of Assumption 3 requires, suppose ℓ is linear or quadratic. Then, $\Delta_{m_L}(0) < 0 \iff \frac{x_E}{x_M} < \frac{F(x_M/2)}{F(x_E/2)}$ or $\frac{x_E(2x_{m_L}-x_E)}{x_M(2x_{m_L}-x_M))} < \frac{F(x_M/2)}{F(x_E/2)}$, respectively. Either inequality holds when x_M is sufficiently low and F is the normal distribution with mean zero and low enough variance.

Assumption 4. *Let $\tilde{x}_I$ satisfy $\Delta_{m_L}(\tilde{x}_I) = 0$. Then, $\frac{d\Delta_{m_L}(x_I)}{dx_I} \geq 0$ for $x_I > \tilde{x}_I$.*[12]

Assumption 4 is a sufficient condition for single-crossing of the median voter's net payoff. It states that once the incumbent's platform is extreme enough for the median primary voter to prefer the extremist, further moves away from the center cannot induce him to switch back to preferring the moderate.

Next, define $\underline{x}_I$ as the incumbent platform that leads to the incumbent's reelection against the moderate challenger M with the same probability as $\tilde{x}_I$ wins against x_E. Formally, let $F\left(\frac{\underline{x}_I+x_M}{2}\right) = F\left(\frac{\tilde{x}_I+x_E}{2}\right)$, if such a platform exists. Otherwise, let $\underline{x}_I \coloneqq 0$. Incumbents with platforms in the interval $(\underline{x}_I, \tilde{x}_I)$ 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 1. *When the median primary voter's preferences satisfy Assumptions 3 and 4, 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 1 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. This is visualized in Figure 2, which plots the incumbent's reelection probability as a function of her platform. To see that this can indeed cause the incumbent to pursue platforms more extreme than her ideal point, observe that

[12] As the incumbent moves away from the center, both the value and the probability of defeating the incumbent increase. The latter, $\left(F\left(\frac{x_I+x_E}{2}\right) - F\left(\frac{x_I+x_M}{2}\right)\right)\ell'(x_I - x_{mL})$, term goes to zero as x_I increases, because the probability each challenger defeats the incumbent goes to one as the incumbent becomes sufficiently extreme. What remains, $\frac{1}{2}\left(\ell(x_I - x_{mL}) - \ell\left(|x_E - x_{mL}|\right)\right) F'\left(\frac{x_I+x_E}{2}\right) - \frac{1}{2}\left(\ell(x_I - x_{mL}) - \ell\left(|x_M - x_{mL}|\right)\right) F'\left(\frac{x_I+x_M}{2}\right)$, is monotone increasing in x_I by the log-concavity of F and the first part of Assumption 3. Although difficult to establish for generic ℓ and F, this is true for linear and quadratic losses with uniform and normal distributions.

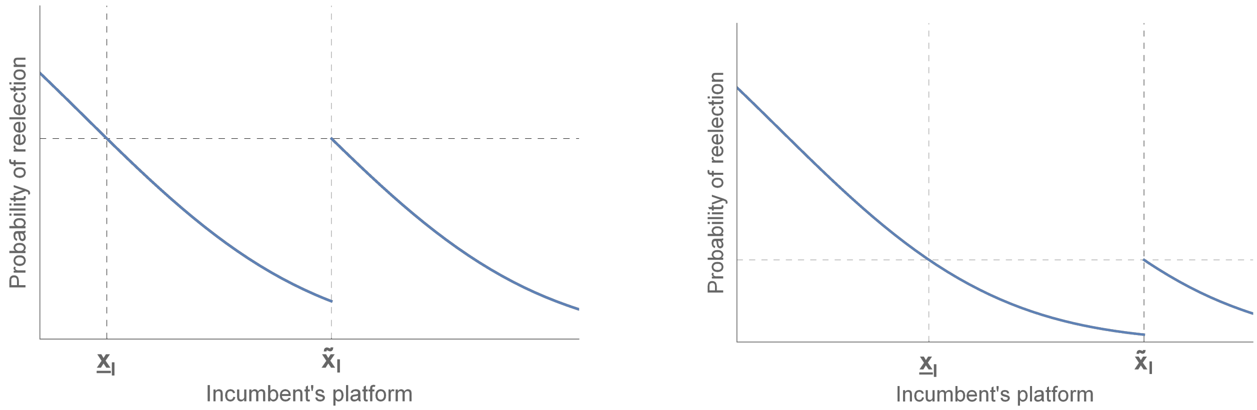


Figure 2: Incumbent's probability of reelection is plotted as a function of her platform. The median voter's ideal point is standard normally distributed. Losses are linear on the left plot and quadratic on the right. Parameter values are $x_E = -3$, $x_M = -0.5$, and $x_{m_L} = -5$.

her expected payoff is

$$\max_{x_I \in \mathbb{R}_{++}} \mathrm{EU}_I(x_I) = \max_{x_I \in \mathbb{R}_{++}} \left(1 - F\left(\frac{x_I + x_J}{2}\right)\right)(B - \ell(|x_I - t|)). \tag{5}$$

where J is the challenger chosen by m_L in equilibrium and t denotes the incumbent's ideal point. Solving the incumbent's problem reveals the conditions under which she provokes the opposition into nominating an extremist by adopting a platform more extreme than her ideal point. Specifically, for incumbents whose ideal points lie in $t \in \left(\frac{\underline{x}_I + \tilde{x}_I}{2}, \tilde{x}_I\right)$, the optimal platform induces the extremist to win the opposition party primary whenever B is in a bounded interval of office rents.[13]

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 a bounded 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 the median primary voter chooses the challenger from a set of party elites

[13] When B is below this interval, the incumbent does not find it worthwhile to give up on policy to improve her reelection probability; when it is above, reelection is sufficiently important that the incumbent moderates.

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 chances. She does this solely to weaken her appeal to the median voter. This emboldens the primary voters in the opposition, inducing them to nominate the extremist E who gathers the votes of all voters to the left of the median primary voter and defeats M. However, because E's ideology is further from the general election median voter than M, 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.

5 Conclusion

In this paper, I propose a model of sequential elections that shows how extremism on one side can give rise to extremism on the other, leading to a spiral of polarization. This explains the observed polarization in US Congress, starting with why incumbent politicians move to extreme positions towards the end of their term. They do so to assist opposition extremists in the primaries, therefore improving the chances they face a weaker opponent. The observation that electoral penalty for extremism for incumbents has been on the decline is in line with this logic. When this gambit backfires, extremist challengers win in the general election — despite being weaker than their more moderate counterparts — explaining the rise of extremists in US Congress. Importantly, the model does not hinge on mass polarization, nationalization of politics, or gerrymandering — it is driven entirely by partisan sorting and constraints parties face in primaries.

The model is based on the idea that an incumbent with a more extreme platform makes it both more *important* and more *likely* for the opposition to defeat her. The first follows because the platform that would be implemented if an extreme incumbent were reelected is disliked more by the members of the opposition. Thus, the opposition has more to gain by defeating an extremist incumbent, which pushes them towards moderation to increase their appeal to the median voter and therefore their probability of beating the incumbent. On the other hand, an extreme platform is further from the ideal point of the median voter, which increases the probability that any given challenger can defeat the incumbent. This emboldens the extremist factions within the opposition party, who see a window of opportunity to pursue their agenda. Empirical work suggests this second effect dominates: parties nominate more congruent challengers against extreme incumbents (Lockhart and Hill, 2023). I present conditions under which this is true. I further establish when incumbents choose to move to platforms more extreme than their ideal points, resulting in the opposition nominating a weaker challenger in the general election, thus improving incumbents' chances of reelection.

Provoking the opposition can work via two distinct mechanisms. In the main text, I show how the incumbent can induce the opposition primary voters to support extremists whose policies they like better than centrists. In Appendix A, I present a modified model between the incumbent and the opposition elites and show that the incumbent can provoke extremists to run for office, possibly driving moderates out. Thus, incumbents can provoke extremist candidates to run or primary voters to support them. A natural next question is how these two forces interact. In Appendix C, I present simulations that combine these two mechanisms. These suggest that the results are robust to when both primary elections and candidates' entry decisions are endogenous and that the effects on primary voters and candidates complement each other. My simulations show that more extreme incumbents push voters towards supporting extremist challengers in the primary. Because they are more likely to win the primary, extremists then run for office, while moderates stay out. Thus, the effects of the incumbent's move away from the center on primary voters and challengers reinforce each other, leading to a stronger overall effect on the incumbent's chances of reelection.

Whether the incumbent can successfully provoke the opposition and improve her reelection chances depends on the set of candidates parties choose from. I show that when the opposition party can freely choose the ideology of their candidate, they always improve the probability of winning against a more extreme incumbent. But primary elections constrain parties' choice by limiting their control over the nomination process. It is then possible for an extremist incumbent to win reelection with a higher probability than a moderate one. This highlights a novel implication of reforms that democratized the nomination process: the inability of the opposition party to coordinate on challengers from a richer set of ideologies enables incumbents to provoke the opposition.

Primaries have been identified as a factor that can exacerbate political polarization by giving ideologically extreme partisans greater say in the nomination process. In this paper, I demonstrate how primaries can also drive moderates in office to adopt extreme positions, leading to further polarization. Thus, primaries can contribute to the observed proliferation of ideological extremists in contemporary politics in three ways: encouraging extremists to run for office, primary voters to support them, and incumbent moderates to adopt more extreme positions.

More generally, the mechanism extends to any sequential Calvert-Wittman contest with asymmetric commitment power.[14] The first mover (*e.g.*, the incumbent) may find it optimal to commit to an extreme platform to induce the second mover (*e.g.*, the opposition party) to also pick an extreme platform. When the second mover is constrained to choose from a coarse set of platforms, like in the *Party Elites* section, such a move can benefit the first mover. In particular, if a move away from the center will induce the second mover to do the same, and if this move is larger than the first mover's, both candidates gain at the expense of the median voter. Put differently, politicians "collude" by pursuing platforms closer to their ideal points than if they had to moderate under the standard Calvert-Wittman model, while maintaining their probabilities of winning. This explains a polarizing elite leapfrogging voters.

Acknowledgements

I thank Scott Abramson, Ekrem Baser, Brendan Cooley, Jean Guillaume Forand, Dan Gibbs, German Gieczewski, Sona Golder, Matias Iaryczower, Federica Izzo, Gleason Judd, Patricia Kirkland, Andrew Little, Asya Magazinnik, Nolan McCarty, Noam Reich, Arturas Rozenas, seminar participants at New York University, New York University Abu Dhabi, Princeton University, Vanderbilt University, Washington University in St. Louis, APSA 2019 Annual Meeting, MPSA 2019 Annual

[14]I thank an anonymous referee for this point.

Meeting, and four anonymous referees for helpful comments and feedback.

Declaration of competing interest

No conflict of interest.

Data availability

No data was used for the research described in the article.

References

Abramowitz, Alan I. (1989). "Viability, Electability, and Candidate Choice in a Presidential Primary Election: A Test of Competing Models". *The Journal of Politics* 51.4, pp. 977–992.

Abramson, Paul R. et al. (1992). "'Sophisticated' Voting in the 1988 Presidential Primaries". *The American Political Science Review* 86.1, pp. 55–69.

Adams, James and Samuel Merrill (2008). "Candidate and party strategies in two-stage elections beginning with a primary". *American Journal of Political Science* 52.2, pp. 344–359.

Agranov, Marina (2016). "Flip-Flopping, Primary Visibility, and the Selection of Candidates". *American Economic Journal: Microeconomics* 8.2, pp. 61–85.

Ascencio, Sergio (2023). "Retaining political talent: A candidate-centered theory of primary adoption". *American Journal of Political Science.*

Bafumi, Joseph and Michael C. Herron (2010). "Leapfrog representation and extremism: A study of American voters and their members in Congress". *American Political Science Review* 104.3, pp. 519–542.

Bagnoli, Mark and Ted Bergstrom (2005). "Log-Concave Probability and Its Applications". *Economic Theory*, pp. 445–469.

Banks, Jeffrey S. and John Duggan (2006). "A Social Choice Lemma on Voting Over Lotteries with Applications to a Class of Dynamic Games". *Social Choice and Welfare* 26.2, pp. 285–304.

Barber, Michael J., Brandice Canes-Wrone, and Sharece Thrower (2017). "Ideologically Sophisticated Donors: Which Candidates Do Individual Contributors Finance?" *American Journal of Political Science* 61.2, pp. 271–288.

Barberá, Pablo (2015). "Birds of the same feather tweet together: Bayesian ideal point estimation using Twitter data". *Political Analysis* 23, pp. 76–91.

Bonica, Adam and Gary W. Cox (2018). "Ideological Extremists in the U.S. Congress: Out of Step but Still in Office". *Quarterly Journal of Political Science* 13.2, pp. 207–236.

Brady, David W., Hahrie Han, and Jeremy C. Pope (2007). "Primary Elections and Candidate Ideology: Out of Step with the Primary Electorate?" *Legislative Studies Quarterly* 32.1, pp. 79–105.

Broockman, David E. et al. (2021). "Why Local Party Leaders Don't Support Nominating Centrists". *British Journal of Political Science* 51.2, pp. 724–749.

Canes-Wrone, Brandice, David W. Brady, and John F. Cogan (2002). "Out of Step, out of Office: Electoral Accountability and House Members' Voting". *The American Political Science Review* 96.1, pp. 127–140.

Canes-Wrone, Brandice and Michael R Kistner (2022). "Out of Step and Still in Congress? Electoral Consequences of Incumbent and Challenger Positioning Across Time". *Quarterly Journal of Political Science* 17.3, pp. 389–420.

Chen, Kong-Pin and Sheng-Zhang Yang (2002). "Strategic Voting in Open Primaries". *Public Choice* 112.1, pp. 1–30.

Cohen, Marty et al. (2009). *The Party Decides: Presidential Nominations Before and After Reform.* University of Chicago Press. 418 pp.

— (2016). "Party versus faction in the reformed presidential nominating system". *PS: Political Science & Politics* 49.4, pp. 701–708.

Coleman, James S. (1971). "Internal processes governing party positions in elections". *Public Choice* 11.1, pp. 35–60.

Dal Bó, Ernesto et al. (2017). "Who Becomes A Politician?" *The Quarterly Journal of Economics* 132.4, pp. 1877–1914.

Duggan, John (2014). "Majority Voting Over Lotteries: Conditions for Existence of a Decisive Voter". *Economics Bulletin* 34.1, pp. 263–270.

Fiorina, Morris P., Samuel A. Abrams, and Jeremy C. Pope (2008). "Polarization in the American Public: Misconceptions and Misreadings". *The Journal of Politics* 70.2, pp. 556–560.

Gerber, Alan S, Gregory A Huber, et al. (2017). "Why don't people vote in US primary elections? Assessing theoretical explanations for reduced participation". *Electoral Studies* 45, pp. 119–129.

Gerber, Elisabeth R and Rebecca B Morton (1998). "Primary election systems and representation". *The Journal of Law, Economics, and Organization* 14.2, pp. 304–324.

Grofman, Bernard, Orestis Troumpounis, and Dimitrios Xefteris (2019). "Electoral competition with primaries and quality asymmetries". *The Journal of Politics* 81.1, pp. 260–273.

Hall, Andrew B. (2015). "What happens when extremists win primaries?" *American Political Science Review* 109.1, pp. 18–42.

— (2019). *Who Wants to Run?: How the Devaluing of Political Office Drives Polarization.* University of Chicago Press.

Hall, Andrew B. and James M. Snyder Jr (2015). *Candidate Ideology and Electoral Success.*

Hill, Seth J and Chris Tausanovitch (2018). "Southern realignment, party sorting, and the polarization of American primary electorates, 1958–2012". *Public Choice* 176.1-2, pp. 107–132.

Hill, Seth J. (2015). "Institution of Nomination and the Policy Ideology of Primary Electorates". *Quarterly Journal of Political Science* 10.4, pp. 461–487.

Hill, Seth J. and Gregory A. Huber (2017). "Representativeness and Motivations of the Contemporary Donorate: Results from Merged Survey and Administrative Records". *Political Behavior* 39.1, pp. 3–29.

Hirano, Shigeo and James M Snyder Jr (2019). *Primary elections in the United States.* Cambridge University Press.

Hummel, Patrick (2010). "Flip-flopping from primaries to general elections". *Journal of Public Economics* 94.11, pp. 1020–1027.

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

Jackson, Matthew O., Laurent Mathevet, and Kyle Mattes (2007). "Nomination processes and policy outcomes". *Quarterly Journal of Political Science* 2.1, pp. 67–92.

Kartik, Navin, SangMok Lee, and Daniel Rappoport (2023). "Single-Crossing Differences in Convex Environments". *The Review of Economic Studies*, rdad103. eprint: https://academic.oup.com/restud/advance-article-pdf/doi/10.1093/restud/rdad103/54092237/rdad103.pdf.

King, Aaron S., Frank J. Orlando, and David B. Sparks (2016). "Ideological Extremity and Success in Primary Elections: Drawing Inferences From the Twitter Network". *Social Science Computer Review* 34.4, pp. 395–415.

Levendusky, Matthew (2009). *The partisan sort: How liberals became Democrats and conservatives became Republicans.* University of Chicago Press.

Lockhart, Mackenzie and Seth J. Hill (2023). "How Do General Election Incentives Affect the Visible and Invisible Primary?" *Legislative Studies Quarterly* 48.4, pp. 833–867.

McCarty, Nolan, Keith T. Poole, and Howard Rosenthal (2016). *Polarized America: The dance of ideology and unequal riches.* mit Press.

McCarty, Nolan and Eric Schickler (2018). "On the theory of parties". *Annual Review of Political Science* 21, pp. 175–193.

McGhee, Eric et al. (2014). "A primary cause of partisanship? Nomination systems and legislator ideology". *American Journal of Political Science* 58.2, pp. 337–351.

Meirowitz, Adam (2005). "Informational Party Primaries and Strategic Ambiguity". *Journal of Theoretical Politics* 17.1, pp. 107–136.

Mirhosseini, Mohammad Reza (2015). "Primaries with strategic voters: trading off electability and ideology". *Social Choice and Welfare* 44.3, pp. 457–471.

Owen, Guillermo and Bernard Grofman (2006). "Two-stage electoral competition in two-party contests: persistent divergence of party positions". *Social Choice and Welfare* 26.3, pp. 547–569.

Patty, John W and Elizabeth Maggie Penn (2019). "Are moderates better representatives than extremists? A theory of indirect representation". *American Political Science Review* 113.3, pp. 743–761.

Polsby, Nelson W (1983). *Consequences of party reform*. Oxford University Press, USA.

Serra, Gilles (2011). “Why primaries? The party’s tradeoff between policy and valence”. *Journal of Theoretical Politics* 23.1, pp. 21–51.

Shafer, Byron E. (2014). *Bifurcated Politics, Evolution and Reform in the National Party Convention*. Reprint 2014. Cambridge: Harvard University Press.

Sides, John et al. (2020). “On the representativeness of primary electorates”. *British Journal of Political Science* 50.2, pp. 677–685.

Snyder, James M. and Michael M. Ting (2011). “Electoral Selection with Parties and Primaries”. *American Journal of Political Science* 55.4, pp. 782–796.

Steger, Wayne P. (2000). “Do Primary Voters Draw from a Stacked Deck? Presidential Nominations in an Era of Candidate-Centered Campaigns”. *Presidential Studies Quarterly* 30.4, pp. 727–753.

Tausanovitch, Chris and Christopher Warshaw (2018). “Does the Ideological Proximity Between Candidates and Voters Affect Voting in U.S. House Elections?” *Political Behavior* 40.1, pp. 223–245.

Thomsen, Danielle M. (2014). “Ideological Moderates Won’t Run: How Party Fit Matters for Partisan Polarization in Congress”. *The Journal of Politics* 76.3, pp. 786–797.

— (2017). *Opting out of congress: Partisan polarization and the decline of moderate candidates*. Cambridge University Press.

— (2023). “Competition in Congressional Elections: Money versus Votes”. *American Political Science Review* 117.2, pp. 675–691.

Utych, Stephen M. (2020). “Man bites blue dog: are moderates really more electable than ideologues?” *The Journal of Politics* 82.1, pp. 392–396.

Woon, Jonathan (2018). “Primaries and Candidate Polarization: Behavioral Theory and Experimental Evidence”. *American Political Science Review* 112.4, pp. 826–843.

Appendices

Appendix A Endogenous Entry

To show that provoking the opposition can also happen on the supply side, here I study the strategic considerations of party elites with a model of costly entry. I assume that party elites care both about office rents and policies. They take into account their probabilities of winning the primary and the general election, as well as the effect their entry has on the outcome. To isolate the effect of the incumbent's position on the candidates' considerations about the general election, I assume here 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, M followed by E.[15] 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 reelected. Candidates run when indifferent. The cost of running for office is $c \geq 0$ and I assume this cost is low enough so that each candidate prefers to run when in equilibrium the other is not running: $c \leq F\left(\frac{x_J}{2}\right)(B + \ell(-x_J))$ for $J \in \{E, M\}$. The game is otherwise identical to the one described in the main text.

The timing of the endogenous entry game is as follows:

1. Incumbent chooses her platform.
2. M announces her running decision.
3. E announces her running decision.
4. If only one candidate runs, she becomes the challenger. If both run, E becomes the challenger with probability p and M with probability $1 - p$.

[15] Sequentiality precludes coordination failures and ensures uniqueness. All substantive results continue to hold when candidates announce simultaneously. Equilibrium running decisions are qualitatively similar when E announces before M — I present simulations in Appendix D .

5. The general election is held between the challenger and the incumbent.

6. 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 model in the main text, so I skip to candidates' entry decisions. Because it is simpler, I start with the limiting case of no cost of running, $c = 0$.

In deciding whether to enter the primary race when the other candidate is running, each candidate evaluates the benefits of running — implementing their ideal 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. Because the moderate always has a higher probability of beating the incumbent than the extremist, M's entry decreases the probability the incumbent is reelected. It follows that when the cost of running is zero, M always enters. 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:

$$\Delta_M(x_I) \coloneqq F\left(\frac{x_I + x_M}{2}\right)(B + \ell(x_I - x_M)) - F\left(\frac{x_I + x_E}{2}\right)(\ell(x_I - x_M) - \ell(x_M - x_E)).$$

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 the incumbent faces E in the general election. Because the probability M beats the incumbent is higher, $\Delta_M(x_I)$ must be positive.

The same does not hold for E: when the moderate is running, the extremist's entry increases the probability the incumbent is reelected. Thus, E 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:

$$\Delta_E(x_I) \coloneqq F\left(\frac{x_I + x_E}{2}\right)(B + \ell(x_I - x_E)) - F\left(\frac{x_I + x_M}{2}\right)(\ell(x_I - x_E) - \ell(x_M - x_E)).$$

As argued above, this expression can be positive or negative. Notice that it must be positive for sufficiently high values of x_I: As $F\left(\frac{x_I+x_M}{2}\right) - F\left(\frac{x_I+x_E}{2}\right)$ goes to zero, entering becomes strictly preferred for E.

Next, I make the following modification to the second part of Assumption 3 for the case of candidate entry. Assumption A.1 ensures that M's probability of beating a moderate incumbent is sufficiently higher than E so that E prefers to stay out when x_I is low. This means that there exists an incumbent platform $\tilde{x}_I'$ that leaves the extremist indifferent between entering and staying out:

Assumption A.1. $\Delta_E(0) < 0$.[16]

Similarly, analogous to Assumption 4, Assumption A.2 guarantees this platform $\tilde{x}_I'$ is unique:

Assumption A.2. $\frac{d\Delta_E(x_I)}{dx_I} \geq 0$ *for* $x_I > \tilde{x}_I'$.

Let $\underline{x}_I'(p)$ be the platform that leads to the same probability of incumbent's reelection as $\tilde{x}_I'$, if such a point exists; otherwise let $\underline{x}_I'(p) = 0$. We can then prove an analogue of Lemma 1 in the main text: despite being more moderate, incumbents with platforms in the interval $(\underline{x}_I'(p), \tilde{x}_I')$ are reelected with a lower probability than incumbents with the platform $\tilde{x}_I'$. This is visualized in Figure 3.

Lemma A.1. *When Assumptions A.1 and A.2 hold, there exists an interval of platforms that result in the incumbent facing the moderate opponent and winning reelection with a lower probability than if she had the more extreme platform* $\tilde{x}_I'$ *and faced the extremist with probability* p.

This observation extends to positive costs of running for office. Consider first the case of p close to $1/2$. When $c > 0$ is small, there exists an incumbent platform that induces the extremist

[16] This holds whenever x_M is sufficiently low and F is sufficiently precise.

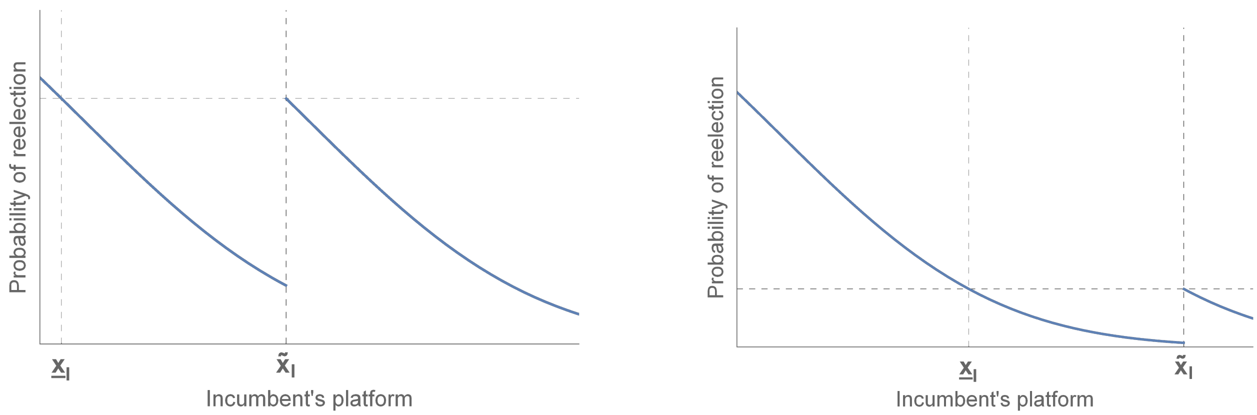


Figure 3: Incumbent's probability of reelection is plotted as a function of her platform. The median voter's ideal point is standard normally distributed. Losses are linear on the left plot and quadratic on the right. Parameter values are $x_E = -3$, $x_M = -0.5$, $p = 0.5$, $B = 5$, and $c = 1$.

opponent to enter the race, resulting in a higher probability of reelection than slightly more moderate platforms. 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 p is close to zero, meaning that the moderate is very likely to win a competitive primary. 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. It follows that when the primary field is slanted towards moderates E does not run, regardless of the incumbent's platform. The general election is held between the moderate challenger and the incumbent.

Finally, suppose there is an extremist advantage in the primary: p is 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 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. Against a sufficiently extreme incumbent, a strong primary advantage induces the extremist to enter the race — even if M also runs. That in equilibrium the extremist enters the race regardless drives the moderate out. In contrast, when M has a much

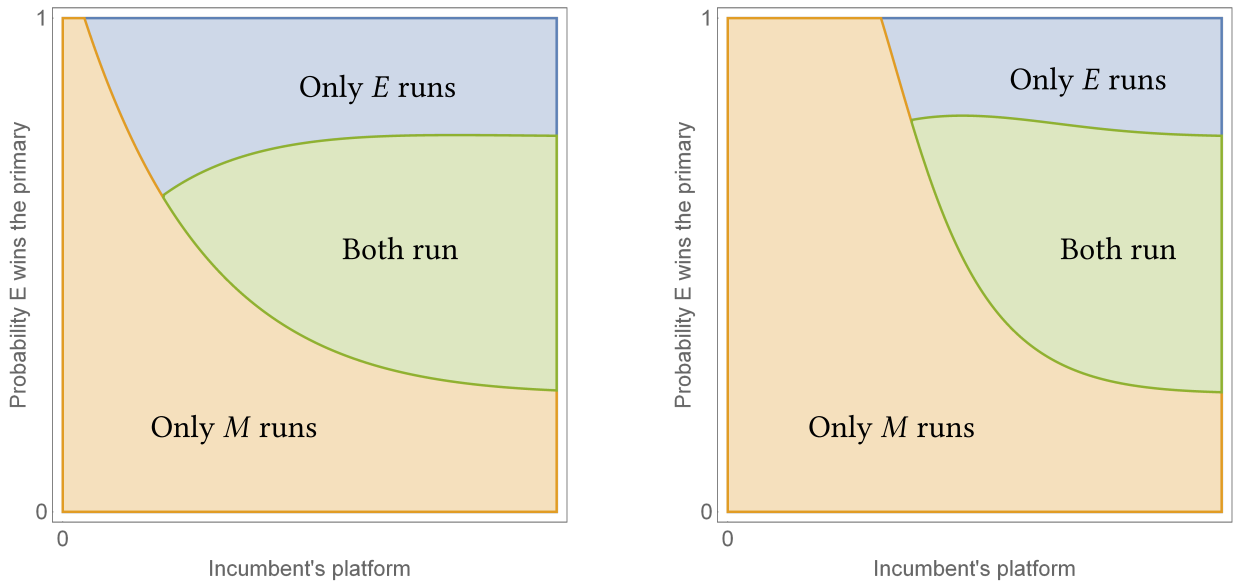


Figure 4: Orange (SW) and blue (NE) regions respectively correspond to parameter values where only M and E run in equilibrium. In green (E) both candidates run. The median voter's ideal point is standard normally distributed. Losses are linear on the left plot and quadratic on the right. Parameter values are $x_E = -1.5$, $x_M = -0.5$, $c = 1.2$, and $B = 4$. In both plots, M announces first. Plots for equilibria when E announces first are presented in the Appendix.

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, M runs if and only if E prefers to stay out when he is running, $\Delta_E(x_I) < \frac{c}{p}$. Thus, when there is a primary advantage for extremists, 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 A.1. *Suppose Assumptions A.1 and A.2 hold and the cost of running is small but strictly positive.*

1. *When neither the extremist nor the moderate has an advantage in the primary (p close to $1/2$), there exists an incumbent platform that forces a competitive primary. With probability p, E wins the primary, and the incumbent is reelected with a higher probability than if she had a more moderate platform and faced M for sure;*

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), there exists a platform for the incumbent that leads to her facing E in the general election for sure and winning reelection with a higher probability than if she had a more moderate platform and faced M.*

Proposition A.1 finds that against a sufficiently moderate incumbent, E stays out and M runs alone. When the primary 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. E only enters if the incumbent is sufficiently extreme. Finally, when the primary is slanted in favor of extremists, there cannot be a competitive primary. Against a sufficiently extreme incumbent, E runs alone, whereas M runs alone against a moderate incumbent. 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.[17]

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 this platform. 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. 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, as can be seen in Figure 5. 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.

Proposition A.2. *Suppose that either there is a primary advantage for extremists or there is no primary advantage for either candidate. Then, given the ideal points of the moderate and the extremist,*

[17]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 primaries as measured by vote share and probability of winning.

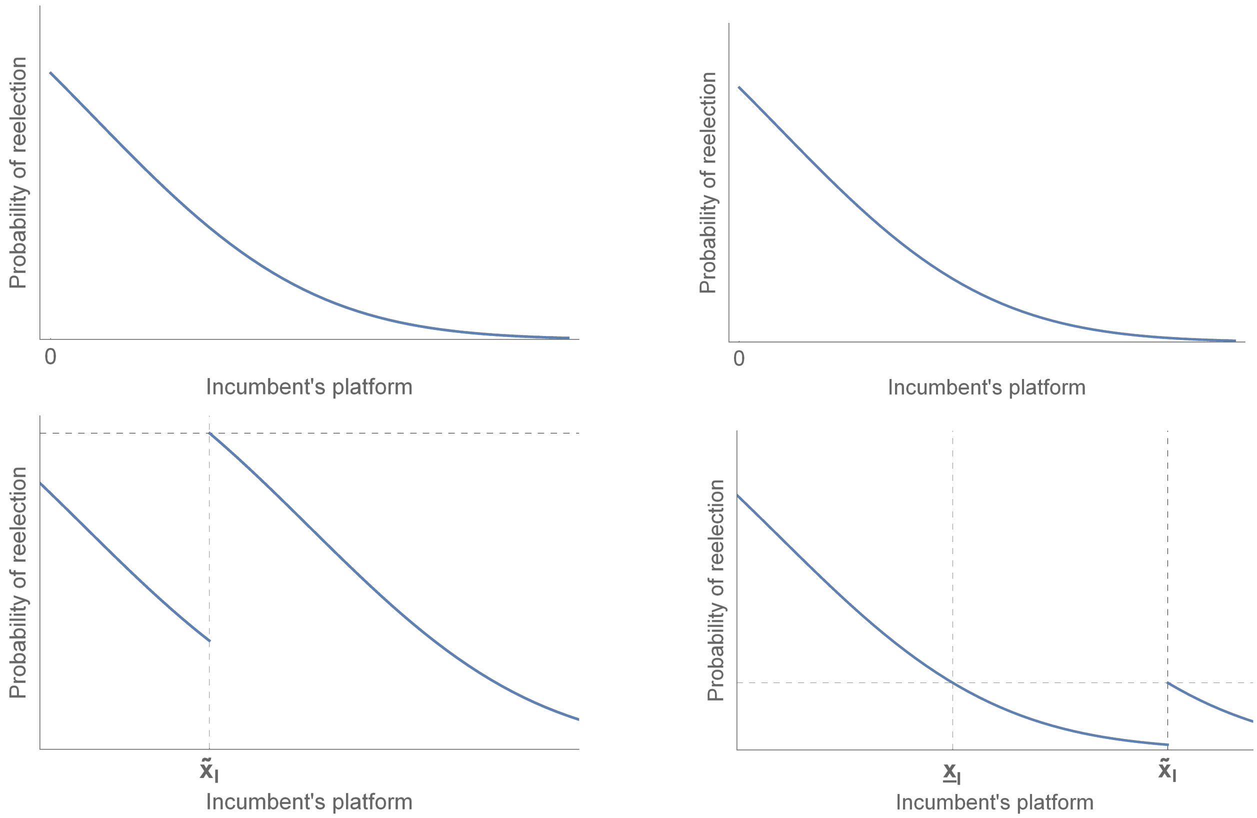


Figure 5: Incumbent's probability of reelection is plotted as a function of her platform. The median voter's ideal point is standard normally distributed. Losses are linear on the left plot and quadratic on the right. Parameter values are $x_E = -3$, $x_M = -0.5$, $p = 0.5$, $B = 5$, and $c = 1$. There is a primary advantage for moderates in the top plots ($p = 0.05$), and for extremists in the bottom ($p = 0.95$).

x_M and x_E, and the cost of running, c; there exist bounded intervals of office rents B and incumbent ideal points t such that the incumbent provokes the opposition by choosing platforms more extreme than her ideal point. Provoking the opposition cannot occur when there is a primary advantage for moderates.

Proposition A.2 shows that, similar to the model of primary elections, provoking the opposition is possible in a model of candidate entry. Specifically, moderate incumbents can benefit from hurting themselves electorally by decreasing their appeal to the general election median voter. 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. For incumbents with ideal points close to the threshold that induces extremist's entry in particular, provoking the opposition leads to a large enough boost to make up for the decreased appeal caused by going more extreme.

Appendix B Proofs

Proof of Proposition 1

Proof. I begin by proving that the optimal challenger platform for a primary voter must be between the voter's ideal point and the incumbent's platform:

Lemma B.1. *For primary voter i, the optimal challenger has an ideal point between his own and the incumbent's platform: $x_i \leq x^* < x_I$.*

Proof of Lemma B.1. Start by observing that the challenger cannot be someone whose platform primary voter i likes less than that of the incumbent, because a candidate whose platform is the same as i's ideal point always yields a higher expected payoff than such a candidate. Restricting attention to platforms i prefers to that of the incumbent: $|x_i - x_I| > |x_i - x|$, from Equation ((3))

$$\text{EU}_i(x, x_I) = -\ell(|x_i - x|)F\left(\frac{x_I + x}{2}\right) - \ell(x_I - x_i)\left(1 - F\left(\frac{x_I + x}{2}\right)\right).$$

The derivative of this is

$$\frac{d\,\text{EU}_i}{dx} = \text{sgn}(x_i - x)\ell'(|x_i - x|)F\left(\frac{x_I + x}{2}\right) + \frac{1}{2}f\left(\frac{x_I + x}{2}\right)(\ell(x_I - x_i) - \ell(|x_i - x|))$$

where $\text{sgn}(x - y) \coloneqq \frac{|x-y|}{(x-y)}$ for all $x \neq y$ and zero otherwise. Notice that the second term in this expression is always positive in this region. The first term is also positive, and thus i's payoff increasing in x for $x < x_i$ because $\frac{d\,\text{EU}_i}{dx} > 0$. This implies that $\text{EU}_i(x_i, x_I) > \text{EU}_i(x, x_I)$ for all $x < x_i$. Thus, as argued in the text we can restrict attention to $x \in [x_i, x_I)$.

Next, notice that the first order condition of the primary voter's problem is:

$$\frac{F\left(\frac{x_I + x^*}{2}\right)}{f\left(\frac{x_I + x^*}{2}\right)} = \frac{\ell(x_I - x_i) - \ell(x^* - x_i)}{2\ell'(x^* - x_i)}. \tag{6}$$

Both sides of Equation ((6)) are positive for $x \in [x_i, x_I)$. Log-concavity of f implies that the left-hand side is increasing in x. The right-hand side is decreasing in x because ℓ is increasing and convex. Thus, there can be at most one solution to Equation ((6)) in $x_i \leq x < x_I$. Notice also

$$\frac{d^2\,\mathrm{EU}_i}{dx^2} = \frac{1}{4} f'\left(\frac{x_I + x}{2}\right)\left(\ell\left(x_I - x_i\right) - \ell\left(x - x_i\right)\right) - f\left(\frac{x_I + x}{2}\right)\ell'\left(x - x_i\right) - F\left(\frac{x_I + x}{2}\right)\ell''\left(x - x_i\right).$$

Evaluating this expression at the solution of Equation ((6)) yields:

$$\frac{1}{2}\frac{f'\left(\frac{x_I+x^*}{2}\right)}{f\left(\frac{x_I+x^*}{2}\right)} - \frac{f\left(\frac{x_I+x^*}{2}\right)}{F\left(\frac{x_I+x^*}{2}\right)} - \frac{\ell''\left(x^* - x_i\right)}{\ell'\left(x^* - x_i\right)} < 0$$

because log-concavity of f implies for all x we have $f'(x)F(x) < f(x)^2$. Therefore, if x solves Equation ((6)), it is a maximum. $\square$

Thus, the median primary voter's optimal platform must be between his ideal point and the incumbent's platform: $x^*_{m_L} \in [x_{m_L}, x_I)$. 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 $x^*_{m_L}$ to any candidate whose ideology is to her right (left). Formally, a sufficient condition for median primary voter's pivotality is that we have $\mathrm{EU}_i(x^*_{m_L}, x_I) \geq \mathrm{EU}_i(x, x_I)$ for both all $x_i \leq x_{m_L}$ and $x \geq x^*_{m_L}$, and all $x_i \geq x_{m_L}$ and $x \leq x^*_{m_L}$. Notice that

$$\mathrm{EU}_i(x^*_{m_L}, x_I) \geq \mathrm{EU}_i(x, x_I) \iff \frac{\ell(x_I - x_i) - \ell(|x^*_{m_L} - x_i|)}{\ell(x_I - x_i) - \ell(|x - x_i|)} \geq \frac{F\left(\frac{x_I+x}{2}\right)}{F\left(\frac{x_I+x^*_{m_L}}{2}\right)}.$$

By the optimality of $x^*_{m_L}$, we know that for all x:

$$\frac{\ell(x_I - x_{m_L}) - \ell(x^*_{m_L} - x_{m_L})}{\ell(x_I - x_{m_L}) - \ell(|x - x_{m_L}|)} \geq \frac{F\left(\frac{x_I+x}{2}\right)}{F\left(\frac{x_I+x^*_{m_L}}{2}\right)}.$$

So a sufficient condition for the existence of a Condorcet winner is that for $(x_i - x_{m_L})(x - x^*_{m_L}) \leq 0$ we have

$$\frac{\ell(x_I - x_i) - \ell(|x^*_{m_L} - x_i|)}{\ell(x_I - x_i) - \ell(|x - x_i|)} \geq \frac{\ell(x_I - x_{m_L}) - \ell(x^*_{m_L} - x_{m_L})}{\ell(x_I - x_{m_L}) - \ell(|x - x_{m_L}|)}.$$

To show that the log-concavity of ℓ' is sufficient for the above equality to hold, we first need to prove the following lemma:

Lemma B.2. *Suppose that ℓ' is log-concave. Then, for any $x_i < x_0 < x_1 < x_2$:*

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

Proof of Lemma B.2. Suppose that ℓ' 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 ℓ', ℓ'', and ℓ''' respectively refer to the first, second, and third derivatives of ℓ. This in turn implies that the cross-partial of the logarithm of ℓ' with respect to any x and x_i is positive, because

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

where the last inequality follows from the log-concavity of ℓ'. This implies that for any $x_1 > x_0$:

$$\frac{\partial \ln(\ell'(x_1-x_i))}{\partial x_i} - \frac{\partial \ln(\ell'(x_0-x_i))}{\partial x_i} \geq 0 \iff \frac{\partial \ln\left(\frac{\ell'(x_1-x_i)}{\ell'(x_0-x_i)}\right)}{\partial x_i} \geq 0 \iff \frac{\partial \frac{\ell'(x_1-x_i)}{\ell'(x_0-x_i)}}{\partial x_i} \geq 0.$$

Define $s(x) = \frac{\ell'(x-x_j)}{\ell'(x-x_i)}$ for $x > x_i > x_j$. 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

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

where the inequality follows from the fact that $s(x)$ is lower everywhere it is evaluated in the numerator than everywhere in the denominator. Thus

$$\frac{\partial \frac{\ell(x_2-x_i)-\ell(x_1-x_i)}{\ell(x_1-x_i)-\ell(x_0-x_i)}}{\partial x_i} \geq 0 \iff \frac{\partial \frac{\ell(x_1-x_i)-\ell(x_0-x_i)}{\ell(x_2-x_i)-\ell(x_1-x_i)}}{\partial x_i} \leq 0.$$

Finally, add 1 to the above expression. Because this is a constant, the derivative does not change, and we get

$$\frac{\partial \frac{\ell(x_1-x_i)-\ell(x_0-x_i)}{\ell(x_2-x_i)-\ell(x_1-x_i)}}{\partial x_i} = \frac{\partial \frac{\ell(x_1-x_i)-\ell(x_0-x_i)}{\ell(x_2-x_i)-\ell(x_1-x_i)}+1}{\partial x_i} = \frac{\partial \frac{\ell(x_2-x_i)-\ell(x_0-x_i)}{\ell(x_2-x_i)-\ell(x_1-x_i)}}{\partial x_i} \leq 0.$$

□

We can now use Lemma B.2 to show that for $x \geq x^*_{m_L}$, we have

$$\frac{d \frac{\ell(x_I-x_i)-\ell(x^*_{m_L}-x_i)}{\ell(x_I-x_i)-\ell(x-x_i)}}{dx_i} \leq 0.$$

Similarly, for $x \leq x^*_{m_L}$ we have

$$\frac{d \frac{\ell(x_I-x_i)-\ell(x^*_{m_L}-x_i)}{\ell(x_I-x_i)-\ell(x-x_i)}}{dx_i} \geq 0.$$

It follows that

$$\frac{\ell(x_I-x_i)-\ell(|x^*_{m_L}-x_i|)}{\ell(x_I-x_i)-\ell(|x-x_i|)} \geq \frac{\ell(x_I-x_{m_L})-\ell(x^*_{m_L}-x_{m_L})}{\ell(x_I-x_{m_L})-\ell(|x-x_{m_L}|)}$$

for both $x_i \leq x_{m_L}$, $x \geq x^*_{m_L}$ and $x_i \geq x_{m_L}$, $x \leq x^*_{m_L}$.

Notice that the above arguments apply to any pair of platforms x and x'. Take any x and x' such that $\mathrm{EU}_i(x, x_I) \geq \mathrm{EU}_i(x', x_I)$. Then, by Lemma B.2, all primary voters on one side of the median primary voter also prefer x to x', because

$$\frac{\ell(x_I-x_i)-\ell(|x-x_i|)}{\ell(x_I-x_i)-\ell(|x'-x_i|)} \geq \frac{\ell(x_I-x_{m_L})-\ell(x-x_{m_L})}{\ell(x_I-x_{m_L})-\ell(|x'-x_{m_L}|)}$$

for all voters $x_i \leq x_{m_L}$ or $x_i \geq x_{m_L}$.

Thus, if the median primary voter prefers a more moderate candidate to a more extreme one, all primary voters who are more moderate than her also prefer the more moderate candidate.

Similarly, if the median primary voter prefers a more extreme candidate to a more moderate one, all primary voters who are more extreme also prefer the extreme candidate. It must then be that the median primary voter's preferred candidate is the Condorcet winner between any pair of candidates. □

Proof of Corollary 1

Proof. Replacing ℓ with either absolute or exponential loss in Equation (6) gives

$$\frac{F\left(\frac{x_I+x^*}{2}\right)}{f\left(\frac{x_I+x^*}{2}\right)} = \frac{\ell(x_I) - \ell(x^*)}{2\ell'(x^*)}. \tag{7}$$

This must have a solution with $x < 0$ by the log-concavity of f, that ℓ 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 ((7)) does not depend on x_i.

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

Proof of Proposition 2

Proof. 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. We need to show $F\left(\frac{x_I'+x^{**}}{2}\right) \geq F\left(\frac{x_I+x^*}{2}\right)$. Recall from Lemma B.1 that the optimal candidate must be in $[x_i, x_I)$. Notice first that if $x^* = x_i$, it must be that $x^* \leq x^{**}$, and the proposition follows immediately. Thus we only need to prove the proposition for $x^* \in$

(x_i, x_I), which means

$$F\left(\frac{x_I + x^*}{2}\right) 2\ell'(x^* - x_i) = f\left(\frac{x_I + x^*}{2}\right)(\ell(x_I - x_i) - \ell(x^* - x_i)).$$

must hold with equality. On the other hand, x^{**} could be on a corner or in the interior. $x^{**} \in [x_i, x_I)$ requires

$$F\left(\frac{x_I' + x^{**}}{2}\right) 2\ell'(x^{**} - x_i) \geq f\left(\frac{x_I' + x^{**}}{2}\right)\left(\ell(x_I' - x_i) - \ell(x^{**} - x_i)\right).$$

Suppose for a contradiction that $F\left(\frac{x_I'+x^{**}}{2}\right) < F\left(\frac{x_I+x^*}{2}\right)$. By the log-concavity of F,

$$\ell'(x^{**} - x_i)(\ell(x_I - x_i) - \ell(x^* - x_i)) > \ell'(x^* - x_i)(\ell(x_I' - x_i) - \ell(x^{**} - x_i)).$$

This requires either $\ell(x^{**} - x_i) > \ell(x^* - x_i)$, or $\ell'(x^{**} - x_i) > \ell'(x^* - x_i)$. Both of these imply $x^{**} > x^*$; former because ℓ is increasing, and latter because ℓ is convex. But, $x^{**} > x^*$ leads to a contradiction with the premises $x_I' > x_I$ and $F\left(\frac{x_I'+x^{**}}{2}\right) < F\left(\frac{x_I+x^*}{2}\right)$. Thus, it must be that $F\left(\frac{x_I'+x^{**}}{2}\right) \geq F\left(\frac{x_I+x^*}{2}\right)$. □

Proof of Lemma 1

Proof. 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 is $\Delta_{m_L}(x_I)$.

Recall that for all x_I, x_E, and x_M, we have $F\left(\frac{x_I+x_M}{2}\right) \geq F\left(\frac{x_I+x_E}{2}\right)$. Moreover, as $x_I \to \infty$, by Chebyshev's Inequality we know that the $\ell(x_I - x_{m_L})\left(F\left(\frac{x_I+x_M}{2}\right) - F\left(\frac{x_I+x_E}{2}\right)\right)$ term in $\Delta_{m_L}(x_I)$ goes to zero, meaning that $\lim_{x_I \to \infty} \mathrm{EU}_{m_L}(x_E, x_I) - \mathrm{EU}_{m_L}(x_M, x_I) = \ell(x_M - x_{m_L}) - \ell(|x_{m_L} - x_E|)$. If this term is negative, the median primary voter always prefers M to E, and E never wins the primary election. The first part of Assumption 3 in the main text ensures that against sufficiently weak incumbents the median primary voter prefers to vote for E.

To rule out the other case where the median primary voter always votes for E, notice that when the incumbent's platform is equal to zero, the median primary voter's net expected utility is

$$\Delta_{m_L}(0) = \left(\ell(-x_{m_L}) - \ell(|x_{m_L} - x_E|)\right) F\left(\frac{x_E}{2}\right) - \left(\ell(-x_{m_L}) - \ell(x_M - x_{m_L})\right) F\left(\frac{x_M}{2}\right).$$

It follows that against a very moderate incumbent, the median primary voter votes for M whenever the second part of Assumption 3 holds. Therefore when Assumption 3 holds 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. Assumption 4 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. Assumption 4 states that the derivative of the expected payoff of the median primary 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 3, and 4 hold, there is a unique incumbent platform $\tilde{x}_I$ that leaves the median primary voter indifferent between E and M. This platform satisfies $\Delta_{m_L}(\tilde{x}_I) = 0$. 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 $1 - F\left(\frac{x_I + x_E}{2}\right)$. Because $x_E < x_M$, we know that for all x_I, $F\left(\frac{x_I+x_E}{2}\right) < F\left(\frac{x_I+x_M}{2}\right)$. 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 see that the lower bound of this interval is $\underline{x}_I$, formally define it as $\underline{x}_I \coloneqq \ell^{-1}(\max\{0, \tilde{x}_I + x_E - x_M\})$, and notice that for any platform in the interval $x_I \in (\underline{x}_I, \tilde{x}_I)$, the challenger is M. Because $F\left(\frac{\tilde{x}_I+x_E}{2}\right) < F\left(\frac{x_I+x_M}{2}\right)$ for all $x_I \in (\underline{x}_I, \tilde{x}_I)$, the lemma obtains. □

Proof of Proposition 3

Proof. Consider an incumbent whose ideal point is $t \in \left(\frac{\underline{x}_I+\tilde{x}_I}{2}, \tilde{x}_I\right)$. Taking the derivative of incumbent's payoff in Equation ((5)) yields

$$-\frac{1}{2}f\left(\frac{x_I + x_J}{2}\right)(B - \ell(|x_I - t|)) - \operatorname{sgn}(x_I - t)\ell'(|x_I - t|)\left(1 - F\left(\frac{x_I + x_J}{2}\right)\right), \tag{8}$$

such that $J = E$ if and only if $x_I \geq \tilde{x}_I$. To eliminate potential regions and narrow the set of possible solutions, let us study this derivative separately in the following two regions: $x_I < \tilde{x}_I$ and $x_I \geq \tilde{x}_I$.

1. For $x_I \geq \tilde{x}_I$, expression ((8)) becomes negative:

$$-\frac{1}{2}f\left(\frac{x_I + x_E}{2}\right)(B - \ell(x_I - t)) - \ell'(x_I - t)\left(1 - F\left(\frac{x_I + x_E}{2}\right)\right).$$

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 $\tilde{x}_I$.

2. For $x_I < \tilde{x}_I$, we can rewrite expression ((8)) as

$$-\frac{1}{2}f\left(\frac{x_I + x_M}{2}\right)(B - \ell(t - x_I)) + \ell'(t - x_I)\left(1 - F\left(\frac{x_I + x_M}{2}\right)\right). \tag{9}$$

First, notice that for $x_I \in (t, \tilde{x}_I)$, this expression is always negative. This means that the optimal platform cannot be in this region. Thus, we can restrict focus to $x_I \leq t$.

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, t. The intuition is that when office rents are low and the incumbent is very likely to be reelected with her ideal point, she does not find it worthwhile to moderate her platform to improve her reelection chances. In contrast, for sufficiently high B, expression ((9)) may be negative for all $x_I \leq t$. This implies that the optimal platform for the incumbent is the one that maximizes her probability of reelection at zero. Here, 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

$$B - \ell(t - x_I^{\text{int}}) = 2\ell'(t - x_I^{\text{int}}) \frac{1 - F\left(\frac{x_I^{int} + x_M}{2}\right)}{f\left(\frac{x_I^{int} + x_M}{2}\right)}. \tag{10}$$

This interval can contain at most one interior solution. 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 ((10)), 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 improves her probability of reelection by moving to the center. Also notice that the cross-partial of the incumbent's payoff with respect to x_I and B is given by $-\frac{1}{2} f\left(\frac{x_I + x_M}{2}\right)$. It must then be that x_I^* is continuously decreasing in B. We can then define $b\colon (0, t) \to \mathbb{R}_+$ as a surjection that maps incumbent platforms to office rents B that make them optimal for the incumbent, subject to the constraint $x_I \leq t$.

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

Next, take $B = b(\hat{x}_I)$, where $\hat{x}_I \coloneqq 2t - \tilde{x}_I$. Again, by construction of b we have $\mathrm{EU}_I(\hat{x}_I) \geq \mathrm{EU}_I(x_I)$ for all $x_I \leq t$. Notice that $\hat{x}_I > \underline{x}_I$, which implies both that $b(\hat{x}_I) \leq b(\underline{x}_I)$ because b is decreasing, and that $\hat{x}_I$ leads to a strictly lower probability of reelection for the incumbent than $\underline{x}_I$. Because $\underline{x}_I$ 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 $\tilde{x}_I$ 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 $b(\underline{x}_I)$ and $b(\hat{x}_I)$ the incumbent's optimal strategy is to provoke the opposition by playing $\tilde{x}_I$. The first derivative of the incumbent's payoff with respect to B is $1 - F\left(\frac{x_I + x_M}{2}\right) \geq 0$. 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 $B \in [b(\hat{x}_I), b(\underline{x}_I)]$. □

Proof of Lemma A.1

Proof. Under Assumption A.1, all arguments from the proof of Lemma 1 involving the existence of an incumbent platform that leaves the extremist indifferent between entering and staying out carry through. When Assumptions A.1 and A.2 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 that 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 $pF\left(\frac{x_I+x_E}{2}\right)+(1-p)F\left(\frac{x_I+x_M}{2}\right)$, if $x_I \geq \tilde{x}_I'$, and $F\left(\frac{x_I+x_M}{2}\right)$ otherwise. In particular, when $x_I=\tilde{x}_I'$, the incumbent is reelected with probability

$$pF\left(\frac{\tilde{x}_I'+x_E}{2}\right)+(1-p)F\left(\frac{\tilde{x}_I'+x_M}{2}\right). \tag{11}$$

Because $F\left(\frac{\tilde{x}_I'+x_M}{2}\right) > F\left(\frac{\tilde{x}_I'+x_E}{2}\right)$ for all x_I, it follows by continuity of ℓ 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 $F\left(\frac{x_M}{2}\right)$. If this is less than the probability in expression ((11)), then $\tilde{x}_I'$ leads to the highest possible reelection probability. If it is larger, then by continuity there must exist a platform $\underline{x}_I'(p)$ such that

$$F\left(\frac{\underline{x}_I'(p)+x_M}{2}\right)=pF\left(\frac{\tilde{x}_I'+x_E}{2}\right)+(1-p)F\left(\frac{\tilde{x}_I'+x_M}{2}\right).$$

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

Proof of Proposition A.1

Proof. I prove each part of this proposition in the order they are presented. Throughout, I use $\overline{\Delta}_E \coloneqq \lim_{x_I\to\infty}\Delta_E(x_I)$, $\overline{\Delta}_M \coloneqq \sup\{\Delta_M(x_I)\}$, and $\underline{\Delta}_M \coloneqq \min\{\Delta_M(x_I)\}$. Let $2c < \min\{\underline{\Delta}_M, \overline{\Delta}_E\}$. We know that $\underline{\Delta}_M, \overline{\Delta}_E \in (0,\infty)$, 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 establish that Lemma A.1 extends to small positive costs of running, notice that because Δ_M is bounded away from zero, we can find some p that satisfies $p < 1-\frac{c}{\underline{\Delta}_M}$. This means that for such values of p, M always runs. Also notice that from Assumption A.1, $\Delta_E(0) < \frac{c}{p}$ immediately follows for any $c, p > 0$. Finally, because f has finite variance and $x_E < x_M$, we can find some $p > \frac{c}{\overline{\Delta}_E}$. Then, by continuity and Assumption A.2 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 $p \in \left(\frac{c}{\overline{\Delta}_E}, 1 - \frac{c}{\underline{\Delta}_M}\right)$, we have our result.

2. Let $\underline{p} = \frac{c}{\overline{\Delta}_E}$. Then, we have $\Delta_E(x_I) < \frac{c}{p}$ for all $x_I \geq 0$ and for any $p \in [0, \underline{p})$. This means that E never enters the race when M runs. Furthermore, for any $p \in [0, \underline{p})$ we have $\frac{c}{1-p} < \underline{\Delta}_M \leq \Delta_M(x_I)$ 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 $\overline{p} = 1 - \frac{c}{\overline{\Delta}_M}$, and suppose first that M announces his decision to run, followed by E. Then, for all $p \in (\overline{p}, 1)$ and x_I, we have $c < \underline{\Delta}_M < \overline{\Delta}_M < \frac{c}{1-p}$. This means that M enters if and only if E will not join him, and 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 $\Delta_E(x_I) > \frac{c}{p}$ for $p \in (\overline{p}, 1)$ sufficiently high x_I because $1 - \frac{c}{\overline{\Delta}_M} > 1 - \frac{c}{\underline{\Delta}_M} > \frac{c}{\overline{\Delta}_E}$. Furthermore, from Assumption A.1 it follows that $\Delta_E(0) < 0 < \frac{c}{p}$. By continuity it must be then for some intermediate values of $\tilde{x}_I'$ such that $\Delta_E(\tilde{x}') = \frac{c}{p}$. When the incumbent's platform is $\tilde{x}_I'$, E is indifferent between entering and staying out, and enters. It follows that for $p \in (\overline{p}, 1)$, 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.

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 $p \in (\overline{p}, 1)$ we have $c < \underline{\Delta}_M < \overline{\Delta}_M < \frac{c}{1-p}$, 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 $2c < \overline{\Delta}_E$. Again, from Assumption A.1 it follows that $\Delta_E(0) < 0 < c$. Then, by Assumption A.2 there exists a unique incumbent platform $\tilde{x}_I'$ such that when $p \in (\overline{p}, 1)$, 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.

□

Proof of Proposition A.2

Proof. When there is a primary advantage for extremists, we know from Lemma A.1 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 $\underline{x}_I \coloneqq \ell^{-1}(\max\{0, \tilde{x}_I' + x_E - x_M\})$, 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 $\tilde{x}_I$.

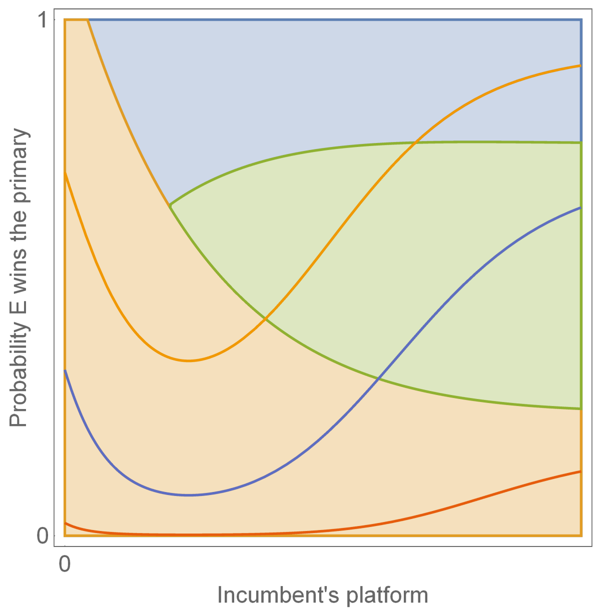
Suppose now there is no primary advantage for either side. Then, we know from Lemma A.1 that M always runs regardless of x_I, and that there exists a unique incumbent platform $\tilde{x}_I'$ such that E enters the race alongside M if and only if $x_I \geq \tilde{x}_I'$. When E enters, she wins the primary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

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. □

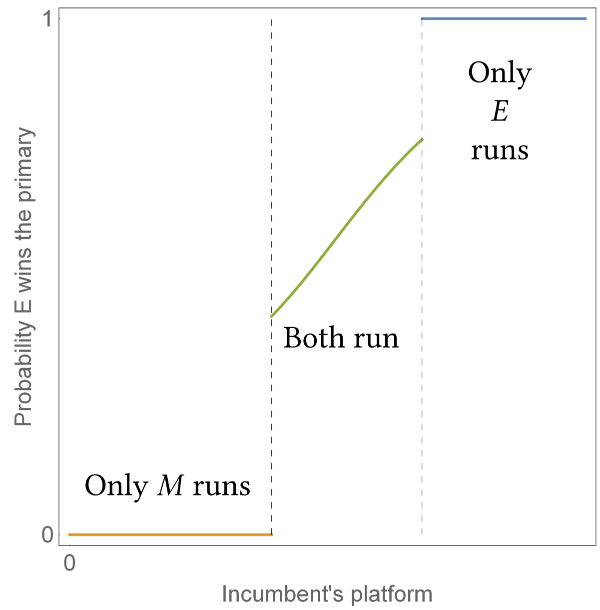
Appendix C Combining the two Models

Here, I present 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 for some parameter values, I assume here that there is also uncertainty about the ideal point of the median primary voter. Specifically, I assume that the ideal point of the median primary voter is drawn from some log-concave distribution with finite variance.

Figure 6a features solid lines that portray the probability the extremist candidate would win a competitive primary as a function of the incumbent platform for three different values of $\mathbb{E}[x_{m_L}]$; the regions correspond to which candidates run in the primary, as before. Thus, the overlap of a line and a region captures which candidates run for office for the corresponding interval



(a) Orange, blue, and red lines refer to the probability the extremist would win a *competitive* primary when the ideology of the median primary voter is drawn from a normal distribution with standard deviation 1, and means −1.3, −1.1, or −0.8 respectively.



(b) The lines refer to the *actual* probability the extremist wins the primary, taking into account candidates' entry decisions, as a function of the incumbent's position. The ideology of the median primary voter is drawn from $\mathcal{N}(-1.3, 1)$.

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. The median voter's ideal point is standard normally distributed and losses are linear. Parameter values are $x_E = -1.5$, $x_M = -0.5$, $c = 1.2$, and $B = 4$.

of incumbent platforms. An example of the induced probabilities of the incumbent facing the extremist as a function of her platform is produced in Figure 6b: the extremist stays out when the line is in the orange region, the moderate stays out when it is in the blue region, and there is a competitive primary — hence an interior probability — when the line is in the green region.

Simulations presented on Figure 6 show that against very moderate incumbents, primary voters support the moderate candidate. When the incumbent is more extreme, the probabilities M and E would beat her start converging. This leads the primary voters with ideal points sufficiently to the left to start supporting the extremist. 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: extremist primary voters and candidates 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 the general election, and the probability she wins the primary election. Thus, the effect of the incumbent's platform on the primary voters' calculus reinforces the extremist candidate's entry decision.

Appendix D Supplementary Figures

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.

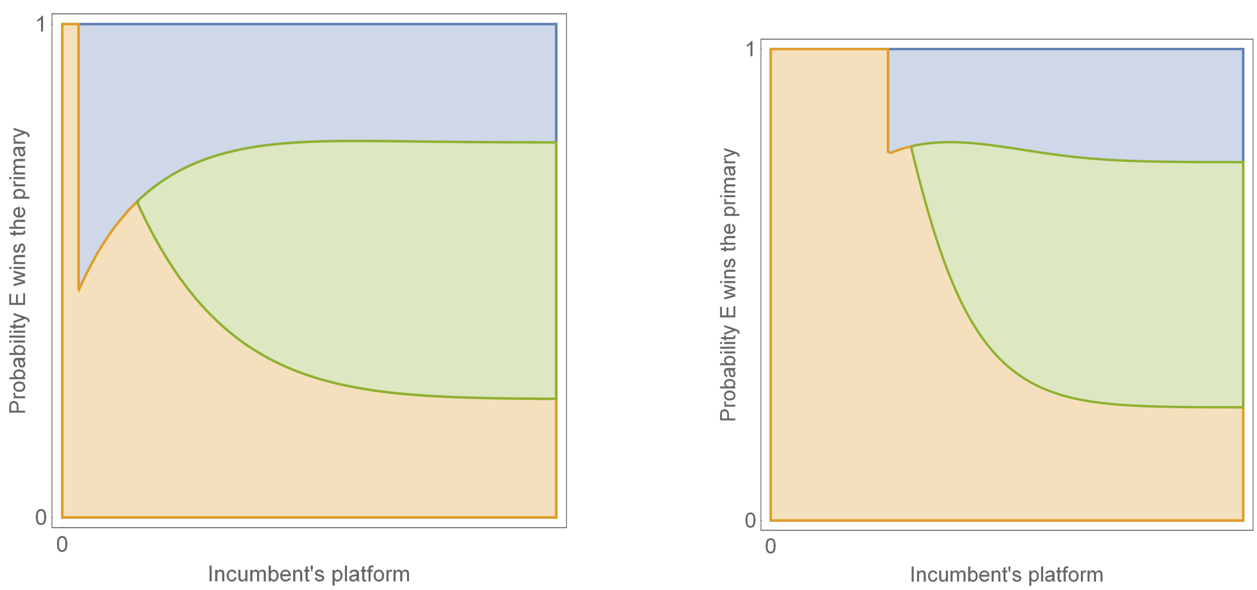


Figure 7: Orange (SW) and blue (NE) regions respectively correspond to parameter values where only M and E run in equilibrium. In green (E) both candidates run. The median voter's ideal point is standard normally distributed. Losses are linear on the left plot and quadratic on the right. Parameter values are $x_E = -1.5$, $x_M = -0.5$, $c = 1.2$, and $B = 4$. E moves first.