Optimal Timing of Policy Announcements in Dynamic Election Campaigns

Yuichiro Kamada and Takuo Sugaya*

March 27, 2019

Abstract

We construct a dynamic model of election campaigns. In the model, opportunities for candidates to refine/clarify their policy positions are limited and arrive stochastically along the course of the campaign until the predetermined election date. We show that this simple friction leads to rich and subtle campaign dynamics. We first demonstrate these effects in a series of canonical static models of elections that we extend to dynamic settings, including models with valence, a multi-dimensional policy space, policy motivated candidates, campaign spending, and incomplete information. We then present general principles that underlie the results from those examples. In particular, we establish that candidates spend a long time using ambiguous language during the election campaign in equilibrium.

*Kamada: Haas School of Business, University of California, Berkeley, Berkeley, CA 94720, e-mail: y.cam.24@gmail.com. Sugaya: Stanford Graduate School of Business, Stanford, CA, 94305, e-mail: tsugaya@stanford.edu; We thank Ernesto Dal Bo, Itay Fainmesser, Drew Fudenberg, Michihiro Kandori, Akitada Kasahara, Kei Kawai, Fuhito Kojima, David Laibson, John B. Londregan, Shih-En Lu, Francisco Martínez-Mora, Adam H. Meirowitz, Stephen Morris, Kristopher Ramsay, Andrzej Skrzypacz, Ken Shepsle, Yuki Takagi, Satoru Takahashi, and the seminar participants at the Harvard University Department of Economics and Department of Government, the Princeton University Department of Economics and Department of Political Science, University of California, Berkeley, Yale University, the 21st Stony Brook International Conference on Game Theory, and the Urrutia Elejalde Workshop on Information, Dynamics and Political Decision Making for helpful comments. In addition, Steven Callander and Josh Gottlieb read through the previous versions of this paper and gave us very detailed comments, which significantly improved the paper. Judith Levi provided a superb professional editorial assistance, which again significantly improved the paper. We thank Emily Her, Douglas Hong, Omair Butt, and especially Lingxuan Wu for their excellent research assistantship. This paper subsumes and supersedes part of our unpublished working paper "Valence Candidates and Ambiguous Platforms in Policy Announcement Games" and its earlier version circulated as "Policy Announcement Game: Valence Candidates and Ambiguous Policies," which Sugaya first presented at the 21st Stony Brook International Conference on Game Theory in 2010.

"Timing is Everything"

—Joe Slade White, the 2014 National Democratic Strategist of the Year

1 Introduction

Election campaigns are inherently dynamic. Joe Slade White, the 2014 National Democratic Strategist of the Year, states "timing makes the difference between winning and losing," in one of his "9 Principles of Winning Campaigns" (White, 2012). Despite the apparent importance of election campaigns on the electoral outcomes and the fact that the campaigns are dynamic in nature, there seem to be no theoretical models of dynamic campaigns in the literature, to the best of our knowledge.[1] One possible reason is that there is no obvious way to model campaigns in a way that would give rise to nontrivial dynamic strategic considerations.[2] The objective of this paper is to fill this gap by proposing a model in which candidates face nontrivial dynamic strategic considerations.

The paper proposes a "policy announcement timing game" in which candidates strategically choose the optimal timing of their policy announcements over a campaign period. Each announcement corresponds to restricting the set of available policies—that is, each candidate may clarify a policy to implement, while she cannot go back to a policy that she has ruled out before—and the final policy announcements before the predetermined election day determine the result of the election.

In our model, opportunities for policy announcements are limited and stochastic. Specifically, we assume that opportunities arrive according to a Poisson process over a campaign period. The assumption of Poisson opportunities is a simple way to represent frictions present in the communication process between candidates and voters. For example, administrative procedures to obtain internal approval for a change of how candidates announce their policies may not always be successful, or candidates may not always be able to communicate with the voters about such a change even if these procedures go through. Moreover, voters may not be convinced that the candidate has changed her policy position.[3] Those frictions cause uncertainty regarding the availability of

[1] By a model of dynamic election campaigns, we mean a model with a single election; in particular, when we speak of "models of dynamic election campaigns," we are excluding models that have primaries and the general election.

[2] The empirical research shows that candidates do react to each other during election campaign (cf. Banda (2013, 2015)).

[3] A richer modeling of administrative procedure or dynamics of voter beliefs would generate a more accurate prediction, but we assume these away and try to concentrate on key effects by investigating what we can say in our

future opportunities, and we capture such uncertainty by using Poisson processes.[4]

After laying out the general model, we present a number of applications to demonstrate that introducing this simple friction to the model generates interesting dynamic strategic considerations and equilibrium dynamics consistent with election dynamics in reality. The first issue we consider is ambiguous policy announcements, which we often see in real election campaigns. For example, in the context of a US presidential election, Nicholas Biddle, the manager of William Henry Harrison's campaign for the US presidency in 1840-1841, advised Harrison in these words: "Let him say not one single word about his principles, or his creed - let him say nothing - promise nothing. Let no Committee, no convention—no town meeting ever extract from him a single word, about what he thinks now, or what he will do hereafter."[5]

We find that our model leads to a new interpretation of ambiguous policy announcements: In our applications, if candidates are purely office-motivated and there is no Condorcet winner in the set of available policies, then the candidates have tentative preferences for the ambiguous policy statement during the course of the election campaign and, in fact, spend most of the campaign time keeping their policy statements ambiguous—not announcing a specific policy.[6] The incentive comes from a dynamic consideration. The candidates' announcements remain ambiguous because the absence of a Condorcet winner implies that it is less favorable to be the first mover than to be the second mover.

There are two leading examples with such "first mover disadvantage": Candidates have valence, or the policy space is multi-dimensional. When there is valence (Section 3.1), we show that, in equilibrium, the weak candidate will not make his policy clear in the early stages of the election campaign, and his policy announcement can possibly occur only close to the election date. This is because if he clarifies his policy too early, then the strong candidate will have enough time to simply copy that policy afterward, so that the weak candidate will certainly lose. The result may

simplest framework. As it turns out, the results from our simple setting are quite rich.

[4]Calvo (1983) uses a Poisson process to model uncertainty about future opportunities of changing prices. This approach offers a tractable way of modeling sticky prices and analyzing the effect of fiscal and monetary policies. At the same time, the literature goes forward to offer a micro-foundation of Calvo (1983) by fixed costs of changing prices, rational inattention, and so on (see Klenow and Malin (2010)). Here, we also show that this Poisson approach is useful to analyze campaign dynamics, and leave micro-foundation to the future research.

[5]McGrane, Reginald Charles C. (1919). The quote appears in Shepsle (1972).

[6]In contrast, we show that if candidates are purely office-motivated and there is a Condorcet winner, each candidate announces it as soon as possible. The formal definition of Condorcet winner for this result will be provided in Section 4.2.

explain the dynamics of the election campaign in the 2014 gubernatorial election for Tokyo, Japan, in which Yoichi Masuzoe and Morihiro Hosokawa fought a close campaign. Although Masuzoe had been seen as the strongest candidate from the outset of the campaign, Hosokawa became popular near the election day when he clarified his stance by announcing opposition to the restart of nuclear power generation. Then Masuzoe, who originally had not specified his policy about nuclear power generation, clarified his position to aim for a gradual phase-out of nuclear power. As a result, Masuzoe won against Hosokawa.[7]

When the policy space is multi-dimensional (Section 3.2), generically there does not exist a Condorcet winner (so there does not exist a pure-strategy Nash equilibrium in the static environment). In our policy announcement timing game, however, we can pin down both the equilibrium probability distribution of times at which candidates make policy announcements and winning probabilities even in such a setting. In the absence of a Condorcet winner, once a candidate commits to any policy platform and then the opponent optimally responds to it, the former candidate will lose. Hence, each candidate, upon making an announcement, "becomes a weak candidate" in that being best-responded afterward will bring the worst outcome. In contrast, the payoff structure is such that, if candidates knew that the current opportunity is the last one and there will be no opportunity for either candidate in the future, they would prefer to clarify their policy. Hence, in equilibrium, when a candidate obtains an opportunity near the election day, they clarify their policies.

This analysis on the multi-dimensional policy space, however, does not give us a precise prediction regarding the policies that candidates announce due to its excessive simplicity of pure office motivation. To show that such indeterminacy is not a consequence of the way our general dynamic model is specified, Section 3.3 introduces policy motivation to the model with a multi-dimensional policy space. Again, we show that ambiguous language is used for a long time in equilibrium, and pin down the policies that candidates announce. Interestingly, in equilibrium, a candidate may announce a policy that is Pareto inefficient among both candidates with positive probability. The reason is that announcing such a policy will make it incentive compatible for the other candidate to announce a policy that is not too unfavorable for the candidate in the event that the other can-

[7] Sankei News (2013) argued on December 24, 2013 that Masuzoe was seen as the strongest among the candidates, Asahi Shimbun Digital (2014a) reported on January 9, 2014 that Hosokawa clarified his policy about nuclear power, and Asahi Shimbun (2014b) reported on January 15, 2014 that Masuzoe showed support to the opposition to nuclear power.

didate obtains a chance of a policy announcement afterward. Announcing a policy position with a motive to influence the opposition's policy, though possibly sounding unrealistic, did actually happen in real campaigns. During the Democratic Party presidential primaries in 2016, for example, the far-left Bernie Sanders called for a $15-an-hour minimum wage (more than twice as much as the $7.25 standard back then) and Medicare-for-all health care, and proposed to end TPP. After losing Pennsylvania, Maryland, Delaware, and Connecticut in a row, Sanders declared in a town hall meeting: "But if we do not win, we intend to win every delegate that we can so that when we go to Philadelphia in July, we are going to have the votes to put together the strongest progressive agenda that any political party has ever seen."[8] An article in Vox (Stein, 2016) writes: "Bernie Sanders moved Democrats to the left. The platform is proof. [...] Hillary Clinton may have won the Democratic Party's presidential nomination, but Bernie Sanders has still left an outsize mark on its future."[9]

Another topic which attracts much attention with regard to campaign dynamics is political campaign advertisements (Section 3.4). We provide a simple model in which we reinterpret our policy announcement timing game to encompass dynamic spending in election campaigns. To make the reinterpretation work, we notice that the cumulative spending for advertisements cannot decrease over time; hence any spending is only restricting the set of possible cumulative spending. Supposing that the probability of winning the election only depends on the ratio of the cumulative spending of the two candidates and money can be used for purposes other than the campaign as well and is sufficiently important, we show that, in equilibrium, candidates do not spend as much money as they can in the early stages of the campaign, and make additional spending close to the election date if they can. The intuition is as follows. If two candidates spend as much as possible, both candidates will have a 50% chance of winning (given symmetry). There is little incentive to spend a lot at the early stages because that means that the opponent can cancel out its effect by later spending equally much with a high probability. Rather, candidates would save their money in the early stages and try to spend them later, leaving only a small probability for the opponent to cancel it out later. Our prediction provides a novel explanation for the empirical evidence, which

[8]The facts and the quote appear in Strauss (2016) and Gurciullo and Debenedetti (2016).

[9]The specific interpretation we give to the policy space we study may not be consistent with this episode of Sanders vs. Clinton. We provide this example to make a point that an entry to a policy platform can happen with a motive to influence the opposition's platform.

suggests that candidates often spend monetary resources gradually over time.[10]

Although the general model is set up as a complete information game, one can extend it to include a wider class of settings. To illustrate, our final application allows for incomplete information by considering a model in which candidates' types can be either normal or extreme, and their preferences are such that it is a dominant strategy for the normal types to always announce a median position, while each extremist prefers a non-median policy and dislikes the bliss policy of the extreme opposition (Section 3.5). We first show that, if it is common knowledge that both candidates are extreme, then candidates would be indifferent between announcing their policy as the median and using ambiguous language. Then we show that, if there is a possibility of the candidates being normal, then in any symmetric equilibrium the extreme candidates keep being ambiguous for a long time over the course of the election campaign although the belief that the opposition is extreme can become very close to (but less than) one when the campaign is long. An example of the situation where this model can potentially fit is the Japanese House of Councillors election in 2014, in which Prime Minister Shinzo Abe avoided making the constitutional reform the main issue of the election and did not specify his plan of how to reform the constitution.[11] Nonetheless, the press argued that he had a particular preference for reforming the constitution such as specifying a foundation of the Self Defense Forces.[12] In fact, he started the process of summarizing the issues about the constitutional reform centered around the Self Defense Forces, once he won the election (Ota, 2016).

After discussing the applications, we present general principles that underlie the results from those applications (Section 4). The general results are concerned with the main topics in the applications, i.e., ambiguity, Condorcet winner, and office-motivated candidates, and show that the implications of our model are more robust than being valid in specific examples. First, Section 4.1 presents a result that we call the *long ambiguity theorem*. We formally define a "first-mover disadvantage" condition and show that, under that condition (together with some genericity conditions),

[10] Gerber et al. (2011) empirically show that the effect of campaign spending declines over time (earlier spending has a weaker effect). Although it might be reasonable to assume such depreciation, and it might result in gradual spending, we do not assume depreciation. Our main point is that the candidates have incentives to spend gradually even if the effect of depreciation is absent. Depreciation would strengthen the incentive to spend gradually.

[11] To reform the constitution, no less than two thirds of the members of the parliament have to agree on the reform at both the House of Representatives and the House of Councillors.

[12] For example, an article titled "Shinzo Abe's Constitution Quest" is published in Wall Street Journal (Harris, 2013).

candidates spend most of the time keeping their policy statements ambiguous in equilibrium, provided the campaign period is long enough. This result generalizes the results for valence candidates and a multi-dimensional policy space.

Second, Section 4.2 examines the robustness of the celebrated median-voter theorem to an extension to our dynamic setting. We show the *dynamic median-voter theorem*. Specifically, we consider a general model in which there is a Condorcet winner in the static version of the model, and show that each candidate makes an announcement corresponding to a Condorcet winner as soon as possible. This result generalizes results under some parameter specifications in our applications where the long ambiguity theorem fails. For example, we show that candidates announce their policies as soon as possible in the absence of valence, and spend their money as soon as possible if there are only two levels of spending. These results are implied by the dynamic-median voter theorem. The third general result presented in Section 4.3 pertains to the cases where candidates are purely office-motivated. Specifically, we analyze a general model with constant-sum payoffs, and prove that any perfect Bayesian equilibrium has the Markov property. We call this the *constant-sum Markov theorem*. The theorem generalizes the results from the model with valence and the one with multi-dimensional policy space. It is also used in proving a result for the model with incomplete information.

Section 5 concludes. The Appendix provides main proofs for the general results presented in Section 4. All the proofs not provided in the main text or in the Appendix are provided in the Online Appendix.

1.1 Literature Review

Ambiguity:

Ambiguous policy announcements have long been discussed in the politics and economics literature. The mechanism that generates ambiguity in our model is starkly different from those presented in the existing literature. For example, Shepsle (1972) and Aragonès and Postlewaite (2002) assume that candidates choose their policy positions simultaneously and once and for all. In their models, ambiguity occurs because voters are assumed to possess convex utility functions and therefore prefer ambiguity.[13] In our model, in contrast, ambiguity arises from dynamic strategic in-

[13] Callander and Wilson (2008) also consider a simultaneous-move voting game, and show that candidates' policy

teractions in an election campaign: Each candidate's strategic concern about the opponent's future play causes ambiguity. In particular, we do not assume convexity; rather, in one of the variants of our model discussed in the Online Appendix, we show that ambiguity occurs even when voters have concave utility functions.

Page (1976, 1978) proposes a theory that attributes ambiguity to the fact that candidates have limited resources to make their policy positions precise, and to voters' limited capacity to understand these positions. In our model, however, voters are capable of understanding what the candidates are announcing. Candidates do have a positive probability of not being able to have any chance to make a policy announcement, but we obtain ambiguity even in the limit as this probability shrinks to zero.

Glazer (1990) argues that ambiguity may occur if candidates are uncertain about the median voter's preferences. In his model, fixing a candidate's opponent's announcement, the candidate would prefer ambiguity. In our model, in contrast, our applications include cases where ambiguity is a suboptimal static response for any fixed announcement by the opponent. In other words, we obtain ambiguity due to dynamic strategic consideration.

Alesina and Cukierman (1990) and Aragonès and Neeman (2000) show that ambiguity occurs in elections if candidates prefer to keep the freedom to choose their policies after being elected, even though voters would prefer their candidates to commit themselves to precise policies before the election. That is, the driving force of ambiguity is different from office motivation. In contrast, the long ambiguity theorem in our model can be obtained with pure office motivation.

When the selection of candidates consists of more than one step, as is true for the US presidential election with its primaries and general elections, Meirowitz (2005) shows that candidates announce ambiguous policies in earlier stages if voter preferences are unknown at the beginning but are revealed by the result of the earlier stages. In our model, no new information arrives about voter preferences, and ambiguous policies are purely the result of strategic interactions between candidates.[14]

statements are ambiguous in equilibrium if voters' voting behavior is based on preferences exhibiting a taste for ambiguity due to their context dependence.

[14]Alesina and Holden (2008) show that candidates announce ambiguous policies even without primaries if (i) candidates have policy motivation, (ii) the policy motivation is their private information unknown to the voters, and (iii) campaign contributions from the voters to the candidates affect the electoral outcomes. In contrast, none of these assumptions are necessary to obtain the long ambiguity theorem in our model.

Valence: In the standard simultaneous-move Hotelling-Downs model with valence candidates, there exists no pure-strategy equilibrium: The strong candidate always wants to copy the weak candidate's policy, while the weak candidate does not want to be copied, just as in the "matching pennies" game. There are two approaches to addressing this issue in the literature. The first approach is to assume that the strong candidate is the incumbent and the weak candidate is the entrant (Bernhardt and Ingberman (1985), Berger et al. (2000), and Carter and Patty (2015)). In this approach, a typical result is that the strong candidate positions her policy close to the median voter and the weak candidate positions his policy at a slight distance from the strong candidate's policy, where the distance between the two policies is determined by the degree of asymmetry between candidates' valences.[15] The second approach is that of Aragonès and Palfrey (2002), who consider the simultaneous-move game seriously and characterize a mixed equilibrium.[16] They show that the strong candidate assigns high probabilities to the platforms which are close to the location of the median voter with high probabilities while the weak candidate assigns small probabilities to such platforms. Although these two approaches give us an understanding of what the equilibrium behavior looks like in an electoral situation with valence candidates, in both these models the order of policy announcements is exogenously given by the modelers. In contrast, we view our model with valence candidates as *endogenizing* the order of policy announcements.[17]

Multi-dimensional policy spaces: It is well known that the Downsian model with a multi-dimensional policy space does not have a pure-strategy Nash equilibrium unless a strong assumption about symmetry of the distribution of voters over the policy space is satisfied. As in the case with valence, one way to respond to the nonexistence is to consider a sequential game where the incumbent moves first and the challenger moves second. However, as Roemer (2001) argues, there may be no natural order, and we again view our approach as endogenizing the order of moves. Other approaches to deal with the nonexistence include that of Lindback and Weibull (1987), who allow the voters' behavior to be probabilistic and derive a sufficient condition for the existence of a pure-strategy equilibrium in a one-shot simultaneous-move game (see also Coughlin (1992)), and

[15]See also Ansolabehere and Snyder (2000) and Groseclose (2001) who consider pure-strategy equilibria in models with valence candidates.

[16]More specifically, Aragonès and Palfrey (2002) characterize the unique equilibrium in a discrete policy space and consider a limit as the discrete space approximates the standard continuous policy space.

[17]This provides a possible answer to the question posed by Aragonès and Palfrey (2002), who ask "What is the correct sequential model."

that of Roemer (2001), who obtains existence using a weaker equilibrium concept in which the set of feasible deviations for each candidate is restricted. In contrast to these papers, our analysis in Section 3.2 keeps the basic structure of the Downsian model.

Campaign spending: Most of the models in the theoretical literature about campaign spending specify how candidates use campaign funds in order to affect voters' behavior.[18] For example, Potters et al. (1997), Prat (2002a,b), and Coate (2004) consider models in which political campaigns can signal the candidate's private information. Also, Bailey (2002) assumes that one candidate chooses the policy position prior to the other, and that contributions can be used to target the campaign at selected people. In the current paper, in contrast, we are agnostic about why campaign spending helps, and focus on the timing of spending. There is also a large strand of empirical literature that analyzes the timing of the campaign spending. We refer interested readers to Gerber et al. (2011) and the reference therein.

Incomplete information: If candidates announce their policies simultaneously and the median voter exists, then it is a unique Nash equilibrium that both candidates announce the policy corresponding to the median voter given that the policy announcement is a credible commitment, regardless of policy preferences or knowledge about them. In models with incomplete information about the candidates' policy preferences, Banks (1990) and Harrington (1992) consider the case in which the policy announcement is not a credible commitment, while there is a cost of implementing a policy different from the announcement. Such a setting is later used by Kartik and McAfee (2007) and Callander and Wilke (2007) to analyze the incentive of telling a lie in elections. In our model, the policy announcement is a credible commitment, while its timing is endogenous (all the papers cited here assume exogenous (simultaneous) timing).

Dynamic games: To formally model the dynamics of policy announcements, we employ a framework with continuous time, a finite horizon, and a Poisson revision process. This modeling device has been extensively explored recently. The revision games in Kamada and Kandori (2017a,b) and Calcagno et al. (2014) consider settings in which players obtain opportunities to revise their preparation of actions according to Poisson processes, and the finally-revised action profile is implemented at the predetermined deadline.[19] In the models of these papers, revisions

[18]In addition, there are papers about the interaction between the lobbyists and politicians. See, for example, Austen-Smith (1987), Baron (1994), and Grossman and Helpman (1996).

[19]Ambrus and Lu (2015) consider a bargaining model in a similar fashion.

of actions are not restricted, in the sense that players can freely choose their actions from a fixed action space at each opportunity to move, as opposed to our assumption that once candidates make their policy platform clear, they cannot change it afterward.

Given the nature of the game analyzed—an election campaign where there is a clear winner and loser—some of our analysis pertains to constant-sum games. While the aforementioned papers mainly consider the situations where cooperation or coordination is important, Gensbittel et al. (2017) analyze general zero-sum revision games in which revisions are not restricted. Our long ambiguity theorem is similar to the "wait and wrestle" property that they find. The difference is that we do not restrict ourselves to constant-sum games, and we consider the case where revisions are restricted.[20]

The policy announcement timing game can be regarded as a stochastic game. Lovo and Tomala (2015) analyze general revision games with payoff-relevant states and show existence of Markov perfect equilibria.[21] In contrast, our focus is on the unique prediction of players' behavior in perfect Bayesian equilibria in the context of election campaigns.

We use a Poisson process to model frictions in the election campaign. Another way to model such frictions is to introduce switching costs. In general, switching costs result in different implications on equilibrium behavior from a Poisson process. See Lipman and Wang (2000) and Caruana and Einav (2008) for models with switching costs in finite-horizon games.

As for the idea of using ambiguous language or not spending their funds in expectation of future events, Gale's (1995, 2001) model of "monotone games" also considers a similar problem. In his model, at each period, players can only (weakly) increase their actions. In effect, players commit to a smaller and smaller subset of their action spaces as time passes, and they will never be able to "expand" that subset (thus, the revisions are restricted). The main difference is that he analyzes "games with positive spillover" played over an infinite horizon and shows that collusive outcomes can be achieved, while we analyze a game with a conflict of interests played over a finite horizon and are interested in uniqueness of an equilibrium outcome.

Another related theoretical literature is about commitment games, where each player simulta-

[20]Note that, although restricted revisions imply that there is no cycling choice of actions as in Gensbittel et al. (2017), it is still not trivial that candidates wait for a long time. Gensbittel et al. (2017) also discuss a comparison between the two models.

[21]Moroni (2018) also provides an existence proof for revision games, allowing for imperfect and incomplete information.

neously commits to a subset of the entire set of actions at the first stage and then plays the game with the restricted set of actions at the second stage (see, for example, Hamilton and Slutsky (1990, 1993); van Damme and Hurkens (1996); Romano and Yildirim (2005); Renou (2009)). These models sometimes yield multiple equilibria. In our analysis, in contrast, the revision opportunities arrive stochastically and asynchronously, and as a result, we obtain an (essentially) unique equilibrium prediction.

2 The Model – Policy Announcement Timing Game

There are two candidates, A and B. Whenever we say candidates i and j, we assume $i \neq j$.[22] There is a set of policies, X. For each candidate $i = A, B$, there is a collection of nonempty subsets of X, denoted $\mathcal{X}_i \subseteq 2^X \setminus \{\emptyset\}$, with a property that $X \in \mathcal{X}_i$ for each $i = A, B$. Each element in $\mathcal{X}_i$ is called i's "policy set." Here, we interpret announcing X as announcing the "ambiguous policy" while announcing other sets in $\mathcal{X}_i$ is seen as (at least partially) specifying a policy platform. Given a profile of policy sets $(X_A, X_B) \in \mathcal{X}_A \times \mathcal{X}_B$, let $v_i(X_i, X_j)$ be candidate i's payoff for each $i = A, B$.

In our policy announcement timing game, time flows continuously from $-T < 0$ to 0. Imagine that 0 is the fixed election date and the campaign starts at $-T$. For each $-t \in [-T, 0]$, according to the Poisson process with arrival rate $\lambda_i > 0$, each candidate $i = A, B$ obtains opportunities to announce her policy set. We assume that the Poisson processes are independent between the candidates. In particular, this implies that policy announcements are asynchronous with probability one. To simplify the exposition, we often use "enter" to denote the act of announcing a singleton set. The result of the election only depends on (X_A, X_B), where X_i with $i \in \{A, B\}$ is candidate i's most recently announced policy set at time 0 (the election date).

In what follows, we analyze perfect Bayesian equilibria of this game. To formally define strategies in our setting, we first define history. A **history** for candidate i is denoted by

$$\left(\left(t_i^k, X_i^k\right)_{k=1}^{k_i}, \left(t_j^l, X_j^l\right)_{l=1}^{l_j}, t, z_i\right),$$

where $-T \leq -t_i^1 < \ldots < -t_i^{k_i} < -t$; $X_i^k \in \mathcal{X}_i$ for all k; $-T \leq -t_j^1 < \ldots < -t_j^{l_j} < -t$; $X_j^l \in \mathcal{X}_j$ for all l; and $z_i \in \{yes, no\}$. The interpretation is that $-t_i^k$ is the time at which candidate i

[22] For ease of exposition, we use feminine pronouns to refer to A and i and masculine pronouns to refer to B and j.

receives his or her k'th revision opportunity, and X_i^k is the policy set that i has chosen at time $-t_i^k$. We assume that candidate i cannot observe whether candidate j receives an opportunity, but can observe candidate j's choice of a policy set whenever it changes.[23] That is, t_j^l is the l'th time that candidate j changes his or her policy set from the previous one, and X_j^l is the policy set that j has chosen at time $-t_j^l$. We let $X_i^0 = X_j^0 = X$, that is, the policy set at time $-T$ is exogenously given to be X. The third element t denotes the current remaining time, and the indicator z_i expresses whether there is an opportunity for candidate i at time $-t$. By $H_i^{k_i,l_j}$, we denote the set of histories in which candidate i for $i = A, B$ has received k_i opportunities in the past and in which candidate j has changed policy sets l_j times. The set of all histories for candidate i is $H_i := \bigcup_{k_i=0}^{\infty} \bigcup_{l_j=0}^{\infty} H_i^{k_i,l_j}$.

A **strategy** for candidate i is denoted by $\sigma_i : H_i \to \Delta(\mathcal{X}_i)$, with three restrictions: First, $\sigma_i(h_i) = X_i^{k_i}$ where k_i is specified in the first element of h_i if the fourth element in h_i specifies $z_i = no$. That is, if there is no opportunity at $-t$ for i, then for notational convenience, we specify that the candidate takes the same policy set as specified in the last opportunity. Second, if $z_i = yes$, then the strategy $\sigma_i(h_i)$ must assign probability zero to $X_i \in \mathcal{X}_i$ if $X_i \not\subseteq X_i^{k_i}$. Thus, the set of candidate i's possible announcements at time $-t$ depends on i's own past policy announcement: If i has already announced $X_i \in \mathcal{X}_i$ in the past, then i can only announce a (weak) subset of X_i. Thus, once a candidate rules out some of the potential platforms, then she cannot go back to them later. The third requirement is technical. To guarantee that candidates' payoffs are integrable with respect to the distribution of the final outcome given the strategy profile, we require that $\sigma_i(h_i)$ puts a positive probability only on a countable subset of $\mathcal{X}_i$.

Let Σ_i be the set of all strategies of candidate i. Let $u_i(\sigma|h_i, h_j)$ be candidate i's continuation payoff given history profile $(h_i, h_j) \in H_i \times H_j$ and the strategy profile $\sigma \in \Sigma_A \times \Sigma_B$.[24] Let $H_j(h_i)$ be the set of candidate j's feasible histories given h_i, and let $\beta(\cdot|\cdot) : H_i \to \Delta(H_j)$ be candidate i's belief about candidate j's history such that $\int_{h_j \in H_j(h_i)} d\beta(h_j|h_i) = 1$ for each $h_i \in H_i$ for each $i = A, B$. Given a belief β, let $u_i^\beta(\sigma|h_i) = \int_{h_j \in H_j} u_i(\sigma|h_i, h_j) d\beta(h_j|h_i)$ be candidate i's expected continuation payoff given h_i. A strategy profile (σ_A^*, σ_B^*) is a **perfect Bayesian equilibrium (PBE)** if there exists a belief β such that, for each $i \in \{A, B\}$, (i) $\sigma_i^* \in \arg\max_{\sigma_i \in \Sigma_i} u_i^\beta(\sigma_i, \sigma_j^*|h_i)$

[23]Except for the incomplete information model in Section 3.5, the prediction of the model will be the same even if candidate i can observe all of candidate j's opportunities, in the sense formalized in the Constant-Sum Markov Theorem (Theorem 3) and Remark 5.

[24]This is well defined because H_i is a countable union of subsets of a finite-dimensional space.

holds for every $h_i \in H_i$ and (ii) β is derived from Bayes rule whenever possible.[25]

3 Examples

We first offer various examples to show that the model of policy announcement timing game enables us to analyze rich strategic considerations when it is applied to otherwise well-known and canonical models of elections.

3.1 Valence Election Campaign

We consider the case in which one candidate is stronger than the other, in the sense that if two of them choose the same policy set, then the former candidate wins. Section 3.1.1 introduces the model. In Section 3.1.2, we establish that if two candidates are perfectly symmetric, then both candidates would want to be clear as soon as possible. In Section 3.1.3, we show that if one candidate is slightly stronger than the other, then there are rich strategic considerations driving the incentive for each candidate to make an ambiguous policy announcement. The incentive for ambiguity follows from the "first-mover disadvantage": The strong candidate wants to copy the weak candidate's policy after the weak candidate enters, while the weak candidate does not want to be the first mover as being copied is the worst outcome. This result presents a novel connection between ambiguity and valence.

3.1.1 The Model

In the language of the general model, candidate A is the strong candidate S, and candidate B is the weak candidate W. We keep the model simple, so as to highlight the complexity introduced by the campaign phase into an election model with valence candidates. In particular, the policy space is assumed to be $X = \{0, 1\}$, and each candidate i's available policy sets are $\mathcal{X}_i = \{X, \{0\}, \{1\}\}$. The Online Appendix presents a general version of the model that involves many other cases, such as a continuous policy space.

If a candidate enters at 0 (or 1) and the other enters at 1 (or 0) or does not enter, then

[25] Although each information set at any time after $-T$ has probability zero, one cay apply Bayes rule to calculate relevant conditional probabilities because any Poisson process has a countable number of arrivals with probability one. We formally define Bayes rule for our context in the Online Appendix.

the former wins with probability p (or $1-p$); if the two candidates enter at the same policy or neither of them enters, then the strong candidate wins with probability one. Candidates are purely office-motivated. That is, we have $v_S(\{0\},\{0\}) = v_S(\{1\},\{1\}) = v_S(X,X) = 1$, $v_S(\{0\},\{1\}) = v_S(\{0\},X) = v_S(X,\{1\}) = p$, $v_S(\{1\},\{0\}) = v_S(\{1\},X) = v_S(X,\{0\}) = 1-p$, and $v_W(X_W,X_S) = 1 - v_S(X_S,X_W)$ for each (X_S,X_W). We assume $p \in (0,\frac{1}{2})$.

This utility function can be micro-founded in the following manner. Suppose that there are a continuum of voters, located at policy 0 and policy 1. The distribution of the voters' locations is stochastic, and with probability p, policy 0 has more voters. During the campaign, the locations of the voters are unknown.

If a candidate $i \in \{S,W\}$ wins the election and implements policy $x \in \{0,1\}$, then a voter with position $y \in \{0,1\}$ obtains a payoff of

$$u(|x-y|) + \delta \cdot \mathbb{I}_{i=S},$$

where $u(0) > u(1)$ and $0 \leq \delta < (u(0)-u(1))/2$, with δ representing the advantage of candidate S due to her charisma or other asymmetries between candidates' characteristics that are unrelated to the policy choices.[26] The voters believe that, if candidate i has specified a policy $x \in \{0,1\}$ and wins, then x will be implemented. If candidate i with the ambiguous policy $X_i = \{0,1\}$ wins, then the voters believe that the policies $\{0\}$ and $\{1\}$ will be implemented with equal probability $\frac{1}{2}$.[27] The voters are sincere, that is, they each vote for the candidate who, if elected, maximizes their expected payoff. The candidate with more votes wins. In the case of a tie, each candidate wins with probability 1/2.

The candidate who obtains more votes wins, and obtains a payoff of 1, while the other candidate obtains a payoff of 0; these are the only payoffs that they receive in this model, i.e., candidates are purely office-motivated. Each candidate's objective is to maximize the expected payoff, that is, their objective is to maximize their probability of winning. The payoff function (v_S, v_W) that we

[26] One way to interpret δ in a "policy related" manner would be to consider a model as in Krasa and Polborn (2010), in which candidates choose one policy out of two for each of multiple policy issues. If candidates make policy announcements for some issues first, they then would compete by choosing policies on remaining issues, where asymmetry between candidates may exist depending on the relative popularity of the policies that each candidate has chosen already. We note that, if $\delta > (u(0)-u(1))/2$, it will be straightforward to show that S wins the election with probability 1 in any PBE.

[27] It is not crucial that the probability is exactly $\frac{1}{2}$. For an open set of probabilities for tie-breaking, our main results are unchanged.

(X_S, X_W) at the deadline	Voters at 0 vote for	Voters at 1 vote for	S's expected utility	W's expected utility
$(\{0,1\},\{0,1\})$	S	S	1	0
$(\{0,1\},\{0\})$	W	S	$1-p$	p
$(\{0,1\},\{1\})$	S	W	p	$1-p$
$(\{0\},\{0,1\})$	S	W	p	$1-p$
$(\{0\},\{0\})$	S	S	1	0
$(\{0\},\{1\})$	S	W	p	$1-p$
$(\{1\},\{0,1\})$	W	S	$1-p$	p
$(\{1\},\{0\})$	W	S	$1-p$	p
$(\{1\},\{1\})$	S	S	1	0

Table 1: Voter behaviors and the expected payoffs for the valence election campaign.

specified above can be obtained by assuming $\delta > 0$. We summarize in Table 1 the voters' behaviors and the resulting expected payoffs for the candidates, given these specifications and $\delta > 0$. Note that, without valence ($\delta = 0$), the environment just specified is the one in which we can apply the median voter theorem in the static version of the model; that is, it is each candidate's dominant action to announce $\{1\}$.

We let $\lambda_S = \lambda_W =: \lambda$.[28] We call this dynamic game a *valence election campaign.* It is characterized by a tuple (p, T, λ).

3.1.2 The Benchmark Case: Perfectly Symmetric Candidates

Before analyzing the model with valence, we analyze the model with symmetric candidates as a benchmark case. The only difference from the model with valence is that, if two candidates end up announcing the same policy set, both of them win with probability $\frac{1}{2}$ (this corresponds to setting $\delta = 0$ in the micro-foundation). Call this game a *no-valence election campaign.* It turns out that there are no incentives to announce the ambiguous policy $\{0, 1\}$.

The following proposition gives us a stark result:

Proposition 1 *In any no-valence election campaign, in any PBE, each candidate announces* $\{1\}$ *as soon as possible.*

To see why this holds, fix time $-t$ and suppose that at any time $-s > -t$, if each candidate has

[28]This assumption is generalized in the Online Appendix.

an opportunity to enter, then he/she enters at 1. Then, at time $-t$, if no candidate has entered, entering at 1 gives the payoff strictly greater than $\frac{1}{2}$, entering at 0 gives $p < \frac{1}{2}$, and not entering gives a payoff of $\frac{1}{2}$ by symmetry of the supposed continuation strategies. Thus, entering at 1 is a unique best response. Therefore, by continuity of the probability in time which implies continuity of the continuation payoff in time, for sufficiently small $\varepsilon > 0$, it is uniquely optimal to enter at 1 for all time in $(-t - \varepsilon, -t]$ if no one has entered. Under the history at which the opponent has entered, an analogous argument shows that entering at 1 is uniquely optimal. This establishes the desired result.[29]

In the next subsection, we demonstrate that (i) the above simple argument breaks down once we introduce asymmetry with respect to candidates' valence ($\delta > 0$ in the micro-foundation), and (ii) candidates face complicated dynamic incentive problems, which involve ambiguous policy announcements. Therefore, a small valence (or small $\delta > 0$) matters and is the key for ambiguous policy announcements.

3.1.3 The Cases with Valence Candidates

Let us start with the following lemma. It states that, if S has an opportunity to enter after W has entered at $x \in \{0, 1\}$, then she enters at x and wins for sure. In contrast, if W has an opportunity to enter after S has entered at $x \in \{0, 1\}$, then he is indifferent between announcing $\{0, 1\}$ and entering at $x' \in \{0, 1\} \setminus \{x\}$. These two conclusions imply that, since the median is more likely to be at policy 1 ($p < \frac{1}{2}$), if a candidate enters before the opponent, he/she enters at $\{1\}$.

Lemma 1 *In any valence election campaign with (p, T, λ), in any PBE, the following are true at any time $-t$:*

1. *Given that W has already entered, S enters at the same platform as soon as possible.*

2. *Given that S has already entered, W is indifferent between announcing $\{0, 1\}$ and entering at the platform different from S's.*

3. *If a candidate i enters before the opponent, then i enters at policy 1.*

[29] This last part follows from the "continuous-time backward induction" that we formally present in Appendix B.

The above lemma characterizes the equilibrium behaviors on and off the equilibrium path except when no candidates have yet entered. It also says that if both are still using ambiguous language and a candidate i enters, then i enters at policy 1. Hence, in the following analysis, we consider the incentives to enter at policy 1 when both are still using ambiguous language.

Before presenting the characterization of the behavior in a PBE in such a situation, we first provide the basic intuition, which exploits the idea that being the first mover is disadvantageous. For the time being, consider the case with $p = \frac{1}{2}$.[30] Suppose that at time $-t$, both S and W have previously announced $\{0, 1\}$. On the one hand, if there is no further revision, W's payoff is 0. So W needs to specify his policy to obtain a positive payoff. Thus, W announces $\{0\}$ or $\{1\}$ at some point in $[-t, 0]$, if he can. Since $\{0\}$ and $\{1\}$ are symmetric with $p = \frac{1}{2}$, assume without loss of generality that W announces $\{1\}$ when he clarifies his policy.

On the other hand, S does not have an incentive to specify her policy until W specifies his policy. This is because she gets $\frac{1}{2}$ for sure by specifying her policy, while using ambiguous language at all times in $[-t, 0]$ gives her either $\frac{1}{2}$ or 1 with the latter taking place with positive probability (when W does not enter afterward and when W enters and S copies his policy).

If W announces $\{1\}$ in the early stages of the campaign, then the probability with which S enters afterward is high. So W wants to postpone announcing. But waiting too much is not optimal for W either, since if he does not have a chance to revise his policy set, W gets a payoff of 0. So there should exist a "cutoff," $-t^*$, until which W announces $\{0, 1\}$ and after which W announces $\{1\}$ when he gets an opportunity of a policy announcement.

Recall that we do not have this type of strategic consideration in the no-valence election campaign ($\delta = 0$), even if we extend the model to include the case with $p = \frac{1}{2}$. The simple argument we provided for Proposition 1 breaks down since the continuation payoff after taking each action is different once we introduce valence. For example, W expects a payoff close to zero if he specifies some policy when the deadline is far away in the valence election campaign, as opposed to a payoff of $\frac{1}{2}$ that he gets in the no-valence election campaign.

Next, consider the case with $p = 0$. In this case, S would want to commit to $\{1\}$ as soon as possible, because she can then obtain a payoff of 1, which is the highest possible payoff. Since W can win if and only if he enters at $\{1\}$ and S does not have an opportunity, W also enters at $\{1\}$

[30]Strictly speaking, since $p < \frac{1}{2}$, this is actually outside of the model, but we consider such a case to provide the intuition. The same comment applies to the case $p = 0$ that we consider next.

as soon as possible.

The next proposition fully characterizes the form of PBE for each $p \in (0, \frac{1}{2}) \setminus \{\frac{1}{1+e}\}$. Suppose that the current policy set of each candidate is $\{0, 1\}$. The equilibrium strategy of W is to wait until a finite cutoff time and to enter as soon as possible after that cutoff. In contrast to the case of $p = 0$, the cutoff is finite for any strictly positive p because the probability that the median voter is at 0 is strictly positive. The equilibrium strategy for S depends on the value of p, and the value $p = \frac{1}{1+e}$ corresponds to the cutoff at which S's incentive changes. If p is close to $\frac{1}{2}$ ($p > \frac{1}{1+e}$, considered in part 1 of Proposition 2), S does not enter until W enters for the same reason as in the case of $p = \frac{1}{2}$. In contrast, for small p ($p < \frac{1}{1+e}$, considered in part 2 of Proposition 2), S enters when the deadline is far away as when $p = 0$, but does not do so when the deadline is close. The intuition for the ambiguity near the deadline is as follows: If S obtains an opportunity when the deadline is close, then the probability with which W has a chance to announce his policy afterward is small. So it is likely that W uses ambiguous language at the deadline. Thus, keeping ambiguous language is profitable for S, because by doing so, S gets a payoff of 1 with a high probability.

Proposition 2 *Consider the valence election campaign with (p, T, λ). There exists a PBE. Moreover, there exist $t^* := \frac{1}{\lambda}$, t_S, and t_W (the latter two depend on p) that are independent of T such that, for any PBE, the following are satisfied if the previous policy sets are both $\{0, 1\}$.*[31]

1. *If $p > \frac{1}{1+e}$, the following hold:*[32]
   - *(a) S announces $\{0, 1\}$ for all $-t \in (-\infty, 0]$.*
   - *(b) W announces $\{0, 1\}$ for all $-t \in (-\infty, -t^*)$ and $\{1\}$ for all $-t \in (-t^*, 0]$.*
2. *If $p < \frac{1}{1+e}$, then the following hold:*
   - *(a) S announces $\{1\}$ for all $-t \in (-\infty, -t_S)$ and $\{0, 1\}$ for all $-t \in (-t_S, 0]$.*
   - *(b) W announces $\{0, 1\}$ for all $-t \in (-\infty, -t_W)$ and $\{1\}$ for all $-t \in (-t_W, 0]$.*
   - *(c) Moreover, $-t_W < -t_S$, $\frac{dt_S}{dp} > 0$ and $\frac{dt_W}{dp} < 0$.*

[31] If $p = \frac{1}{1+e}$, then there is indeterminacy about S's equilibrium strategy at all $-t < -t^*$ since she is indifferent.

[32] Although the entire game lasts for the time interval $[-T, 0]$, we state results for all times in $(-\infty, 0]$ as the results do not rely on whether the cutoff times (at which equilibrium actions change) are earlier or later than $-T$. Any statement about time interval $K \subseteq (-\infty, 0]$ should be interpreted as a statement about the time interval $K \cap [-T, 0]$. The same caution applies to all other formal statements involving time intervals.

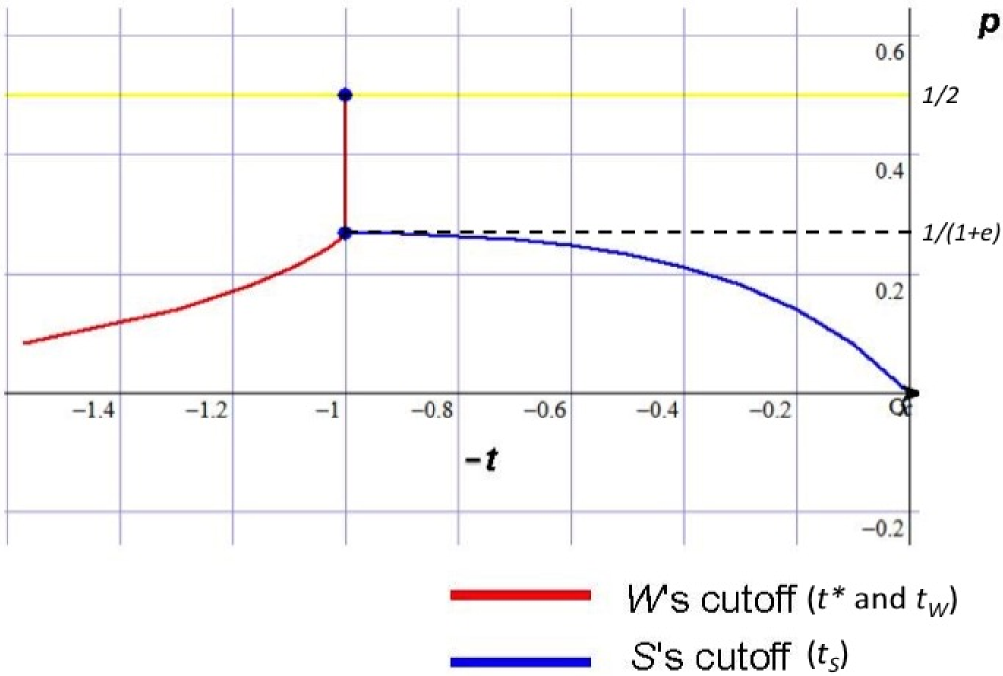


Figure 1: Cutoff times for the valence election campaign ($\lambda = 1$).

Note that the cutoffs are independent of T. Hence, when T and p are large, we expect that candidates use ambiguous language for most of the campaign period. Note also that stretching T and enlarging λ with the same ratio are equivalent. Hence, this also implies that for a fixed length of campaign period T, if we consider the situation in which the opportunities arrive frequently, candidates spend most of the time in $[-T, 0]$ using ambiguous language.

In Figure 1, we depict the times t^*, t_S, and t_W that appear in Proposition 2, for different values of p for the case of $\lambda = 1$. For example, $p = .4$ ($> \frac{1}{1+e}$) corresponds to part 1 of the proposition. In this case, there is one point at which the graph in the figure intersects with the $p = .4$ line, so as a result, the time spectrum is divided into two regions: In the left region, no candidate enters. In the right region, S does not enter while W enters. When $p = .2$ ($< \frac{1}{1+e}$), there are two intersections, and as a result the time spectrum is divided into three regions: In the left-most region, S enters while W does not enter. In the middle region, both candidates enter. Finally, in the right-most region, S does not enter while W enters.

Notice that this particular model predicts that when p is small ($p < \frac{1}{1+e}$) and $T > t_S$, S enters as soon as possible, so if T is large, then there would be almost no ambiguity in equilibrium. This hinges on our assumption that even if W enters after S, S does not incur any loss. In the Online Appendix, we show that if there is even a small loss, S prefers to use ambiguous language until some point in time that does not depend on the horizon length T, and so the modified model is consistent with ambiguity even if p is low. Despite this feature, we believe that the simple model in this section provides a basic intuition about the dynamic incentives that candidates face. The basic takeaway is that the nature of the election game with valence leads candidates to strategically "time" their announcements, since the benefit and cost of maintaining flexibility of choice vary over time. Consider W's incentive, for example. On the one hand, the benefit comes from the fact that the election game is constant-sum, so avoiding being the first mover is a good thing. On the other hand, the cost comes from the difference in valence. He does not want to end up making the same choice as S (that is, taking $\{0, 1\}$). This is the general trade-off of timing strategies faced by electoral candidates, and our model succinctly captures such a trade-off.

Remark 1 (Empirical implication) Note that Proposition 2 applies to any PBE. This uniqueness property enables us to conduct meaningful comparative statics, which one can potentially test empirically. The analysis shows that ambiguity is likely when the probability distribution of the median voter's position is close to uniform (p is close to 1/2). This is consistent with Campbell (1983) who suggests that opinion dispersion has a strong positive effect on the ambiguity in candidates' language.[33] Also, a researcher would be able to infer which candidate is stronger, given the information about the timing of entry or the final policy profiles announced. More detailed accounts of these claims are in the Online Appendix.

Remark 2 (Robustness of the prediction) The basic structure of the equilibrium is robust even if the two candidates have different arrival rates, although the fine details change. One can show that a relatively higher arrival rate makes the candidate better off. This is due to the fact that the underlying game is constant-sum, and is in a stark contrast to the results for coordination games in Calcagno et al. (2014) that having a higher arrival rate makes the player worse off since it

[33] Specifically, we have in mind a situation where n voters are independently distributed over $\{0, 1\}$ where the probability on the policy 0 is $q < \frac{1}{2}$. A higher q suggests more option dispersion (a higher standard deviation of the preferred policies among the voters. Campbell (1983) also considers standard deviation), and corresponds to a higher p.

decreases his/her commitment power. More detailed discussions about heterogeneous arrival rates and a general model with heterogeneous arrival rates and a general class of payoff functions are provided in the Online Appendix.

Remark 3 (Welfare implications) One may be tempted to conduct a welfare analysis resorting to the micro-foundation we provided, but there is a caveat in doing so: The distribution of the median voter does not necessarily pin down the voter distribution at each realized state of the world. With additional assumptions about the voter distribution, one can conduct welfare analysis. For example, suppose that there is a single voter. It is then necessary that this single voter's ideal policy is 0 with probability p and 1 with probability $1-p$. Then, one can show by a calculation that the voter's expected payoff in our model is smaller than under a unique mixed Nash equilibrium model in which each candidate chooses between 0 and 1 as in Aragonès and Palfrey (2002) when our model predicts long ambiguity ($p > \frac{1}{1+e}$), the valence term $\delta > 0$ is sufficiently small, and T is sufficiently large.[34] Although we acknowledge that welfare analysis is an interesting direction of research, we do not explore it under other assumptions about the voter distribution in the valence election campaign, or in other examples we present. This is because such an exercise necessities additional assumptions about voter distributions which are not necessary when considering our main focus on candidates' timing problems and resulting policies.

3.2 Multi-dimensional Policy Space – the Case with Purely Office-Motivated Candidates

When the policy space is multi-dimensional, there does not generally exist a Condorcet winner, and a pure-strategy Nash equilibrium does not exist in a static model. In this section, in contrast, we show that our policy announcement timing game admits existence of a PBE. The ambiguity again results from a disadvantage of being the first-mover, which follows from the nonexistence of a Condorcet winner.

Suppose that the policy space X is a full-dimensional connected subset of $\mathbb{R}^n$ for some $n \in \mathbb{N}$. The voters are distributed according to a measure $\mu \in \Delta(X)$. As a normalization, let $\mu(X) = 1$. We assume that $\mu(Y) = 0$ for any zero-Lebesgue measure set $Y \subseteq \mathbb{R}^n$. Given a policy profile

[34] The calculation is given in the Online Appendix.

$(x_A, x_B) \in X \times X$, we define the set of supporters for each candidate as:

$$S_A(x_A, x_B) = \{x \in X|\ |x - x_A| < |x - x_B|\} \text{ and } S_B(x_B, x_A) = \{x \in X|\ |x - x_A| > |x - x_B|\},$$

where $|\cdot|$ denotes the Euclidian distance. We define the probability of A's winning, $P_A(x_A, x_B)$, to be 1 if $\mu(S_A(x_A, x_B)) > \mu(S_B(x_B, x_A))$, $\frac{1}{2}$ if $\mu(S_A(x_A, x_B)) = \mu(S_B(x_B, x_A))$, and 0 otherwise. Let the probability of B's winning be $P_B(x_B, x_A) = 1 - P_A(x_A, x_B)$. An interpretation is that each voter receives a utility that is strictly decreasing in the Euclidian distance between her bliss point and a policy, and supports the candidate with the policy that would give rise to a strictly higher utility.

Each candidate is purely office-motivated: She receives payoff 1 if elected, and 0 otherwise. Each candidate's objective is to maximize the expected payoff.

Each problem is characterized by a pair (X, μ). Let the set of all problems be $\mathcal{P}$. Define

$$\mathcal{M} = \left\{(X, \mu) \in \mathcal{P} \mid \exists x^* \in X \text{ s.t. } \forall y \in X \setminus \{x^*\},\ \mu\left(\{z \in X | (y - x^*) \cdot (z - x^*) > 0\}\right) = \frac{1}{2}\right\}.$$

Notice first that if $X \subseteq \mathbb{R}$, that is, when X is uni-dimensional, then $(X, \mu) \in \mathcal{M}$ holds for any μ. Notice second that for multi-dimensional X, $\mathcal{M}$ is imposing a severe symmetry condition. For example, if μ is the uniform distribution over X, then $(X, \mu) \in \mathcal{M}$ is equivalent to X being point symmetric. Also, for a given multi-dimensional X, $(X, \mu) \notin \mathcal{M}$ holds generically in the space of μ.[35] Third, when $(X, \mu) \in \mathcal{M}$, the x^* satisfying the condition in the definition of $\mathcal{M}$ is uniquely determined because X is connected. We denote this unique x^* by $x^*(X, \mu)$.

In the policy announcement timing game, for each i, the available policy sets are all the singletons and the entire set X, so $\mathcal{X}_i = \{X\} \cup \left(\bigcup_{x \in X} \{\{x\}\}\right)$ for each $i = A, B$. To define the vote share and winning probabilities for the case in which some candidate does not enter, we expand the domain of S_i with a restriction that $X \setminus (S_i(X_i, X_j) \cup S_j(X_j, X_i))$ has measure zero for any (X_i, X_j) such that $X_i \neq X_j$, and $S_i(X, X) = \emptyset$ for each $i = A, B$. That is, unless the two candidates specify the same policy set, the set of indifferent voters has measure zero. The domain of P_i is expanded accordingly for each $i = A, B$.[36] The payoffs are the same as in the static model:

[35] See Theorem 7.2 of Roemer (2001) for the detail.

[36] That is, $P_A(X_A, X_B)$ is 1 if $\mu(S_A(X_A, X_B)) > \mu(S_B(X_B, X_A))$, it is $\frac{1}{2}$ if $\mu(S_A(X_A, X_B)) = \mu(S_B(X_B, X_A))$, and is 0 otherwise. We let the probability of B's winning be $P_B(X_B, X_A) = 1 - P_A(X_A, X_B)$.

$v_i(X_i, X_j) = P_i(X_i, X_j)$ for each $X_i \in \mathcal{X}_i$, $X_j \in \mathcal{X}_j$, and $i = A, B$. We assume that there exists $\bar{x} \in X$ such that $v_i(\{\bar{x}\}, X) = 1$ for each $i = A, B$, which implies that clarifying a policy position is better when the other candidate is ambiguous. Moreover, once a candidate enters, she prefers the opponent not to enter, that is, $\mu(S_i(x_i, X)) > \inf_{x' \in X} \mu(S_i(x_i, x'))$ holds for each $x_i \in X$, $i = A, B$. This assumption is satisfied if voters believe that candidates, without specifying a policy, take a policy randomly upon being elected, and the voter utility functions are strictly concave. Finally, $S_i(X, X) = \emptyset$ implies that if no one has entered, then the winning probabilities are half-half. We call this dynamic game a *symmetric office-motivated election campaign*. It is characterized by a tuple $(X, \mu, T, \lambda_A, \lambda_B)$.

Proposition 3 *Consider a symmetric office-motivated election campaign with* $(X, \mu, T, \lambda_A, \lambda_B)$*. There exists a PBE, and the following are true.*

1. *Suppose that* $(X, \mu) \in \mathcal{M}$*. Then, in any PBE, conditional on any history, each candidate* i *announces* $x^*(X, \mu)$*.*

2. *Suppose that* $(X, \mu) \notin \mathcal{M}$*. Then, there exist* $t_A^*, t_B^* \in (0, \infty)$ *such that, in any PBE, if no one has entered at time* $-t$*, candidate* i *does not enter if* $-t \in (-\infty, t_i^*)$*, and does enter at some policy if* $-t \in (t_i^*, 0]$*. It must be the case that* $\mathrm{sign}(\lambda_A - \lambda_B) = \mathrm{sign}(t_A^* - t_B^*)$*.*

Remark 4 (Existence of a pure-strategy Nash equilibrium) Note that $(X, \mu) \in \mathcal{M}$ if and only if there exists a pure-strategy Nash equilibrium in the static game in which each candidate chooses a policy in X.[37] Hence, the proposition shows that ambiguity emerges in a PBE if and only if there is no pure-strategy Nash equilibrium in such a static game.

Part 1 implies that, if there exists a Condorcet winner, then it is optimal to announce that policy as soon as possible. In part 2, intuitively, each candidate's strategic situation is similar to that of the weak candidate in the valence election campaign (Section 3.1): If the other candidate cannot enter, she prefers entering to not entering (the former gives a payoff of 1 while the latter gives $\frac{1}{2}$). However, if the other candidate can enter, then she prefers not entering to entering (the

[37]The reason for the "if" direction is that there always exists at least one candidate, say i, who receives no more than half of the entire vote share in a Nash equilibrium, and $(X, \mu) \notin \mathcal{M}$ implies that there exists a policy close to j's policy such that i always has an incentive to deviate to it to receive a vote share strictly higher than $1/2$ (see, for example, Theorem 7.1 of Roemer (2001) for a related result).

former gives a positive expected payoff while the latter gives 0). As a result, it is optimal not to enter if the deadline is far because the probability that the other candidate can enter afterward is large. If the deadline is close, however, since the probability of such an event is small, it is optimal to enter.

If candidate B can only move slower ($\lambda_B < \lambda_A$), then the proposition predicts that he is more likely to be ambiguous at the election date, and conditional on entering, the expected entry time is later. For the entry time, there are two opposing forces: On the one hand, since candidate B cannot move fast, the risk of him not being able to enter afterward is substantial. This force would make him willing to enter early. On the other hand, candidate B knows that candidate A is likely to obtain an opportunity later, and this would make him willing to wait until the last moment. Since the loss from the latter is particularly large, he does not want to enter until the last moment ($t_B^* < t_A^*$).

An implication of Proposition 3 is that the faster candidate is more likely to win:

Proposition 4 *Consider a symmetric office-motivated election campaign with $(X, \mu, T, \lambda_A, \lambda_B)$. If $\lambda_A > \lambda_B$, then for any PBE, candidate A's expected payoff is strictly greater than that of candidate B.*

The proposition is straightforward if $(X, \mu) \in \mathcal{M}$. The case of $(X, \mu) \notin \mathcal{M}$ may seem subtle, but there is a simple intuition: Given the previous proposition, candidate B does not enter until $-t_B^*$. Candidate A can obtain a higher payoff than candidate B by simply waiting until time $-t_B^*$ because A receives opportunities more frequently than B does after $-t_B^*$. Such a strategy is suboptimal but provides a lower bound of candidate A's PBE payoff.

Finally, we state an implication of Proposition 3 on the relationship between the dynamics in PBE and the dimensionality of the policy set.

Corollary 1 *Fix X. The following are true:*

1. *If X is uni-dimensional ($n = 1$), then for any $(\mu, T, \lambda_A, \lambda_B)$, the following is true: In the symmetric office-motivated election campaign with $(X, \mu, T, \lambda_A, \lambda_B)$, there exists a PBE. Moreover, in any PBE, conditional on any history, each candidate i announces $x^*(X, \mu)$.*

2. *If X is multi-dimensional ($n \geq 2$), then for generic μ, for any $(T, \lambda_A, \lambda_B)$, the following are true: There exists a PBE. Moreover, there exist $t_A^*, t_B^* \in (0, \infty)$ such that, in any PBE, if no one has entered at time $-t$, candidate i does not enter if $-t \in (-\infty, t_i^*)$, and does enter at some policy if $-t \in (t_i^*, 0]$.*

Notice that the results in this section show the uniqueness of a distribution of entry times in any PBE, while it does not show the uniqueness of the policies to which the candidates enter. In fact, there may exist multiple PBE due to the fact that there may exist multiple policies at which each candidate can win if the opponent does not enter and multiple policies at which each candidate can win when she enters after her opponent. In the next section, we consider the case with a multi-dimensional policy space with policy-motivated candidates and show that, with policy motivation, uniqueness of policies may obtain.

3.3 Multi-dimensional Policy Space – the Case with Policy-Motivated Candidates

We again consider the policy announcement timing game with a multi-dimensional policy space, but now with policy-motivated candidates. We show that, in a PBE, if a candidate cares about the policy implemented by the winner of the election, then she may announce a Pareto-inefficient policy to influence a later announcement by the opposition party. By announcing such a policy, she can induce the opponent to implement a policy that is not too undesirable even in the event that she loses.

Specifically, we consider the following setting of Persson and Tabellini (2000): $X = \{(x_1, x_2) \in [0,1]^2 : x_1 + x_2 \leq 1\}$. Here, a higher x_1 is interpreted as a more conservative economic policy and a higher x_2 is interpreted as a more aggressive military policy. There are three voters: Voter 1's ideal policy is $(1, 0)$ and her utility from policy x is $-(1 - x_1)$. That is, she is right-wing and only cares about the economic policy. Voter 2's ideal policy is $(0, 1)$ and her utility from policy x is $-(1 - x_2)$. That is, she is also right-wing and only cares about the military policy. Finally. voter 3's ideal policy is $(0, 0)$ and her utility from policy x is $-x_1 - x_2$. That is, she generally likes a left-wing policy.

There are two candidates L and R, whose ideal policies are $(0, 0)$ and $\left(\frac{1}{2}, \frac{1}{2}\right)$, respectively.[38]

[38]When we need to distinguish between the two candidates, we use a masculine pronoun for L and a feminine

Their ideal policies are common knowledge, and the voters correctly believe that the candidate who wins without specifying a policy will implement her ideal policy. If a candidate wins with a specified policy x, then she must implement x. The voters vote for the candidate who brings the higher utility, with a tie broken equally in favor of the entrant if there is only one candidate who enters, and in favor of the candidate who enters later if both enter.[39] This in particular implies that R collects two votes when no candidate enters. The candidate collecting two or three votes wins. Since R has an ideal policy that is preferred by two voters (voters 1 and 2), she has a chance of winning with probability 1 if no candidate specifies a policy. In this sense, candidate R is similar to the "strong candidate" in the valence election campaign analyzed in Section 3.1. We will show, however, that the distribution of entry times differs from the one for that model because the payoff from the entry is specified differently.

If a candidate $k \in \{L, R\}$ wins the election and implements policy x, the payoff of candidate $i \in \{L, R\}$ is

$$\mathbb{I}_{i=k} + \varepsilon u_i(x),$$

where $u_L(x) = -\max_{n\in\{1,2\}} x_n$ and $u_R(x) = \min_{n\in\{1,2\}} x_n$ are the utility functions to represent candidates' policy preferences, and $\varepsilon > 0$.[40] The payoff function v_i for each $i = L, R$ is specified accordingly. Persson and Tabellini (2000) show that there is no Condorcet winner (no median voter) and there is no pure-strategy Nash equilibrium in the simultaneous-move game in which choosing X is not allowed.

In the policy announcement timing game, for each i, the available policy sets are again all the singletons and the entire set X, so $\mathcal{X}_i = \{X\} \cup \left(\bigcup_{x\in X} \{\{x\}\}\right)$. As a tie-breaking rule, we assume that if it is optimal for a candidate to enter and $\bar{X}$ is the set of all policies such that entering at any policy in $\bar{X}$ generates the maximum continuation payoff, then she enters at a policy in

pronoun for R.

[39] This tie-breaking rule is consistent with considering a limit of unique PBEs in models with discrete policy spaces. Palfrey (1984) conducts the same exercise of taking a limit in a game where best responses do not exist with a continuous policy space. To be precise, this tie-breaking rule violates the assumption that the payoffs to the candidates are determined solely by the functions $(v_i)_{i=L,R}$ which only depend on the profile of policy sets. One could define a value function that depends both on the policy-set profile and on the times at which they are announced, but we do not write out such formalization in light of the justification due to discrete policy spaces and for the sake or readability. The same comment applies to later sections where we introduce tie-breaking rules (Section 3.5 and the Online Appendix).

[40] To avoid confusion, we use n for the index of a dimension of the policy space; and i, j and k for the indices of the candidates.

$\arg\min_{(x_1,x_2)\in\bar{X}} |x_1 - x_2|$. That is, each candidate enters at a policy that is the most equally right-wing in both dimensions. Call this game a *policy-motivated election campaign.* It is characterized by a tuple $(\varepsilon, T, \lambda_L, \lambda_R)$.

Suppose a candidate has entered at x. Since the tie is broken in favor of the last mover and there is no Condorcet winner, there exists a closed set $X(i,x)$ such that the remaining candidate i wins if she enters at a policy in $X(i,x)$. Let $y_i(x)$ be the unique minimizer of $|x'_1 - x'_2|$ among all $x' \in \arg\max_{x''\in X(i,x)} u_i(x'')$, that is, it is the policy that candidate i enters.[41]

Proposition 5 *Fix λ_L and λ_R such that $\lambda_L \neq 2\lambda_R$. There exists $\bar{\varepsilon} > 0$ such that, for any $T < \infty$ and $\varepsilon \in (0, \bar{\varepsilon})$, any PBE of the policy-motivated election campaign with $(\varepsilon, T, \lambda_L, \lambda_R)$ satisfies the following: Each candidate enters at $y_i(x)$ as soon as possible, once the other candidate enters at x. If the other candidate has not entered, the following hold:*

1. *Candidate R does not enter at any $-t \in (-\infty, 0]$ for any (λ_L, λ_R).*

2. *Candidate L's strategy depends on the parameters (λ_L, λ_R).*

   (a) *If $\frac{\lambda_L}{\lambda_R} > 2$, then there exists $t_L \in (0,\infty)$ such that L does not enter at $-t \in (-\infty, -t_L)$ and does enter at $\left(\frac{1}{2}, \frac{1}{2}\right)$ for $-t \in (-t_L, 0]$.*

   (b) *If $\frac{\lambda_L}{\lambda_R} < 2$, then there exist $t^*_L, t^{**}_L \in (0,\infty)$ such that L does not enter at $-t \in (-\infty, -t^{**}_L)$, enters at either $\left(\frac{2}{3}, 0\right)$ or $\left(0, \frac{2}{3}\right)$ at $-t \in (-t^{**}_L, -t^*_L)$, and enters at $\left(\frac{1}{2}, \frac{1}{2}\right)$ at $-t \in (-t^*_L, 0]$.*

The proof in the Online Appendix provides an explicit expression of $\bar{\varepsilon}$. The bound ensures that it is a dominated strategy for candidate i to enter at a policy x such that i loses at a policy set profile $(\{x\}, X)$.

On the one hand, since both voters $(1,0)$ and $(0,1)$ prefer candidate R's ideal policy, R wins with probability 1 if no candidate specifies a policy. Moreover, if candidate R enters and then candidate L can enter, R will lose for sure. These facts turn out to imply that candidate R does not have an incentive to enter unless candidate L enters.

On the other hand, candidate L has to enter at some point to receive a positive payoff. If the deadline is very far, then since candidate R will enter with a very high probability once L enters,

[41] The proof of Proposition 5 shows uniqueness of the minimizer.

it is optimal for him not to enter. If the deadline is very close, then the probability that candidate R will enter is very small. Therefore, L enters at the policy he prefers the most among those with which he can win, namely, $\left(\frac{1}{2}, \frac{1}{2}\right)$. In the middle, his optimal policy depends on the relative arrival rates of opportunities. If candidate L is a relatively fast mover ($\frac{\lambda_L}{\lambda_R} > 2$), then the risk of not being able to enter at all is small. Hence, he waits until the probability of candidate R entering after L becomes sufficiently small, and then enters at $\left(\frac{1}{2}, \frac{1}{2}\right)$. If L is relatively slow ($\frac{\lambda_L}{\lambda_R} < 2$), it is too risky for him to wait until the probability of candidate R entering becomes small. Hence, he enters even when there is a significant probability of candidate R entering after L. Taking this event into account, he does not enter at the policy he prefers the most among those with which he can win, but at $\left(\frac{2}{3}, 0\right)$ or $\left(0, \frac{2}{3}\right)$. This narrows down the set of policies with which candidate R can win after L's entry, so L can make R's policy more left-wing.

We note that the consideration in this last part (leading L to entering at $\left(\frac{2}{3}, 0\right)$ or $\left(0, \frac{2}{3}\right)$) does not occur if L does not care about what policy R picks when R wins. For example, candidates may care about the utility from being in the office and the cost of persuading the voters that they implement a policy far from their bliss points, while they do not derive any utility from the implemented policy per se. In the Online Appendix, we formalize such a model, and show that the equilibrium dynamics in such a model are simpler.

Remark 5 (Outcome-equivalence for a public-monitoring model) The PBE we characterize in this section (as well as the PBE characterized in the Online Appendix) is Markov-perfect (except for measure-zero sets of times). Hence this equilibrium is outcome-equivalent to a Markov perfect equilibrium in the "public monitoring" model where candidates observe the other candidate receiving opportunities even when the policy set does not change.[42] Moreover, we solve the equilibrium by backward induction. This means that any SPE under public monitoring is outcome-equivalent to a PBE in our main model where the opponent's opportunities are not observable. The same remark applies to Section 3.4.[43]

Remark 6 (Flexibility in office) We believe that there are various reasons for ambiguous announcements in real election campaigns. It is not our intention to capture all of those reasons in our general model, but to focus on those that relate to candidates' dynamic incentives. In

[42] See Section 4.3 for the formal description of such a model.

[43] In Sections 3.1 and 3.2, those claims are a consequence of Theorem 3 in Section 4.3.

the valence election campaign (Section 3.1) and the symmetric office-motivated election campaign (Section 3.2), ambiguity is present because each candidate does not want to be the first mover. In the policy-motivated election campaign in this section, this effect is still present, while there is another reason to be ambiguous: Not specifying a policy gives a flexibility in choosing a preferred policy after being elected. This same reason will be present in the analysis of the situation with incomplete information in Section 3.5.

3.4 Dynamic Campaign Spending Model

The empirical evidence suggests that campaign spending has nontrivial effects on the election outcome, and candidates often spend monetary resources gradually over time.[44] Since the spending can only increase over time, we can represent such a situation using our policy announcement timing game. We specify $X = \{0, L, H\}$ with $0 < L < H$. The interpretation is that there are two levels of positive campaign spending, where L is the lower level of spending and H is the higher level. The available policy sets are $\mathcal{X}_i = \{X, \{L, H\}, \{H\}\}$ for each $i = A, B$. The interpretation is that $\{L, H\}$ implies that candidate i has spent L by the current time and the total spending at the deadline can be either L (if she does not spend more) or H (if she spends more). To focus on issues regarding campaign spending, assume that the election outcome depends solely on the amount of campaign spending.[45] In particular, for each candidate $i = A, B$, the probability of i winning the election under the policy set profile (X_i, X_j) is $w_i(X_i, X_j) := \frac{\min_{x \in X_i} x}{\left(\min_{x \in X_i} x\right) + \left(\min_{x \in X_j} x\right)}$ with a convention that $\frac{0}{0+0} = \frac{1}{2}$. Note that this implies that we are assuming no depreciation of the effect of campaign spending over time. For simplicity, we assume that the arrival rates of opportunities are symmetric: $\lambda_A = \lambda_B := \lambda$.

Assuming that the remaining fund is useful to a candidate, candidate i's payoff is

$$v_i(X_i, X_j) := \alpha w_i(X_i, X_j) + (1 - \alpha)\left(H - \min_{x \in X_i} x\right),$$

where

$$\alpha \in \left(\max\left\{\frac{L}{L + \frac{1}{2}}, \frac{H + L}{H + L + \frac{1}{2}}\right\}, 1\right). \tag{1}$$

[44]See, for example, Gerber et al. (2011).
[45]With this definition, we capture both positive and negative advertising.

Here, $H - \min_{x \in X_i} x$ in the second term of $v_i(X_i, X_j)$ is the remaining cash holdings. The restriction on α implies that α is sufficiently large so that a static best response to $X_j = X$ is $\{L, H\}$ and that to $X_j = \{L, H\}$ or $\{H\}$ is $\{H\}$.[46] We call this game a *dynamic campaign spending game*. It is characterized by a tuple $(H, L, \alpha, T, \lambda)$.

A typical pattern found in the empirical literature (Gerber et al., 2011) is that candidates spend more money near the deadline than far from it. With depreciation, such a pattern follows from a mechanical reason: Earlier spending is less effective, holding the other candidate's spending fixed. We show that, even without depreciation, it is possible that the candidates spend more near the deadline. This follows from a strategic reason: Suppose candidate B has spent L and the deadline is far. If candidate A spends H now, then candidate B will match up with a very high probability. If she spends L, then the race of matching up will start immediately as well, resulting in both spending H with a very high probability. In contrast, if she waits at 0, then the opponent will wait at L as well because both of them know that, once the former spends at least L or the latter spends H, then the race of matching up will start immediately. Hence they can avoid a wasteful competition of matching up with the opponent's high spending until the deadline becomes near.

Depending on how important winning is relative to keeping the money, such an incentive may be alleviated due to the risk that the winning probability will be 0 if candidate A does not have an opportunity to spend more later. Specifically, if α is large, that is, if winning the election is sufficiently important compared to keeping the money, then the risk is prominent and thus they spend H as soon as possible far from the deadline. In contrast, if α is small, then the benefit of avoiding escalation when the deadline is far is sufficiently large. Hence, candidates stay at a spending profile $(L, 0)$ or $(0, L)$ for a long time.

Proposition 6 *Fix the dynamic campaign spending game with $(H, L, \alpha, T, \lambda)$.*

1. If

$$\alpha > \frac{H+L}{H+L+\frac{1}{4}} \tag{2}$$

holds, then there exists a PBE. In any PBE, the following hold:

[46] Formally, these conditions are expressed as: $\alpha + (1-\alpha)(H-L) > \max\{\alpha, \alpha\frac{1}{2} + (1-\alpha)H\}$, $\alpha\frac{H}{H+L} > \max\{\alpha\frac{L}{H+L} + (1-\alpha)(H-L), (1-\alpha)H\}$, and $\alpha\frac{1}{2} > \max\{\alpha\frac{L}{H+L} + (1-\alpha)L, (1-\alpha)H\}$, which are equivalent to $\alpha > \max\left\{\frac{L}{L+\frac{1}{2}}, \frac{H+L}{H+L+\frac{1}{2}}\right\}$.

(a) *For each* $-t < -\frac{1}{\frac{\alpha}{2(1-\alpha)(H+L)}-1}\frac{1}{\lambda}$, *each candidate spends* H *after each history.*

(b) *For each* $-t > -\frac{1}{\frac{\alpha}{2(1-\alpha)(H+L)}-1}\frac{1}{\lambda}$,[47] *each candidate spends* H *if the other candidate has spent* L *or* H, *and spends* L *otherwise.*

2. *If*

$$\alpha < \frac{H+L}{H+L+\frac{1}{4}} \tag{3}$$

holds, then there exists a PBE. In any PBE, the following hold:

(a) *For each* $-t < -\frac{1}{\lambda}$, *each candidate spends* L *if both candidates have spent* 0; *each candidate does not increase the spending if one candidate has spent* L *and the other has spent* 0; *and each candidate spends* H *otherwise.*

(b) *For each* $-t > -\frac{1}{\lambda}$, *each candidate spends* H *if the other candidate has spent* L *or* H, *and spends* L *otherwise.*

Note that it is straightforward to show that the set of parameters (H, L, α) satisfying both (1) and (2) is nonempty, and the one satisfying both (1) and (3) is also nonempty. When the parameters satisfy (1) and (3), the proposition shows that candidates spend a long time not spending the highest possible amount for the campaign.

3.5 Incomplete Information Model

The general model setup in Section 2 features a complete information game. To show the potential of the model to include a wider class of settings, here we allow for incomplete information. We find that ambiguity still prevails in our incomplete-information game.

Suppose that the set of policies is $X = \mathbb{R}$, and there are two candidates L (he) and R (she). Each of the candidates has two types, Normal or Extreme. The normal type has the ideal policy of 0, while the extreme type has the ideal policy of x^i with $i \in \{L, H\}$, depending on the index of the candidate. We assume $x^L < 0 < x^R$ and $x^L = -x^R$. Let $p \in (0, 1]$ be the probability that a candidate is extreme, and we assume that types are independently distributed between candidates. We extend the perfect Bayesian equilibrium defined in Section 2 in a straightforward way, where the associated belief specifies the belief about the other candidate's type.

[47] Given $\alpha > \frac{H+L}{H+L+\frac{1}{4}}$, we have $\frac{\alpha}{2(1-\alpha)(H+L)} - 1 > 0$.

There is a continuum of voters whose ideal policies are distributed on $\mathbb{R}$, with the median position being 0. Given the ideal policy y and policy x, the voter's utility is $-|x-y|$. A candidate's utility given her ideal policy y and implemented policy x is $-|x-y|$. That is, candidates are purely policy-motivated.

In the policy announcement timing game, we let $\mathcal{X}_i = \{X\} \cup \left(\bigcup_{x \in X} \{\{x\}\}\right)$ for each $i = L, R$, and assume that the arrival rates of opportunities are symmetric: $\lambda_L = \lambda_H = \lambda$. The policy commitment is irreversible and credible: Even if a candidate enters at a platform different from her ideal point, she must implement it. However, if a candidate does not enter, then the voters and the other candidate believe that she will implement her ideal policy. Given this commitment and belief, each voter votes for a candidate who brings the higher expected utility (ties are broken randomly with equal probability). The candidate with a higher vote share will win. The payoff function v_i for each type of each $i = L, R$ is specified accordingly. We break ties in favor of a candidate who enters if only one candidate enters; in favor of the second entrant if both candidates enter; randomly with equal probability in all other cases. We call this dynamic game an *election with incomplete information*. It is characterized by a tuple (p, T, λ). We focus on symmetric PBE in this game.[48]

3.5.1 Benchmark: Complete Information

Before fully analyzing the case with $p < 1$, we analyze the case where both candidates are extreme for sure ($p = 1$). The following proposition says that, with $p = 1$, there is a continuum of equilibria when the horizon is sufficiently long.

To state the result formally, given the opponent's policy x, we define $BR_R(x)$ and $BR_L(x)$ to be candidate R's and L's static best responses, respectively, when they are extreme:

$$BR_R(x) := \begin{cases} x^R & \text{if } |x| > x^R \\ x & \text{if } x \in (0, x^R] \\ -x & \text{if } x \in [-x^R, 0] \end{cases}, \qquad BR_L(x) := \begin{cases} -x^R & \text{if } |x| > x^R \\ -x & \text{if } x \in [0, x^R] \\ x & \text{if } x \in [-x^R, 0) \end{cases}. \tag{4}$$

[48]Some results in this section apply to any (possibly asymmetric) PBE, and for those results we do not restrict ourselves to symmetric PBE in stating them.

Proposition 7 *In any election with incomplete information with* $(1, T, \lambda)$*,* σ *is a pure PBE if and only if the following hold under* σ*.*

1. *If the opponent has not entered, then, the following hold.*

   (a) *If* i *enters at time* $-t \in (-\infty, -\frac{1}{\lambda}\ln 2)$*, she enters at 0.*

   (b) *If* L *enters at time* $-t = -\frac{1}{\lambda}\ln 2$*, he enters at a policy in* $[x^L, 0]$*. If* R *enters at time* $-t = -\frac{1}{\lambda}\ln 2$*, she enters at a policy in* $[0, x^R]$*.*

   (c) *Each candidate* $i \in \{L, R\}$ *enters at* x^i *for* $-t \in (-\frac{1}{\lambda}\ln 2, 0]$*.*

2. *If the opponent has entered at* x*, then each candidate* $i \in \{L, R\}$ *enters at* $BR_i(x)$ *as soon as possible.*

Intuitively, if the deadline is sufficiently far, it is likely that the opponent will have an opportunity later. Hence, if candidate R enters at $x \geq 0$ then $-x \leq 0$ will be implemented with a high probability, and if L enters at $x \leq 0$ then $-x \geq 0$ will be implemented with a high probability. Thus, it is better for each candidate to enter at 0 if she ever enters.

If she skips an opportunity, then by symmetry and the constant-sum nature of payoff functions, the expected payoff is the same between the candidates. Since each candidate enters at some policy in $\left[-x^R, x^R\right]$ when entering (if a candidate enters outside of this interval, then she will certainly lose since the median voter will prefer the opponent's ideal policy), this symmetry, together with piecewise-linearity of the utility function, means that each candidate's expected payoff is the same as the one from entering at 0.

In total, each candidate is indifferent between entering at 0 and not entering when the deadline is far. When the deadline is close, since it is likely that the opponent cannot enter afterward, it is optimal to enter at her own ideal policy. The cutoff time turns out to be $-\frac{1}{\lambda}\ln 2$.

As will be seen, in the model where $p \in (0, 1)$, for sufficiently large T, the extreme candidates do not enter until $-t$ near $-\frac{1}{\lambda}\ln 2$. Intuitively, there is an option value of not entering and figuring out the opponent's type. More precisely, the normal type enters at 0 as soon as possible, and hence waiting allows a candidate to learn about the opponent's type.

3.5.2 Strategy of the Normal Type

The next lemma formally pins down the strategy of the normal type:

Lemma 2 *In any PBE of the election with incomplete information with (p, T, λ) with $p \in (0,1)$, each normal-type candidate enters at 0 as soon as possible at any history.*

The intuition is that, if a normal-type candidate enters at 0, then with probability 1 the winning policy is 0, which is her ideal policy.

3.5.3 Strategy of the Extreme Candidate

Since the candidates are symmetric and we focus on symmetric PBE, without loss, we consider candidate R's incentive.

We first analyze what each extreme candidate does once the opponent enters. Given the definition of the best response, in any PBE, once the opponent enters at x, the extreme type of each candidate $i \in \{L, R\}$ enters at $BR_i(x)$ as soon as possible.

Hence, given an arbitrary conditional probability $\tilde{p}$ of candidate L being extreme, the expected payoff of extreme candidate R entering at x at $-t$ when candidate L has not entered and is extreme depends only on $(t, x, \tilde{p})$ in any PBE. Let $v_t(\tilde{p}, x)$ be this expected payoff of extreme candidate R and let $v_t(\tilde{p}, \text{enter}) = \max_x v_t(\tilde{p}, x)$ be the expected payoff of entering.[49] The next lemma characterizes the optimal policy to enter:

Lemma 3 *For each $\tilde{p} \in (0,1)$ and $t \geq 0$, we have*

$$\arg\max_x v_t(\tilde{p}, x) = \begin{cases} \{\tilde{p}x^R\} & \text{if } \frac{e^{-\lambda t}}{1-e^{-\lambda t}} > \tilde{p} \\ [0, \tilde{p}x^R] & \text{if } \frac{e^{-\lambda t}}{1-e^{-\lambda t}} = \tilde{p} \\ \{0\} & \text{if } \frac{e^{-\lambda t}}{1-e^{-\lambda t}} < \tilde{p} \end{cases}.$$[50]

Intuitively, if the deadline is far (that is, $\frac{e^{-\lambda t}}{1-e^{-\lambda t}} < \tilde{p}$), then it is likely that the opponent L will have an opportunity to enter and flip the policy if he is extreme: If extreme candidate R enters at x, then $-x$ will be implemented with a high probability if the opponent is extreme. Hence it is optimal to enter at 0 (if she ever enters). On the other hand, if the deadline is near (that is, $\frac{e^{-\lambda t}}{1-e^{-\lambda t}} > \tilde{p}$), then it is unlikely that the opponent will have an opportunity. Hence she enters

[49] As will be seen, the maximizer always exists.
[50] We use the convention that $\frac{1}{0} = +\infty$ (which applies when $t = 0$).

at a policy close to her ideal policy. The value of entering, $v_t(\tilde{p}, \text{enter}) = \max_x v_t(\tilde{p}, x)$, can be computed by using this lemma.[51]

Since we have pinned down the strategy of the normal type and the continuation strategy of the extreme type after the opponent has entered, we are left to specify extreme candidate R's strategy at the histories where candidate L has not entered. Let $\bar{H}_t^R$ be the set of candidate R's histories such that no candidate has entered by $-t$.

Fix any symmetric PBE σ. Given the history $h_R^t \in \bar{H}_t^R$, let $p\left(h_R^t\right)$ be the posterior probability that candidate L is extreme. Since candidate R's opportunity and candidate L's opportunity are independently distributed and it is possible that candidate L has not obtained any opportunity by the current time, the posterior $p\left(h_R^t\right)$ depends only on the public history—the event that neither candidate has entered by $-t$. Hence, we write it as a function of t, by setting $p\left(h_R^t\right) = p(t)$. Moreover, candidate R and the voters share the same posterior about candidate L being extreme. Thus, $p(t)$ is the voters' posterior about L's type as well.

The next lemma states that candidate R does not enter at $-t$ with $\frac{e^{-\lambda t}}{1-e^{-\lambda t}} \leq p(t)$:

Lemma 4 *In any election with incomplete information with $(p.T, \lambda)$ with $p \in (0,1)$, under any symmetric PBE, for each time $-t$ and $h_R^t \in \bar{H}_R^t$, candidate R does not enter if $\frac{e^{-\lambda t}}{1-e^{-\lambda t}} \leq p(t)$.*

To see the intuition, suppose candidate R has an opportunity at $-t$. Since candidate L and the voters cannot observe if candidate R received an opportunity, if candidate R does not enter, then the situation is the same as the case in which no candidate receives an opportunity at $-t$. Since we focus on symmetric equilibria and the implemented policy is in $\left[-x^R, x^R\right]$, if the opponent is extreme, then candidate R obtains a payoff of $-x^R$ (corresponding to policy 0) if she does not enter. If the opponent is normal, then she obtains a payoff greater than $-x^R$ by not entering. This is because, since the opponent enters at 0, she can obtain a payoff that is at least $-x^R$ at any history; and if no candidate receives further opportunities, then she obtains a payoff of $-\frac{1}{2}x^R$. In contrast, Lemma 3 implies that, if candidate R enters at $-t$ with $\frac{e^{-\lambda t}}{1-e^{-\lambda t}} \leq p(t)$, she obtains $-x^R$. Hence, it is uniquely optimal for her not to enter.

Given this lemma, we now characterize the equilibrium dynamics. Fix any prior $p > 0$. If T is sufficiently large, then we have $\frac{e^{-\lambda t}}{1-e^{-\lambda t}} \leq p \leq p(t)$ for all $-t \in \left[-\frac{T}{2}, -T\right]$. Hence, the extreme candidates do not enter for any $-t \in \left[-\frac{T}{2}, -T\right]$.

[51] The calculation is provided in the proof of Lemma 3.

Since normal types enter as soon as they obtain an opportunity and extreme types do not enter in any symmetric PBE, by a standard argument $p(t)$ can be shown to evolve as follows for $-t \in \left[-\frac{T}{2}, -T\right]$:

$$\frac{d}{dt}p(t) = -\lambda p(t)(1 - p(t)).$$

Given this evolution, for large T, we have $p\left(\frac{T}{2}\right) \approx 1$. Since the normal candidates always enter, we have $p(t) \geq p\left(\frac{T}{2}\right) \approx 1$ for all $-t \geq -\frac{T}{2}$.

At $-t = -\frac{T}{2}$, candidate R does not enter. In contrast, for $-t$ sufficiently close to 0, she enters. To see this, consider the following two scenarios: If she enters at $p(t)x^R \approx x^R$, she will win and obtain a payoff near 0 if and only if candidate L cannot enter, which happens with probability close to 1. If candidate R does not enter, then again with a probability close to 1, no candidate obtains a further opportunity. Hence her payoff is near $-x^R$. Thus, for $-t$ sufficiently close to 0, it is optimal for R to enter at $p(t)x^R$.

In fact, we can show that there exists a unique cutoff time $-t^*$ such that candidate R enters at $p(t)x^R$ if $-t > -t^*$ and does not enter if $-t < -t^*$, where $p(t)$ evolves according to $\frac{d}{dt}p(t) = -\lambda p(t)(1 - p(t))$ for $-t \in [-T, -t^*]$ (the extreme candidate does not enter for $-t \in [-T, -t^*))$ and $p(t) = p(t^*)$ for $-t \geq -t^*$ (the extreme candidate enters for $-t \in (-t^*, 0]$ and so there is no belief update). Moreover, we can show that t^* converges to the maximum cutoff for the complete-information case as the horizon becomes long ($T \to \infty$), providing one possible refinement of the set of PBE in the case of complete information (cf. Proposition 7). We summarize our results as follows.

Proposition 8 *For each $p \in (0,1)$ and $\lambda > 0$, there exists $\bar{T}_{p,\lambda} < \infty$ such that, for each $T \geq \bar{T}_{p,\lambda}$, in any election with incomplete information with (p, λ, T), there exists a symmetric PBE, and there exists $t^*(p, \lambda, T)$ such that any symmetric PBE satisfies the following equilibrium dynamics: For each $-t < -t^*(p, \lambda, T)$, extreme candidates do not enter and $p(t)$ evolves according to $\frac{d}{dt}p(t) = -\lambda p(t)(1 - p(t))$ over $-t = [-T, -t^*]$; and for each $-t > -t^*(p, \lambda, T)$, extreme candidate $i = L, R$ enters at $p(t^*(p, \lambda, T))x^i$ and $p(t) = p(t^*(p, \lambda, T))$. Moreover, for any $p \in (0,1)$ and $\lambda > 0$,*

$-\lambda t^{*}(p,\lambda,T)$ converges to $-\ln 2$ as $T\to\infty$:

$$\lim_{T\to\infty}\left|\lambda t^{*}(p,\lambda,T)-\ln 2\right|=0.$$

The proposition shows that candidates use ambiguous language at times before $-t^{*}(p,\lambda,T)$ in any symmetric PBE. This implies that extreme candidates spend most of the campaign time keeping their policies ambiguous, provided the campaign is sufficiently long.

4 General Predictions

In Section 3, we have seen that the policy announcement timing game can be applied to analyses of various examples. In those examples, we showed results that match observations in real election campaigns (cf. discussions in the Introduction). Now we present general principles that underlie those results. This helps us understand the logic behind various results in Section 3, as well as shows the robustness of those results to wider classes of environments.

To recap, our discussion of the applications have the following in common: Candidates use ambiguous language (or do not use up all the campaign funds) when the election date is not close if entering before the opponent is disadvantageous, while they enter as soon as possible if a Condorcet winner exists. Moreover, we obtained uniqueness of the entry times in many results, and in particular, we obtain uniqueness in the models in which candidates are purely office-motivated. In this section, we aim to generalize those results.

In Section 4.1, we offer a general condition for candidates to use ambiguous language. The key condition is what we call the "first-mover disadvantage," which roughly corresponds to the non-existence of a Condorcet winner. In contrast, Section 4.2 shows that if there is a Condorcet winner, then candidates announce the policy corresponding to the Condorcet winner as soon as possible. Finally, Section 4.3 offers a general implication of the candidates being purely office-motivated.

For each application, Table 2 represents which general prediction is applicable to which application. In some applications, the corresponding general theorem only applies to part of the claims made there. In the subsections that follow, we explain which part of each of those applications is covered by each general theorem.

Model	Long ambiguity	Dynamic median-voter	Constant-sum Markov
Section 3.1: Valence Candidates	Yes	Yes	Yes
Section 3.2: Multi-dimensional policy space, purely office-motivated	Yes	Yes	Yes
Section 3.3: Multi-dimensional policy space, policy-motivated	No	No	No
Section 3.4: Spending	No	Yes	No
Section 3.5: Incomplete information	No	No	Yes

Table 2: General Predictions and Applications: "Yes" means that the corresponding result is used in a proof for the corresponding section, while "No" means it is not.

In Sections 4.1 and 4.2, we assume that, for each candidate $i = A, B$,

$$\mathcal{X}_i = \{\{x\}|x \in X\} \cup \{X\}. \tag{5}$$

That is, the choice of a policy set is either to specify a single policy or not to specify any policy at all.

4.1 The Long Ambiguity Theorem

In this section, we are going to prove the following claim:

Long Ambiguity Theorem: *Under certain conditions, for each candidate i, there exists t_i such that i does not enter for any $-t \in [-T, -t_i)$.*

The actual statement of this result is rather complicated and thus is provided at the end of this section after a presentation of the analysis (Section 4.1.4).

To start the analysis, let (x, X) denote the set of histories at which candidate A has entered at x and candidate B has not entered. Other sets of histories are denoted in an analogous manner. Abuse notation to write "x_i" to mean $\{x_i\}$ as part of the argument of v_i. For each $x_i \in X$, let $BR_j(x_i)$ be the set of candidate j's best responses against candidate i's policy x_i:

$$BR_j(x_i) = \arg\max_{X_j \in \mathcal{X}_j} v_j(X_j, x_i),$$

and suppose that it is non-empty. To simplify the notation, we sometimes write $x_j \in BR_j(x_i)$ to mean $\{x_j\} \in BR_j(x_i)$.

We say that $X_i^* \subseteq X$ is candidate i's **optimal set** if the following hold for each $x_i^* \in X_i^*$.

1. $v_i^{BR_j} := \sup_{x_j \in BR_j(x_i^*)} v_i(x_i^*, x_j) \geq \sup_{x_i \notin X_i^*, x_j \in BR_j(x_i)} v_i(x_i, x_j)$.
2. $v_i(x_i^*, X) = \sup_{x_i \in \mathcal{X}_i} v_i(x_i, X)$.

Note that the equality in part 1 holds if (v_i, v_j) is constant-sum, which happens for example if candidates are purely office-motivated. For general (v_i, v_j), it is straightforward that the definition of the optimal set ensures that there exits a unique largest optimal set (the optimal set that is a superset of all other optimal sets). Hereafter, let X_i^* be the largest optimal set for candidate i.[52]

Assumption 1 *For each candidate i, the largest optimal set X_i^* is non-empty, and satisfies the following properties.*

1. *For any $x_i^* \in X^*$ and $x_j, x_j' \in BR_j(x_i^*)$, $v_i(x_i^*, x_j) = v_i\left(x_i^*, x_j'\right)$ holds.*
2. *For any $x_i^*, x_i^{**} \in X_i^*$, $v_j(x_i^*, X) = v_j(x_i^{**}, X)$ and $\max_{\{x_j\} \in \mathcal{X}_j} v_j(x_i^*, x_j) = \max_{\{x_j\} \in \mathcal{X}_j} v_j(x_i^{**}, x_j)$.*

Assumption 1 ensures that X_i^* is non-empty. Note that once i enters at x_i, j enters at some $x_j \in BR_j(x_i)$ in any PBE. Hence, conditional on any history such that i's opponent has not entered, if i enters, then she enters at some $x_i^* \in X_i^*$. In addition, i's expected payoff when she enters is uniquely pinned down. Moreover, fixing j's strategy, if i enters, then j's payoff is also pinned down uniquely. Assumption 1 thus implies that any $x_i \in X_i^*$ gives the same continuation payoff to both candidates i and j in any PBE.

Assumption 2 *For each candidate i and any $x_i^* \in X_i^*$, $v_i(x_i^*, X) \geq v_i^{BR_j}$.*

This assumption implies that, after i's entry, i cannot be better off by the opponent's subsequent entry. Define:

$$v_{i,t}(\text{enter}) = e^{-\lambda_j t} v_i(x_i^*, X) + \left(1 - e^{-\lambda_j t}\right) v_i^{BR_j}.$$

Assumption 1 implies that this is candidate i's expected payoff at time $-t$ when she enters (in any PBE), and Assumption 2 implies that $v_{i,t}(\text{enter})$ is weakly decreasing in t.

We consider the following three cases, depending on the incentives at the deadline.

[52] For each $x_i \notin X_i^*$, we have either "$v_i^{BR_j} > v_i(x_i, x_j)$ for all $x_j \in BR_j(x_i)$" or "$v_i(x_i^*, X) > v_i(x_i, X)$."

- Case 1: $v_i(X,X) > v_i(x_i^*, X)$ for each i.

- Case 2: $v_i(X,X) < v_i(x_i^*, X)$ for each i.

  - More generally, there exist $t_0 \geq 0$ and a number $v_{i,t_0}(X,X)$ such that the continuation payoff at time $-t_0$ given any history in (X,X) is equal to $v_{i,t_0}(X,X)$ in any PBE, and that $v_{i,t_0}(\text{enter}) > v_{i,t_0}(X,X)$ holds for each i.[53]

- Case 3: $v_A(X,X) > v_A(x_A^*, X)$ and $v_B(X,X) < v_B(x_B^*, X)$.[54]

4.1.1 Case 1: No Candidate Enters at the Deadline

In this case, uniqueness and long ambiguity hold without additional assumptions, as follows.

Proposition 9 *Consider Case 1. Under Assumptions 1 and 2, there exists a PBE. In any PBE, at histories in (X,X), candidate i does not enter at any $-t \in (-\infty, 0]$.*

The intuition is simple: Candidate i's entry at time $-t$ results in either $v_i(x_i^*, X)$ if the opponent j does not enter afterward, or $v_i^{BR_j}$ if j does. Given that no candidate enters at histories in (X,X) after time $-t$, the former payoff is lower than the payoff from not entering, $v_i(X,X)$, by the definition of Case 1, and the latter is weakly lower due to Assumption 2.

4.1.2 Case 2: Both Candidates Enter at the Deadline

Fix t_0 that defines Case 2. For $t > t_0$, define $\bar{v}_{i,t}(\text{not})$ as candidate i's expected continuation payoff at time $-t$ when she does not enter, assuming that each candidate will enter at times in $(-t, -t_0)$ upon receiving an opportunity. Such a payoff is well defined due to Assumption 1.[55] Let

$$t_i^* \equiv \inf\{t > t_0 : \bar{v}_{i,t}(\text{not}) \geq v_{i,t}(\text{enter})\}.$$

Given the continuity of the continuation payoffs in time, $-t_i^*$ is the time closest to the deadline at which candidate i is indifferent between entering and not entering.

[53] "$v_i(X,X) < v_i(x_i^*, X)$ for each i" corresponds to taking $t_0 = 0$.

[54] The case with $v_A(X,X) < v_A(x_A^*, X)$ and $v_B(X,X) > v_B(x_B^*, X)$ is symmetric.

[55] The formal expression of this payoff is complicated, so we relegate it to the Online Appendix.

Assumption 3 (Genericity) *At least one of the following holds:* $v_i^{BR_j} < \sup_{\{x_i\}\in\mathcal{X}_i} v_i(x_i, X)$ *for each* i*, or* $t_A^* \neq t_B^*$*, or* $t_A^* = t_B^* = \infty$*.*

This assumption is a genericity assumption in the sense that the environment in which it is violated constitutes a degenerate (non-full-dimensional) space in the space of payoff functions.[56]

Proposition 10 *Consider Case 2. Under Assumptions 1, 2, and 3, there exists a PBE. There exists a profile* $(t_A, t_B) \in (\mathbb{R}_{++} \cup \{\infty\})^2$ *such that, for any PBE, at any histories in* (X, X)*, candidate* i *does not enter at any* $-t \in (-\infty, -t_i)$*, and enters at every time* $-t \in (-t_i, 0]$*. Moreover, if* $t_i^* \leq t_j^*$*, then* $t_i \leq t_j$ *and* $t_i = t_i^*$*.*

If $t_i < \infty$, then candidate i does not enter when the deadline is sufficiently far. The following condition, which is stronger than the condition for candidate i in Assumption 2, is a sufficient condition for $t_i < \infty$:

$$\text{First-mover disadvantage for } i \quad \begin{cases} v_i\left(X, x_j^*\right), v_i\left(x_i^*, X\right), v_i\left(X, X\right) \geq v_i^{BR_j} \\ \sup_{x_i \in X} v_i\left(x_i, x_j^*\right) > v_i^{BR_j} \end{cases}. \tag{6}$$

The second line of this condition states that, if both candidates have to enter and the order of the moves is known, then being the first mover is worse than being the second mover. The first line further requires that the disadvantage of being the first mover is so large, that it is the worst option even if we include the possibility of some candidates not specifying a policy. Intuitively, when it is the worst for candidate i to be best-responded by her opponent, i has little incentive to enter when the election day is far away. This is because when the election day is far away, the probability of the opponent best-responding in the future is high. In Section 4.1.4, we explain that this condition holds in various applications.

Proposition 11 *For each* i*, Proposition 10 holds with* $t_i < \infty$ *if we additionally require first-mover disadvantage for* i *to hold.*[57]

[56] To see why $t_A^* \neq t_B^*$ or $t_A^* = t_B^* = \infty$ holds generically, notice that, for each $i = A, B$ and $t < \infty$, $\bar{v}_{i,t}(\text{enter})$ is independent of $v_i(X, X)$, while $\bar{v}_{i,t}(\text{not})$ is strictly increasing in it. Hence, if there exists $w \in \mathbb{R}$ such that $t_A^* = t_B^* < \infty$ holds for some payoff function (v_A, v_B) such that $v_A(X, X) = w$, then $t_A^* \neq t_B^*$ holds for any payoff function that is the same as (v_A, v_B) except that $v_A(X, X) \neq w$.

[57] This result is not inconsistent with the case with $t_A^* = t_B^* = \infty$ which is allowed in Assumption 3 because the proof shows that if first-mover disadvantage for i holds then $t_i^* < \infty$.

4.1.3 Case 3: Only One Candidate Enters at the Deadline

We define $\bar{v}_{i,t}^{A}(\text{not})$ as candidate i's expected payoff at time $-t$ when she does not enter, assuming that only candidate B will enter at times in $(-t, 0]$ upon receiving an opportunity.[58] Such a payoff is well defined due to Assumption 1.

Let

$$\begin{aligned} \hat{t}_A &\equiv \inf\left\{t > 0 : \bar{v}_{A,t}^{A}(\text{not}) \leq v_{A,t}(\text{enter})\right\}; \\ \hat{t}_B &\equiv \inf\left\{t > 0 : \bar{v}_{B,t}^{A}(\text{not}) \geq v_{B,t}(\text{enter})\right\}. \end{aligned}$$

Given the continuity of the continuation payoffs in time, $\hat{t}_i$ is the time closest to the deadline at which i is indifferent between entering and not entering, respectively, assuming that only candidate B will enter afterward.

Assumption 4 (Genericity) *$\hat{t}_A \neq \hat{t}_B$ or $\hat{t}_A = \hat{t}_B = \infty$ holds.*

Lik Assumption 3, this assumption is again a genericity assumption. If $\hat{t}_A = \hat{t}_B = \infty$, then for each time $-t$ in any PBE, candidate A does not enter and candidate B enters. Hence we focus on the case in which $\hat{t}_A \neq \hat{t}_B$.

Proposition 12 *Consider Case 3. Under Assumptions 1, 2, and 4, there exists a PBE, and the following hold.*

1. *If $\hat{t}_A < \hat{t}_B$, then there exists $\bar{\varepsilon} > 0$ such that for all $\varepsilon \in (0, \bar{\varepsilon})$, in any PBE σ and its associated belief β, at any history at time $-(\hat{t}_A + \varepsilon)$ in (X, X), each candidate strictly prefers to enter under the continuation strategy given by σ and the belief β.*

2. *If $\hat{t}_A > \hat{t}_B$, then for any PBE, at any history in (X, X), no candidate enters at any $-t \in \left(-\infty, -\hat{t}_B\right)$.*

If $\hat{t}_A < \hat{t}_B$, we can use Proposition 10 in Case 2 to characterize any PBE, with a substitution that time t_0 is set to be equal to $\hat{t}_A + \varepsilon$ where $\varepsilon > 0$ is sufficiently small (and with an additional requirement of a genericity assumption (Assumption 3)). If $\hat{t}_A > \hat{t}_B$, in contrast, no candidate enters at any $-t \in \left(-\infty, -\hat{t}_B\right)$.

[58] The superscript denotes the candidate who does not enter close to the deadline in Case 3.

Corresponding to (6), define:

$$\text{Strong first-mover disadvantage for } i \begin{cases} v_i\left(x_i^*, X\right) > v_i^{BR_j} \\ \text{(6) holds if } \hat{t}_A < \hat{t}_B \end{cases}.$$

Putting the two parts of Proposition 12 together, we can show the following result:[59]

Proposition 13 *Consider Case 3 and suppose that strong first-mover disadvantage for i holds. Under Assumptions 1, 2, and 4, there exists a PBE, and for any PBE, there exists $t_i < \infty$ such that candidate i does not enter at any $-t \in (-\infty, -t_i)$.*

4.1.4 Summary

We are now ready to state our first general prediction:

Theorem 1 (Long Ambiguity) *Under Assumptions 1 and 2, the following claims are true.*

1. *Suppose $v_i(X, X) > v_i(x_i^*, X)$ for each i. Then, there exists a PBE, and in any PBE, candidate i does not enter at any history in (X, X) at any $-t \in (-\infty, 0]$.*

2. *Suppose $v_i(X, X) < v_i(x_i^*, X)$ for each i. Then, with additionally requiring Assumption 3, there exists a PBE. Moreover, if first-mover disadvantage for i holds, then there exists $t_i < \infty$ such that, for any PBE, candidate i does not enter at any history in (X, X) at any $-t \in (-\infty, -t_i)$.*

3. *Suppose that $v_i(X, X) > v_i(x_i^*, X)$ and $v_j(X, X) < v_j\left(x_j^*, X\right)$ for $i \neq j$. Then, with additionally requiring Assumption 4, there exists a PBE. Moreover, fix an arbitrary $k \in \{i, j\}$ and suppose that $\hat{t}_i > \hat{t}_j$ or strong first-mover disadvantage for candidate k holds. Then, there exists $t_k < \infty$ such that, for any PBE, candidate k does not enter at any history in (X, X) at any $-t \in (-\infty, -t_k)$.*

Although the conditions referred to in the theorem involve evaluation of variables that are endogenously determined in equilibrium (such as $\hat{t}_i$), they are fairly easy to check. For example, in the valence election campaign, the environment of Proposition 2 corresponds to part 3 of Theorem

[59] The proof shows that, if strong first-mover disadvantage for B holds, then $\hat{t}_B < \infty$ must hold.

1, where $i = W$ and $j = S$. It satisfies Assumptions 1, 2, and 4, and first-mover disadvantage for W. In addition, for Proposition 3, any symmetric office-motivated election campaign model with $(X, \mu) \notin \mathcal{M}$ satisfies Assumptions 1, 2, 3, and first-mover disadvantage for each candidate. Hence, part 2 of Theorem 1 applies.[60]

Since some of the assumptions in the above theorem are genericity conditions, we can also restate part of the theorem in a way that is easier to interpret, as follows.

Corollary 2 *Under Assumptions 1 and 2, the following claims are true.*

1. *Suppose $v_i(X, X) > v_i(x_i^*, X)$ for each i. Then, there exists a PBE. In any PBE, candidate i does not enter at any history in (X, X) at any $-t \in (-\infty, 0]$.*

2. *Suppose $v_i(X, X) < v_i(x_i^*, X)$ for each i. Then, generically in the space of payoff functions, the following holds. There exists a PBE, and if first-mover disadvantage for i holds, then there exists $t_i < \infty$ such that, for any PBE, candidate i does not enter at any history in (X, X) at any $-t \in (-\infty, -t_i)$.*

Note that the corollary states that we expect long ambiguity in many cases, but does not identify conditions under which we expect it. Theorem 1, in contrast, pins down the sufficient condition for when we expect long ambiguity.

4.2 The Dynamic Median-Voter Theorem

In this section, we consider an extension of the median voter theorem, which has an implication on several of our examples in Section 3. To this end, we focus on symmetric elections, that is, given any $\bar{X}, \bar{X}' \in \mathcal{X}_i = \mathcal{X}_j$, $v_i(\bar{X}, \bar{X}') = v_j(\bar{X}, \bar{X}')$.[61] A policy $x^* \in X$ is a **Condorcet winner** if, for each i, $x^* \in BR_i(x^*)$, $x^* \in BR_i(X)$, $v_i(x^*, X) > v_i(X, X)$, and for each $X_i \neq x^*$, $v_i(x^*, x_j) > v_i(X_i, x_j')$ for some $x_j \in BR_j(x^*)$ and $x_j' \in BR_j(X_i)$.[62] For example, in a uni-dimensional Downsian model

[60] One might think that part 2 of Theorem 1 can be applied to the analysis of the policy-motivated election campaign. However, Assumption 1 fails because the optimal set is empty in that example. Specifically, for candidate L, the intersection of the set of best responses to X, $\{(\frac{1}{2}, \frac{1}{2})\}$, and the set of best responses assuming the opponent's subsequent best response, $\{(\frac{2}{3}, 0), (0, \frac{2}{3})\}$, is disjoint. The example hence demonstrates that even outside the environment in which our assumptions hold, long ambiguity can be an equilibrium phenomenon, showing the robustness of the result.

[61] An extension to non-symmetric elections is straightforward, where one would define a profile of policies as Condorcet winners. Analogous results to the ones we present here as well as in Appendix D would go through.

[62] A policy is in the largest optimal set if it is a Condorcet winner and the stage game is constant sum.

in which (i) a candidate wins with probability one if the vote share is strictly greater than 1/2 and with probability 1/2 if the vote share is 1/2, and (ii) entering at the median voter ensures winning when the opponent does not enter (for example, the voters are risk averse and think that there is uncertainty about what policy a candidate announcing X would implement), the policy corresponding to the median voter is the unique Condorcet winner. In addition, the policy 1 with $\delta = 0$ in Proposition 1 and the policy $x^*(X, \mu)$ in the symmetric office-motivated election campaign with $(X, \mu) \in \mathcal{M}$ (part 1 of Proposition 3) are Condorcet winners. Note that, by definition, there is at most one Condorcet winner.

The following theorem extends the median voter theorem to a dynamic environment.

Theorem 2 (Dynamic Median-Voter) *Suppose that X is finite, (v_A, v_B) is a symmetric constant-sum game, and there exists a Condorcet winner. Then, there exists a PBE, and in any PBE, at any time $-t$, conditional on any history, (i) if the opponent's action is X, each candidate i announces the Condorcet winner and (ii) otherwise, each candidate i best-responds to the opponent's current policy.*

The theorem can be applied to prove that, in the examples mentioned above, candidates enter at the Condorcet winner specified above as soon as possible.

To see the intuition, note first that if a candidate obtains an opportunity at the deadline, then the assumption on the payoff function implies that she enters at the Condorcet winner. To show that this holds for all time $-t$, we resort to the continuous-time backward induction (formally presented in Appendix B), which in particular shows that it is not possible in any PBE that candidates keep using ambiguous language for a long time and try to enter to win at the last moment.

In Appendix D, we generalize the theorem to cover the case with non-constant-sum games. We show the existence of a PBE in which each candidate announces the Condorcet winner. We also show the uniqueness of a PBE when we further require that there is a policy that is strictly dominant for each i, while not requiring condition (5). This last result in particular implies that, in the dynamic campaign spending game, in subgames in which each candidate has already spent L, each candidate spends H as soon as possible.

4.3 The Constant-Sum Markov Theorem

In some of the examples we consider in Section 3, candidates are purely office-motivated, and thus their utility functions are constant-sum since the winning probabilities must add up to one. In this section, we provide a characterization of the equilibrium dynamics for constant-sum elections by showing that, in constant-sum elections, candidates' continuation payoffs at any history is determined only by the remaining time and the current policy set profile. Moreover, we show that it is irrelevant whether each candidate observes the arrival of the opponent's opportunities. More specifically, as specified in Section 2, we assume throughout the paper that each candidate cannot observe the arrivals of opportunities to the opponent but only the changes of the policy set. We compare such a setting with the model in which each candidate can observe the arrivals of the opponent's opportunities, including those that do not involve changes in the policy set. We call the former and the latter setups "private monitoring" and "public monitoring," respectively.

To define the setup of "public monitoring" formally, let $h^t = \left(\left(t_A^k, X_A^k\right)_{k=1}^{k_A}, \left(t_B^k, X_B^k\right)_{k=1}^{k_B}, t\right)$ be the entire history at $-t$, where $-t_j^k < -t$ is the time at which candidate j receives his or her k'th revision opportunity; X_j^k is the policy set that j has chosen at time $-t_j^k$; and t denotes the current remaining time. Let H be the set of all histories. We say that a history for candidate i at time $-t$, denoted h_i^t, is consistent with h^t if the former is given by deleting information about j's opportunities at which j did not change the policy set. Let $\theta\left(h^t\right) = \left(X_A^{k_A}, X_B^{k_B}\right)$ be the most recent policy profile at time $-t$; and $\theta_i\left(h^t\right) = X_i^{k_i}$ be candidate i's most recent policy at $-t$. Note that $\theta\left(h^t\right) = \theta\left(h_i^t\right)$ for each i and t. We allow the available policy set to depend on the current policy sets. Formally, let $\mathcal{X}_i(\theta) \subseteq 2^{\theta_i} \setminus \{\emptyset\}$ be the collection of available policy sets under $\theta \in \mathcal{X}_A \times \mathcal{X}_B$. Candidate i's strategy is a map $\sigma_i : H \to \cup_{\theta\in\mathcal{X}_A\times\mathcal{X}_B}\Delta\left(\mathcal{X}_i(\theta)\right)$, with a restriction that $\sigma_i\left(h^t\right) \in \Delta\left(\mathcal{X}_i\left(\theta\left(h^t\right)\right)\right)$. Let Σ_i be the space of i's strategies, and $\Sigma = \Sigma_A \times \Sigma_B$. With σ_i, candidate i takes $\sigma_i\left(h^t\right)$ if she has an opportunity at time $-t$ and takes $\theta_i\left(h^t\right)$ otherwise. A subgame-perfect equilibrium (SPE) can be defined in the standard manner. We call this setup "public monitoring."

In the "private monitoring" setup, the definition of PBE naturally extends to the case in which each candidate i's feasible announcements depend on the current announcement through the $\mathcal{X}_i(\cdot)$ function. Recall that a strategy profile (σ_A^*, σ_B^*) is a PBE if there exists a belief β such that, for each $i \in \{A, B\}$, (i) $\sigma_i^* \in \arg\max_{\sigma_i\in\Sigma_i} u_i^\beta(\sigma_i, \sigma_j^*|h_i^t)$ holds for every $h_i^t \in H_i$ and (ii) β is derived from Bayes rule whenever possible.

First, take an arbitrary PBE σ in private monitoring, and let $w_t^i\left(\sigma, h_i^t, X_i\right)$ be candidate i's continuation payoff of taking $X_i \in \mathcal{X}_i\left(\theta\left(h_i^t\right)\right)$ when her private history is h_i^t and she receives an opportunity at $-t$. Similarly, let $\hat{w}_t^i\left(\sigma, h_i^t, X_j\right)$ be candidate i's continuation payoff when her private history is h_i^t and candidate j receives an opportunity and takes $X_j \in \mathcal{X}_j\left(\theta\left(h_j^t\right)\right)$; and let $w_t^i\left(\sigma, h_i^t, no\right)$ be candidate i's continuation payoff when her private history is h_i^t and no candidate receives an opportunity at $-t$.

Second, take an arbitrary SPE $\bar{\sigma}$ in public monitoring, and let $W_t^i\left(\bar{\sigma}, h^t, X_i\right)$ be candidate i's continuation payoff of taking policy $X_i \in \mathcal{X}_i\left(\theta\left(h^t\right)\right)$ when the public history is h^t and she receives an opportunity at $-t$. Similarly, let $\hat{W}_t^i\left(\bar{\sigma}, h^t, X_j\right)$ be candidate i's continuation payoff when the public history is h^t and candidate j receives an opportunity and takes $X_j \in \mathcal{X}_j\left(\theta\left(h^t\right)\right)$; and let $W_t^i\left(\bar{\sigma}, h^t, no\right)$ be candidate i's continuation payoff when the public history is h^t and no candidate receives an opportunity at $-t$.

We can show that the continuation payoff of choosing policy X_i and not receiving an opportunity depends only on the current time $-t$ and the current policy set of the opponent $\theta_j\left(h^t\right)$ (recall that $\theta_j\left(h^t\right) = \theta_j\left(h_i^t\right) = \theta_j\left(h_j^t\right)$).

Theorem 3 (Constant-Sum Markov) *Suppose $v_A(X_A, X_B) + v_B(X_A, X_B) = 1$ for each $(X_A, X_B) \in \mathcal{X}_A \times \mathcal{X}_B$. Then, there exists $v_{i,t} : \mathcal{X}_i \times \mathcal{X}_j \to \mathbb{R}$ such that, for any PBE σ under private monitoring, SPE $\bar{\sigma}$ under public monitoring, public history h^t, private history h_i^t consistent with h^t, and $(X_i, X_j) \in \mathcal{X}_i \times \mathcal{X}_j$, we have*

$$w_t^i\left(\sigma, h_i^t, X_i\right) = W_t^i\left(\bar{\sigma}, h^t, X_i\right) = v_{i,t}\left(X_i, \theta_j\left(h^t\right)\right); \tag{7}$$

$$\hat{w}_t^i\left(\sigma, h_i^t, X_j\right) = \hat{W}_t^i\left(\bar{\sigma}, h^t, X_j\right) = v_{i,t}\left(\theta_i\left(h^t\right), X_j\right) \tag{8}$$

and

$$w_t^i\left(\sigma, h_i^t, \theta_i\left(h_i^t\right)\right) = \hat{w}_t^i\left(\sigma, h_i^t, \theta_j\left(h_i^t\right)\right) = w_t^i\left(\sigma, h_i^t, no\right) \tag{9}$$

$$= W_t^i\left(\bar{\sigma}, h^t, \theta_i\left(h^t\right)\right) = \hat{W}_t^i\left(\bar{\sigma}, h^t, \theta_j\left(h^t\right)\right) = W_t^i\left(\bar{\sigma}, h^t, no\right) = v_{i,t}\left(\theta\left(h^t\right)\right).$$

In the revision games with public monitoring, Gensbittel et al. (2017) show that the minimax theorem holds. In addition, Lovo and Tomala (2016) show the existence of Markov perfect equilibrium (MPE) with a finite equilibrium payoff after each profile of current policy sets. Putting them together, we obtain the results for public monitoring.

The theorem's contribution is to show that the equilibrium continuation payoff under private monitoring is the same as the one under public monitoring, and hence depends only on the current policy set profile. In the private monitoring case, for any Markov strategy of candidate i (where i's Markov strategies refer to those that depend only on the current policy set profile, the current time, and whether i receives an opportunity), there exists a best response by j that is Markov. This implies that i can guarantee her minimax value as in the public monitoring case. Since the symmetric argument implies that candidate j can guarantee his minimax value too, the equilibrium continuation payoff is uniquely determined.[63]

In the valence election campaign and symmetric office-motivated election campaign, candidates are office-motivated, so the payoffs are constant-sum. Also, the case with $p = 1$ in the election with incomplete information (the case with no normal types presented in Proposition 7) can be thought of as a constant-sum game after an elimination of strictly dominated strategies. Moreover, these models satisfy (5). Hence, in each of those models, the outcome characterized under private monitoring is outcome-equivalent to the one under public monitoring, and the continuation payoffs are uniquely determined for each time $-t$.

5 Conclusion

We have introduced the first model of dynamic campaigns into the literature on elections, which we call "policy announcement timing game." In the model, candidates cannot always announce their policies, but stochastically obtain opportunities to announce their policies or spend their funds. We applied the model to various examples, demonstrating that the introduction of such a simple friction to the model generates interesting dynamic strategic considerations and equilibrium dynamics consistent with election dynamics in reality. In particular, we showed that it is useful to analyze the candidates' motivations to defer a clear announcement of policies, depending on the opponent's latest announcement and the time left until the election; and to keep the budget for later use, depending on the opponent's cumulative spending and the time left. Depending on the

[63] Although it is intuitive, we do not know whether the result extends to the case with non-constant sum games. The (generic) uniqueness of continuation payoffs is an open question in the revision-games literature.

environment that the candidates face, they may or may not have such incentives for ambiguity. The insights from the examples are generalized in the Long Ambiguity Theorem, the Dynamic Median Voter Theorem, and the Constant-Sum Markov Theorem. Our work raises a wide range of new questions.

First, except for Sections 3.4 and 4.3, we restricted ourselves to the case in which policies are either perfectly ambiguous or perfectly precise. One could allow for "intermediate language" and analyze how gradually candidates shift from ambiguous to clear language over the course of the campaign. For example, in a uni-dimensional Downsian model, one could let the candidates choose any subintervals of $[0, 1]$ for the initial opportunity, and from the next opportunity, let them choose any subintervals included in their most recent announcements. In the multi-dimensional case, we can consider a model where a candidate commits to a policy in one dimension first, and then commits to a policy in another dimension. We consider a model of such a case in the Online Appendix.

Second, it would be more realistic to assume that policy announcements are sometimes synchronous and sometimes asynchronous. Although this problem seems nontrivial as Ishii and Kamada (2011) show in their analysis of revision games with synchronous and asynchronous revisions, we conjecture that there should remain the incentive to announce an ambiguous policy when the deadline is far. We consider a model of such a case in the Online Appendix.

Third, we restricted ourselves to the case in which, once a candidate commits to a particular policy, he or she cannot overturn it later. Although we believe that this is a reasonable starting point for analysis, one could also assume that candidates can change their policies if they are willing to incur a "reputational cost" for announcing "inconsistent" policies. The idea is that if a candidate overturns his or her opinion, voters would infer that it is likely that the candidate would change policies even after the election.

Fourth, it would be interesting to enrich the model by assuming that the median voter's position gets gradually revealed over the course of the campaign (for example, because of polls), so that candidates have an additional reason to wait. Our analysis in Section 3.5 shows that our general model can be extended to cover the cases involving incomplete information.

Fifth, the model of dynamic campaign spending could be enriched to test hypotheses for the "June Puzzle." This is a puzzle that asks why the Obama campaign significantly outspent the

Romney campaign in June 2012, even though the election was in November and the effect of TV advertisements on voter's preferences is known to be short-lived.[64] An explanation for this puzzle argues that popularity in the early stages may help with gathering more donations. Another explanation claims that if the opponent's popularity is below a certain level, then that opponent will "never come back." It will be interesting to enrich the model of Section 3.4 to analyze these hypotheses.

Sixth, we have considered two-candidate elections, but it would be interesting to consider more than two candidates.[65] In such an environment, there is no pure-strategy equilibrium in a static election game, while we can hope for the existence of an (essentially) unique pure-strategy PBE in a corresponding election campaign game, just as in the case with the multi-dimensional policy space.

Finally, our work raises empirical questions as well. For example, first, our model predicts different patterns of the timing of policy clarification/campaign spending for different parameter values. For example, in the valence election campaign, p, which measures how much uncertainty candidates face with respect to the position of the median voter, affects the timing of policy announcements. One may want to test whether this prediction is supported by the data.[66] The second example is about the case with a multi-dimensional policy space. In that model, we obtained a unique prediction about the entry timing and announced policies (when candidates are policy-motivated). The uniqueness may be useful in empirically testing the model.

References

[1] **Alesina, Alberto, and Alex Cukierman (1990)**, "The Politics of Ambiguity," *Quarterly Journal of Economics*, 4, 829–50.

[2] **Alesina, Alberto, and Richard Holden (2008)**, "Ambiguity and Extremism in Elections," unpublished.

[64] See footnote 10 for the discussion on the depreciation of the effect of earlier spending. We thank Avidit Acharya for sharing the story of June Puzzle, who attributes the story to Seth Hill, Brett Gordon, and Michael Peress.

[65] We thank Alessandro Lizzeri for pointing this out.

[66] As mentioned in Remark 1, this pattern is roughly consistent with the empirical finding in Campbell (1983).

[3] **Ambrus, Attila, and Shih-En Lu (2015)**, "A Continuous-Time Model of Multilateral Bargaining," *AEJ: Microeconomics*, 7, 208–249.

[4] **Ansolabehere, Stephen, and James M. Snyder Jr. (2000)**, "Valence Politics and Equilibrium in Spatial Election Models," *Public Choice*, 103, 327–336.

[5] **Aragonès, Enriqueta, and Zvika Neeman (2000)**, "Strategic Ambiguity in Electoral Competition," *Journal of Theoretical Politics*, 12, 183–204.

[6] **Aragonès, Enriqueta, and Thomas R. Palfrey (2002)**, "Mixed Equilibrium in a Downsian Model with a Favored Candidate," *Journal of Economic Theory*, 103, 131–161.

[7] **Aragonès, Enriqueta, and Andrew Postlewaite (2002)**, "Ambiguity in Election Games," *Review of Economic Design*, 7, 233–255.

[8] **Asahi Shimbun Digital (2014a)**, "Hosokawa-Koizumi 'Ex Prime Minister Union,' If Realized, May Change the Scene for Tokyo Gubernatorial Election" (in Japanese), January 14, https://web.archive.org/web/20140121185951/http://www.asahi.com/articles/ASG187WCQG18UTIL04R.html?ref=reca, accessed on 04/23/2018.

[9] **Asahi Shimbun Digital (2014b)** "The Focus is on Anti-Nuclear Power in the Tokyo Gubernatorial Election: Hosokawa and Masuzoe Declared to Run for Election" (in Japanese), January 15, https://web.archive.org/web/20140228074009/http://www.asahi.com/articles/ASG1G5R5SG1GUTIL03Y.html?ref=reca, accessed 04/23/2018.

[10] **Austen-Smith, David (1987)**, "Interest Groups, Campaign Contributions, and Probabilistic Voting," *Public Choice*, 54, 123–139.

[11] **Bailey, Michael A. (2002)**, "Money and Representation: An Exploration in Multiple Dimensions with Informative Campaigns," unpublished.

[12] **Banda, Kevin K. (2013)**, "The Dynamics of Campaign Issue Agendas," *State Politics & Policy Quarterly*, 13, 446–470.

[13] **Banda, Kevin K. (2015)**, "Competition and the Dynamics of Issue Convergence," *American Politics Research*, 43, 821–845.

[14] **Banks, Jeffrey S. (1990)**, "A Model of Electoral Competition with Incomplete Information," *Journal of Economic Theory*, 50, 309–325.

[15] **Baron, David P. (1994)**, "Electoral Competition with Informed and Uninformed Voters," *American Political Science Review*, 88, 33–47.

[16] **Berger, Mark M., Michael C. Munger, and Richard F. Potthoff (2000)**, "With Uncertainty, the Downsian Model Predicts Divergence," *Journal of Theoretical Politics*, 12, 262–268.

[17] **Bernhardt, Daniel, and Daniel Ingberman (1985)**, "Candidate Reputations and the Incumbency Effect," *Journal of Public Economics*, 27, 47–67.

[18] **Calcagno, Riccardo, Yuichiro Kamada, Stefano Lovo, and Takuo Sugaya (2014)**, "Asynchronicity and Coordination in Common and Opposing Interest Games," *Theoretical Economics*, 9, 409–434.

[19] **Callander, Steven, and Simon Wilkie (2007)**, "Lies, Damned Lies, and Political Campaigns," *Games and Economic Behavior*, 60, 262–286.

[20] **Callander, Steven, and Catherine H. Wilson (2008)**, "Context-Dependent Voting and Political Ambiguity," *Journal of Public Economics*, 92, 565–581.

[21] **Calvo, Guillermo A. (1983)**, "Staggered Prices in a Utility-Maximizing Framework," *Journal of Monetary Economics*, 12, 383–398.

[22] **Campbell, James E. (1983)**, "Ambiguity in the Issue Positions of Presidential Candidates: A Causal Analysis," *American Journal of Political Science*, 27, 284–293.

[23] **Carter, Jennifer and John W. Patty (2015)**, "Valence and Campaigns." *American Journal of Political Science* 59, 825–840.

[24] **Caruana, Guillermo and Liran Einav (2008)**, "A Theory of Endogenous Commitment," *Review of Economic Studies*, 75, 99–116.

[25] **Coate, Stephen (2004)**, "Political Competition with Campaign Contributions and Informative Advertising," *Journal of the European Economic Association*, 2, 772–804.

[26] **Coughlin, Peter J. (1992)**, *Probabilistic Voting Theory*, New York: Cambridge University Press.

[27] **van Damme, Eric, and Sjaak Hurkens (1996)**, “Commitment Robust Equilibria and Endogenous Timing,” *Games and Economic Behavior*, 15, 290–311.

[28] **Gale, Douglas (1995)**, “Dynamic Coordination Games,” *Economic Theory*, 5, 1–18.

[29] **Gale, Douglas (2001)**, “Monotone Games with Positive Spillovers,” *Games and Economic Behaviors*, 37, 295–320.

[30] **Gensbittel, Fabien, Stefano Lovo, Jérôme Renault, and Tristan Tomala (2017)**, “Zero-Sum Revision Games, *Games and Economic Behavior*, Forthcoming.

[31] **Glazer, Amihai (1990)**, “The Strategy of Candidate Ambiguity,” *American Political Science Review*, 84, 237–41.

[32] **Gerber, Alan, James G. Gimpel, Donald P. Green, and Daron R. Shaw (2011)**, “How Large and Long-lasting Are the Persuasive Effects of Televised Campaign Ads? Results from a Randomized Field Experiment,” *American Political Science Review*, 105, 135–150.

[33] **Groseclose, Tim (2001)** “A Model of Candidate Location When One Candidate Has a Valence Advantage,” *American Journal of Political Science*, 45, 862–886.

[34] **Grossman Gene M., and Elhanan Helpman (1996)**, “Electoral Competition and Special Interest Politics,” *Review of Economic Studies*, 63, 256–286.

[35] **Gurciullo, Brianna, and Gabriel Debenedetti (2016)** “Sanders: ‘We are in this campaign to win’ ” in *POLITICO*, https://www.politico.com/blogs/2016-dem-primary-live-updates-and-results/2016/04/bernie-sanders-jeff-weaver-victory-222536, accessed on 04/19/2018.

[36] **Hamilton, Jonathan H., and Steven M. Slutsky (1990)**, “Endogenous Timing in Duopoly Games: Stackelberg or Cournot Equilibria,” *Games and Economic Behavior*, 2, 29–46.

[37] **Hamilton, Jonathan H., and Steven M. Slutsky (1993)**, "Endogenizing the Order of Moves in Matrix Games," *Theory and Decision*, 34, 47–62.

[38] **Harrington, Jr. , Joseph E. (1992)**, "The Revelation of Information Through the Electoral Process: An Exploratory Analysis," *Economics and Politics*, 4, 255–275.

[39] **Harris, Tobias (2013)** "Shinzo Abe's Constitution Quest: Instead of focusing on economics, the Japanese leader risks it all to change the rules of governance" in *Wall Street Journal*, http://jp.wsj.com/articles/SB10001424127887324582304578488313061482222, accessed 05/26/2018.

[40] **Hummel, Patrick (2010)**, "On the Nature of Equilibria in a Downsian Model with Candidate Valence," *Games and Economic Behavior*, 70, 425–445.

[41] **Ishii, Yuhta, and Yuichiro Kamada (2011)**, "The Effect of Correlated Inertia on Coordination," unpublished.

[42] **Kamada, Yuichiro, and Michihiro Kandori (2017a)**, "Revision Games, Part I: Theory," unpublished.

[43] **Kamada, Yuichiro, and Michihiro Kandori (2017b)**, "Revision Games, Part II: Applications and Robustness" unpublished.

[44] **Karazas, Ionnis, and Steven E. Shreve (1988)**, *Brownian Motion and Stochastic Calculus*, New York: Springer-Verlag.

[45] **Kartik, Navin, and R. Preston McAfee (2007)**, "Signaling Character in Electoral Competition," *American Economic Review*, 97, 852–870.

[46] **Klenow, Peter J. and Benjamin A. Malin (2010)**, "Microeconomic Evidence on Price-Setting," *Handbook of Monetary Economics*, 3, 231–284.

[47] **Krasa, Stefan and Mattias Polborn (2010)**, "The Binary Policy Model," *Journal of Economic Theory*, 145, 661–688.

[48] **Lindbeck, Assar, and Jorgen Weibull (1987)**, "Balanced Budget Redistribution as the Outcome of Political Competition," *Political Choice*, 52, 273-297.

[49] **Lipman, Barton L. and Ruqu Wang (2000)**, "Switching Costs in Frequently Repeated Games," *Journal of Economic Theory*, 93, 149–190.

[50] **Lovo, Stefano, and Tristan Tomala (2015)**, "Markov Perfect Equilibria in Stochastic Revision Games," unpublished.

[51] **McGrane, Reginald C. (ed.) (1919)**, *The Correspondence of Nicholas Biddle*. Boston, New York: The Houghton Mifflin Company.

[52] **Meirowitz, Adam (2005)**, "Informational Party Primaries and Strategic Ambiguity," *Journal of Theoretical Politics*, 17, 107–136.

[53] **Moroni, Sofia (2018)**, "Existence of Trembling Hand Perfect and Sequential Equilibrium in Games with Stochastic Timing of Moves," unpublished.

[54] **Ota, Shinji (2016)** "2016 House of Councillors Election, Prospect of Constitutional Reform and Politics" (in Japanese), *Jiji-Koron*, http://www.nhk.or.jp/kaisetsu-blog/100/248861.html, accessed on 04/19/2018.

[55] **Page, Benjamin I. (1976)**, "The Theory of Political Ambiguity," *American Political Science Review*, 70, 742–752.

[56] **Palfrey, Thomas R. (1984)**, "Spatial Equilibrium with Entry," *Review of Economic Studies*, 51, 139–156.

[57] **Page, Benjamin I. (1978)**, *Choices and Echoes in Presidential Elections*, Chicago: University of Chicago Press.

[58] **Persson, Torsten, and Guido Tabellini (2000)**, *Political Economics: Explaining Economic Policy*, Cambridge: MIT Press.

[59] **Potters, Jan, Randolph Sloof, and Frans van Winden (1997)**, "Campaign Expenditures, Contributions and Direct Endorsements: The Strategic Use of Information and Money to Influence Voter Behavior," *European Journal of Political Economy*, 13, 1–31.

[60] **Prat, Andrea (2002a)**, "Campaign Spending with Office-Seeking Politicians, Rational Voters, and Multiple Lobbies," *Journal of Economic Theory*, 103, 162–189.

[61] **Prat, Andrea (2002b)**, "Campaign Advertising and Voter Welfare," *Review of Economic Studies*, 69, 999–1018.

[62] **Renou, Ludovic (2009)**, "Commitment games," *Games and Economic Behavior*, 66, 488–505.

[63] **Roemer, John E. (2001)**, *Political Competition: Theory and Applications*, Cambridge: Harvard University Press.

[64] **Romano, Richard, and Huseyin Yildirim (2005)**, "On the endogeneity of Cournot–Nash and Stackelberg equilibria: games of accumulation," *Journal of Economic Theory*, 120, 73–107.

[65] **Sankei News (2013)**, "Survey by the Liberal Democratic Party Reveals Strong Supports for Masuzoe and Ikegami for Tokyo Gubernatorial Election" (in Japanese), December 24, https://web.archive.org/web/20140629090111/http://sankei.jp.msn.com/politics/news/131224/elc13122421110000-n1.htm, accessed 04/23/2018.

[66] **Strauss, Daniel (2016)** "Sanders vows to go all the way, but signals a new focus on influencing Clinton" in *POLITICO*, https://www.politico.com/blogs/2016-dem-primary-live-updates-and-results/2016/04/bernie-sanders-election-results-convention-222506, accessed on 04/19/2018.

[67] **Shepsle, Kenneth (1972)**, "The Strategy of Ambiguity: Uncertainty and Electoral Competition," *American Political Science Review*, 66, 555–568.

[68] **Stein, Jeff (2016)** "Bernie Sanders moved Democrats to the left. The platform is proof." in *Vox*, https://www.vox.com/2016/7/25/12281022/the-democratic-party-platform, accessed on 05/07/2018.

[69] **White, Joe Slade (2012)**, "The 9 Principles of Winning Campaigns," http://joesladewhite.com/wp-content/uploads/2015/01/9-principles-blue-jsw.pdf, accessed on 05/07/2018.

A Structure of the Appendix

We first state and prove the continuous-time backward induction, which turns out to be useful in many proofs. Second, we offer the proofs of the results. Although we present the applications before the general theorems in the main text to highlight the applicability of the model, since the general theorems are useful for proving the results in the applications, here we prove the general theorems. The proofs of the results in the applications can be found in the Online Appendix.

B Continuous-Time Backward Induction

The following result, which we call continuous-time backward induction, is due to Calcagno et al. (2014), and is repeatedly used in the proofs of this paper. We reproduce its statement and proof for reader's convenience.

Lemma 5 *Suppose that for any $t \in [0, \infty)$, there exists $\varepsilon > 0$ such that statement $A_{t'}$ is true for all $t' \in [t, t+\varepsilon)$ if statement $A_{t''}$ is true for any $t'' < t$. Then, for any $t \in [0, \infty)$, statement A_t is true.*

Proof. Suppose that the premise of the lemma holds. Let $-t^*$ be the supremum of $-t$ such that A_t is false. If $t^* = \infty$, we are done. So suppose that $t^* < \infty$. Then it must be the case that for any $\varepsilon > 0$, there exists $-\tau \in (-t^* - \varepsilon, -t^*]$ such that A_τ is false. But by the definition of t^*, there exists $\tilde{\varepsilon} > 0$ such that statement A_τ is true for all $-\tau \in (-t^* - \tilde{\varepsilon}, -t^*]$ because the premise of the lemma is true. This is a contradiction. ■

C Proof of Theorem 1

The most important proofs for Theorem 1 are here. Other proofs for the theorem can be found in the Online Appendix.

Proof of Proposition 11.

Suppose $t_j = \infty$. On the one hand, since candidate j enters whenever she has an opportunity, if candidate i does not enter until j enters, the payoff i obtains at time $-t$ converges to $\sup_{x_i} v_i\left(x_i, x_j^*\right)$ as $t \to \infty$. On the other hand, we have $v_{i,t}(\text{enter}) \to v_i^{BR_j}$ as $t \to \infty$ since candidate j will have an opportunity afterward with a probability converging to 1 as $t \to \infty$. Hence, for sufficiently large t, not entering is the unique best response at time $-t$ in any PBE given that first-mover disadvantage for i holds.

Suppose next $t_j < \infty$. On the one hand, for each $-t < -t_j$, if candidate i does not enter until $-t_j$, i obtains the payoff $\bar{v}_{i,t_j}(\text{not})$, which is a convex combination of $v_i\left(X, x_j^*\right)$, $v_i\left(x_i^*, X\right)$, $v_i\left(X, X\right)$, $\sup_{x_i} v_i\left(x_i, x_j^*\right)$, and $v_i^{BR_j}$. Moreover, since j enters at times sufficiently close to time $-t_0$ (given the continuity of the continuation payoff in time), we have a strictly positive weight on $\sup_{x_i} v_i\left(x_i, x_j^*\right)$ in the convex combination. On the other hand, we have $v_{i,t}(\text{enter}) \to v_i^{BR_j}$ as $t \to \infty$. Hence, for sufficiently large t, not entering is the unique best response at time $-t$ in any PBE given that first-mover disadvantage for i holds. ■

Proof of Proposition 13.

Suppose first that strong first-mover disadvantage for A holds. In this case, if $\hat{t}_A = \hat{t}_B = \infty$, then by the definition of $\hat{t}_A$, A never has an incentive to enter at any time. Hence the conclusion of the proposition holds with $t_A = 0$. If $\hat{t}_A > \hat{t}_B$, then part 2 of Proposition 12 implies that A never has an incentive to enter at all times strictly before $-\hat{t}_B$. Hence the conclusion of the proposition holds with $t_A = \hat{t}_B$. If $\hat{t}_A < \hat{t}_B$, then part 1 of Proposition 12 implies that we can use the argument for Case 2. Hence, Proposition 11 implies that A never has an incentive to enter at all times strictly before $-\tilde{t}$ where $\tilde{t}$ is equal to t_A that we take in Proposition 10. Hence the conclusion of the proposition holds with $t_A = \tilde{t}$.

Next, suppose that strong first-mover disadvantage for B holds. First, we show $\hat{t}_B < \infty$. To see why this holds, observe the following. On the one hand, the payoff from entering at $-t$ converges to $v_B^{BR_A}$ as $t \to \infty$. On the other hand, for an arbitrary fixed $\bar{t} \in (0, \infty)$, not entering until $-\bar{t}$ (and entering when an opportunity arrives at time $-t \geq -\bar{t}$) gives the payoff that is a convex combination of $v_B\left(x_B^*, X\right)$, $v_B^{BR_A}$, and $v_B\left(X, X\right)$ with a strictly positive weight on $v_B\left(x_B^*, X\right)$ (recall that $\hat{t}_B$ is calculated assuming that candidate A never enters unless B has entered). Hence, for sufficiently large t, not entering is better at time $-t$ in any PBE given strong first-mover disadvantage for B. Thus, we have $\hat{t}_B < \infty$. With this condition, we obtain the desired result as in the case of strong first-mover disadvantage for A, using part 2 of Proposition 12 for the case of $\hat{t}_A > \hat{t}_B$ and Proposition 11 for the case of $\hat{t}_A < \hat{t}_B$. ■

D Proof of a Generalized Version of Theorem 2

In in part 3 of the following theorem, we consider general games in which it is not necessarily the case that $\mathcal{X}_i = \{\{x\}|x \in X\} \cup \{X\}$ holds. In such a game, when $\{x_i^*\} \in \mathcal{X}_i$, we say that x_i^* is a **strictly dominant policy** if for all $X_i \in \mathcal{X}_i \setminus \{x_i^*\}$, $v_i(\{x_i^*\}, X_j) > v_i(X_i, X_j)$ for all $X_j \in \mathcal{X}_j$.

We note that, even though we require (v_A, v_B) to be symmetric in the following theorem, it is straightforward to extend the result to the case with non-symmetric cases, as described in footnote 61.

Theorem 4 (General Dynamic Median-Voter) 1. *Suppose that (v_A, v_B) is symmetric and, for each $i = A, B$, x^* is a Condorcet winner. Then, there exists a PBE in which, at any time $-t$, conditional on any history, (i) if the opponent's current policy set is X, each candidate i announces $\{x^*\}$, and (ii) otherwise, each candidate i chooses a static best response to the opponent's current policy.*

2. *Suppose that X is finite, (v_A, v_B) is a symmetric constant-sum game, and there exists a Condorcet winner. Then, there exists a PBE, and in any PBE, at any time $-t$, conditional on any history, (i) if the opponent's current policy set is X, each candidate i announces the Condorcet winner and (ii) otherwise, each candidate i chooses a static best response to the opponent's current policy.*

3. *Suppose that X is finite and (v_A, v_B) is symmetric. Consider an environment with a general $(\mathcal{X}_A, \mathcal{X}_B)$. Suppose that x_i^* is a strictly dominant policy for each $i = A, B$. Then, in any PBE, at any time $-t$, conditional on any history, each candidate i announces $\{x^*\}$.*

As will be seen in the Online Appendix, part 3 of Theorem 4 does not hold if we replace "strictly dominant policy" with "weakly dominant policy."

Proof. Part 1: Let x^* be a Condorcet winner. First, note that after the opponent enters at x^*, the given strategy specifies a best response. Second, we show that it is optimal for each candidate to announce x^* at time $-t$ if j's current policy set is X. Take an arbitrary $\bar{x}_j(x_i) \in \max_{x_j \in X} v_i(x^*, x_j)$. By the assumption that x^* is a Condorcet winner, it follows that for any $x_i \in X$, there exists $\bar{x}_j(x_i) \in BR_j(x_i)$ such that $v_i(x^*, \bar{x}_i(x^*)) > v_i(x_i, \bar{x}_j(x_i))$. We consider a strategy profile in which, once i enters at $x_i \in X$, j enters at $\bar{x}_j(x_i)$; and also if no one has entered, each i enters at x^*.

Take any time $-t \in [-T, 0]$. Suppose that for every time in $(-t, 0]$, conditional on any history, each candidate i announces x^*. Under the strategy profile specified above, the following are true.

If i announces x^*, her payoff is

$$e^{-\lambda_j t} v_i(x^*, X) + (1 - e^{-\lambda_j t}) v_i(x^*, \bar{x}_j(x^*)). \tag{10}$$

If i announces $x_i \neq x^*$, her payoff is

$$e^{-\lambda_j t} v_i(x_i, X) + (1 - e^{-\lambda_j t}) v_i(x_i, \bar{x}_j(x_i)). \tag{11}$$

If i announces X, her payoff is

$$\begin{aligned}&e^{-\lambda_j t}\left(e^{-\lambda_i t} v_i(X, X) + (1 - e^{-\lambda_i t}) v_i(x^*, X)\right) + (1 - e^{-\lambda_j t})\\&\left(e^{-\lambda_i t} v_i(X, x^*) + (1 - e^{-\lambda_i t})\left(w_i v_i(x^*, \bar{x}_j(x^*)) + (1 - w_i)\, v_i(\bar{x}_i(x^*), x^*)\right)\right),\end{aligned} \tag{12}$$

where $w_i \in (0, 1)$.

Note that the first term in (10) is weakly larger than the first terms in (11) and (12) due to the assumption that x^* is a Condorcet winner. In addition, the second term in (10) is weakly larger than the second term in (11) by the construction of the function $\bar{x}_j(\cdot)$. Finally, the second term in (10) is weakly larger than the second term in (12) due to $x^* \in BR_i(x^*)$ (implied by x^* being a Condorcet winner), and its implication that $v_i(\bar{x}_i(x^*), x^*) = v_i(x^*, x^*)$. This implies that, for any $t \geq 0$, (10) is weakly larger than both (11) and (12). Hence, it is optimal for each candidate to announce x^* at time $-t$ if j's current policy set is X.

Part 2: Let x^* be the Condorcet winner. For each $X_i \in \mathcal{X}_i$, we have $v_i(x_i^*, x_j) > v_i(X_i, x_j')$ for some $x_j \in BR_j(x^*)$ and $x_j' \in BR_j(X_i)$. Since (v_A, v_B) is constant-sum, we have $v_i(x^*, x_j) > v_i(X_i, x_j')$ for each $x_j \in BR_j(x^*)$ and $x_j' \in BR_j(X_i)$. This implies that, since $x^* \in BR_j(x^*)$, we have $v_i(x^*, x^*) = \min_{X_j \in \mathcal{X}_j} v_i(x^*, X_j) > \min_{X_j \in \mathcal{X}_j} v_i(X_i, X_j)$ for each $X_i \neq \{x^*\}$.

Take any time $-t \in [-T, 0]$. Suppose that, for every time in $[-t, 0]$, conditional on any history, (i) if the opponent's current policy set is X, each candidate i announces x^* and (ii) otherwise, each candidate i takes a static best response to the opponent's current policy. We show that, conditional on any history, if the opponent's current policy set is X at time $-t$, each candidate i has a strict incentive to announce x^* over announcing X or any singleton policy set $\{x_i\} \neq \{x^*\}$. Given the constant-sum assumption, if i announces x^*, her payoff is

$$e^{-\lambda_j t} v_i(x^*, X) + (1 - e^{-\lambda_j t}) v_i(x^*, x^*). \tag{13}$$

If i announces $x_i \neq x^*$, her payoff is

$$e^{-\lambda_j t} v_i(x_i, X) + (1 - e^{-\lambda_j t}) \min_{X_j \in \mathcal{X}_j} v_i(x_i, X_j). \tag{14}$$

If i announces X, her payoff is

$$e^{-\lambda_j t} \left(e^{-\lambda_i t} v_i(X, X) + (1 - e^{-\lambda_i t}) v_i(x^*, X) \right) + (1 - e^{-\lambda_j t}) \left(e^{-\lambda_i t} v_i(X, x^*) + (1 - e^{-\lambda_i t}) v_i(x^*, x^*) \right). \tag{15}$$

Given that x^* is the Condorcet winner and $x_i \neq x^*$, (13) is strictly larger than (14) and (15). Moreover, by the assumption that X is finite, there exits $\varepsilon > 0$ such that, for all $x_i \neq x^*$, the value in (13) is no less than the sum of ε and the value in (14), and also no less than the sum of ε and the value in (15). By continuity of the continuation payoff in time, this implies that there exists $\varepsilon' > 0$ such that i strictly prefers announcing x^* to announcing X or any singleton policy set $\{x_i\} \neq \{x^*\}$ at times in $(-t - \varepsilon', -t]$ if j's current policy set is X. Therefore, by the continuous-time backward induction, in any PBE, at any time $-t$, conditional on any history, each candidate i announces x^* if j's current announcement is X.

Part 3: Fix time $-t$, suppose that at all time strictly after $-t$, each candidate i enters at x_i^* conditional on any history. Then, if i announces x_i^* when the current policy set is (X_i, X_j), then her payoff is

$$e^{-\lambda_j t} v_i(x_i^*, X_j) + (1 - e^{-\lambda_j t}) v_i(x_i^*, x_j^*).$$

If i announces $X_i \neq \{x_i^*\}$ when the current policy set is (X_i, X_j), then her payoff is

$$e^{-\lambda_j t} \left(e^{-\lambda_i t} v_i(X_i, X_j) + (1 - e^{-\lambda_i t}) \bar{v}_i \right) + (1 - e^{-\lambda_j t}) \left(e^{-\lambda_i t} v_i(X_i, x_j^*) + (1 - e^{-\lambda_i t}) \bar{\bar{v}}_i \right),$$

where $\bar{v}_i, \bar{\bar{v}}_i \leq v_i(x_i^*, X_j)$. Note that $\bar{v}_i$ and $\bar{\bar{v}}_i$ are equal to $v_i(x_i^*, X_j)$ and $v_i(x_i^*, x_j^*)$, respectively, if $x_i^* \in X_i$, but they are respectively strictly less than those values otherwise, due to the definition of a strictly dominant policy.

Since $v_i(X_i, X_j) < v_i(x_i^*, X_j)$ and $v_i(X_i, x_j^*) < v_i(x_i^*, x_j^*)$ by the definition of a strictly dominant policy, the payoff from announcing $\{x_i^*\}$ is strictly greater than the payoff from announcing $X_i \neq \{x_i^*\}$. Hence, by the continuous-time backward induction, we obtain the desired result. ■

E Proof of Theorem 3

We first prove that conditions (7)–(9) hold for public monitoring. Using this result, we prove that conditions (7)–(9) hold for private monitoring.

E.1 Public Monitoring

We prove that conditions (7)–(9) hold for public monitoring:

Lemma 6 *Suppose $v_A(X_A, X_B) + v_B(X_B, X_A) = 1$ for each $(X_A, X_B) \in \mathcal{X}_A \times \mathcal{X}_B$. There exists $v_{i,t}(\theta)$ for each $\theta \in \mathcal{X}_i \times \mathcal{X}_j$ such that, for any SPE $\bar{\sigma}$ and h^t, we have*

$$W_t^i\left(\bar{\sigma}, h^t, X_i\right) = v_{i,t}\left(X_i, \theta_j\left(h^t\right)\right);$$

$$\hat{W}_t^i\left(\bar{\sigma}, h^t, X_j\right) = v_{i,t}\left(\theta_i\left(h^t\right), X_j\right);$$

and

$$W_t^i\left(\bar{\sigma}, h^t, \theta_i\left(h^t\right)\right) = \hat{W}_t^i\left(\bar{\sigma}, h^t, \theta_j\left(h^t\right)\right) = W_t^i\left(\bar{\sigma}, h^t, no\right) = v_{i,t}\left(\theta\left(h^t\right)\right).$$

Proof. In the revision games with public monitoring, the minimax theorem holds if there exists an equilibrium with a finite equilibrium payoff after any history (see Gensbittel et al. (2017)). Hence, we are left to show that there exists an equilibrium with a finite equilibrium payoff after each profile of policy sets, θ. Lovo and Tomala (2016) show the existence of Markov perfect equilibrium (MPE) with a finite equilibrium payoff after each θ, as desired. ■

E.2 Private Monitoring

Fix any σ_j (not necessarily an equilibrium strategy). Given h_i^t and $\theta\left(h_i^t\right) = (X, X)$, calculate candidate i's belief about h_j^t. For any time $-t$ and history h_i^t, $\theta_j\left(h_i^t\right) = X$ happens with a positive probability given (σ_i, σ_j) for any strategy σ_i of candidate i since it is possible that no candidate receives any opportunity in the time interval $[-T, -t)$. Hence the belief given $\theta_j\left(h_i^t\right) = X$ should satisfy Bayes rule. In particular, we can show that, since the arrivals of opportunities are independent between candidates, for any two histories of candidate i such that no commitment has been made at $-t$, candidate i's belief about h_j^t is the same. Denote by β^{σ_j} a belief to be explicit about the fact that the belief is solely determined by σ_j:

Lemma 7 *For any σ_j, there exists β^{σ_j} such that, for each h_i^t and $\tilde{h}_i^t$ with $\theta\left(h_i^t\right) = \theta\left(\tilde{h}_i^t\right) = (X, X)$, we have $\beta^{\sigma_j}\left(h_j^t|\tilde{h}_i^t\right) = \beta^{\sigma_j}\left(h_j^t|h_i^t\right) =: \beta^{\sigma_j}\left(h_j^t\right)$.*

Proof. Note that, if $\theta\left(h_i^t\right) = (X, X)$, then we have $\left(t_i^l, X_i^l\right)_{l=1}^{l_i} = \{\emptyset\}$,[67] that is, candidate i never changes her policy announcement. Let $H_j^{\sigma_j}\left(h_i^t\right)$ be the set of candidate j's histories compatible with h_i^t and σ_j.[68] Define $H_j^{\sigma_j}\left(\tilde{h}_i^t\right)$ analogously. Note that $H_j^{\sigma_j}\left(h_i^t\right)$ and $f_i(h_j^t|h_i^t)$ depend on h_i^t only through $\left(t_i^l, X_i^l\right)_{l=1}^{l_i} = \{\emptyset\}$. Hence, $H_j^{\sigma_j}\left(h_i^t\right) = H_j^{\sigma_j}\left(\tilde{h}_i^t\right)$ and $f_i(h_j^t|h_i^t) = f_i(h_j^t|\tilde{h}_i^t)$ for each h_i^t and $\tilde{h}_i^t$ with $\theta\left(h_i^t\right) = \theta\left(\tilde{h}_i^t\right) = (X, X)$. Thus, the result follows from (20). ■

Using this independence of the belief, we can show that candidate i's continuation payoff does not depend on h_i^t. Take any strategy profile σ (not necessarily an equilibrium). Let

$$\tilde{w}_t^i\left(\sigma_i, \sigma_j, h_i^t, X\right) = \int_{h_j \in H_j^{\sigma_j}\left(h_i^t\right)} u_i\left(\sigma_i, \sigma_j | \left(h_i^t, X\right), h_j^t\right) d\beta^{\sigma_j}\left(h_j^t | h_i^t\right)$$

be candidate i's payoff when she takes X given h_i^t, given that (i) candidate i takes a continuation strategy determined by σ_i and history $\left(h_i^t, X\right)$ for $(-t, 0]$,[69] and (ii) if candidate j has never received an opportunity before time $-t$ in h_j^t, she takes a continuation play determined by σ_j and history h_j^t for $(-t, 0]$. Note that (i) candidate i's decision X does not affect candidate i's belief $\beta^{\sigma_j}\left(\cdot|h_i^t\right)$; and (ii) the belief $\beta^{\sigma_j}\left(\cdot|h_i^t\right)$ does not depend on whether candidate i obtains an opportunity at time $-t$ by the independence of the Poisson processes.

[67] We follow the convention that, with $l_j < 1$, we define $\left(t_i^l, X_i^l\right)_{l=1}^{l_i} = \{\emptyset\}$.

[68] The formal definition of $H_j^{\sigma_j}\left(h_i^t\right)$ is provided in the Online Appendix when we formally define Bayes rule.

[69] Recall that, in Section 2, we define $u_i(\sigma_i, \sigma_j | h_i^t, h_j^t)$. Here, we define $u_i(\sigma_i, \sigma_j | (h_i^t, X), h_j^t)$ analogously, conditional on the event that candidate i takes X at time $-t$.

By Lemma 7, for each h_i^t and $\tilde{h}_i^t$ with $\theta\left(h_i^t\right) = \theta(\tilde{h}_i^t) = (X, X)$, we have

$$\begin{aligned}
\sup_{\sigma_i} \tilde{w}_t^i\left(\sigma_i, \sigma_j, h_i^t, X\right) &= \sup_{\sigma_i} \int_{h_j} u_i(\sigma_i, \sigma_j | \left(h_i^t, X\right), h_j^t) d\beta^{\sigma_j}\left(h_j^t\right) \\
&= \sup_{\sigma_i} \int_{h_j} u_i(\sigma_i, \sigma_j | (\tilde{h}_i^t, X), h_j^t) d\beta^{\sigma_j}\left(h_j^t\right) \\
&= \sup_{\sigma_i} \tilde{w}_t^i\left(\sigma_i, \sigma_j, h_i^t, X\right).
\end{aligned}$$

The second last line follows since the distribution of the final outcome that candidate i can induce depends only on $\beta^{\sigma_j}\left(h_j^t\right)$ and $\theta\left(h_i^t\right)$. Hence, we can write

$$\tilde{w}_t^i\left(\sigma_j, X\right) = \sup_{\sigma_i} \tilde{w}_t^i\left(\sigma_i, \sigma_j, h_i^t, X\right)$$

for each h_i^t with $\theta\left(h_i^t\right) = (X, X)$.

Similarly, let $\tilde{w}_t^i\left(\sigma_i, \sigma_j, h_i^t, no\right)$ be candidate i's payoff given that she does not receive an opportunity at time $-t$. We also have

$$\tilde{w}_t^i\left(\sigma_j, X\right) = \sup_{\sigma_i} \tilde{w}_t^i\left(\sigma_i, \sigma_j, h_i^t, X\right) = \sup_{\sigma_i} \tilde{w}_t^i\left(\sigma_i, \sigma_j, h_i^t, no\right) \tag{16}$$

since, given h_j^t, candidate j's continuation play after (h_i^t, X) and that after (h_i^t, no) are the same (candidate j's history will be the same after (h_i^t, X) and after (h_i^t, no)).

Together with the constant-sum assumption, we can show that $\tilde{w}_t^i\left(\sigma_j, X\right) + \tilde{w}_t^j\left(\sigma_i, X\right) = 1$ for any PBE σ.

Lemma 8 *Suppose $v_A(X_A, X_B) + v_B(X_B, X_A) = 1$ for each $(X_A, X_B) \in \mathcal{X}_A \times \mathcal{X}_B$. For any PBE σ, the following holds: Fix $v_i \in [0, 1]$ and $t \geq 0$. Then, the following two claims hold:*

1. *If we have $\tilde{w}_t^i\left(\sigma_j, X\right) > v_i$, then we have $\tilde{w}_t^j\left(\sigma_i, X\right) < 1 - v_i$.*

2. *If we have $\tilde{w}_t^i\left(\sigma_j, X\right) < v_i$, then we have $\tilde{w}_t^j\left(\sigma_i, X\right) > 1 - v_i$.*

Proof. By symmetry, we only prove Claim 1. The ex-ante continuation payoff for candidate i from period t given $\theta\left(h_i^t\right) = (X, X)$ is, by Bayes rule,

$$\begin{aligned}
\frac{\int_{h_i^t:\theta\left(h_i^t\right)=(X,X)} \tilde{w}_t^i\left(\sigma, h_i^t, no\right) d\beta\left(h_i^t\right)}{\int_{h_i^t:\theta\left(h_i^t\right)=(X,X)} d\beta\left(h_i^t\right)} &= \frac{\int_{h_i^t:\theta\left(h_i^t\right)=(X,X)} \tilde{w}_t^i\left(\sigma_j, X\right) d\beta\left(h_i^t\right)}{\int_{h_i^t:\theta\left(h_i^t\right)=(X,X)} d\beta\left(h_i^t\right)} \quad \text{(by the equilibrium condition)} \\
&= \tilde{w}_t^i\left(\sigma_j, X\right) > v_i.
\end{aligned}$$

Similarly, the ex-ante continuation payoff for candidate j from period t given $\theta\left(h_j^t\right) = (X, X)$ is $\tilde{w}_t^j\left(\sigma_i, X\right)$. Since the ex ante continuation payoffs should add up to one, we have $\tilde{w}_t^j\left(\sigma_i, X\right) < 1 - v_i$. ■

Given Lemma 6, together with Lemmas 7 and 8, proving the following lemma will be sufficient for conditions (7)–(9) to hold in private monitoring:

Lemma 9 *Suppose $v_A(X_A, X_B) + v_B(X_B, X_A) = 1$ for each $(X_A, X_B) \in \mathcal{X}_A \times \mathcal{X}_B$. Take $v_{i,t}(\theta)$ that satisfies conditions stated in Lemma 6. Then, for any h_i^t, we have*

$$w_t^i\left(\sigma, h_i^t, X_i\right) = v_{i,t}\left(X_i, \theta_j\left(h_i^t\right)\right);$$

$$\hat{w}_t^i\left(\sigma, h_i^t, X_j\right) = v_{i,t}\left(\theta_i\left(h_i^t\right), X_i\right);$$

and

$$w_t^i\left(\sigma, h_i^t, \theta_i\left(h^t\right)\right) = \hat{w}_t^i\left(\sigma, h_i^t, \theta_j\left(h^t\right)\right) = w_t^i\left(\sigma, h_i^t, no\right) = v_{i,t}\left(\theta\left(h^t\right)\right).$$

Proof. Once a candidate takes $x \in X$, then the other candidate takes a static best response against x whenever he receives an opportunity. In the constant-sum game, this continuation strategy uniquely pins down the equilibrium payoff once a candidate takes $x \in X$, and the payoff does not depend on whether candidates observe the opponent's arrivals of the Poisson process. Hence, we will focus on the case $\theta\left(h_i^t\right) = \theta\left(h_j^t\right) = (X, X)$. By (16) and the equilibrium condition, we can write $w_t^i\left(\sigma, h_i^t, X\right) = \hat{w}_t^i\left(\sigma, h_i^t, X\right) = w_t^i\left(\sigma, h_i^t, no\right) = \tilde{w}_t^i\left(\sigma_j, X\right)$.

Suppose that there exists a PBE $\tilde{\sigma} \in \Sigma$ such that, for some $i \in \{A, B\}$ and h_i^t, we have $\tilde{w}_t^i\left(\tilde{\sigma}_j, X\right) \neq v_{i,t}(X, X)$. Without loss,[70] we can assume

$$\tilde{w}_t^i\left(\tilde{\sigma}_j, X\right) > v_{i,t}(X, X). \tag{17}$$

[70] If $\tilde{w}_t^i\left(\tilde{\sigma}_j, X\right) < v_{i,t}(X, X)$, then since the game is constant-sum, we have $v_{j,t}(X, X) = 1 - v_{i,t}(X, X)$. From Lemma 8, we have $\tilde{w}_t^j\left(\tilde{\sigma}_i, X\right) > v_{j,t}(X, X)$. The following lemma goes through with indices i and j being reversed.

From Lemma 8, for each $\tilde{h}_j^t$ with $\theta\left(\tilde{h}_j^t\right) = (X, X)$, candidate j's expected payoff is less than $1 - v_{i,t}(X, X)$.

First, candidate i's Markov strategy is a map $\sigma_i : \mathcal{X}_i \times \mathcal{X}_j \times [0, T] \to \Delta(\mathcal{X}_i)$. Let M_i be the the space of i's Markov strategies. Note that the space for Markov strategies in public monitoring is the same as the space for Markov strategies in private monitoring. Since Markov strategies are constant with respect to the part of the histories other than the current policy sets and the current time, we write $M_i \subseteq \Sigma_i$ for each i.

Note that, in the model with public monitoring, there exists a Markov perfect equilibrium (MPE), where each candidate's strategy depends only on t, $\theta\left(h^t\right)$, and whether he or she receives an opportunity at the current time (see Gensbittel et al. (2017) for the proof). Fix a MPE $(\sigma_i, \sigma_j) \in M_i \times M_j$. We have

$$W_t^i\left(\sigma_i', \sigma_j, h^t, X\right) = \hat{W}_t^i\left(\sigma_i', \sigma_j, h^t, X\right) \leq v_{i,t}(X, X) \tag{18}$$

for each $\sigma_i' \in \Sigma_i$ since σ_i must designate a best response at every h^t in public monitoring.

Since this strategy σ_j is Markov, candidate j can take this strategy in private monitoring. We will show $w_t^i\left(\tilde{\sigma}_i, \sigma_j, \hat{h}_i^t, X\right) > v_{i,t}(X, X)$ for some $\hat{h}_i^t$ with $\theta\left(\hat{h}_i^t\right) = (X, X)$ by (17) since otherwise candidate j would like to deviate to σ_j from $\tilde{\sigma}_j$ given each h_j^t with $\theta\left(h_j^t\right) = (X, X)$ in private monitoring and obtain the expected payoff no less than $1 - v_{i,t}(X, X)$. Then, by (16), we have $\tilde{w}_t^i(\sigma_j, X) > v_{i,t}(X, X)$.

Thus, for each $\tilde{h}_i^t$ with $\theta\left(\tilde{h}_i^t\right) = (X, X)$, we have

$$\begin{aligned} \tilde{w}_t^i(\sigma_j, X) &= \sup_{\sigma_i} \tilde{w}_t^i\left(\sigma_i, \sigma_j, \tilde{h}_i^t, X\right) \\ &= \sup_{\sigma_i} \int_{h_j^t} u_i\left(\sigma_i, \sigma_j | \left(\tilde{h}_i^t, X\right), h_j^t\right) d\beta^{\sigma_j}\left(h_j^t\right). \end{aligned}$$

Since $\sigma_j \in M_j$, candidate j's continuation strategy depends only on $\theta_t = (X, X)$. Hence,[71] for each

[71] This is a standard result in dynamic programming. See Gensbittel, et al. (2017) for the application of this result to the game with Poisson arrivals. Although their paper assumes public monitoring, since σ_j is Markov, the observability of the Poisson arrivals does not affect the formulation of Bellman equations.

h_j^t, we can write

$$\sup_{\sigma_i} u_i\left(\sigma_i, \sigma_j | \left(\tilde{h}_i^t, X\right), h_j^t\right) = \sup_{\sigma_i \in M_i} u_i(\sigma_i, \sigma_j | \theta_t = (X, X)).$$

Therefore,

$$\tilde{w}_t^i\left(\sigma_j, X\right) = \sup_{\sigma_i \in M_i} u_i(\sigma_i, \sigma_j | \theta_t = X) = v_{i,t}\left(X, X\right).$$

This is a contradiction. Thus, for each PBE σ, we have $w_t^i\left(\sigma, X\right) = v_{i,t}\left(X, X\right)$. ■

Online Supplementary Appendix to: "Optimal Timing of Policy Announcements in Dynamic Election Campaigns"

Yuichiro Kamada[†] Takuo Sugaya[‡]

March 27, 2019

F Definition of Bayes Rule

Fix candidate i's history $h_i^t = \left(\left(t_i^k, X_i^k\right)_{k=1}^{k_i}, \left(t_j^l, X_j^l\right)_{l=1}^{l_j}, t, z_i \right)$ arbitrarily. If $t = T$, then candidate i believes that $h_j^t = (\emptyset, \emptyset, T, no)$ with probability one. Hence, we focus on $t < T$. Let $\left(t_i^l, X_i^l\right)_{l=1}^{l_i}$ be what candidate j can observe and is compatible with $\left(t_i^k, X_i^k\right)_{k=1}^{k_i}$: Let t_i^1 be the smallest time $t \in \{t_i^1, ..., t_i^{k_i}\}$ such that, for k with $t = t_i^k$, $X_i^k \neq X_i^0$ holds (that is, $-t_i^1$ is the first time for candidate i to change her policy set); given t_i^1, let t_i^2 be the smallest time $t \in \{t_i^1, ..., t_i^{k_i}\}$ such that $t > t_i^1$ and for k with $t = t_i^k$, $X_i^k \neq X_i^{k-1}$ holds (that is, $-t_i^2$ is the second time for candidate i to change her policy set), and so on. Fix $\left(t_i^l, X_i^l\right)_{l=1}^{l_i}$. Suppose that there exists $\left(t_j^k, X_j^k\right)_{k=1}^{k_j}$ with which $\left(t_j^l, X_j^l\right)_{l=1}^{l_j}$ is compatible, such that

$$h_j^t = \left(\left(t_i^l, X_i^l\right)_{l=1}^{l_i}, \left(t_j^k, X_j^k\right)_{k=1}^{k_j}, t, no \right)$$

happens with a positive probability by σ_j^* conditional on the realization of $\left(t_j^k\right)_{k=1}^{k_j}$, $\left(t_i^l, X_i^l\right)_{l=1}^{l_i}$, and t:[1] At each time t_j^k for $k = 1, ..., k_j$, given candidate j's history $h_j^{t_j^k} = \left(\left(t_i^l, X_i^l\right)_{l=1}^{l\left(t_j^k\right)}, \left(t_j^{k'}, X_j^{k'}\right)_{k'=1}^{k-1}, t_j^k, yes \right)$ with $l\left(t_j^k\right)$ being the largest l with $t_i^l < t_j^k$ (that is, $h_j^{t_j^k}$ is the history compatible with h_j^t), $\sigma_j^*(h_j^{t_j^k})\left(X_j^k\right) > 0$. Let $H_j^{\sigma_j^*}\left(h_i^t\right)$ be the set of candidate j's history satisfying this condition.

If $H_j^{\sigma_j^*}\left(h_i^t\right) \neq \emptyset$, then for each $h_j^t = \left(\left(t_i^l, X_i^l\right)_{l=1}^{l_i}, \left(t_j^k, X_j^k\right)_{k=1}^{k_j}, t, no \right) \in H_j^{\sigma_j^*}\left(h_i^t\right)$, we define the

[†] Haas School of Business, University of California Berkeley, 2220 Piedmont Avenue, Berkeley, CA 94720-1900, USA, e-mail: `y.cam.24@gmail.com`

[‡] Stanford Graduate School of Business, Stanford, CA, 94305, e-mail: `tsugaya@stanford.edu`

[1] Given candidate i's history, she believes that candidate j does not receive an opportunity at $-t$ when candidate i receives an opportunity. Hence, we require that the last element is *no*.

density f as follows:

$$f(h_j^t|h_i^t) = e^{-(T-t)\lambda}\frac{((T-t)\lambda)^{k_j}}{k_j!}\prod_{k=1}^{k_j}\sigma_j^*(h_j^{t_j^k})\left(X_j^k\right) \tag{19}$$

for $k_j \neq 0$ and $f(h_j^t|h_i^t) = e^{-(T-t)\lambda}$ for $k_j = 0$. Note that $e^{-(T-t)\lambda}\frac{((T-t)\lambda)^{k_j}}{k_j!}$ is the probability that candidate j receives k_j opportunities between $-T$ and $-t$. Conditional on this event, every $\left(t_j^k\right)_{k=1}^{k_j}$, has the same density.

Using the density of h_j^t defined in (19),[2] we define

$$d\beta_i\left(h_j^t|h_i^t\right) = \frac{f(h_j^t|h_i^t)}{\int_{\tilde{h}_j^t \in H_j^{\sigma_j^*}(h_i^t)} f(\tilde{h}_j^t|h_i^t)d\tilde{h}_j^t}dh_j^t. \tag{20}$$

If $H_j^{\sigma_j^*}\left(h_i^t\right) = \emptyset$ (h_i^t cannot be explained without j's deviation), then $d\beta_i\left(h_j^t|h_i^t\right)$ is arbitrary, as long as $\int_{h_j^t \in H_j(h_i^t)} d\beta_i\left(h_j^t|h_i^t\right) = 1$.

Now, given a history h_i^t (note that this determines $(t_i^l, X_i^l)_{l=1}^{l_i}$ and $\left(t_j^l, X_j^l\right)_{l=1}^{l_j}$) and a set $\hat{H}_j\left(h_i^t\right) \subseteq H_j^{\sigma_j^*}\left(h_i^t\right)$, we can classify $h_j^t \in \hat{H}_j\left(h_i^t\right)$ into the following subsets: Given h_i^t and $\hat{H}_j\left(h_i^t\right)$, let KX_j be the set of k_j and $\left(X_j^k\right)_{k=1}^{k_j}$ such that there exists $(t_j^k)_{k=1}^{k_j}$ such that $\left((t_i^l, X_i^l)_{l=1}^{l_i}, \left(t_j^k, X_j^k\right)_{k=1}^{k_j}, t, no\right) \in \hat{H}_j\left(h_i^t\right)$. Given h_i^t, $\hat{H}_j\left(h_i^t\right)$, and $(k_j, \left(X_j^k\right)_{k=1}^{k_j}) \in KX_j$, let T_j^1, $T_j^2\left(t_j^1\right)$, ..., $T_j^{k_j}(t_j^1, ..., t_j^{k_j-1})$ be, respectively, the set of t_j^1 such that there exists $(t_j^2, ..., t_j^{k_j})$ such that $\left(t_j^k, X_j^k\right)_{k=1}^{k_j}$ is compatible with $\left(t_j^l, X_j^l\right)_{l=1}^{l_j}$ and $\left((t_i^l, X_i^l)_{l=1}^{l_i}, \left(t_j^k, X_j^k\right)_{k=1}^{k_j}, t, no\right) \in \hat{H}_j\left(h_i^t\right)$; the set of t_j^2 such that, given t_j^1, there exists $(t_j^3, ..., t_j^{k_j})$ such that $\left(t_j^k, X_j^k\right)_{k=1}^{k_j}$ is compatible with $\left(t_j^l, X_j^l\right)_{l=1}^{l_j}$ and $\left((t_i^l, X_i^l)_{l=1}^{l_i}, \left(t_j^k, X_j^k\right)_{k=1}^{k_j}, t, no\right) \in \hat{H}_j\left(h_i^t\right)$; and so on, up to the set of $t_j^{k_j}$ such that, given $t_j^1, t_j^2, \ldots, t_j^{k-1}$, $\left(t_j^k, X_j^k\right)_{k=1}^{k_j}$ is compatible with $\left(t_j^l, X_j^l\right)_{l=1}^{l_j}$ and $\left((t_i^l, X_i^l)_{l=1}^{l_i}, \left(t_j^k, X_j^k\right)_{k=1}^{k_j}, t, no\right) \in \hat{H}_j\left(h_i^t\right)$.[3] Given h_i^t and $H_j^{\sigma_j^*}\left(h_i^t\right)$, define KX_j^*, $T_j^{1,*}$, $T_j^{2,*}\left(t_j^1\right)$, ..., $T_j^{k_j,*}(t_j^1, ..., t_j^{k_j-1})$ in a similar manner. Then, given h_i^t and $\hat{H}_j\left(h_i^t\right)$, for

[2] Since the Poisson process has a density, we directly define the conditional probability using the ratio of the density functions. Since the Poisson process is right-continuous, we can micro-found this definition by a measure-theoretic definition as well. See Karazas and Shreve (1988, Proposition 1.13) for the details.

[3] For simple notation, we suppress the dependence of KX_j on h_i^t and $\hat{H}_j\left(h_i^t\right)$ and the dependence of T_j^1, $T_j^2\left(t_j^1\right)$, ..., $T_j^{k_j}(t_j^1, ..., t_j^{k_j-1})$ on h_i^t, $\hat{H}_j\left(h_i^t\right)$, and $(k_j, (X_j^k)_{k=1}^{k_j})$.

any $g\left(h_j^t\right)$, we define

$$\int_{h_j^t \in \hat{H}_j\left(h_i^t\right)} g\left(h_j^t\right) d\beta_i\left(h_j^t | h_i^t\right)$$

$$:= \frac{\sum_{\left(k_j,(X_j^k)_{k=1}^{k_j}\right)\in KX_j} \int_{t_j^1\in T_j^1}\int_{t_j^2\in T_j^2(t_j^1)}\cdots\int_{t_j^{k_j}\in T_j^{k_j}(t_j^1,\ldots,t_j^{k_j-1})} \begin{pmatrix} g\left(h_j^t\right) e^{-(T-t)\lambda}\frac{((T-t)\lambda)^{k_j}}{k_j!} \\ \times\prod_{k=1}^{k_j}\sigma_j^*(h_j^{t_j^k})\left(X_j^k\right) dt_j^{k_j}\cdots dt_j^1 \end{pmatrix}}{\sum_{\left(k_j,(X_j^k)_{k=1}^{k_j}\right)\in KX_j^*} \int_{t_j^1\in T_j^{1,*}}\int_{t_j^2\in T_j^{2,*}(t_j^1)}\cdots\int_{t_j^{k_j}\in T_j^{k_j,*}(t_j^1,\ldots,t_j^{k_j-1})} \begin{pmatrix} g\left(h_j^t\right) e^{-(T-t)\lambda}\frac{((T-t)\lambda)^{k_j}}{k_j!} \\ \times\prod_{k=1}^{k_j}\sigma_j^*(h_j^{t_j^k})\left(X_j^k\right) dt_j^{k_j}\cdots dt_j^1 \end{pmatrix}}.$$

For example, given a fixed continuation strategy profile σ and h_i^t, the function $g\left(h_j^t\right)$ can be candidate i's continuation payoff $u_i\left(\sigma | h_i^t, h_j^t\right)$.

G Proofs Omitted in Appendix C

G.1 Proof of Proposition 9

Fix any $t \in (-\infty, 0]$, and suppose that no candidate enters at any $-\tau \in (-t, 0]$. On the one hand, if candidate i enters at $-t$, her payoff is $v_{i,t}(\text{enter})$. By Assumption 2, $v_{i,t}(\text{enter}) \leq v_{i,0}(\text{enter})$. Since $v_{i,0}(\text{enter}) = v_i(x_i^*, X)$ by definition and $v_i(x_i^*, X) < v_i(X, X)$ as we are in Case 1, we have $v_{i,t}(\text{enter}) < v_i(X, X)$. On the other hand, if she does not enter, then her payoff is $v_i(X, X)$. Hence, it is uniquely optimal not to enter at $-t$. Since the payoffs are continuous in time, there exists $\varepsilon > 0$ such that no candidate enters for any time in $(-t-\varepsilon, -t]$. Hence the continuous-time backward induction implies the desired result.

G.2 The Formal Definition of $\bar{v}_{i,t}(\text{not})$

Formally, $\bar{v}_{i,t}(\text{not})$ is defined by the following:

$$\begin{aligned} \bar{v}_{i,t}(\text{not}) \quad = \quad & e^{-(\lambda_i+\lambda_j)t}v_i(X,X) + e^{-\lambda_i t}(1-e^{-\lambda_j t})v_i(X,x_j^*) + (1-e^{-\lambda_i t})e^{-\lambda_j t}v_i(x_i^*,X) \\ & + \left(1-e^{-\lambda_i t}\right)\left(1-e^{-\lambda_j t}\right)\left(\frac{\lambda_i}{\lambda_i+\lambda_j}v_i^{BR_j} + \frac{\lambda_j}{\lambda_i+\lambda_j}\sup_{x_i} v_i(x_i, x_j^*)\right). \end{aligned}$$

G.3 Proof of Proposition 10

By the definition of t_0, there exists $\varepsilon > 0$ such that for all time in $(-t_0 - \varepsilon, -t_0]$, each candidate i enters under any PBE. Hence, if $t_i^* = t_j^* = \infty$, each candidate enters at all times in $(-\infty, -t_0]$. For the rest of the proof, we focus on the case in which at least one of t_A^* and t_B^* is less than ∞. Without loss, we assume $t_A^* \leq t_B^*$.

The following lemma shows that, for any PBE, candidate A does not enter at any time $-t < -t_A^*$:

Lemma 10 *Fix any σ_B such that (i) $\sigma_B(h_B^t) = x_B^*$ for any h_B^t with $\theta_B(h_B^t) = X$ for each $-t \in (-t_A^*, -t_0]$ and (ii) $\sigma_B(h_t) = BR_B(x_A)$ for any h_B^t with $\theta_A(h_B^t) = x_A$ for each $-t \in [-T, 0]$.*[4] *If σ_A is a best response to σ_B, then for any $h_t \in (X, X)$ with $-t < -t_A^*$, we have $\sigma_A(h_t)(X) = 1$.*

The proof of the lemma is complicated, so we first assume that the lemma holds and show the proposition, and then prove the lemma. If $t_A^* = t_B^*$, then Lemma 10 implies Proposition 10 with $t_i = t_i^*$ for each i. Hence, we assume $t_A^* < t_B^*$.

Fix a PBE and, for each $i = A, B$, let $v_{i,t}(\text{not})$ be candidate i's continuation payoff at time $-t$ when i does not enter. Given Lemma 10, for $t \in [t_A^*, t_B^{**}]$ with t_B^{**} defined below, we calculate $v_{i,t}(\text{not})$ assuming that only candidate B enters in the time interval $(-t, -t_A^*)$ and both candidates enter in the time interval $[-t_A^*, -t_0]$. For $\tau \geq t$, Lemma 10 implies that candidate A does not enter at times in $(-\tau, -t)$. Hence, we have $v_{B,\tau}(\text{not}) \geq v_{B,t}(\text{not})$ for $\tau \geq t$ because candidate B at $-\tau$ can receive $v_{B,t}(\text{not})$ by committing to a strategy in which he keeps skipping opportunities from $-\tau$ to $-t$. Let

$$t_B^{**} \equiv \inf \{t > t_0 : v_{B,t}(\text{not}) \geq v_{B,t}(\text{enter})\}.$$

There are the following two cases: $t_B^{**} < \infty$ or $t_B^{**} = \infty$. The following lemma is useful:

Lemma 11 *If $t_B^{**} < \infty$, then $v_2(x_B^*, X) > v_B^{BR_A}$.*

Proof. Suppose otherwise. Then, Assumption 2 implies $v_B(x_B^*, X) = v_B^{BR_A}$. Then, $v_{B,t}(\text{enter})$ is constant in $t \in [t_0, \infty)$. At time $-t_B^{**} < -t_A^*$, there are the following three cases:

1. Candidate A has the next opportunity at time $-t \in (-t_B^{**}, -t_A^*]$. Conditional on this event, candidate B obtains a payoff of $v_B^{BR_A} = v_B(x_B^*, X) = v_{B,t}(\text{enter})$ when he enters at $-t$

[4] Recall the definition of $\theta_i(\cdot)$ from Section 4.3.

and a payoff of $v_{B,t}(\text{not})$ when he does not. Since t_B^{**} is the infimum of $t > t_0$ such that $v_{B,t}(\text{not}) \geq v_{B,t}(\text{enter})$, candidate B prefers to enter in this event.

2. Candidate B has the next opportunity at time $-t \in (-t_B^{**}, -t_A^*]$. Conditional on this event, since candidate B receives $v_{B,t}(\text{enter})$ at any $-t$ upon entering, candidate B is indifferent between entering and not entering.

3. No candidate has an opportunity at any time $-\bar{t} \in (-t_B^{**}, -t_A^*]$. Conditional on this event, candidate B strictly prefers to enter since $v_{B,t_A^*}(\text{enter}) > v_{B,t_A^*}(X, X)$.

Hence, it is uniquely optimal to enter at $-t_B^{**}$, which is a contradiction. ■

Given this lemma, consider the following two cases:

1. $t_B^{**} < \infty$: In this case, we are left to prove Lemma 10. To see why, once we have shown Lemma 10, then for $t > t_B^{**}$, $v_{B,t}(\text{not}) \geq v_{B,t_B^{**}}(\text{not})$ since candidate B can skip opportunities until $-t_B^{**}$ without the opponent entering. Together with the fact that $v_{B,t}(\text{enter})$ is strictly decreasing in t by Lemma 11, we can conclude that candidate B does not enter at times in $(-\infty, -t_B^{**})$ in any PBE.

2. $t_B^{**} = \infty$: This means that candidate B enters at times in $(-\infty, 0]$ in any PBE given Lemma 10.

We now prove Lemma 10:

Proof of Lemma 10. Suppose now candidate A receives an opportunity at time $-\bar{t} < -t_A^*$ at a history in (X, X).

Fix candidate B's strategy arbitrarily. Once we fix his strategy, conditional on the event that candidate A does not enter at any time in $(-\infty, -t_A^*]$ and that candidate B has at least one opportunity in $(-\bar{t}, -t_A^*]$, we can define a random variable t that is the largest $\tau \in [t_A^*, \bar{t})$ such that candidate B enters at time $-\tau$.

From candidate A's perspective at time $-\bar{t}$, there are the following two possible events:

1. $t \leq t_A^*$ or candidate B does not have any opportunities in $(-\bar{t}, -t_A^*]$. Conditional on this event, not entering at times in $[-\bar{t}, -t_A^*)$ ensures candidate A the value of $v_{A,t_A^*}(\text{not})$. Since $v_{A,t_A^*}(\text{not}) = v_{A,t_A^*}(\text{enter})$ and $v_{A,t}(\text{enter})$ is weakly decreasing in t by Assumption 2, it is weakly better for candidate A not to enter at time $-\bar{t}$.

2. $t > t_A^*$. Conditional on this event, at time $-\bar{t}$, candidate A's continuation payoff from entering is weakly less than her continuation payoff from not entering if and only if

$$v_A^{BR_B} \leq (1 - e^{-\lambda_A t}) \left(\max_{X_A \in \mathcal{X}_A} v_A(X_A, x_B^*) \right) + e^{-\lambda_A t} v_A(X, x_B^*) =: \hat{v}_{A,t}.$$

To see why, note first that the left-hand side is the payoff from entering at time $-\bar{t}$, while the right-hand side is the payoff at time $-t$ when the current policy profile is (X, x_B^*). The payoff from not entering at time $-\bar{t}$ is a convex combination of the following two payoffs, where the weight on the latter payoff is strictly positive.

- The payoff under the event that candidate A receives at least one opportunity at which she enters in the time interval $(-\bar{t}, -t)$.
- The payoff under the event that candidate A does not receive any opportunity at which she enters in the time interval $(-\bar{t}, -t)$.

Note that the former payoff is equal to the left-hand side of the expression ($v_A^{BR_B}$), while the latter payoff is the same as the right-hand side of the expression. This implies the desired equivalence.

To compare the two values, it is instructive to examine why candidate A at $-t_A^*$ is indifferent between entering and not entering at histories in (X, X). Suppose now that candidate B has not entered at $-t_A^*$. There are following three events that can happen with positive probability until the deadline:

(a) Candidate A receives the next opportunity at $-\tau > -t_A^*$: In this case, candidate A receives $v_{A,t}(\text{enter})$ at $-\tau$ regardless of candidate A's choice at $-t_A^*$. Note that, even if candidate A has entered before $-\tau$, since we assume that candidate A enters at some policy in X_A^*, the situation is that candidate A enters at some policy in X_A^* and candidate B has not at $-\tau$ (note that all the policies in X_A^* give rise to the same payoff).

(b) Candidate B receives the next opportunity at $-\tau > -t_A^*$: Candidate A receives a payoff of $v_A^{BR_B}$ (candidate B best-responds to x_A^* at $-\tau$) if she enters before $-t_A^*$; and $\hat{v}_{A,\tau}$ (candidate B enters while candidate A has not at $-\tau$) if she does not enter before $-t_A^*$.

(c) No candidate receives any opportunity in the time interval $(-t_A^*, -t_0]$: Candidate A receives $v_{A,t_0}(\text{enter})$ if she enters at $-t_1^*$; and $v_{A,t_0}(\text{not})$ if she does not at $-t_A^*$. We have assumed that $v_{A,t_0}(\text{enter}) > v_{A,t_0}(\text{not})$.

Note that candidate A is indifferent between entering and not entering at $-t_A^*$ in case (2a) and strictly prefers entering in case (2c). Since case (2c) happens with positive probability, it must be the case that candidate A strictly prefers not entering to entering in case (2b), in order for her to be indifferent between entering and not entering at $-t_A^*$. This implies that there exists $-\tilde{t} \in (-t_A^*, t_0]$ such that $v_A^{BR_B} < \hat{v}_{A,\tilde{t}}$. Since candidate A can always skip opportunities between $-t$ and $-\tilde{t}$, we have $\hat{v}_{A,\tilde{t}} \leq \hat{v}_{A,t}$, implying $v_A^{BR_B} < \hat{v}_{A,t}$.

Hence, conditional on the event that candidate B has an opportunity and enters at $-t$, candidate A at $-\bar{t}$ strictly prefers not entering to entering.[5]

Now we prove that candidate A does not enter at any history in (X, X) at any time before $-t_A^*$ in any PBE. There are the following two cases:

1. If "$v_i^{BR_j} < \sup_{\{x_i\}\in\mathcal{X}_i} v_i(x_i, X)$ for each i" holds in Assumption 3, then $v_{A,t}(\text{enter})$ is strictly decreasing in t. Hence, conditional on the event that $t \leq t_A^*$ or candidate B does not have any opportunities in $(-\bar{t}, -t_A^*]$, candidate A at $-\bar{t}$ strictly prefers not entering to entering. Since it happens with a positive probability that candidate B does not receive any opportunity in $(-\bar{t}, -t_A^*]$, candidate A does not enter before $-t_A^*$ in any PBE.

2. If $t_A^* \neq t_B^*$ holds in Assumption 3, then $t_A^* < t_B^*$. By continuity of the continuation payoff in time, there exists $\varepsilon > 0$ such that candidate B enters for each $-t \in (-t_A^* - \varepsilon, -t_A^*]$. Hence, the event that candidate B has an opportunity and enters at some time in this time interval happens with a positive probability. Therefore, candidate A does not enter before $-t_A^*$ in any PBE.

■

[5] Here we are using Assumption 1 which implies that $v_A^{BR_B}$ must be independent of the time at which candidate A chooses x_A^*.

G.4 Proof of Part 1 of Proposition 12

By continuity of the continuation payoff in time, for times $-t < -\hat{t}_A$ sufficiently close to $-\hat{t}_A$, candidate B enters, and thus we focus on candidate A's incentive at those times. Let

$$\hat{v}_{A,t} := (1 - e^{-\lambda_A t}) \left(\max_{X_A \in \mathcal{X}_A} v_A(X_A, x_B^*) \right) + e^{-\lambda_A t} v_A(X, x_B^*)$$

be candidate A's payoff when she has not entered and candidate B has at time $-t$. The straightforward algebra shows that $\bar{v}_{A,t}^A(\text{not})$ satisfies

$$\begin{aligned} \bar{v}_{A,t}^A(\text{not}) &= \int_0^t \lambda_B e^{-\lambda_B \tau} \hat{v}_{A,t-\tau} d\tau \\ &= \left(e^{-\lambda_A t} - e^{-(\lambda_1+\lambda_2)t}\right) v_A\left(X, x_B^*\right) + \left(1 - e^{-2\lambda_B t} - 2e^{-\lambda_B t}\right) \max_{x_A} v_A\left(x_A, x_B^*\right). \end{aligned}$$

In contrast, we have

$$v_{A,t}(\text{enter}) = e^{-\lambda_B t} v_A\left(x_A^*, X\right) + \left(1 - e^{-\lambda_B t}\right) v_A^{BR_B}.$$

Hence, $v_{A,t}(\text{enter})$ and $\bar{v}_{A,t}^A(\text{not})$ are differentiable in t. Since $\hat{t}_A$ is the infimum of t with $\bar{v}_{A,t}^A(\text{not}) \leq v_{A,t}(\text{enter})$, we have

$$\left.\frac{d}{dt}\bar{v}_{A,t}^A(\text{not})\right|_{t=\hat{t}_A} < \left.\frac{d}{dt}v_{A,t}(\text{enter})\right|_{t=\hat{t}_A}.$$

Consider candidate A's incentive at time $-\hat{t}_A$. For any $\varepsilon > 0$, there are the following three cases (assuming that candidate B enters as soon as she obtains an opportunity):

1. Candidate A has the next opportunity at time $-\bar{t} \in (-t, -(t-\varepsilon)]$. Conditional on this event, since we fix candidate B's strategy at histories in (X, X), candidate A is indifferent between entering and not entering.

2. Candidate B has the next opportunity at time $-\bar{t} \in (-t, -(t-\varepsilon)]$. Conditional on this event, candidate A obtains a payoff of $v_A^{BR_B}$ when she enters at $-t$ and a payoff of $\hat{v}_{A,\bar{t}}$ when she does not.

3. No candidate has an opportunity at any time $-\bar{t} \in (-t, -(t-\varepsilon)]$. Conditional on this event, candidate A obtains a payoff of $v_{A,t-\varepsilon}(\text{enter})$ when she enters at $-t$ and a payoff of $\bar{v}_{A,t-\varepsilon}^1(\text{not})$

when she does not.

Since candidate A is indifferent between entering and not entering at time $-\hat{t}_A$, for any $\varepsilon > 0$, we have

$$\begin{aligned}
&\int_{\tau=0}^{\varepsilon} \underbrace{\lambda_B e^{-(\lambda_A+\lambda_B)\tau}}_{\text{Candidate } B \text{ has the next opportunity at time } -t+\tau} \left(v_A^{BR_B} - \hat{v}_{A,t-\tau}\right) d\tau \\
= \;& e^{-(\lambda_A+\lambda_B)\varepsilon} \left(\bar{v}^1_{A,t-\varepsilon}(\text{not}) - v_{A,t-\varepsilon}(\text{enter})\right).
\end{aligned}$$

Dividing both sides by ε and taking the limit as $\varepsilon \downarrow 0$, we have

$$v_A^{BR_B} - \hat{v}_{A,\hat{t}_A} = \left(\frac{d}{dt}\bar{v}^1_{A,t}(\text{not})\Big|_{t=\hat{t}_A} - \frac{d}{dt}v_{A,t}(\text{enter})\Big|_{t=\hat{t}_A} \right) < 0.$$

By continuity of the continuation payoff in time, there exists $\bar{\eta} > 0$ such that, for each $\eta \in [0, \bar{\eta})$, we have

$$v_A^{BR_B} - \hat{v}_{A,\hat{t}_A+\eta} < 0. \tag{21}$$

We now consider candidate A's incentive at $-t < -\hat{t}_A$. There are again following three cases:

1. Candidate A has the next opportunity at time $-\bar{t} \in (-t, -\hat{t}_A]$. Conditional on this event, since we fix candidate B's strategy at histories in (X, X), candidate A is indifferent between entering and not entering.

2. Candidate B has the next opportunity at time $-\bar{t} \in (-t, -\hat{t}_A]$. Conditional on this event, candidate A obtains a payoff of $v_A^{BR_B}$ when she enters at $-t$ and a payoff of $\hat{v}_{A,\bar{t}}$ when she does not.

3. No candidate has an opportunity at any time $-\bar{t} \in (-t, -\hat{t}_A]$. Conditional on this event, candidate A is indifferent between entering and not entering.

Hence, (21) implies that, for $-t \in (-\hat{t}_A - \bar{\eta}, -\hat{t}_A)$, candidate A strictly prefers to enter in any PBE, as desired.

G.5 Proof of Part 2 of Proposition 12

Since candidate A does not enter for each $-t \in \left[-\hat{t}_B, 0\right]$, the fact that candidate B strictly prefers to enter at time 0 and becomes indifferent between entering and not entering at $-\hat{t}_B$ implies that he is strictly worse off if candidate A enters than if she does not enter, after candidate B enters:

Lemma 12 *$\hat{t}_B \leq \hat{t}_A$ implies $v_B\left(x_B^*, X\right) > v_B^{BR_A}$.*

Proof. Suppose otherwise. Then, by Assumption 2, we have $v_B\left(x_B^*, X\right) = v_B^{BR_A}$ and so we have

$$v_{B,t}(\text{enter}) = v_B\left(x_B^*, X\right).$$

At time $-\hat{t}_B$, consider the following three cases:

1. Candidate A has the next opportunity at time $-t \in (-\hat{t}_B, 0]$. Conditional on this event, candidate B obtains a payoff of $v_B^{BR_A} = v_B\left(x_B^*, X\right)$ when he enters at $-t$ and a payoff of $\bar{v}_{B,t}^1(\text{not})$ when he does not.

2. Candidate B has the next opportunity at time $-t \in (-\hat{t}_B, 0]$. Conditional on this event, since we fix candidate A's strategy at histories in (X, X), candidate B is indifferent between entering and not entering.

3. No candidate has an opportunity at any time $-\bar{t} \in (-\hat{t}_B, 0]$. Conditional on this event, candidate B strictly prefers to enter since $v_B\left(x_B^*, X\right) > v_B\left(X, X\right)$.

Hence, candidate B strictly prefers to enter at $-\hat{t}_B$, which is a contradiction. ■

Given this lemma, we are left to show that, at each $-t \in (-\infty, -\hat{t}_B]$, given that no candidate enters for $-\tau \in (-t, -\hat{t}_B)$, each candidate strictly prefers not to enter at $-t < -t_B$.

On the one hand, if candidate i enters at $-t$, her payoff is $v_{i,t}(\text{enter})$. By Lemma 12, $v_{i,t}(\text{enter}) < v_{i,\hat{t}_B}(\text{enter})$. On the other hand, if she does not enter, then her payoff is $\bar{v}_{i,\hat{t}_B}^1(\text{not})$. Hence, it is indeed uniquely optimal not to enter at $-t$.

H Example for Playing the Weakly Dominated Action

In Section D, we claimed that the conclusion of part 3 of Theorem 4 does not hold if we replace strictly dominant policy with weakly dominant policy. This appendix provides an example to illustrate this. Let $X = \{0, 1\}$ and define (v_A, v_B) by the payoff matrix as in Table 3.

	0	1
0	$0, 1$	$2, 1$
1	$1, 2$	$3, 0$

Table 3: Payoff matrix for an example with multiple PBE with weakly dominant policies

In addition, for each $i = A, B$, we define $v_i(X, a_j) = \sum_{a_i \in X} \frac{1}{2} v_i(a_i, a_j)$ for each $a_j \in X$; $v_i(a_i, X) = \sum_{a_j \in X} \frac{1}{2} v_i(a_i, a_j)$ for each $a_i \in X$; and $v_i(X, X) = \sum_{(a_i, a_j) \in X \times X} \frac{1}{4} v_i(a_i, a_j)$.

Notice that $(1, 0)$ is the weakly (but not strictly) dominant policy profile, meaning that the defining inequality for a strictly dominant policy is required to hold only weakly for all $X_i \in \mathcal{X}_i \backslash \{x_i^*\}$ except at least one $X_i \in \mathcal{X}_i \setminus \{x_i^*\}$ for which the inequality needs to hold strictly. However, if candidate B announces $\{1\}$ after candidate A announces $\{0\}$ — he rewards her by taking the weakly dominated strategy, then it is possible in a PBE that candidate A announces $\{0\}$ for some time interval.

Specifically, suppose $\lambda_A = \lambda_B$, let $t^* = -\frac{1}{\lambda} \ln 2$, and consider the strategy profile as follows: Candidate A chooses $\{1\}$ except when B's current policy set is X and the time is in the interval $[-T, -t^*)$, at which she takes $\{0\}$. Player B chooses $\{0\}$ except when candidate A has already chosen $\{0\}$, in which case he takes $\{1\}$. We show that this strategy profile is a PBE.

It is straightforward to check that candidate B is taking a best response. We check candidate A's incentive. Suppose first that $-t \in [-t^*, 0]$. The expected payoff from taking $\{1\}$ is

$$e^{-\lambda t} 2 + (1 - e^{-\lambda t}). \tag{22}$$

The expected payoff from taking $\{0\}$ is $e^{-\lambda t} 1 + (1 - e^{-\lambda t}) 2$, and this is no more than (22) if $-t \in [-t^*, 0]$. Next, the expected payoff from taking $\{0, 1\}$ is

$$\begin{aligned}
&\int_0^t e^{-2\lambda(t-s)} 2\lambda \frac{\left(e^{-\lambda s} + (1 - e^{-\lambda s})2\right) + \left(e^{-\lambda s} 0.5 + (1 - e^{-\lambda s})\right)}{2} ds + e^{-2\lambda t} 1.5 \\
&= 3 - 3e^{-\lambda t} + 1.5 e^{-2\lambda t}.
\end{aligned}$$

Given (22), the payoff from $\{1\}$ is larger than the payoff from $\{0,1\}$ if and only if

$$e^{-\lambda t}2+(1-e^{-\lambda t})>3-3e^{-\lambda t}+1.5e^{-2\lambda t} \iff 4e^{-\lambda t}>2+1.5e^{-2\lambda t}.$$

This holds if and only if $-t\in(-t^*,0]$.

Second, suppose that $-t\in(-\infty,-t^*)$. The expected payoff from taking $\{0\}$ and the one from taking $\{1\}$ have the same expressions as before, and the former is strictly greater than the latter if $-t\in(-\infty,-t^*)$. The expected payoff from taking $\{0,1\}$ is at most a strict convex combination of (i) the expected payoff at time $-t^*$ from the continuation strategy profile that coincides with the specified strategy profile, (ii) the expected payoff from taking $\{0\}$ at time $-t$, and (iii) the expected payoff from the opponent taking $\{0\}$ at time $-t$. Since we have shown that (i) is less than (ii) for any $-t=-t^*$ and (ii) is increasing in t, (i) is less than (ii) for any $-t\in(-\infty,-t^*)$. Hence, it suffices to show that (iii) is no more than (ii), which is equivalent to

$$e^{-\lambda t}+(1-e^{-\lambda t})2>e^{-\lambda t}0.5+(1-e^{-\lambda t}),$$

and this holds for any $t\geq 0$.

Overall, we have shown that candidate A is taking a best response conditional on any history.

I A Proof and Additional Discussions for Section 3.1

This section provides discussions of the valence election campaign model. First, Section I.1 provides a proof of Proposition 2. Next, Section I.2 derives empirical implications of our model. Although we see these findings as only suggestive, they are consistent with the empirical findings such as those presented in Campbell (1983). Then, Section I.3 conducts a welfare analysis, comparing our model with that of Aragonès and Palfrey (2002).

The dynamic model we have analyzed in Section 3.1 was kept as simple as possible to highlight the complexity added by the fact that candidates face dynamic incentive problems in the presence of valence. In Appendix I.4, we extend and modify this model to examine robustness of our prediction that candidates use ambiguous language at the early stages of the campaign.

Appendix I.5 considers a model with synchronous opportunities for policy announcements, and

Appendix I.6 considers the case with partial commitment.

I.1 Proof of Proposition 2

Note that Assumptions 1 and 2 in Section 4.1 are satisfied given $X_i^* = \{1\}$. Moreover, we have $v_i^{BR_j} < \sup_{x_i \in \mathcal{X}_i} v_i(x_i, X)$ (Assumption 3) and first-mover disadvantage is satisfied for $i = W$.

Fix a PBE σ arbitrarily. Given candidate i's history $h_i^t = \left(\left(t_i^k, X_i^k\right)_{k=1}^{k_i}, \left(t_j^l, X_j^l\right)_{l=1}^{l_j}, t, z_i \right)$ at $-t$, let $w_t^i\left(\sigma, h_i^t\right)$ be candidate i's continuation payoff at time $-t$ given σ and h_i^t. In addition, let $\theta(h_i^t) = (X_i^{k_i}, X_j^{l_j})$ be the profile of policy sets that are chosen most recently, where we always write S's current policy set first in this proof. Since the most recently chosen policy sets are observable, we have $\theta(h_S^t) = \theta(h_W^t)$. For simple notation, we write $\theta(h_S^t) = \theta(h_W^t) = \theta\left(h^t\right)$. By Theorem 3, there exists $v_{i,t}\left(\theta\left(h^t\right)\right)$ such that $w_t^i\left(\sigma, h_i^t\right) = v_{i,t}\left(\theta\left(h^t\right)\right)$ in any PBE σ.

From Lemma 1, the following statements are true:

- If $\theta(h^t) = (\{x\}, \{0,1\})$ with $x \in \{0,1\}$ and if W can move, then W is indifferent between entering at $x' \in \{0,1\}$ with $x' \neq x$ and announcing $\{0,1\}$. S wins if and only if the median voter is located at x.
- If $\theta(h^t) = (\{0,1\}, \{x\})$ with $x \in \{0,1\}$ and if S can move, then S enters at x and wins.

Hence, we have

$$\begin{aligned} v_{S,t}\left(\theta\left(h^t\right)\right) &= 1 - (1-p)\, e^{-\lambda t} \text{ if } \theta(h^t) = (\{0,1\}, \{1\}) \\ v_{W,t}\left(\theta\left(h^t\right)\right) &= (1-p)\, e^{-\lambda t} \text{ if } \theta(h^t) = (\{0,1\}, \{1\}). \end{aligned}$$

Similarly, we have

$$\begin{aligned} v_{S,t}\left(\theta\left(h^t\right)\right) &= 1 - p \text{ if } \theta\left(h^t\right) = (\{1\}, \{0,1\}) \\ v_{W,t}\left(\theta\left(h^t\right)\right) &= p \text{ if } \theta\left(h^t\right) = (\{1\}, \{0,1\}). \end{aligned}$$

When $-t$ is sufficiently close to the deadline 0, then at any h^t with $\theta(h^t) = (\{0,1\}, \{0,1\})$, the following are true:

- If W can move, then W enters at 1. Note that, since $-t$ is sufficiently close to zero, with a probability close to 1, there is no more opportunity to announce a policy. Hence, $\{1\}$ gives W the payoff close to $1-p$, $\{0\}$ gives W the payoff close to p, and $\{0,1\}$ gives W the payoff close to zero. S wins if and only if the median voter is located at 0.

- If S can move, then S does not enter. Note that, since $-t$ is sufficiently close to zero, with a probability close to 1, there is no more opportunity to announce a policy. Hence, $\{1\}$ gives S the payoff close to $1-p$, $\{0\}$ gives S the payoff close to p, and $\{0,1\}$ gives S the payoff close to 1.

Hence, we are in Case 3 for Theorem 1 (with candidate A being S), and using the notation of Section 4.1, we have

$$\begin{aligned} \bar{v}^S_{S,t}(\text{not}) &= 1-(1-p)\,\lambda t e^{-\lambda t}; \\ v_{S,t}(\text{enter}) &= 1-p; \end{aligned}$$

and

$$\begin{aligned} \bar{v}^S_{W,t}(\text{not}) &= (1-p)\,\lambda t e^{-\lambda t}; \\ v_{W,t}(\text{enter}) &= (1-p)\,\lambda t e^{-\lambda t}. \end{aligned}$$

Hence, $\hat{t}_S$ and $\hat{t}_W$, whose notation is introduced in Section 4.1, are characterized, respectively, by

$$1-(1-p)\,\lambda \hat{t}_S e^{-\lambda \hat{t}_S} = 1-p \Leftrightarrow 1 > \frac{p}{1-p} = \lambda \hat{t}_S e^{-\lambda \hat{t}_S} \tag{23}$$

and

$$(1-p)\,\lambda t e^{-\lambda \hat{t}_W} = (1-p)\,e^{-\lambda \hat{t}_W} \Leftrightarrow \hat{t}_W = \frac{1}{\lambda}. \tag{24}$$

To fully characterize the candidates' strategies, we examine the following three possible cases.

Case (1): $\frac{p}{1-p} > e^{-1}$. In this case, we have $\frac{1}{\lambda} = \hat{t}_W < \hat{t}_S$. Hence, Proposition 12 ensures that both S and W announce $\{0,1\}$ for all time in $(-t,-t^*)$ with $t^* := \hat{t}_W$. By Lemma 5, we have shown the claims.

Case (2): $\frac{p}{1-p} < e^{-1}$**.** In this case, we have $\frac{1}{\lambda} = \hat{t}_W > \hat{t}_S$. Moreover, by the implicit function theorem, we have

$$\frac{d\hat{t}_S}{dp} = -\frac{\frac{d\lambda\hat{t}_S e^{-\lambda\hat{t}_S}}{d\hat{t}_S}}{\frac{d\left(\frac{p}{1-p}\right)}{dp}} = -(1-p)^2 \lambda e^{-\lambda\hat{t}_S}\left(1-\lambda\hat{t}_S\right) < 0. \tag{25}$$

Recall that the definition of $-\hat{t}_S$ implies that, at time $-\hat{t}_S$, S becomes indifferent between entering at 1 and announcing $\{0,1\}$ given the continuation play in which S does not enter and W enters at times in $(-\hat{t}_S, 0]$. The definition implies that this indifference holds in any PBE. By part 1 of Proposition 12, there exists $\bar{\varepsilon} > 0$ such that both S and W strictly prefer entering at 1 for each $-t \in [-\hat{t}_S - \bar{\varepsilon}, \hat{t}_S)$. Therefore, we are in Case 2 for Theorem 1 with $t_0 = -\hat{t}_S - \bar{\varepsilon}$.

We will show that candidate S always enters at 1 for $-t < -\hat{t}_S$. Suppose S always enters at 1 for all time in $(-t, -\hat{t}_S)$. If S announces $\{0,1\}$ at $-t$, there are following three subcases to consider.

1. If W can move next by $-\hat{t}_S$, then one strategy that W can take is to announce $\{0,1\}$. The following two cases are possible: If S enters at $\{1\}$ by $-\hat{t}_S$, W gets p. If S does not enter by $-\hat{t}_S$, by the definition of $-\hat{t}_S$ (that is, S is indifferent between $\{1\}$ and $\{0,1\}$ at $-\hat{t}_S$), S gets $1-p$ and W gets p. In both cases, W gets at least p. Furthermore, if W can get the first revision opportunity sufficiently close to $-\hat{t}_S$, W gets strictly more than p since W strictly prefers entering at 1 to announcing $\{0,1\}$. Overall, W gets strictly more than p, which means S gets strictly less than $1-p$.

2. If S can move next by $-\hat{t}_S$, S enters and gets $1-p$.

3. If no candidate can move by $-\hat{t}_S$, then by definition, S gets $1-p$.

Therefore, the payoff from announcing $\{0,1\}$ is strictly less than $1-p$. This implies that it is uniquely optimal for S to enter at 1, as desired. Hence, $t_S = \infty$ in Proposition 10.

We will now examine candidate W's incentives. Since first-mover disadvantage for W holds, there exists

$$t_W^* > \hat{t}_S \tag{26}$$

such that it is uniquely optimal for W not to enter at $-t < -t_W^*$ and uniquely optimal for W to enter at $-t \in (-t_W^*, 0]$.[6]

[6]This notation of t_W^* is introduced in Section 4.1.

Moreover, $t_W^* \equiv \inf \{t > t_0 : \bar{v}_{W,t}(\text{not}) \geq v_{W,t}(\text{enter})\}$ implies

$$(1-p)\, e^{-\lambda t} = \int_0^{t-\hat{t}_S} e^{-2\lambda\tau}\lambda\,(1-p)\, e^{-\lambda(t-\tau)}d\tau + p\left(1 - \int_0^{t-\hat{t}_S} \lambda e^{-2\lambda\tau}d\tau\right)$$

$\Leftrightarrow$

$$e^{-\lambda\left(2t_W^* - \hat{t}_S\right)} = \frac{p}{1-p}\frac{1}{2}\left(1 + e^{-2\lambda\left(t_W^* - \hat{t}_S\right)}\right).$$

Since $\frac{p}{1-p} = \lambda\hat{t}_S e^{-\lambda\hat{t}_S}$ by the definition of $\hat{t}_S$, this inequality is equivalent to

$$e^{-\lambda\left(2t_W^* - \hat{t}_S\right)} = \lambda\hat{t}_S e^{-\lambda\hat{t}_S}\frac{1}{2}\left(1 + e^{-2\lambda\left(t_W^* - \hat{t}_S\right)}\right) \Leftrightarrow e^{-2\lambda t_W^*} = \frac{\frac{1}{2}\lambda\hat{t}_S}{1 - \frac{1}{2}\lambda\hat{t}_S}e^{-2\lambda\hat{t}_S}.$$

Taking the log of both sides and rearranging, we obtain

$$t_W^* = \hat{t}_S - \frac{1}{2\lambda}\log\left(\frac{\frac{1}{2}\lambda\hat{t}_S}{1 - \frac{1}{2}\lambda\hat{t}_S}\right).$$

Hence, we have

$$\frac{dt_W}{dp} = \frac{dt_W^*}{d\hat{t}_S}\frac{d\hat{t}_S}{dp} = \left(1 - \frac{1}{\lambda\hat{t}_S\left(2 - \lambda\hat{t}_S\right)}\right)\frac{d\hat{t}_S}{dp}.$$

Recalling that $\lambda\hat{t}_S \in (0,1)$, we have

$$\sqrt{\lambda\hat{t}_S\left(2 - \lambda\hat{t}_S\right)} < \frac{1}{2}\left(\lambda\hat{t}_S + \left(2 - \lambda\hat{t}_S\right)\right) = 1,$$

and so

$$\frac{1}{\lambda\hat{t}_S\left(2 - \lambda\hat{t}_S\right)} > 1.$$

Therefore, together with (25), we have

$$\operatorname{sign}\frac{dt_W}{dp} = \operatorname{sign}\left(1 - \frac{1}{\lambda\hat{t}_S\left(2 - \lambda\hat{t}_S\right)}\right)\operatorname{sign}\frac{d\hat{t}_S}{dp} = 1. \tag{27}$$

The inequalities (25), (26), and (27) prove part 2(c) of Proposition 2.

Case (3): $\frac{p}{1-p} = e^{-1}$. At time $-t^* = \frac{1}{\lambda}$, for each h^{t^*} with $\theta(h^{t^*}) = (\{0,1\},\{0,1\})$, S is indifferent between "announcing $\{1\}$ and thereby ensuring $1-p$," and "announcing $\{0,1\}$." At the same time,

W is indifferent between announcing $\{1\}$ and $\{0,1\}$.

For $-t < -t^*$, on the one hand, when W can move, his payoff from not entering is at least p since he gets p if S enters at 1 by $-t^*$. If S does not enter by $-t^*$, by the definition of $-t^*$, S gets $1-p$ and W gets p. On the other hand, entering at 1 gives W a payoff of $1-p$ times the probability of S not having any future revision opportunity, which is equal to $(1-p)e^{-\lambda t} < (1-p)e^{-\lambda t^*} = p$. Therefore, W strictly prefers not entering.

Given this, S is always indifferent between "announcing $\{1\}$ and thereby ensuring $1-p$," and "announcing $\{0,1\}$."

I.2 Empirical Implications

In this section, we derive empirical implications of the results from the model of valence election campaign. We see these implications as only suggestive, but as will be seen in Appendix I.4, it is possible to enrich the model by incorporating various features (such as heterogenous arrival rates and general utilities from the outcomes). This suggests that, if one wants to conduct empirical research, then it will be possible to extend the model to incorporate more characteristics and to derive testable implications from such a general model, as we do here for the base model.

First, we show that ambiguity is likely when the probability distribution of the median voter's position is close to uniform, that is, when p is close to $\frac{1}{2}$. Specifically, fix a horizon length $T \in (\frac{1}{\lambda}, \infty)$. Let p^W be the p such that $t_W = T$.[7] By definition, $p^W < \frac{1}{1+e}$. Proposition 2 implies the following:

1. For $p \in (0, \frac{1}{2}) \setminus \{\frac{1}{1+e}\}$, the probabilities of W and S announcing the ambiguous policy are both nondecreasing in p.

2. For $p \in (0, p^W)$, the probability of W announcing the ambiguous policy is constant in p, and that of S announcing the ambiguous policy is strictly increasing in p.

3. For $p \in (p^W, \frac{1}{1+e})$, the probabilities of W and S announcing the ambiguous policy are both strictly increasing in p.

4. For $p \in (\frac{1}{1+e}, \frac{1}{2})$, the probabilities of W and S announcing the ambiguous policy are constant in p.

[7] Such p^W exists and is unique due to Proposition 2 2(c) and $t^* = \frac{1}{\lambda}$.

Hence, roughly, as the position of the median voter becomes more unpredictable, the probability of ambiguous policy announcement at the election date increases. This is consistent with Campbell (1983) who suggests that opinion dispersion has a strong positive effect on the ambiguity in candidates' language.[8]

Next, suppose that there are two candidates A and B, and outside researchers know $p > \frac{1}{1+e}$ but do not know which candidate is strong and which candidate is weak. They have a prior that assigns a positive probability to both candidate A's being strong and candidate B's being strong. If the researchers can observe the campaign phase, the first entrant can be inferred to be weak (and if there is no entrance, then the posterior about valence is the same as the prior). In contrast, if they cannot observe the campaign phase but only the final policy choices by the candidates, then if only one candidate enters, such a candidate can be inferred to be weak. Otherwise, the posterior about valence is the same as the prior.

I.3 Welfare Comparison with the Static Model

As mentioned in Remark 3, conducting a welfare analysis necessitates us to impose some specific assumption about the voter distribution. Here, we assume that there is a single voter. It is then necessary that this voter's ideal policy is 0 with probability p and 1 with probability $1-p$. we focus on the case in which $p > \frac{1}{1+e}$. Normalize the voter's payoff so that $u(0) - u(1) = 1$. With this normalization, if a candidate $i \in \{S, W\}$ with the ideal policy $y \in \{0, 1\}$ wins and implements a policy $x \in \{0, 1\}$, the voter's payoff can be written as $\mathbb{I}_{x=y} + \delta \cdot \mathbb{I}_{i=S}$.

Aragonès and Palfrey (2002) consider the one-shot game where each of candidates S and W simultaneously chooses a policy. Here we consider a version of their model adopted to our environment in which the policy space is $\{0, 1\}$. That is, each candidate chooses either 0 or 1, and there is no choice of $\{0, 1\}$.

Since their expected payoffs are represented by the following payoff matrix, the unique mixed-strategy Nash equilibrium is that S takes 0 and 1 with probabilities p and $1-p$, respectively; and W takes 0 and 1 with probabilities $1-p$ and p, respectively:

[8] As discussed in footnote 33 of the main text, we have in mind a situation where n voters are independently distributed over $\{0, 1\}$ where the probability on the policy 0 is $q < \frac{1}{2}$. A higher q suggests more option dispersion (a higher standard deviation of the preferred policies among the voters. Campbell (1983) also considers standard deviation), and corresponds to a higher p.

S's payoff \ W's payoff	0	1
0	$1, 0$	$p, 1-p$
1	$1-p, p$	$1, 0$

Given this equilibrium strategy, the expected welfare of the voter is

$$\begin{aligned} & \underbrace{p}_{y=0}\left(\underbrace{p}_{x_S=0}(1+\delta)+\underbrace{(1-p)^2}_{x_S=1 \text{ and } x_W=0}+\underbrace{p(1-p)}_{x_S=1 \text{ and } x_W=1}\delta\right) \\ & +\underbrace{(1-p)}_{y=1}\left(\underbrace{(1-p)}_{x_S=1}(1+\delta)+\underbrace{p^2}_{x_S=0 \text{ and } x_W=1}+\underbrace{(1-p)p}_{x_S=0 \text{ and } x_W=0}\delta\right) \\ = \; & (1+\delta)\left(1-p+p^2\right), \end{aligned}$$

where x_i for $i = S, W$ denotes the realized policy choice by candidate i. This expected payoff converges to $W(p) := 1-p+p^2$ in the limit as δ goes to 0.

Next, consider our model of valence election campaign. Since $p > \frac{1}{1+e}$, given Proposition 2, in any PBE, W does not enter for each $-t < -t_W = -\frac{1}{\lambda}$, and enters at $x = 1$ for each $-t > -\frac{1}{\lambda}$, while S never enters unless W enters. Hence, (i) with probability $e^{-\lambda \cdot \frac{1}{\lambda}} = e^{-1}$, no candidate enters; (ii) with probability $\int_0^{\frac{1}{\lambda}} \lambda e^{-\lambda s} e^{-\lambda\left(\frac{1}{\lambda}-s\right)} ds = e^{-1}$, W enters at policy 1 but S does not enter; and (iii) with probability $1-2e^{-1}$, both candidate enter at policy 1. In the respective cases, (i) if no candidate enters, then the voter's expected payoff is $\frac{1}{2}+\delta$ (recall that we assume that a candidate without specifying her policy takes each policy with probability $\frac{1}{2}$); (ii) if W enters at 1 while S does not enter, then the expected payoff is $1-p+p(1+\delta)$; and (iii) if both candidates enter at 1, then the expected payoff is $1-p+\delta$. In total, the expected payoff is

$$\begin{aligned} & e^{-1}\left(\frac{1}{2}+\delta\right)+e^{-1}(1-p+p(1+\delta))+\left(1-2e^{-1}\right)(1-p+\delta) \\ & =1-\frac{1}{2}e^{-1}-p+2pe^{-1}+\delta-(1-p)\delta e^{-1}. \end{aligned}$$

This expected payoff converges to $V(p) := 1-\frac{1}{2}e^{-1}-\left(1-2e^{-1}\right)p$ in the limit as δ goes to 0.

Finally, we compare the two expected payoffs.

$$\begin{aligned} W(p) > V(p) &\iff 1 - p + p^2 > 1 - \frac{1}{2}e^{-1} - \left(1 - 2e^{-1}\right)p \\ &\iff p^2 + 2pe^{-1} + \frac{1}{2}e^{-1} > 0, \end{aligned}$$

which holds for any p. Hence, in particular, we obtain $W(p) > V(p)$ for $p > \frac{1}{1+e}$.

Hence, the voter's expected payoff in our model is smaller than under a unique mixed Nash equilibrium model in which each candidate chooses between 0 and 1 as in Aragonès and Palfrey (2002) when $p > \frac{1}{1+e}$, $\delta > 0$ is sufficiently small, and T is sufficiently large.

I.4 A Generalized Model with Valence Candidates

I.4.1 Heterogeneous Arrival Rates

This section discusses the effect of heterogeneous arrival rates. Let the arrival rate for candidate i be $\lambda_i > 0$, and allow for the possibility that $\lambda_S \neq \lambda_W$. We define $r = \frac{\lambda_S}{\lambda_W}$ as the relative frequency of the opportunities to enter between the candidates.

First, it is straightforward to show that the basic structure of the equilibrium does not change even if $\lambda_S \neq \lambda_W$: The equilibrium behaviors after some candidate has already entered are the same as before. When both candidates are announcing the ambiguous policy, there exist p^* and t^* such that if $p > p^*$, then W enters if $-t < -t^*$, he does not if $-t > -t^*$, and S never enters in any PBE. If $p < p^*$, then W enters after some cutoff and S enters as soon as possible until another cutoff. The former cutoff for W to start entering precedes in time the latter for S to stop entering.

When $r \neq 1$, the cutoff p^* can be calculated as $p^* = r^{\frac{r}{1-r}}/(1 + r^{\frac{r}{1-r}})$, and the expected payoff profile for S and W when $p > p^*$ is $\left(1 - r^{\frac{r}{1-r}}, r^{\frac{r}{1-r}}\right)$. Note that these values converge to the ones in the base model as $r \to 1$.

Since $r^{\frac{r}{1-r}}$ is decreasing in $r = \frac{\lambda_S}{\lambda_W}$, it follows that p^* is decreasing in r and S's payoff is increasing in r. Thus, having a relatively higher arrival rate makes the candidate better off. This is intuitive. With W's strategy being fixed, if S has a higher arrival rate, she has a greater chance to copy W's position. In contrast, with S's strategy being fixed, if W has a higher arrival rate, then he can wait longer at the policy profile $\{0, 1\}$ to reduce the probability of being copied afterward. Of course W's strategy is not constant in the former case and S's is not in the latter, so determination of the equilibrium strategy profile is more complicated, but these are the main driving forces of the comparative statics.

Note that Calcagno et al. (2014) show that having a higher arrival rate makes the player worse off in their analysis of battle-of-the-sexes games. This follows because having a higher arrival rate decreases his/her commitment power. The difference from our result is due to the nature of the stage game being analyzed. In a battle of the sexes, player i's ability to commit to an action a_i can help induce his or her opponent to take a_j such that (a_i, a_j) constitutes player i's favorite Nash equilibrium. In contrast, in the valence election campaign, the game is a constant-sum game, so being unable to change an action over a longer time means that the player can react to the opponent less quickly and suffers a low payoff with a larger probability.

I.4.2 Model with General Payoff Functions

Model

The simple model of "valence election campaign" presented in Section 3.1.1 was intended to provide a basic intuition for the dynamic incentive problems faced by candidates. This section extends this base model to more general cases. The policy space is X, and available policy sets are $\mathcal{X} = \{X\} \cup \left(\bigcup_{x\in X} \{\{x\}\}\right)$ for each i. The two candidates, S and W, correspond to the strong and the weak candidates, respectively. The candidates are purely office-motivated, so $v_S(X_S, X_W) + v_W(X_W, X_S) = 1$ for any $(X_S, X_W) \in \mathcal{X}_S \times \mathcal{X}_W$. The strong candidate's payoff when only the weak candidate enters is $\min_{x\in X} v_S(X, \{x\}) =: \alpha$.[9] We assume that the policy to which the strong candidate enters does not depend on the time of the entry. Formally, we assume: $\arg\max_{x\in X} v_S(\{x\}, X) \cap \arg\max_{x\in X} \min_{y\in X} v_S(\{x\}, \{y\}) \neq \emptyset$, and let an (arbitrary) element of this intersection be x^*. With this assumption, we let the weak candidate's payoff when only the strong candidate enters be $v_W(X, \{x^*\}) =: \beta$. Also, the weak candidate's payoff when the strong candidate enters and then the weak candidate enters is $\max_{y\in\mathcal{X}} v_W(\{y\}, \{x^*\}) =: \gamma$. Finally, the strong candidate wins for sure if the two candidates announce the same policy set, so

[9] As will be seen, we assume that W loses for sure if S enters after W enters. Hence, when W chooses his policy to enter, he maximizes his payoff from his entry, that is, he minimizes S's payoff, conditional on the event that S will not enter afterward.

$v_S(X_S, X_W) = 1$ if $X_S = X_W$. To summarize, the payoffs are represented as follows:

$$(S\text{'s payoff}, W\text{'s payoff}) = \begin{cases} (\alpha, 1-\alpha) & \text{if only } W \text{ enters;} \\ (1-\beta, \beta) & \text{if only } S \text{ enters;} \\ (1-\gamma, \gamma) & \text{if } S \text{ enters and then } W \text{ enters;} \\ (1, 0) & \text{if } W \text{ enters and then } S \text{ enters, or if neither enters.} \end{cases}$$

We assume $\alpha \in [0, 1)$ and $\beta, \gamma \in [0, 1]$.[10,11] We let S's arrival rate and W's arrival rate be $\lambda_S > 0$ and $\lambda_W > 0$, respectively. Call this model the *generalized valence election campaign*. It is characterized by a tuple $(\alpha, \beta, \gamma, \lambda_S, \lambda_W)$.

Note that the crucial assumptions that we make here are (i) the payoff from the game is determined solely by the policy sets at the election, (ii) S wins for sure if S and W choose the same policy, and (iii) the position in the policy space that S enters does not depend on the timing of entry.[12] These are the only restrictions that we impose. These assumptions are satisfied in our base model, with $\lambda_S = \lambda_W = \lambda$ and $\alpha = \beta = \gamma = p$.

Moreover, the specification fits other cases as well. For example, this general model can be applied to the case of a continuous policy space, the model that the literature on elections often considers. Specifically, $\mathcal{X}_i = \{x\}_{x \in [0,1]} \cup [0, 1]$ for each $i = A, B$, i.e., we allow the candidates to announce either a specific policy $x \in [0, 1]$ or an ambiguous policy $[0, 1]$. Analogous to the base model, the policy set at time $-T$ is $[0, 1]$. If candidate i wins the election and implements policy $x \in [0, 1]$, then the voter's utility with position $y \in [0, 1]$ is defined as $u(x, y) + \delta \cdot \mathbb{I}_{i=S}$, where the utility function u is strictly concave with respect to x (i.e., the voters are risk-averse). If a candidate with the ambiguous policy $[0, 1]$ wins, then the voter believes that the candidate will implement the policies in $[0, 1]$ according to the uniform distribution. Hence, the expected payoff

[10] We assume $\alpha \neq 1$ because otherwise W obtains a payoff of 0 in any equilibrium.

[11] It is not crucial that S receives the payoff exactly equal to 1 when W enters and then S enters. Specifically, with all other parameters fixed, there exists $\bar{\varepsilon} > 0$ such that for all $\varepsilon < \bar{\varepsilon}$, all the cutoffs characterizing the equilibrium behavior change continuously if the payoff profile from W entering and then S entering is $(1-\varepsilon, \varepsilon)$ and if all those cutoffs are distinct from each other at such a payoff profile.

[12] Note that (i) and (ii) imply that the position that W enters is also independent of the timing of his entry. This is because since W loses if S enters afterward by (ii), when W chooses his policy to enter, he can condition on the event that S will not enter afterward. Under such an event, by (i), W's payoff is determined solely by his policy announcement. Hence, the position that W enters is independent of his entry time.

is $\int_0^1 u(x,y)dx + \delta \cdot \mathbb{I}_{i=S}$.[13] The probability distribution of the median voter is uniform over the policy space $[0,1]$. Again, we assume that the valence term is $\delta > 0$, but is sufficiently small so that W at policy $\frac{1}{2}$ beats S with the ambiguous policy.[14,15]

In this model with the continuous policy space, if S enters before W does, she enters at policy $\frac{1}{2}$ regardless of the timing of her entry. This is because (i) this policy uniquely maximizes her payoff if W enters afterward, and (ii) it guarantees a payoff of 1 if W does not enter. If W enters before S does, he enters at a policy around $\frac{1}{2}$ regardless of the timing of her entry. This is because (i) if S enters afterward then S copies W's policy so W loses for sure, and (ii) if S does not enter afterward, policies around $\frac{1}{2}$ guarantee a payoff of 1 since voters are risk-averse.

Since the payoffs are constant-sum, Theorem 3 implies that the model with private monitoring is outcome-equivalent to the one with public monitoring. For simple notation, for the rest of this section, we assume public monitoring. That is, we assume that $h^t = \left(\left(t_S^k, x_S^k\right)_{k=1}^{k_S}, \left(t_W^k, x_W^k\right)_{k=1}^{k_W}, t \right)$ is public and analyze SPE.

Analysis and Equilibrium Dynamics

To state our result, we define three pieces of notation. First, write $Q_t = (E, N)$ if in all SPE, (i) S enters if she receives an opportunity at time $-t$ when W has not entered, and (ii) W does not enter if he receives an opportunity at $-t$ when S has not entered. That is, the first element Q_t denotes S's action at time $-t$ and the second element denotes W's action at the same time. The symbol E stands for "entering" and the symbol N stands for "not entering." Define $Q_t = (E, E)$, $Q_t = (N, E)$, and $Q_t = (N, N)$ analogously.

Second, we define functions

$$f_S(t) \quad : \quad = \begin{cases} \frac{1}{1-r}\left(e^{-\lambda_S t} - e^{-\lambda_W t}\right) - \frac{\beta + \left(1 - e^{-\lambda_W t}\right)\max\{\gamma - \beta, 0\}}{1-\alpha} & \text{if } r \neq 1; \\ \lambda_W t e^{-\lambda_W t} - \frac{\beta + \left(1 - e^{-\lambda_W t}\right)\max\{\gamma - \beta, 0\}}{1-\alpha} & \text{if } r = 1, \end{cases}$$

$$f_W(t) \quad : \quad = \begin{cases} \frac{1}{1-r}\left(e^{-\lambda_S t} - e^{-\lambda_W t}\right) - e^{-\lambda_S t} & \text{if } r \neq 1; \\ \lambda_W t e^{-\lambda_W t} - e^{-\lambda_W t} & \text{if } r = 1, \end{cases}$$

[13]The integral is well defined because u is concave and thus it is measurable.

[14]Specifically, $\int_0^1 u(x,y)dx + \delta < u(\frac{1}{2}, y)$ for all y. Note that such a $\delta > 0$ exists by the strict concavity of u.

[15]As we mentioned in the literature review, if we assume convexity, ambiguity does not need valence: If candidates are symmetric, it is optimal for a candidate to announce $[0,1]$ when the opponent is announcing $\{\frac{1}{2}\}$.

where $r = \frac{\lambda_S}{\lambda_W}$.[16]

Finally, let t_S be the smallest positive solution for $f_S(t) = 0$ (if there is no solution, then define $t_S = \infty$); and let t_W be the smallest positive solution for $f_W(t) = 0$ (since $f_W(t)$ is negative for sufficiently small $t > 0$, positive for sufficiently large t, and continuous, there always exists a positive solution).[17]

The equilibrium behavior is characterized as follows:

Proposition 14 *For the generalized valence election campaign with $(\alpha, \beta, \gamma, \lambda_S, \lambda_W)$, in any SPE, S enters at the same position as W once W has entered but S has not. In addition, the following hold.*

1. *If $\beta \geq \gamma$, then the following are true.*

   (a) *If $-t_S < -t_W$, then $Q_t = (N, E)$ for all $-t \in (-t_W, 0]$; and $Q_t = (N, N)$ for all $-t \in (\infty, -t_W)$.*

   (b) *If $-t_S > -t_W$, then there exists $t_W^* \in (t_S, \infty)$ such that $Q_t = (N, E)$ for all $-t \in (-t_S, 0]$; $Q_t = (E, E)$ for all $-t \in (-t_W^*, -t_S)$; and $Q_t = (E, N)$ for all $-t \in (-\infty, -t_W^*)$.*[18]

2. *If $\beta < \gamma$, then the following are true.*

   (a) *If $-t_S < -t_W$, then $Q_t = (N, E)$ for all $-t \in (-t_W, 0]$; and $Q_t = (N, N)$ for all $-t \in (\infty, -t_W)$.*

   (b) *If $-t_S > -t_W$, then there exists $\varepsilon > 0$ such that $Q_t = (N, E)$ for all $-t \in (-t_S, 0]$; $Q_t = (E, E)$ for $-t \in (-t_S - \varepsilon, -t_S)$. The equilibrium behavior for $-t < -t_S$ depends on the details of the parameters, but the following properties hold:*

      i. *There exists $t_W^{**} \in (t_S, \infty)$ such that W does not enter for all $-t \in (-\infty, -t_W^{**})$; and*

      ii. *There exists $\bar{r} \leq 1$ such that $r \geq \bar{r}$ if and only if there exists $t_S^{**} \in (t_S, \infty)$ such that S does not enter for all $-t \in (-\infty, -t_S)$.*

3. *All the time-cutoffs described above can be taken independent of T.*

[16] One can show that $f_S(t)$ and $f_W(t)$ are continuous in r at $r = 1$.

[17] The smallest positive solutions always exist because f_S and f_W are both continuous. We note that the notation for t_S and t_W defined here is different from the one we introduced in Section 4.1.

[18] t_W^* in the statement is the same as t_W^* in Section 4.1.

This means that, for a sufficiently long election campaign phase, W uses ambiguous language (and for many cases S uses such language as well) for a long time during the early stages of the election campaign, but the candidates' incentive to do so changes as the election date approaches. This basic pattern is common across a wide range of parameter specifications, although the exact way the incentives change varies across different specifications. Notice that in the base model, the parameters satisfy $\beta = \gamma$. In this case, if p is sufficiently small, then S enters as soon as possible. Thus, Proposition 14 claims that, if S expects even the slightest cost of W entering after her own entry (i.e., $\beta < \gamma$), then she will not enter when the election date is far away.[19]

Recall that the model includes the case of a continuous policy space with a concave payoff function. Thus, the proposition implies that the essence of our result is orthogonal to the convexity of payoff functions. This is in contrast to the models of Shepsle (1972) and Aragonès and Postlewaite (2002) in which the convexity of payoff functions is essential to the ambiguous policy announcement.

We now offer comparative statics of the cutoff times with respect to the parameter values:

Proposition 15 *In the generalized valence election campaign with $(\alpha, \beta, \gamma, \lambda_S, \lambda_W)$, the following comparative statics hold:*

1. *For each (α, β, γ), there exists $r^* \in (0, \infty)$ such that $-t_S < -t_W$ if and only if $r^* < r$.*

2. *For each $(\beta, \gamma, \lambda_S, \lambda_W)$, there exists $\alpha^* \in [0, 1)$ such that $-t_S < -t_W$ if and only if $\alpha^* < \alpha$.*

3. *For each $(\alpha, \gamma, \lambda_S, \lambda_W)$, there exists $\beta^* \in [0, 1)$ such that $-t_S < -t_W$ if and only if $\beta^* < \beta$.*

4. *For each $(\alpha, \beta, \lambda_S, \lambda_W)$, there exists $\gamma^* \in [0, 1]$ such that $-t_S < -t_W$ if and only if $\gamma^* < \gamma$.*

5. *For each $(\alpha, \beta, \lambda_S, \lambda_W)$, there exists $\bar{\gamma} \in [0, 1)$ such that, for each $\bar{\gamma} < \gamma$, there exists $-\bar{t}$ such that S does not enter at all $-t < -\bar{t}$.*

Part 1 of this proposition implies that, for sufficiently large r, Case 1(a) or 2(a) in Proposition 14 applies. Intuitively, since S can move quickly compared to W, W enters only if the deadline is very close ($-t_W$ is close to 0).

Parts 2 and 3 imply that for sufficiently large α or β, Case 1(a) or 2(a) in Proposition 14 applies. To see the intuition, notice that high α implies that S gets a high payoff when only W enters, and

[19] Note that $\beta < \gamma$ implies that S's payoff when she is the only one who enters, $1 - \beta$, is strictly greater than her payoff when W enters afterward, which is $1 - \gamma$.

high β implies that S gets a low payoff when only S enters. Hence, in these situations, S has only a small incentive to enter.

If $\beta \geq \gamma$, since W never enters after S enters, the value of γ does not affect the cutoff times. On the other hand, if $\beta < \gamma$, Part 4 implies that for sufficiently large γ, Case 1(a) or 2(a) in Proposition 14 applies. Intuitively, high γ implies that S gets a small payoff when W enters after S's entry. In such a situation, S has only a small incentive to enter.

Part 5 implies that, if γ is sufficiently large, then S does not enter if the election is sufficiently far away. To see this, consider the extreme case with $\gamma = 1$. In this case, S's payoff is zero if S enters first and then W enters afterward. Hence, if S enters when the election is far away, then with a high probability W will enter and S's payoff is close to zero. Therefore, in equilibrium, S does not enter when the election is far away.

Remark 7 (Sufficient condition for $-t_S < -t_W$) The numbers t_S and t_W that appear in Proposition 14 are only implicitly defined as the smallest solutions of $f_S(t) = 0$ and $f_W(t) = 0$, respectively. There is a sufficient condition to ensure that $-t_S < -t_W$. The sufficient condition is that $\phi < 0$, where[20]

$$\phi := \begin{cases} -\frac{\gamma}{1-\alpha} & \text{if } \gamma > \beta \text{ and } r < 1 - \frac{1-\alpha}{\gamma-\beta} \\ e^{\max\left\{\frac{\gamma-\beta}{1-\alpha}, 0\right\}-1} - \frac{\max\{\beta,\gamma\}}{1-\alpha} & \text{if } r = 1 \\ \left(\frac{1}{r} - \frac{1-r}{r}\max\left\{\frac{\gamma-\beta}{1-\alpha}, 0\right\}\right)^{\frac{r}{r-1}} - \frac{\max\{\beta,\gamma\}}{1-\alpha} & \text{otherwise} \end{cases}.$$

In fact, we use the condition in Remark 7 to show that r^* is finite and α^* and β^* are strictly less than 1 in Proposition 15. Moreover, Part 5 of Proposition 15 ensures the existence of $\bar{\gamma}$ such that $\gamma > \bar{\gamma}$ implies S does not enter if the deadline is far. In total, if at least one of these parameters is sufficiently high, then there is a long period of no entry by any candidate.

Recall that in the base model, $r = 1$ and $\alpha = \beta = \gamma = p$. Proposition 2 implies that for sufficiently large p, there is a cutoff time $-t^*$ such that no candidate enters for all $-t < -t^*$. The specification of the base model implies that the three parameters α, β, and γ move simultaneously as p varies, so it is not possible in the base model to examine the effects of individual parameters.

[20] One can show that ϕ is continuous in r at $r = 1$.

Proposition 15 ensures that if at least one of these parameters is sufficiently high, then no candidate enters when the deadline is far, as in the case of high p's. In addition, in Section I.4.1, we define p^* to be a cutoff of p such that $p > p^*$ implies the existence of t^* with which, in any equilibrium, (i) no candidate enters for all $-t < -t^*$ and (ii) W enters and S does not for all $-t > -t^*$. Part 1 of Proposition 15 generalizes the claim that p^* is decreasing in r (and it converges to 0 as $r \to \infty$). Overall, the insight from the base model carries over to the general setting.

I.5 Synchronous Policy Announcements

So far, we have assumed that candidates' policy announcements are asynchronous. In practice, not all the announcements are asynchronous; for example, televised political debates would be better modeled as synchronous policy announcements. To understand the role of the move structure on our ambiguity result, in this section we consider the case in which all the opportunities are synchronous. That is, time flows from $-T$ to 0 and, according to the Poisson process with arrival rate λ, both of the candidates receive opportunities to announce their policy platforms simultaneously. We will show that the ambiguous policy announcements are robust to this setting. The very basic intuition— S wants to wait for W who does not want to be copied, which makes both candidates announce ambiguous policies when the election date is still far away— is the same as in the base model, but the detailed equilibrium structure is different. In particular, candidates use mixed strategies at any time point close to the election date.

We assume the same voter's utility and the same distribution of the median voter as in the original model explained in Section 2. For sufficiently small valence, the payoffs at the deadline 0 are given by the payoff matrix in Figure 2 with $t = 0$.

In this model, it is straightforward to see that parts 1 and 2 of Lemma 1 continue to hold. Therefore, the only relevant state is the state in which no player has entered so far. Assume for now that a Markov perfect equilibrium exists, and fix one of them.[21] Let V_t^i be the value of candidate i when no one has yet entered at $-t$ and an opportunity to enter arrives at $-t$ but actions have not been taken. Note that this value is independent of the other histories since we consider a Markov equilibrium. Suppose neither candidate enters at $-t$. Then, if they have an opportunity at $-\tau > -t$, they will then get (V_τ^S, V_τ^W). Otherwise, $\{0,1\}, \{0,1\}$ will be realized at time 0 and

[21]We will show in Proposition 16 that a subgame perfect equilibrium exists, all subgame perfect equilibria are essentially Markov, and they have a unique continuation payoff at each time.

$S \setminus W$	$\{0\}$	$\{1\}$	$\{0,1\}$
$\{0\}$	1	p	p
$\{1\}$	$1-p$	1	$1-p$
$\{0,1\}$	$1-pe^{-\lambda t}$	$1-(1-p)\,e^{-\lambda t}$	$\int_0^t e^{-\lambda\tau}\lambda V^S_{t-\tau}d\tau + e^{-\lambda t}$

Figure 2: The payoff matrix at time $-t$

they will get $(1,0)$. Hence, the value profile of choosing $\{0,1\},\{0,1\}$ at time $-t$ is[22]

$$\left(\int_0^t \lambda e^{-\lambda\tau}V^S_{t-\tau}d\tau + e^{-\lambda t}, \int_0^t \lambda e^{-\lambda\tau}V^W_{t-\tau}d\tau\right).$$

For other action profiles, parts 1 and 2 of Lemma 1 determine the value profile. As in the base model, the game has a constant sum since the winning probabilities must sum up to 1, so it suffices to keep track of S's payoffs. Specifically, when the candidates have an opportunity at $-t$, S's payoffs for the choices of policy platforms are given by the payoff matrix in Figure 2. and V_t^S is the unique minimax value of this constant-sum game.

Unfortunately, a complete characterization of the equilibria for all parameter values is hard to obtain. However, we can show that a Markov perfect equilibrium exists (and so does a subgame perfect equilibrium), and the Markov perfect equilibrium value V_t^S is unique. Moreover, all the subgame perfect equilibria are essentially Markov, meaning that for each subgame perfect equilibrium σ, there exists σ' such that the following two conditions are satisfied:

1. For each $i \in \{S, W\}$ and h_t, candidate i's continuation payoff at h_t given strategy profile σ coincides with the one given σ'.

2. For each h_t, if the minimax strategy profile is unique in the payoff matrix represented by Figure 2, then $(\sigma_S(h_t), \sigma_W(h_t)) = (\sigma'_S(h_t), \sigma'_W(h_t))$.

Moreover, we provide two analytical results on the basic dynamics of the equilibrium behaviors.

[22]The integration is well-defined because V_t^i is continuous in t for each $i \in \{S, W\}$ for the following reason: Let W_t^S be S's continuation payoff at time $-t$ when no opportunity arrives. Since expected payoffs are continuous in probability, W_t^S is continuous in t.

In Markov equilibria, the continuation play after taking $(\{0,1\},\{0,1\})$ at $-t$ and that after not receiving an opportunity are the same. Hence, we can replace the right-bottom entry of the payoff matrix with W_t^S in Figure 2. Since the minimax value of a constant-sum normal-form game is continuous in its payoff function, this means that the expected payoff from the Nash equilibrium of the game in Figure 2 is also continuous in t. Since by definition V_t^S is the expected payoff from the Nash equilibrium of the game, V_t^S is continuous in t. Since $V_t^W = 1 - V_t^S$, both integrations in these payoffs are well-defined.

Proposition 16 *A Markov perfect equilibrium exists and the Markov perfect equilibrium value V_t^S is unique. Moreover, all the subgame perfect equilibria are essentially Markov. In addition, in each subgame perfect equilibrium, the following are true:*

1. *There exists $t^* > 0$ such that for all time $-t \in (-t^*, 0]$, both candidates use completely mixed strategies conditional on the event that the opponent has not entered.*

2. *There exists $t^{**} < \infty$ such that for all $-t < -t^{**}$, the probability with which a candidate enters at $\{0\}$ or $\{1\}$, conditional on the event that the opponent has not entered, is zero.*

Part 1 of the proposition states that if the election date is close, both candidates have to mix. This is in stark contrast to the asynchronous case, but is a natural consequence of the game representation above. The continuation payoff matrix approaches the original payoff matrix in the one-shot game whose unique equilibrium is completely mixed, and by the upper hemi-continuity of the set of Nash equilibria, the result holds.

Part 2 of the proposition shows the robustness of our ambiguity result with respect to the move structure. The intuition is the same as before. If W enters at $-t$ sufficiently far from the election date with positive probability, then it is optimal for S to wait and try to copy W's policy later. Given this, W does not enter. S gains a lot by copying W's policy, so she has an option value of waiting. Thus S does not enter either, when the election date is sufficiently far away.

As part 1 shows, the equilibrium involves mixing when the election date is close if opportunities arrive simultaneously. The mixing probabilities have to change over time, since the Nash equilibrium of the game matrix above changes as t changes. The transition of mixing probabilities is complicated and the incentive problems faced by the two candidates are subtle. We illustrate its complexity with an example with specific p and λ in Section I.5.2 in the appendix.

I.5.1 Proof of Proposition 16

We first show the result that will be useful for the following proof. Fix t arbitrarily. Suppose that the candidates play the one-shot constant-sum game, where S's payoff is given by

$S \setminus W$	$\{0\}$	$\{1\}$	$\{0,1\}$
$\{0\}$	1	p	p
$\{1\}$	$1-p$	1	$1-p$
$\{0,1\}$	$1-pe^{-\lambda\tau}$	$1-(1-p)\,e^{-\lambda\tau}$	w

Let $V(w)$ be the unique minimax value given w. We will show that

$$\left|V(w) - V(w')\right| \leq \left|w - w'\right| \tag{28}$$

for each w and w'. Without loss, we can assume $w \geq w'$.

We first derive an upper bound for $V(w) - V(w')$. By the minimax theorem, we can assume that W moves first to minimize S's payoff and then S moves to maximize S's payoff. Let $\sigma^W(w)$ be an optimal strategy for W given w. When W takes the same strategy $\sigma^W(w)$ given w', then S can improve her payoff compared to $V(w)$ at most by $w - w'$. Hence, $V(w) - V(w') \leq w - w'$.

We second derive a lower bound for $V(w) - V(w')$. By the minimax theorem, we can assume that S moves first to maximize S's payoff and then W moves to minimize S's payoff. Let $\sigma^S(w)$ be an optimal strategy for S given w. When S takes the same strategy $\sigma^S(w)$ given w', then W can improve his payoff at most by $w - w'$. Hence, $V(w) - V(w') \geq -(w - w')$. In total, we have shown (28).

We now show that Markov equilibria exist for each T. Consider the following functional equation f: Given $v^S : [0,T] \to [0,1]$ such that v^S is continuous in t, $f(v^S)(t)$ is equal to the unique minimax value of the following payoff matrix

$S \setminus W$	$\{0\}$	$\{1\}$	$\{0,1\}$
$\{0\}$	1	p	p
$\{1\}$	$1-p$	1	$1-p$
$\{0,1\}$	$1-pe^{-\lambda t}$	$1-(1-p)\,e^{-\lambda t}$	$\int_0^t e^{-\lambda\tau}\lambda v^S(t-\tau)d\tau + e^{-\lambda t}$

If v^S is continuous, then each element of the payoff matrix is continuous in t, and $f(v^S)(t)$ is also continuous in t. Hence, f is the mapping from the set of continuous functions such that $v^S : [0,T] \to [0,1]$ to itself.

Consider the sup norm: $\left\|v^S - \hat{v}^S\right\| \equiv \sup_{t\in[0,T]} \left|v^S(t) - \hat{v}^S(t)\right|$. Given this norm, the mapping f is contraction. To see why, note that, for each $t \in [0,T]$, we have

$$\begin{aligned}
\left|f(v^S)(t) - f(\hat{v}^S)(t)\right| &\leq \left|\int_0^t e^{-\lambda\tau}\lambda\left(v^S(t-\tau) - \hat{v}^S(t-\tau)\right)d\tau\right| \\
&\leq \sup_{t\in[0,1]} \left|v^S(t) - \hat{v}^S(t)\right| \int_0^t e^{-\lambda\tau}\lambda d\tau \\
&= \left(1 - e^{-\lambda T}\right)\left\|v^S - \hat{v}^S\right\|.
\end{aligned}$$

The first inequality follows from (28). Hence, there exists a unique fixed point $\bar{v}^S$ for the mapping f. When we define $V_t^S = \bar{v}^S(t)$ for each t, such V_t^S is the equilibrium value. Moreover, taking the minimax strategy of the game

$S \setminus W$	$\{0\}$	$\{1\}$	$\{0,1\}$
$\{0\}$	1	p	p
$\{1\}$	$1-p$	1	$1-p$
$\{0,1\}$	$1-pe^{-\lambda t}$	$1-(1-p)e^{-\lambda t}$	$\int_0^t e^{-\lambda\tau}\lambda V_{t-\tau}^S d\tau + e^{-\lambda t}$

(29)

in each period t is an equilibrium. Therefore, the existence is proven.

Next, we will prove that the equilibrium value in the subgame perfect equilibrium is unique. Let $h_{<-t}$ be the history before time $-t$:

$$h_{<-t} = \left(\left(t^k, x_S^k, x_W^k\right)_{k=0}^K\right),$$

where $-T < -t^1 < \ldots < -t^k < -t$ and $x_i^k \in 2^X \setminus \{\emptyset\}$ for all k and $i = S, W$. The interpretation is that $-t^k$ is the time at which the candidates receive their k'th revision opportunity, and x_i^k is the policy set that i has chosen at time $-t_i^k$.

Intuitively, the same proof as in the proof of Proposition 14 establishes the uniqueness, with h_t replaced with $h_{<-t}$. In addition, since the opportunity arrives synchronously, we consider the event such that the candidates receive an opportunity at $-t$ and both of them take $\{0,1\}$, instead of $z_i = yes$ (that is, candidate i receives an opportunity) and i taking N.

The formal proof proceeds as follows. Let $\tilde{W}_t^S(\sigma, h_{<-t})$ be S's payoff when both candidates take $\{0,1\}$ at $-t$ and take a strategy σ such that $\sigma|_{h_{<-t},yes,\{0,1\},\{0,1\}}$ is subgame perfect in the game after $(h_{<-t}, yes, \{0,1\}, \{0,1\})$, where $\sigma|_{h_{<-t},yes,\{0,1\},\{0,1\}}$ denotes a continuation strategy defined for such a subgame given by restriction of σ on such a subgame. That is, $h_{<-t}$ is the record of what has been observed before $-t$, yes means that the candidates receive an opportunity at $-t$, and both of them take $\{0,1\}$ at $-t$. Moreover, let $\bar{W}_t^S$ be the supremum of S's continuation payoff: $\bar{W}_t^S \equiv \sup_{\sigma, h_{<-t}} \tilde{W}_t^S(\sigma, h_{<-t})$, where the supremum is taken over all the possible histories and strategies such that $h_{<-t}$ is the history at $-t$, the candidates receive an opportunity at $-t$, and both of them take $\{0,1\}$, and $\sigma|_{h_{<-t},yes,\{0,1\},\{0,1\}}$ is subgame perfect after $(h_{<-t}, yes, \{0,1\}, \{0,1\})$. Similarly, let $\underline{W}_t^S$ be the infimum of S's continuation payoff: $\underline{W}_t^S \equiv \inf_{\sigma, h_{<-t}} \tilde{W}_t^S(\sigma, h_{<-t})$. Let w_t^S be the difference between the supremum and infimum: $w_t^S \equiv \bar{W}_t^S - \underline{W}_t^S$. Note that $w_0^S = 0$ since the game that the candidates play at time 0 has a unique equilibrium payoff because it is a constant-sum game.

We first show that w_t^S is continuous in t. To this end, as we do in footnote 22, let $W_t^S(\sigma, h_{<-t})$ be S's payoff when there is no opportunity at $-t$ and the candidate takes a strategy σ such that $\sigma|_{h_{<-t},no}$ is subgame perfect in the game after $(h_{<-t}, no)$ (that is, $h_{<-t}$ is what has been observed before $-t$ and no means that the candidates do not receive an opportunity at $-t$). As seen in footnote 22, $W_t^S(\sigma, h_{<-t})$ is continuous in t given σ. Hence, $\bar{W}_t^{S,\text{no}} \equiv \sup_{\sigma, h_{<-t}} W_t^S(\sigma, h_{<-t})$ and $\underline{W}_t^{S,\text{no}} \equiv \inf_{\sigma, h_{<-t}} W_t^S(\sigma, h_{<-t})$ are continuous, where supremum and infimum are taken over all the possible histories and strategies such that there is no opportunity at $-t$ and $\sigma|_{h_{<-t},no}$ is subgame perfect.

To show w_t^S is continuous in t, it suffices to show that $\bar{W}_t^S = \bar{W}_t^{S,\text{no}}$ and $\underline{W}_t^S = \underline{W}_t^{S,\text{no}}$. Let us define $\tilde{\sigma}$ as the strategy such that, after $(h_{<-t}, no)$, the candidates at time $-\tau$ follows the strategy $\sigma|_{h_{<-t},yes,\{0,1\},\{0,1\}}$. That is, they take actions as if there were an opportunity at $-t$ and both of them took $\{0,1\}$ at $-t$. Since the strategic environment is the same between $(h_{<-t}, no)$ and $(h_{<-t}, yes, \{0,1\}, \{0,1\})$, this continuation strategy is subgame perfect after $(h_{<-t}, no)$ if $\sigma|_{h_{<-t},yes,\{0,1\},\{0,1\},h_{t,\tau}}$ is subgame perfect after $(h_{<-t}, yes, \{0,1\}, \{0,1\}, h_{t,\tau})$. Therefore, for each $\tilde{W}_t^S(\sigma, h_{<-t})$ such that $\sigma|_{h_{<-t},yes,\{0,1\},\{0,1\}}$ is subgame perfect, there exists $\tilde{\sigma}$ such that $\tilde{W}_t^S(\sigma, h_{<-t}) = W_t^S(\tilde{\sigma}, h_{<-t})$ and $\tilde{\sigma}|_{h_{<-t},no}$ is subgame perfect. Therefore, we have $\bar{W}_t^S = \bar{W}_t^{S,\text{no}}$ and $\underline{W}_t^S = \underline{W}_t^{S,\text{no}}$.

We will now show that $w_t^S = 0$ for each $t \geq 0$. To this end, fix $-\tau \in (-t, 0]$ arbitrarily. Suppose that the candidates receive an opportunity at $-\tau$ for the first time after $-t$, that is, $z_{\tau'} = no$ for each $-\tau' \in (-t, -\tau)$ and $z_\tau = yes$. Here, $z_t \in \{yes, no\}$ represents whether the candidates receive an opportunity at $-t$. Let $W_t^S(\sigma, h_t \mid \tau)$ be S's continuation payoff from $-\tau$, conditional on that $z_{\tau'} = no$ for each $-\tau' \in (-t, -\tau)$ and $z_\tau = yes$. This S's continuation payoff from $-\tau$, denoted by $W_t^S(\sigma, h_t \mid \tau)$, varies at most by w_τ^S by (28).

On the other hand, let $W_t^S(\sigma, h_t \mid no)$ be candidate S's continuation payoff at 0 given σ and h_t, conditional on that no opportunity comes after $-t$ (that is, $z_\tau = no$ for each $-\tau \in (-t, 0]$). By definition, this difference is equal to $w_0^S = 0$.

Since the probability that the candidates receive the first opportunity after $-t$ at $-\tau \in (-t, 0]$ is $1 - \exp(-\lambda t)$ (that is, $z_{\tau'} = no$ for each $-\tau' \in (-t, -\tau)$ and $z_\tau = yes$ for some $-\tau \in (-\tau, 0]$), we have

$$\begin{aligned} w_t^S &\leq (1 - \exp(-\lambda t)) \times \underbrace{\max_{\tau \leq t} w_\tau^S}_{\substack{\text{Supremum difference} \\ \text{in } W_\tau^S(\sigma, h_{<\tau} \mid \tau)}} + \exp(-\lambda t) \times 0 \\ &= (1 - \exp(-\lambda t)) \max_{\tau \leq t} w_\tau^S. \end{aligned}$$

The same proof as in (17) with $\lambda_S + \lambda_W$ replaced with λ establishes the uniqueness. Let V_t^S be the unique value. Given V_t^S, the candidates at $-t$ play the constant-sum game with payoff matrix (29). Hence, as long as the minimax strategy for (29) is unique, the strategies for the candidates are unique. Hence, the equilibrium is essentially Markov.

Now we prove parts 1 and 2. Part 1 holds since (i) each candidate takes a completely mixed strategy at $-t = 0$ and (ii) the payoff function is continuous in t. Hence, we focus on proving part 2.

In equilibrium, there are following three possibilities:

1. S takes a pure strategy $\{x\}$ at $-t$. W then takes $\{x'\}$ or $\{0, 1\}$, with $x' = \{0, 1\} \setminus \{x\}$. For x to be optimal, it must be the case that $x = 1$. Consider the following two possible subcases:

   (a) If W takes a pure strategy $\{x'\}$, then S takes $\{x'\}$. This is a contradiction.

   (b) If W takes $\{0, 1\}$ with positive probability, then the payoff of S's taking $\{0, 1\}$ is $1 - p$

if W enters at $x' = 0$, and strictly greater than $1-p$ if W takes $\{0,1\}$. To see this, we calculate S's payoff for taking each action when W takes $\{0,1\}$. Conditional on W taking $\{0,1\}$, S's payoffs are given by the following table:

$S \setminus W$	$\{0,1\}$
$\{0\}$	p
$\{1\}$	$1-p$
$\{0,1\}$	$\int_0^t e^{-\lambda\tau}\lambda V_{t-\tau}^S d\tau + e^{-\lambda t}$

Since S can always enter at $\{1\}$ and thereby guarantee payoff $1-p$, it follows that $V_{t-\tau}^S \geq 1-p$ for all τ. Therefore, $\int_0^t e^{-\lambda\tau}\lambda V_{t-\tau}^S d\tau + e^{-\lambda t} \geq \left(1-e^{-\lambda t}\right)(1-p) + e^{-\lambda t} > 1-p$. This means it is a strict best response for S to announce $\{0,1\}$. This is a contradiction.

2. S takes a mixed strategy only over $\{0\}$ and $\{1\}$ at $-t$. It is then a strict best response for W to take $\{0,1\}$ since the probability of S and W entering at the same platform would then be zero. This means it is a strict best response for S to announce $\{0,1\}$ by the same argument as above. This is a contradiction.

3. S takes $\{0,1\}$ with positive probability. In order to show that it is a strict best response for W to take $\{0,1\}$, we compare W's payoff for entering at $\{x\}$ at $-t$ and that of taking $\{0,1\}$ in the following three possible subcases:

   (a) Conditional on the event that S enters at $\{x\}$ at $-t$, W gets zero if W enters at $\{x\}$. Compared to this, announcing $\{0,1\}$ is strictly better for W since that gives him at least $1-p$.

   (b) Conditional on the event that S enters at $\{x'\}$ at $-t$, W gets p by entering at $\{x\}$ if $x=0$, and gets $1-p$ if $x=1$. Announcing $\{0,1\}$ also gives W the same payoff.

   (c) Conditional on the event that S does not enter, W gets at most $p \Pr(S \text{ will not have an opportunity}) = p\exp(-\lambda t)$ by entering at $\{x\}$. On the other hand, consider the strategy in which W announces $\{0,1\}$ until $-\bar{t} = -\frac{1}{\lambda}$. If player S has entered by $-\bar{t}$, W will get at least $1-p$. Otherwise, when the candidates have an opportunity to enter at $-s \geq -\bar{t}$, then the value for S should be less than the minimax value of the following constant-sum game.

$S \setminus W$	$\{0\}$	$\{1\}$	$\{0,1\}$
$\{0\}$	1	p	p
$\{1\}$	$1-p$	1	$1-p$
$\{0,1\}$	$1-pe^{-1}$	$1-(1-p)\,e^{-1}$	1

This is because this payoff matrix is the same as the original payoff matrix except that we replace the payoffs when S takes $\{0,1\}$ with higher payoffs. The value is bounded away from 1, which means the payoff for W is bounded away from 0. Let $\underline{v}$ be this lower bound. When we take into account the probability of the candidates having an opportunity between $-\bar{t}$ and 0, the expected payoff is no less than $\left(1-e^{-1}\right)\underline{v}$. For sufficiently large t, $p\exp(-\lambda t) < \min\left\{1-p, \left(1-e^{-1}\right)\underline{v}\right\}$, which means taking $\{0,1\}$ is strictly better.

To summarize Case 3, since we assume that S takes $\{0,1\}$ with a positive probability, it follows that $\{0,1\}$ is a strict best response for W for sufficiently large t.

Let us consider S's incentive, given that W takes $\{0,1\}$. Recall that S's payoffs given that W takes $\{0,1\}$ for sure are given by the following table:

$S \setminus W$	$\{0,1\}$
$\{0\}$	p
$\{1\}$	$1-p$
$\{0,1\}$	$\int_0^t e^{-\lambda\tau}\lambda V_{t-\tau}^{S} d\tau + e^{-\lambda t}$

For the same reason as in Case 1(b) above, S should take $\{0,1\}$ with probability 1.

I.5.2 An Example of Equilibrium Dynamics with Simultaneous Arrivals

As part 1 of Proposition 16 shows, equilibria in the synchronous announcement model involve mixing when the election date is close. The mixing probabilities have to change over time, since the Nash equilibrium of the game matrix in Figure 2 changes as t changes. The transition of mixing probabilities is complicated. We illustrate its complexity in the numerical results for $p = 0.45$ and $\lambda = 1$. This example illuminates the subtle incentive problems faced by the two candidates.

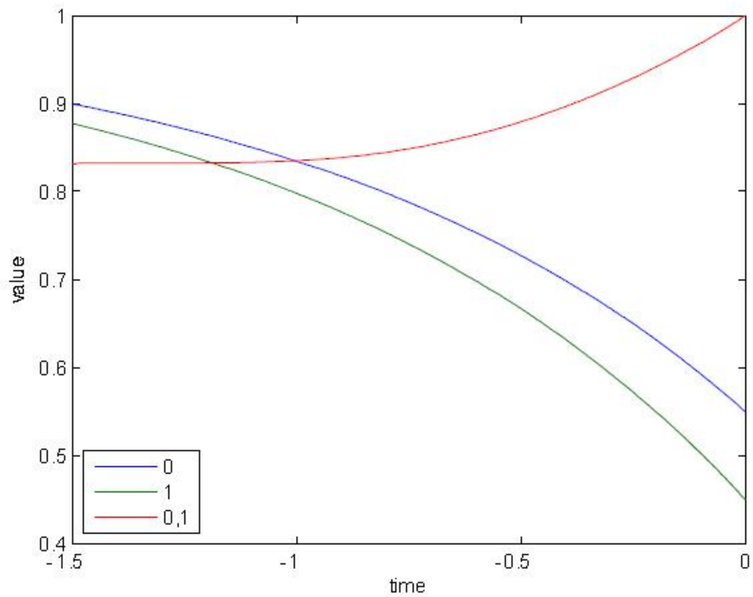


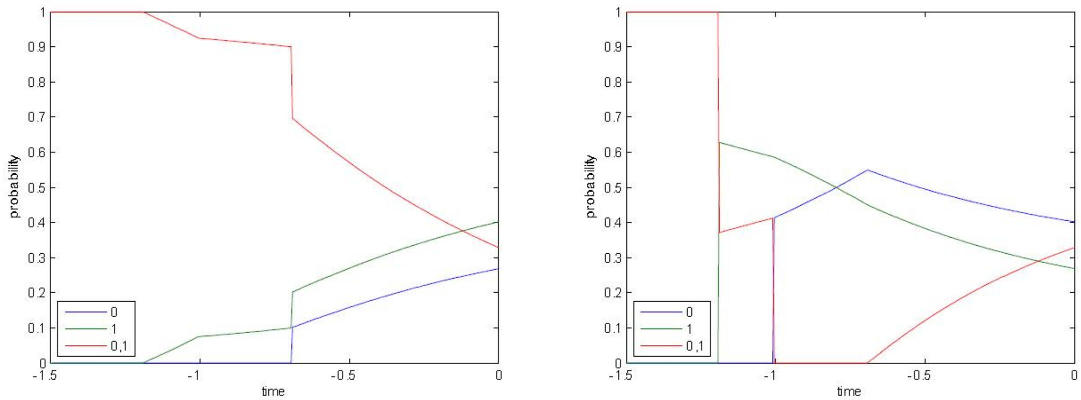
Figure 3: S's value V_t^S in the synchronous model when the candidates do not receive an opportunity at $-t$ given W's most recent announcement, given that S has been taking $\{0, 1\}$. For example, the blue line corresponds to S's value given that W has taken $\{0\}$ and S has been taking $\{0, 1\}$.

The values of S when she takes $\{0, 1\}$ against various announcements of W are as depicted in Figure 3 as a function of $-t$. Note that S's payoffs at policy profiles $(\{0, 1\}, \{0\})$ and $(\{0, 1\}, \{1\})$ at $-t$ increase as $-t$ decreases since the probability with which S can enter afterwards and copy W's policy increases. On the other hand, S's payoff at $(\{0, 1\}, \{0, 1\})$ at $-t$ decreases since the weight for the highest payoff 1 decreases.

Figure 4 depicts S's and W's strategies as functions of $-t$. When $-t$ is sufficiently close to zero, each candidate mixes over all the announcements, as we stated in part 1 of Proposition 16. Now we consider the strategies of the candidates one by one for time $-t$ close to the deadline.

(α) Consider the transition of S's strategy. Since S's mixing probability is determined in order to make W indifferent between his actions, we hypothetically fix S's mixing probability over time and examine how W's incentive changes over time; we then use this transition of W's incentive to determine how S's mixing probability should change over time.

To this end, suppose that S enters at $x \in \{0, 1\}$ with the same probability at time 0 and



S's strategy $\quad\quad$ W's strategy

Figure 4: Strategies in the synchronous model

time $-t < 0$. Then, it must be the case that W's incentive to enter at x is weaker at time $-t$ than at time 0. To show this, we compare W's payoff for taking each action at time 0 with his payoff at time $-t < 0$. At time 0, entering at x gives W a positive payoff if and only if S either enters at the other point or takes $\{0,1\}$; but taking $\{0,1\}$ gives W a positive payoff if and only if S does not take $\{0,1\}$. On the other hand, at time $-t$, entering at x gives W a positive payoff if and only if S either enters at the other point at $-t$ or "takes $\{0,1\}$ and cannot enter until the deadline"; but taking $\{0,1\}$ gives W a positive payoff if S does not take $\{0,1\}$. Furthermore, if both take $\{0,1\}$ at $-t$, then the payoff depends on the continuation play after $-t$ but is weakly higher than the payoff for both candidates taking $\{0,1\}$ at time 0.

To summarize, W's payoff for entering at x is smaller at time $-t < 0$ than at time 0 while W's payoff for taking $\{0,1\}$ is no less at time $-t < 0$ than at time 0, if S entered at x with the same probability over time. Hence, to incentivize W to enter at x, S should reduce the probability of her taking $x \in \{0,1\}$ over time.

(β) Consider the transition of W's strategy. In an approach similar to Argument (α) with the roles of S and W reversed, suppose that W enters at $x \in \{0,1\}$ with the same probability over time $[-t, 0]$. Given this assumption about W's strategy, we will show that S's incentive to enter at x is stronger at time 0 than at time $-t$.

To compare S's payoff for each action at time 0 with her payoff at time $-t$, we first show that S's payoff for entering at $x \in \{0, 1\}$ is the same between time $-t$ and time 0. At time 0, entering at x gives S a positive payoff if and only if either W takes x or the median voter is at x. At time $-t$, entering at x gives S a positive payoff if and only if either W takes x or the median voter is at x. Since we assume that W enters at x with the same probability at both times 0 and $-t$, the two payoffs are the same.

Next, we show that S's payoff for taking $\{0, 1\}$ is lower at time 0 than at time $-t$. Playing $\{0, 1\}$ at time 0 gives S the mixed-strategy equilibrium payoff in the one-shot game. If there is no opportunity after time $-t$, then since we assume that W enters at $x \in \{0, 1\}$ with the same probability between time 0 and time $-t$, S's expected payoff for taking $\{0, 1\}$ is the same as this mixed-strategy equilibrium payoff. If there is an opportunity to enter, S's payoff for taking $\{0, 1\}$ depends on W's realized action at time $-t$. If W takes $\{0, 1\}$ at time $-t$, then S's payoff again corresponds to the mixed-strategy equilibrium payoff at time 0.[23] On the other hand, if W specifies his policy, then S's payoff is 1. Since W's strategy assigns a strictly positive probability to specifying his policy, S's expected payoff for taking $\{0, 1\}$ is lower at time 0 than at time $-t$.

The above comparison implies that S's payoff for entering at x would be constant but S's payoff for taking $\{0, 1\}$ would increase as $-t$ becomes smaller, if W took each action with the same probability between time 0 and time $-t$. Hence, in order to incentivize S to enter at x, W should increase the probability of his taking x as $-t$ becomes smaller, so that both S and W enter at x with a higher probability. Therefore, W puts higher probabilities on $\{0\}$ and $\{1\}$ as $-t$ becomes smaller.

Now we consider the candidates' strategies for times further away from the deadline.

Around $-t = -0.7$, the constraint that the probability of $\{0, 1\}$ is nonnegative binds for W. As $-t$ becomes further away from the deadline than such a cutoff time, W cannot increase the probability of entering both at $\{0\}$ and $\{1\}$. Then, as seen in the comparison of S's payoff above

[23] Here, we assume that this another opportunity to enter is the last opportunity until the deadline because the probability to have one more opportunity is small for $-t$ close to 0.

(Argument (β)), entering becomes less attractive for S. Since the median voter is located with a lower probability at $\{0\}$, S stops entering at $\{0\}$.

Now let us consider the transition of the mixing probabilities in the time interval $(-1, -0.7)$. Again, as seen in the comparison of S's payoff above, W increases the probability of entering at $\{1\}$ as $-t$ becomes smaller in order to incentivize S to enter at $\{1\}$. On the other hand, as seen in the comparison of W's payoff above (Argument (α)), S reduces the probability of taking $\{1\}$ as $-t$ becomes smaller in order to incentivize W to enter at $\{1\}$.

Consider the incentive at $-t < -1$. For each time $-t \in (-1, 0]$, W is indifferent between $\{0\}$ and $\{0, 1\}$. As in the comparison of W's payoff above, if S took each action with the same probability between times -1 and $-t < -1$, W's incentive to enter at $x \in \{0, 1\}$ would decrease as $-t$ becomes smaller. As $-t$ gets smaller, this incentive gets even weaker since if S has not yet specified her policy, then S can enter with a higher probability later and W's risk of being copied by S later goes up. In general, entering at $\{0\}$ is less attractive for W than entering at $\{1\}$ since the median voter is less likely to be at policy 0. Hence, there is a time $-\bar{t}$ such that for each $-t < -\bar{t}$, W strictly prefers $\{0, 1\}$ to $\{0\}$.

Again, as seen in Argument (α) (that is, the comparison of W's payoff), as $-t$ becomes smaller, S reduces the probability of taking $\{1\}$ in order to incentivize W to enter at $\{1\}$. Finally, the probability of S taking $\{0, 1\}$ hits 1. If $-t$ gets further away from the deadline, then no player enters.

To wrap up the discussion, although the exact transition of incentives is complicated, the basic reason for the ambiguous policy announcements with synchronous arrivals is the same as in the case with asynchronous arrivals— S wants to wait for W who does not want to be copied, which makes both candidates announce ambiguous policies when the election date is still far away.

I.6 Partial Commitment

In this section, we extend our baseline model, allowing candidates to gradually clarify their policy announcements. To this end, we extend our baseline valence model to the two dimensional policy space and allow candidates to clarify policies, dimension by dimension. The policy space is $X =$

$\{0,1\}^2$. The voter is located at $(x_1, x_2) \in \{0,1\}^2$ with probability

	$x_2 = 0$	$x_2 = 1$
$x_1 = 0$	p^2	$p(1-p)$
$x_1 = 1$	$p(1-p)$	$(1-p)^2$

When she is located at (x_1, x_2), the voter's utility from a candidate i with policy (y_1, y_2) winning is

$$-|y_1 - x_1| - |y_2 - x_2| + \delta_{\{i=S\}}.$$

We allow candidates to gradually clarify their policy announcements. In particular, when a candidate's most recent policy announcement is $(X_{i,1}, X_{i,2})$ with $X_{i,1}, X_{i,2} \in \{\{1\}, \{0\}, \{1,0\}\}$, she can announce a policy $\left(\tilde{X}_{i,1}, \tilde{X}_{i,2}\right)$ with $\tilde{X}_{i,1} \subseteq X_{i,1}$ and $\tilde{X}_{i,2} \subseteq X_{i,2}$.

Given the policy set $\left(\tilde{X}_{i,1}, \tilde{X}_{i,2}\right)$ announced in the last opportunity before the deadline, the voter believes that the candidate takes a policy in $\left(\tilde{X}_{i,1}, \tilde{X}_{i,2}\right)$ according to uniform distribution, and vote for the candidate who brings the higher expected utility. In particular, with $\bar{X} \equiv \{1,0\}$, candidate S's payoff matrix of the component game (here, we omit W's payoff since the game is constant sum) is given by

$S \backslash W$	$1,1$	$1,0$	$0,1$	$0,0$	$1,\bar{X}$	$\bar{X},1$
$\bar{X},0$	p	p	p	$1-p$	$1-(1-p)^2$	p
$\bar{X},1$	p	$1-p$	$1-p$	$1-p$	$1-(1-p)p$	1
$\bar{X},\bar{X}$	$1-(1-p)^2$	$1-p(1-p)$	$1-p(1-p)$	$1-p^2$	p	p
$1,0$	p	1	$1-p(1-p)$	$1-p$	p	p
$1,1$	1	$1-p$	$1-p$	$1-p^2$	$1-p$	$1-p$
$1,\bar{X}$	p	$1-p$	$1-p$	$1-p$	1	$1-(1-p)p$
$0,0$	$1-(1-p)^2$	p	p	1	p	p
$0,1$	p	$1-p(1-p)$	1	$1-p$	p	p
$0,\bar{X}$	p	p	p	$1-p$	p	$1-(1-p)^2$

$S\backslash W$	$\bar{X},\bar{X}$	$0,\bar{X}$	$\bar{X},0$
$\bar{X},0$	p	$1-p(1-p)$	1
$\bar{X},1$	$1-p$	$1-(1-p)^2$	$1-p$
$\bar{X},\bar{X}$	1	$1-p$	$1-p$
$1,0$	$1-p(1-p)$	$1-p$	$1-p$
$1,1$	$1-p^2$	$1-p$	$1-p$
$1,\bar{X}$	$1-p$	$1-p$	$1-(1-p)^2$
$0,0$	$1-(1-p)^2$	p	p
$0,1$	$1-p(1-p)$	$1-p$	$1-p$
$0,\bar{X}$	p	1	$1-p(1-p)$

For example, when S's policy choice is $\left(\bar{X},\bar{X}\right)$ and W's choice is $(1,1)$, when the voter is located at $(0,0)$, the voter's payoff from S is

$$-\frac{1}{4}(0+1+1+2)+\delta_{\{i=S\}}=-1+\delta,$$

and that from W is -2. The similar calculation shows that S wins if the voter is located at $(1,1)$, $(0,1)$, or $(1,0)$, which happens with probability $1-(1-p)^2$.

We first narrow down the set of policies that candidates announce when the opponent's current policy is $\left(\bar{X},\bar{X}\right)$. For W, entering at (x_1,x_2) is suboptimal. To see why, observe that, for each policy (x_1,x_2), either $\left(x_1,\bar{X}\right)$ or $\left(\bar{X},x_2\right)$ guarantees the same payoff to W if S does not enter until the deadline; and $\left(x_1,\bar{X}\right)$ and $\left(\bar{X},x_2\right)$ allow W more flexibility in the continuation play in case S enters.

Lemma 13 *Given $X_S=\left(\bar{X},\bar{X}\right)$, at each $-t\in(-\infty,0]$, if he enters, W enters at $\left(x_1,\bar{X}\right)$ or $\left(\bar{X},x_2\right)$.*

Proof. It is suboptimal to enter at (x_1,x_2) with $x_1\in\{0,1\}$ and $x_2\in\{0,1\}$. To see why, we compare the strategy of entering at (x_1,x_2) at $-t$ and entering at $\left(1,\bar{X}\right)$ given $X_S=\left(\bar{X},\bar{X}\right)$ (and take a best response against S's strategy if $X_S\neq\left(\bar{X},\bar{X}\right)$ at each $-t'\in[-t,0]$.

If S does not obtain any opportunity in $(-t,0]$, then entering at $\left(x_1,\bar{X}\right)$ is at least as good as entering at (x_1,x_2) since the former obtains the payoff of $1-p$, which is the highest payoff W can obtain after entering at $(x_1, x_2) \in \{0,1\} \times \{0,1\}$.

Otherwise, let $-\tau > -t$ be the first opportunity that S obtains after $-t$. Conditional on the event that S obtains the first opportunity at $-\tau$, if W enters at (x_1, x_2) at $-t$, his expected continuation payoff is 0. If W keeps entering at $\left(1, \bar{X}\right)$ in the interval $[-t, -\tau]$, his continuation payoff is strictly positive since (i) if S does not enter at $\left(1, \bar{X}\right)$ at $-\tau$, then he can win with a positive probability if S does not have further opportunities and (ii) if S enters at $\left(1, \bar{X}\right)$ at $-\tau$, then he can still win with a positive probability (for example, entering at $(1,1)$ at each $-\tau' > -\tau$ allows him to win if S does not have further opportunities and the voter is located at $(1,1)$).

In both cases, entering at $\left(1, \bar{X}\right)$ is better than entering at (x_1, x_2). ■

For S, she enters at $(1,1)$:

Lemma 14 *Given $X_W = \left(\bar{X}, \bar{X}\right)$, S's payoff of entering is given by*

$$e^{-\lambda_W t}\left(1-p^2\right) + \left(1 - e^{-\lambda_W t}\right)(1-p). \tag{30}$$

Proof. By entering at $(1,1)$, S obtains the payoff of $1-p^2$ if W cannot move and $1-p$ if he can. It is straightforward to show that all the other policy announcement gives her the payoff less than (30): If S enters at $\left(\bar{X}, 0\right)$, then S obtains the payoff of p if W cannot move. If he can, since S's feasible announcement in the continuation play is $\left(\bar{X}, 0\right), (1,0), (0,0)$, W can ensure the payoff of p by entering at $(0,1)$. Hence, S's payoff is at most

$$e^{-\lambda_W t} p + \left(1 - e^{-\lambda_W t}\right)(1-p) < e^{-\lambda_W t}\left(1-p^2\right) + \left(1 - e^{-\lambda_W t}\right)(1-p).$$

The verification of other announcements is the same, so is omitted. ■

Next, we pin down the continuation play after W entering at $\left(1, \bar{X}\right)$ and at $\left(0, \bar{X}\right)$, given S's current policy announcement $\left(\bar{X}, \bar{X}\right)$.

Subgame after W enters at $\left(1, \bar{X}\right)$ Given $X_W = \left(1, \bar{X}\right)$, if S always take $\left(1, \bar{X}\right)$, then she obtains the payoff at least

$$e^{-\lambda_W t} + \left(1 - e^{-\lambda_W t}\right) p. \tag{31}$$

This lower bound will turn out to be useful.

Lemma 15 *Suppose the current policy announcement is* $X_W = \left(1, \bar{X}\right)$ *and* $X_S = (\bar{X}, \bar{X})$ *at* $-t \in (-\infty, 0]$*. Then, S's policy announcement at* $-\tau \geq -t$ *satisfies* $X_S \notin \{(0,1), (0,0), (0,\bar{X})\}$*.*

Intuitively, for each $x_2 \in \{0,1\}$, the voter located in $(0, x_2)$ will vote for S if S's policy is $\left(\bar{X}, \bar{X}\right)$ if $X_W \in \{\left(1, \bar{X}\right), (1,1), (1,0)\}$. Hence, S does not have incentive to make a commitment for $x_1 = 0$.

Proof. Suppose S enters at $X_S \in \{(0,1), (0,0), (0,\bar{X})\}$. Given the final realization of the policy announcement at time 0, the payoff matrix is

$S \backslash W$	$(1,1)$	$(1,0)$	$\left(1, \bar{X}\right)$
$(0,1)$	p	$1-(1-p)\,p$	p
$(0,0)$	$1-(1-p)^2$	p	p
$\left(0, \bar{X}\right)$	p	p	p

Here, we omit infeasible actions given $X_W = \left(1, \bar{X}\right)$ and $X_S \in \{(0,1), (0,0), (0,\bar{X})\}$.

W can guarantee S's payoff is bounded by p by always taking $\left(1, \bar{X}\right)$, which is smaller than (31). Hence, taking $X_S \in \{(0,1), (0,0), (0,\bar{X})\}$ is suboptimal. ■

We can also show that entering at $(1,0)$, $\left(\bar{X}, 0\right)$, $\left(\bar{X}, \bar{X}\right)$, or $\left(\bar{X}, 1\right)$ is suboptimal.

1. S does not enter at $(1,0)$ given $X_W = \left(1, \bar{X}\right)$: Given this policy announcement, S's payoff becomes p, which is dominated by (31).

2. S does not enter at $\left(\bar{X}, 0\right)$ given $X_W = \left(1, \bar{X}\right)$. Given this policy announcement, W can guarantee that S's payoff is no more than

$$e^{-\lambda_W t}\left(1-(1-p)^2\right) + \left(1-e^{-\lambda_W t}\right)p,$$

which is dominated by (31).

3. It is suboptimal for S to stay at $\left(\bar{X}, \bar{X}\right)$. To see why, suppose S strictly prefers $\left(1, \bar{X}\right)$ to $\left(\bar{X}, \bar{X}\right)$ given $X_W = \left(1, \bar{X}\right)$ for each $-\tau \in (-t, 0]$. At timing $-t$, there are following three possible events in the continuation play:

(a) The next opportunity comes to S: In this case, S's payoff is the same between her entering at $\left(1,\bar{X}\right)$ and staying at $\left(\bar{X},\bar{X}\right)$.

(b) The next opportunity comes to W at some $-\tau \in (-t,0]$. (i) If W's best response to $X_S = \left(1,\bar{X}\right)$ at $-\tau$ is $(1,1)$, then S's payoff of entering at $\left(1,\bar{X}\right)$ at time $-t$ is $e^{-\lambda_S t}p + \left(1 - e^{-\lambda_S t}\right)$. If S stays at $\left(\bar{X},\bar{X}\right)$ at $-t$, then W's continuation payoff at $-\tau$ is at least $e^{-\lambda_S t}(1-p)$ (for example, he can always stay at $\left(1,\bar{X}\right)$), and S's payoff is bounded by $e^{-\lambda_S t}p$. Hence, S strictly prefers entering at $\left(1,\bar{X}\right)$ at $-t$ conditional on this event. (ii) If W's best response to $X_S = \left(1,\bar{X}\right)$ at $-\tau$ is $\left(1,\bar{X}\right)$, then by inductive hypothesis, S strictly prefers $\left(1,\bar{X}\right)$ to $\left(\bar{X},\bar{X}\right)$ at $-t$ conditional on this event.

(c) No opportunity arrives to any candidate: In this case, S strictly prefers $\left(1,\bar{X}\right)$ to $\left(\bar{X},\bar{X}\right)$ at $-t$ conditional on this event by inductive hypothesis.

Since the payoff is continuous in time and the events (b) and (c) happen with a positive probability, by continuous time backward induction, we can conclude that S prefers at $\left(1,\bar{X}\right)$ to $\left(\bar{X},\bar{X}\right)$ given $X_W = \left(1,\bar{X}\right)$.

4. S prefers $\left(1,\bar{X}\right)$ to $\left(\bar{X},1\right)$. Given $X_S = \left(1,\bar{X}\right)$, the component game with relevant policies is

$S\backslash W$	$1,1$	$1,0$	$1,\bar{X}$
$1,\bar{X}$	p	$1-p$	1
$1,0$	p	1	p
$1,1$	1	$1-p$	$1-p$

and given $X_S = \left(\bar{X},1\right)$, we have

$S\backslash W$	$1,1$	$1,0$	$1,\bar{X}$
$\bar{X},1$	p	$1-p$	$1-p^2,p^2$
$0,1$	p	$1-p(1-p)$	p
$1,1$	1	$1-p$	$1-p$

If we re-label $(0,1)$ as $(1,0)$ in the second matrix, we obtain

$S \backslash W$	$1,1$	$1,0$	$1,\bar{X}$
$\bar{X},1$	p	$1-p$	$1-p^2, p^2$
$1,0$	p	$1-p(1-p)$	p
$1,1$	1	$1-p$	$1-p$

This re-labelling does not affect the strategic incentive, and each element of the payoff matrix is (weakly) smaller for $X_S = (\bar{X},1)$. Since the game is constant sum, S's payoff is (weakly) larger with $X_S = (1,\bar{X})$. Moreover, by backward induction, we can also show that S's payoff is strictly larger with $X_S = (1,\bar{X})$.[24] Hence, we can assume that S enters at $(1,\bar{X})$.

Hence, the relevant choices for S at $X_W = (1,\bar{X})$ is $(1,1)$ and $(1,\bar{X})$. We now figure out S's payoff of entering at $X_W = (1,\bar{X})$. To streamline the analysis, we focus on the case with $p > r^{\frac{r}{1-r}}/(1+r^{\frac{r}{1-r}})$ with $r = \lambda_S/\lambda_W$.

Given $X_S = X_W = (1,\bar{X})$, both candidates select the policy in the second dimension, and the strategic incentive is the same as the baseline model. Hence, with $t^* = \log(\lambda_S/\lambda_W)/(\lambda_S - \lambda_W)$,[25]

1. S announces $(1,\bar{X})$ for each $-t \in (-\infty, 0]$.

2. W announces $(1,\bar{X})$ for each $-t \in (-\infty, -t^*)$ and $(1,1)$ for each $-t \in (-t^*, 0]$.

S's payoff at $((1,\bar{X}),(1,\bar{X}))$ at time $-t$ is

$$v_t^S\left((1,\bar{X}),(1,\bar{X})\right) = \begin{cases} \frac{e^{\lambda_S(-t)}\left(\lambda_W - \lambda_W p + \lambda_W(p-1)e^{t(\lambda_S-\lambda_W)} + (\lambda_S-\lambda_W)e^{\lambda_S t}\right)}{\lambda_S-\lambda_W} & \text{if } t \leq t^*, \\ 1-(1-p)\left(\frac{\lambda_S}{\lambda_W}\right)^{\frac{\lambda_S}{\lambda_W-\lambda_S}} & \text{if } t \geq t^*. \end{cases}$$

Note that S announcing $(1,\bar{X})$ all the time implies that S prefers entering at $(1,\bar{X})$ to $(1,1)$ for each $-t$ given $X_W = (1,\bar{X})$.

Given Lemma 13, W stays at $(1,\bar{X})$ as long as S stays at $(\bar{X},\bar{X})$. Therefore, S's payoff of

[24] We can also show that S strictly prefers $(1,\bar{X})$ to $(\bar{X},1)$. The details are available upon request.

[25] If we consider $p < r^{\frac{r}{1-r}}/(1+r^{\frac{r}{1-r}})$, then we use part 2 of Proposition 2 instead of part 1 to characterize the candidates' behavior.

$(X_S, X_W) = \left(\left(\bar{X}, \bar{X}\right), \left(1, \bar{X}\right)\right)$ is

$$v_t^S\left(\left(\bar{X}, \bar{X}\right), \left(1, \bar{X}\right)\right) = \int_{\tau=0}^{t} \lambda_S e^{-\lambda_S \tau} v_{t-\tau}^S\left(\left(1, \bar{X}\right), \left(1, \bar{X}\right)\right) d\tau + \left(1 - e^{-\lambda_S t}\right) p.$$

Subgame after W enters at $\left(0, \bar{X}\right)$ The symmetric analysis implies that entering at $\left(0, \bar{X}\right)$ is always worse than entering at $\left(1, \bar{X}\right)$. Intuition is simple: The voter is less likely to be located at $(0, x_2)$ than at $(1, x_2)$ for each x_2.

On-Path Outcome at $X_W = X_S = \left(\bar{X}, \bar{X}\right)$ If t is sufficiently close to 0, W enters at $\left(1, \bar{X}\right)$ and S stays at $\left(\bar{X}, \bar{X}\right)$ at time $-t$. Let w_t^S be S's value at time $-t$ given the current policy announcements $X_S = \left(\bar{X}, \bar{X}\right), X_W = \left(\bar{X}, \bar{X}\right)$ and given that W enters at $\left(1, \bar{X}\right)$ and S stays at $\left(\bar{X}, \bar{X}\right)$ at all time $-\tau \in (-t, 0]$:

$$w_t^S = \int_{\tau=0}^{t} \lambda_W e^{-\lambda_W \tau} v_{t-\tau}^S\left(\left(\bar{X}, \bar{X}\right), \left(1, \bar{X}\right)\right) d\tau + \left(1 - e^{-\lambda_W t}\right).$$

Given this continuation payoff, compare S's payoff of entering, (30), with w_t^S. There exists $\bar{r}$ such that, for each $r < \bar{r}$, there exists $\bar{t}_S$ such that w_t^S is greater than (30) for each $t \in [0, \bar{t}_S)$ and smaller for each $t \in (\bar{t}_S, \infty)$; and for each $r > \bar{r}$, w_t^S is greater than (30) for each $t \in [0, \infty)$. Intuitively, if S cannot move fast (compared to W), S incurs the cost when S skips the current opportunity and then W enters at $\left(1, \bar{X}\right)$.

In contrast, compare W's payoff of entering, $1 - v_t^S\left(\left(\bar{X}, \bar{X}\right), \left(1, \bar{X}\right)\right)$, with $1 - w_t^S$. For each r, there exists $\bar{t}_W$ such that $1 - v_t^S\left(\left(\bar{X}, \bar{X}\right), \left(1, \bar{X}\right)\right)$ is greater than $1 - w_t^S$ for each $t \in [0, \bar{t}_W)$ and smaller for each $t \in (\bar{t}_W, \infty)$. Intuitively, since W incurs the cost if S copies his policy, W does not want to reduce the flexibility if the remaining time is long.

There are two cases: $\bar{t}_S > \bar{t}_W$. In this case, on equilibrium path at $X_W = X_S = \left(\bar{X}, \bar{X}\right)$, we have the following:

1. S announces $\left(\bar{X}, \bar{X}\right)$ for each $-t \in (-\infty, 0]$.

2. W announces $\left(\bar{X}, \bar{X}\right)$ for each $-t \in (-\infty, -\bar{t}_W)$ and $\left(1, \bar{X}\right)$ for each $-t \in (-\bar{t}_W, 0]$.

The proof of S not entering once W stops entering is analogous to that of Proposition 2, so is omitted.

In contrast, if $\bar{t}_S < \bar{t}_W$, then there exists $\tilde{t}_W > \bar{t}_S$ such that, on equilibrium path at $X_W = X_S = \left(\bar{X}, \bar{X}\right)$, we have the following:

1. S announces $(1,1)$ for each $-t \in (-\infty, -\bar{t}_S)$ and $\left(1, \bar{X}\right)$ for each $-t \in (-\bar{t}_S, 0]$.

2. W announces $\left(\bar{X}, \bar{X}\right)$ for each $-t \in (-\infty, -\tilde{t}_W)$ and $\left(1, \bar{X}\right)$ for each $-t \in (-\tilde{t}_W, 0]$.

The proof of the existence of such $\tilde{t}_W$ is analogous to that of Proposition 2, so is omitted.

J Proofs for Section 3.2

J.1 Proof of Proposition 3

Part 1: Policy $x^*(X,\mu)$ is a Condorcet winner. To see why, we have $v_i\left(\{x^*(X,\mu)\}, X\right) = 1$ by assumption. Moreover, given the definition of $\mathcal{M}$, Theorem 7.2 of Roemer (2001) implies that $x^*(X,\mu)$ is a best response to $x^*(X,\mu)$, and for each $x_i' \neq x^*(X,\mu)$, $v_i\left(x^*(X,\mu), x_j\right) > v_i(x_i', x_j')$ for each $x_j \in BR_j(x_i)$ and each $x_j' \in BR_j(x_i')$.

Since the game is symmetric and constant-sum, Theorem 2 implies that, in any PBE, each candidate enters at $x_i \in X_i^*$ as soon as possible.

Part 2: There exists a function $y: X \to X$ such that $P_i\left(x, y\left(x\right)\right) < \frac{1}{2}$ for each $x \in X$ for each $i = A, B$. If candidate i has not entered and j has already entered at x, then it is optimal for i to enter at $y(x)$, which gives the highest feasible payoff. If a candidate enters while the other candidate has not yet entered, then she is indifferent among any policy x with $v_i(\{x\}, X) = 1$ (which exists by assumption) since once the other candidate enters later, she will lose for sure. Therefore, Assumptions 1-3 and first-mover disadvantage for i in Section 4.1 are satisfied for each $i \in \{A, B\}$. Moreover, each candidate has a strict incentive to enter at $t = 0$. Hence, we have Case 2 with $t_0 = 0$ for Theorem 1. Hence, Theorem 1 implies that, for each i, there exists t_i such that candidate i enters at all times $-t \in (-t_i, 0]$ and does not enter at all times $-t \in (-\infty, -t_i)$.

In addition, t_i^* in Section 4.1 is calculated as follows: On the one hand, i's expected payoff of entering is the probability that the other candidate will not have an opportunity to enter. That is, $v_{i,t}\left(\text{enter}\right) = e^{-\lambda_j t}$. On the other hand, supposing that each player enters at every time $-\tau \in (-t, 0]$, we have

$$\bar{v}_{i,t}\left(\text{not}\right) = \left(\int_0^t \underbrace{e^{-(\lambda_i+\lambda_j)\tau}}_{\text{no one moves until time } -(t-\tau)} \left(\underbrace{\lambda_i d\tau}_{i \text{ moves at time } -(t-\tau)} \right) \underbrace{e^{-\lambda_j(t-\tau)}}_{j \text{ does not move in } (-(t-\tau),0]} \right)$$

$$+ \left(\int_0^t \underbrace{e^{-(\lambda_i+\lambda_j)\tau}}_{\text{no one moves until time } -(t-\tau)} \left(\underbrace{\lambda_j d\tau}_{j \text{ moves at time } -(t-\tau)} \right) \left(\underbrace{1 - e^{-\lambda_i(t-\tau)}}_{i \text{ can move in } (-(t-\tau),0]} \right) \right) + \underbrace{e^{-(\lambda_i+\lambda_j)t}}_{\text{no one moves until time } 0} \frac{1}{2}.$$

Hence, t_i^* is characterized by $f_i\left(t_i^*\right) = 0$ with

$$f_i(t) := -e^{-\lambda_i t} + \frac{\lambda_j}{\lambda_i + \lambda_j}(1 - e^{-(\lambda_i+\lambda_j)t}) + e^{-(\lambda_i+\lambda_j)t}\frac{1}{2}. \tag{32}$$

Differentiating $f_i(t)$, we get

$$f_i'(t) = \lambda_i(e^{-\lambda_i t} - e^{-(\lambda_i+\lambda_j)t}\frac{1}{2}) + \lambda_j e^{-(\lambda_i+\lambda_j)t}\frac{1}{2} > 0.$$

Since $f_i(t)$ is $-\frac{1}{2}$ at $t = 0$, continuous and strictly increasing in t, and approaches $\frac{\lambda_j}{\lambda_i+\lambda_j} > 0$ as $t \to \infty$, there exists a unique t such that $f_i(t) = 0$. The cutoff t_i^* is such t.

J.2 Proof of Proposition 4

Equation (32) implies that $e^{-\lambda_A t_B^*}$ is strictly more than $1/2$. The reason is that, letting $i = A$, the sum of the second and the third terms is a strict convex combination of $\frac{\lambda_B}{\lambda_A+\lambda_B} < \frac{1}{2}$ and $\frac{1}{2}$. Hence, $f_B\left(t_B^*\right) = 0$ implies that $e^{-\lambda_A t_B} < \frac{1}{2}$. Since this is B's continuation payoff from entering at time $-t_B^*$ and B is indifferent between entering and not entering at time $-t_B^*$, B's continuation payoff from not entering at time $-t_B^*$ is also strictly less than $1/2$. Hence, A's continuation payoff from not entering at time $-t_B^*$ is strictly greater than $1/2$. One strategy A can take is not to enter until time $-t_B^*$ and then enter for all the times in $(-t_B^*, 0]$. This gives a lower bound of A's PBE payoff that is strictly greater than $1/2$ because B does not enter for times in $(-\infty, -t_B^*)$ in any PBE. This implies that A's payoff is strictly greater than B's.

K A Proof and Additional Discussions for Section 3.3

K.1 Proof of Proposition 5

First, we compute a lower bound of the probability of candidate i winning conditional on her being able to move at time $-t$. To calculate such a bound, suppose candidate i does not enter for each time in the time interval $(-t, -\tau)$, and then enters for each time in the time interval $[-\tau, 0]$. A lower bound of the probability of winning when i uses this strategy, denoted by $\bar{p}_\tau$, is given by the following consideration: Since the second entrant can win for sure, the minimum winning probability is given by the assumption that the opponent will not enter until a candidate enters. The bound can be computed as follows:

$$\bar{p}_\tau = \int_0^\tau \underbrace{\lambda_i e^{-\lambda_i s}}_{i \text{ enters at } -(\tau - s)} \times \underbrace{e^{-\lambda_j(\tau-s)}}_{j \text{ cannot enter after } i \text{ enters}} ds = \begin{cases} \frac{\lambda_i \left[e^{-\lambda_i \tau} - e^{-\lambda_j \tau}\right]}{\lambda_j - \lambda_i} > 0 & \text{if } \lambda_i \neq \lambda_i \\ \lambda_i \tau e^{-\lambda_i \tau} & \text{if } \lambda_i = \lambda_j \end{cases}.$$

Another lower bound can be calculated by assuming that i enters at time $-t$, and it is given by $e^{-\lambda_j t}$. Hence, in total, we obtain a bound of $\max\{e^{-\lambda_j t}, \max_{\tau \in [0,t]} \bar{p}_\tau\}$. This implies that, if we take $\varepsilon < \frac{\min_{t \in [0,\infty)} \max\{e^{-\lambda_j t}, \max_{\tau \in [0,t]} \bar{p}_\tau\}}{\max_{x,y \in X} |u_i(x) - u_i(y)|}$, then at every time $-t$, there exists a strictly better strategy for candidate i than entering at a policy with which i will lose for sure.

These bounds can be used to derive an explicit expression of $\bar{\varepsilon}$:

$$\bar{\varepsilon} = \min\left\{1, \min_{i \in \{L,R\}} \frac{\min_{t \in [0,\infty)} \max\{e^{-\lambda_j t}, \max_{\tau \in [0,t]} \bar{p}_\tau\}}{\max_{x,y \in X} |u_i(x) - u_i(y)|}\right\}. \tag{33}$$

Given this definition of $\bar{\varepsilon}$, $\varepsilon < \bar{\varepsilon}$ ensures that it is a dominated strategy for candidate i to enter at a policy x such that i loses at a policy set profile $(\{x\}, X)$.

We next derive the set of policies with which candidate i can win given that candidate j has entered at x, which we denote by $X(i, x)$. If candidate j's policy is $x \in X$, candidate i can win if and only if her policy is x' (including the case where she picks X and her ideal policy is x') satisfying one of the following three conditions:

1. $x_1 \leq x'_1$ and $x_2 \leq x'_2$ (voters at $(1,0)$ and $(0,1)$ vote for her);

2. $x_1 \leq x'_1$ and $x'_1 + x'_2 \leq x_1 + x_2$ (voters at $(1,0)$ and $(0,0)$ vote for her); or

3. $x_2 \leq x_2'$ and $x_1' + x_2' \leq x_1 + x_2$ (voters at $(0,1)$ and $(0,0)$ vote for her).

Second, we derive the set of policies with which candidate L can win if candidate R does not enter. Since the voters believe that candidate R implements $\left(\frac{1}{2}, \frac{1}{2}\right)$ if she does not enter, the set is the same as $X\left(L, \left(\frac{1}{2}, \frac{1}{2}\right)\right)$. Similarly, candidate R can win with policies in $X\left(R, (0,0)\right)$ if candidate L does not enter.

We now consider each candidate's best response to the opponent's entry to x. First, suppose that R has entered at x. Candidate L enters at (x_1, x_2') with $x_2' \leq x_1$ if $x_1 \leq x_2$, and (x_1', x_2) with $x_1' \leq x_2$ if $x_1 \geq x_2$. Given the tie breaking rule, we conclude that candidate L enters at $(\min\{x_1, x_2\}, \min\{x_1, x_2\})$.

Second, suppose that L has entered at x. Given this, suppose that R's entry to x' is a best response.

1. If $x_1 \leq x_1'$ and $x_2 \leq x_2'$, then the following hold.

   (a) If $x_1 \leq \frac{1}{2}$ and $x_2 \leq \frac{1}{2}$, then $x' = \left(\frac{1}{2}, \frac{1}{2}\right)$. In this case, she receives $u_R(x') = \frac{1}{2}$.

   (b) Otherwise, given the tie breaking rule, x' is on the line segment connecting $(0,1)$ and $(1,0)$. In particular, $x_1' = x_1$ and $x_2' = 1 - x_1$ if $x_1 > \frac{1}{2}$; and $x_1' = 1 - x_2$ and $x_2' = x_2$ if $x_2 > \frac{1}{2}$. In this case, she receives $u_R(x') = 1 - \max\{x_1, x_2\}$.

2. If $x_1' + x_2' \leq x_1 + x_2$, then $x' = \left(\frac{x_1+x_2}{2}, \frac{x_1+x_2}{2}\right)$ and she receives $u_R(x') = \frac{x_1+x_2}{2}$.

Hence, for $x \in X\left(L, \left(\frac{1}{2}, \frac{1}{2}\right)\right)$ (L never enters outside of $X\left(L, \left(\frac{1}{2}, \frac{1}{2}\right)\right)$), R enters at $\left(\frac{1}{2}, \frac{1}{2}\right)$ if $x = \left(\frac{1}{2}, 0\right), \left(0, \frac{1}{2}\right)$; she enters at x' with $x_1 \leq x_1'$ and $x_2 \leq x_2'$ if x satisfies $x_1 \leq \frac{1}{2}$ and $x_2 \leq \frac{1}{2}$ or x satisfies

$$\frac{x_1 + x_2}{2} \leq 1 - \max\{x_1, x_2\}; \tag{34}$$

and she enters at $\left(\frac{x_1+x_2}{2}, \frac{x_1+x_2}{2}\right)$ if x satisfies $\frac{x_1+x_2}{2} \geq 1 - \max\{x_1, x_2\}$ and $x \neq \left(\frac{1}{2}, 0\right), \left(0, \frac{1}{2}\right)$. Since $x_1 + x_2 = \max\{x_1, x_2\} + \min\{x_1, x_2\}$, (34) is equivalent to

$$\min\{x_1, x_2\} \leq 2 - 3\max\{x_1, x_2\}.$$

Given this response of the other candidate, the following property holds for L. To formalize, let $X_t^L \subseteq X$ be the set of policies such that $x \in X_t^L$ if and only if L's continuation payoff is maximized if he enters at x at time $-t$ conditional on the event that R has not entered and L enters at $-t$. Let t_L^1 be the solution for

$$e^{-\lambda_R t} = \frac{1}{2}. \tag{35}$$

Lemma 16 $X_t^L = \left\{\left(\frac{1}{2}, \frac{1}{2}\right)\right\}$ *for* $-t \in (-t_L^1, 0]$, *and* $X_t^L = \left\{\left(\frac{2}{3}, 0\right), \left(0, \frac{2}{3}\right)\right\}$ *for* $-t \in \left(-\infty, -t_L^1\right)$.

Proof. First, note that candidate L does not enter at a policy $x \notin X\left(L, \left(\frac{1}{2}, \frac{1}{2}\right)\right)$ since R's best response against such x guarantees L to get payoff $\varepsilon u_L(\frac{1}{2}, \frac{1}{2})$, which is dominated by a payoff from a strategy of entering at $(\frac{1}{2}, \frac{1}{2}) \in X\left(L, \left(\frac{1}{2}, \frac{1}{2}\right)\right)$. Second, $X\left(L, \left(\frac{1}{2}, \frac{1}{2}\right)\right) = \{x \in X | \max\{x_1, x_2\} \geq \frac{1}{2}\}$ holds. Third, we consider the following three exhaustive cases depending on which policy among $X\left(L, \left(\frac{1}{2}, \frac{1}{2}\right)\right)$ candidate L enters at:

1. If L enters at $\left(\frac{1}{2}, 0\right)$ or $\left(0, \frac{1}{2}\right)$, then R will win and implement $\left(\frac{1}{2}, \frac{1}{2}\right)$ if she enters afterward. Hence, L's payoff is

$$\underbrace{e^{-\lambda_R t}}_{\substack{\text{Probability of } R \text{ not} \\ \text{receiving an opportunity}}} + \varepsilon \underbrace{\left(-\frac{1}{2}\right)}_{\text{Utility from the policy is } -\frac{1}{2} \text{ anyway}}.$$

2. If L enters at x with $\min\{x_1, x_2\} \leq 2 - 3\max\{x_1, x_2\}$, then R, if she enters afterward, will win and implement (x_1', x_2') such that $x_1' = x_1$ and $x_2' = 1 - x_1$ if $x_1 > \frac{1}{2}$, and $x_1' = 1 - x_2$ and $x_2' = x_2$ if $x_2 > \frac{1}{2}$. Hence, L's payoff is

$$\underbrace{e^{-\lambda_R t}}_{\substack{\text{Probability of } R \text{ not} \\ \text{receiving an opportunity}}} - \varepsilon \underbrace{\left(-\max\{x_1, x_2\}\right)}_{\substack{\text{Utility from the policy is } -\max\{x_1,x_2\} \text{ anyway} \\ \text{since } \max\{x_1,x_2\}=\max\left\{x_1',x_2'\right\}}}.$$

   Thus, among all x's in this case, L's payoff is maximized if and only if he enters at $\left(\frac{1}{2}, 0\right)$, $\left(0, \frac{1}{2}\right)$, $\left(\frac{1}{2}, \frac{1}{2}\right)$, or any convex combination of them, and his payoff is then

$$e^{-\lambda_R t} + \varepsilon\left(-\frac{1}{2}\right).$$

3. If L enters at x with $\min\{x_1, x_2\} \geq 2 - 3\max\{x_1, x_2\}$, then R will win and implement

$\left(\frac{x_1+x_2}{2}, \frac{x_1+x_2}{2}\right)$. Hence, L's payoff is

$$e^{-\lambda_R t} - \varepsilon e^{-\lambda_R t} (\max\{x_1, x_2\}) - \varepsilon \left(1 - e^{-\lambda_R t}\right) \left(-\frac{x_1 + x_2}{2}\right).$$

If (x_1, x_2) is the optimal policy for L under this case, then the constraint $\min\{x_1, x_2\} \geq 2 - 3\max\{x_1, x_2\}$ has to bind, since otherwise L wants to reduce $\max\{x_1, x_2\}$. The set of x's satisfying $\min\{x_1, x_2\} = 2 - 3\max\{x_1, x_2\}$ is expressed as

$$\left\{\left(\frac{2}{3} - \theta, 3\theta\right) \cup \left(3\theta, \frac{2}{3} - \theta\right) : \text{there exists } \theta \geq 0 \text{ and } \frac{2}{3} - \theta \geq 3\theta\right\}.$$

Given θ, L's payoff is equal to

$$\begin{aligned} &e^{-\lambda_R t} - \varepsilon e^{-\lambda_R t}\left(\frac{2}{3} - \theta\right) - \varepsilon\left(1 - e^{-\lambda_R t}\right)\left(\frac{\frac{2}{3} - \theta + 3\theta}{2}\right) \\ &= e^{-\lambda_R t} - \varepsilon e^{-\lambda_R t}\left(\frac{2}{3} - \theta\right) - \varepsilon\left(1 - e^{-\lambda_R t}\right)\left(\frac{1}{3} + \theta\right). \end{aligned}$$

Hence, if $e^{-\lambda_R t} \geq \frac{1}{2}$, then it is the best for L to enter at $\left(\frac{1}{2}, \frac{1}{2}\right)$; and if $e^{-\lambda_R t} \leq \frac{1}{2}$, then it is the best for him to enter at $\left(\frac{2}{3}, 0\right)$ or $\left(0, \frac{2}{3}\right)$.

In total, for $-t \in (-t_L^1, 0]$, candidate L enters at $\left(\frac{1}{2}, 0\right)$, $\left(0, \frac{1}{2}\right)$, $\left(\frac{1}{2}, \frac{1}{2}\right)$, or any convex combination of them, and obtains a payoff of $e^{-\lambda_R t} - \varepsilon\frac{1}{2}$. Again, by the tie breaking rule, L enters at $\left(\frac{1}{2}, \frac{1}{2}\right)$.

■

In addition, the following property holds for R:

Lemma 17 *For all $-t \in (-\infty, 0]$, if L enters at $\left(\frac{1}{2}, \frac{1}{2}\right)$ for all times in $(-t, 0]$, then R's unique best response at $-t$ is not to enter.*

Proof. Let σ_R^* be the strategy of R such that R does not enter unless L enters, and best-responds to L's policy once L enters. Consider the following two cases:

1. Conditional on the event under which L will have an opportunity at some $-\tau \in (-t, 0]$, (i) if R enters at $-t$, her payoff will be at most $\varepsilon\frac{1}{2}$, but (ii) σ_R^* gives her a payoff strictly greater than $\varepsilon\frac{1}{2}$ (since L enters at $\left(\frac{1}{2}, \frac{1}{2}\right)$ and R can win if she can enter after L enters).

2. Conditional on the event under which L will not enter, both entering at $\left(\frac{1}{2},\frac{1}{2}\right)$ and σ_R^* are optimal for R.

Since the first event happens with strictly positive probability, the proof is complete. ■

We now pin down the candidates' strategies at $-t$ sufficiently close to 0. Let t_L^2 be the unique t satisfying the following.

$$\begin{cases} \frac{\lambda_R e^{-\lambda_R t} - \lambda_L e^{-\lambda_L t}}{\lambda_L - \lambda_R} = 0 & \text{if } \lambda_L \neq \lambda_R, \\ t = \frac{1}{\lambda} & \text{if } \lambda_L = \lambda_R = \lambda. \end{cases} \tag{36}$$

For each $t < \min\{t_L^1, t_L^2\}$, suppose that candidates take the following continuation play for each $-\tau \in (-t, 0]$: R does not enter unless L enters (and takes a static best-response once L enters) and L enters at $\left(\frac{1}{2},\frac{1}{2}\right)$. Then, we show that, at time $-t$, it is optimal for R not to enter at $-t$ and for L to enter at $\left(\frac{1}{2},\frac{1}{2}\right)$.

Given this continuation play, Lemma 17 ensures that R has a strict incentive not to enter at $-t$. Hence, we consider L's incentive. L's payoff when he does not enter at time $-t$ is

$$\int_0^t \lambda_L e^{-\lambda_L \tau}\left(e^{-\lambda_R(t-\tau)} - \varepsilon\frac{1}{2}\right) d\tau - e^{-\lambda_L t}\left(-\varepsilon\frac{1}{2}\right) = \begin{cases} \lambda_L \frac{e^{-\lambda_R t} - e^{-\lambda_L t}}{\lambda_L - \lambda_R} - \varepsilon\frac{1}{2} & \text{if } \lambda_L \neq \lambda_R \\ e^{-\lambda t}\lambda t - \varepsilon\frac{1}{2} & \text{if } \lambda_L = \lambda_R = \lambda \end{cases}.$$

Hence, L strictly prefers to enter at $\left(\frac{1}{2},\frac{1}{2}\right)$ at time $-t$ if the following holds: $t < t_L^1$ and

$$\begin{cases} \lambda_L \frac{e^{-\lambda_R t} - e^{-\lambda_L t}}{\lambda_L - \lambda_R} - \varepsilon\frac{1}{2} > e^{-\lambda_R t} - \varepsilon\frac{1}{2} & \text{if } \lambda_L \neq \lambda_R, \\ \lambda t > 1 & \text{if } \lambda_L = \lambda_R = \lambda \end{cases} \Leftrightarrow t < t_L^2.$$

Moreover, if $t_L^2 \le t_L^1$, then L is indifferent between entering and not entering at time $-t_L^2$.

Therefore, by continuity of the continuation payoffs in time and the continuous-time backward induction, for each $t < \min\{t_L^1, t_L^2\}$, at time $-t$, it is uniquely optimal for R not to enter and for L to enter at $\left(\frac{1}{2},\frac{1}{2}\right)$. In what follows, we consider candidates' incentives at time $-t$ with $t > \min\{t_L^1, t_L^2\}$.

If time $-t_L^2$ is after the time at which L's optimal entering policy switches from $\left(\frac{1}{2},\frac{1}{2}\right)$ to $\left(0,\frac{2}{3}\right)$, that is, if $t_L^2 < t_L^1$, then neither L nor R enters for $-t < -t_L^2$. To see why, suppose this claim holds for $-\tau \in [-t, -t_L^2)$. Note that, on the one hand, L's payoff from entering at time $-t$ is strictly decreasing in t since the probability of candidate R entering afterward increases. On the other hand, given that R does not enter for each $-\tau$ with $\tau \leq t$, L can secure a payoff of

$$\begin{cases} \lambda_L \frac{e^{-\lambda_R t_L^2} - e^{-\lambda_L t_L^2}}{\lambda_L - \lambda_R} - \varepsilon \frac{1}{2} & \text{if } \lambda_L \neq \lambda_R \\ e^{-\lambda t_L^2} \lambda t_L^2 - \varepsilon \frac{1}{2} & \text{if } \lambda_L = \lambda_R = \lambda \end{cases}$$

by not entering in the time interval $[-t, -t_L^2)$. Since candidate L is indifferent between entering and not entering at $-t = -t_L^2$, he strictly prefers not entering for each $-t < -t_L^2$. With the same reasoning, one can show that R strictly prefers not entering for each $-t < -t_L^2$. Hence, by continuity of the continuation payoffs in time and the continuous-time backward induction, neither L nor R enters at any $-t < -t_L^2$ in any PBE.

Hence, we are left to consider the case in which $t_L^2 > t_L^1$. By continuity of the continuation payoff in time, there exists $\varepsilon > 0$ such that candidate L enters at $\left(\frac{2}{3}, 0\right)$ or $\left(0, \frac{2}{3}\right)$ for each $-t \in \left(-t_L^1 - \varepsilon, -t_L^1\right)$. Given this behavior of candidate L, candidate R faces the following trade-off:

1. Conditional on the event under which L will enter after R, either entering at $\left(\frac{1}{2}, \frac{1}{2}\right)$ or not entering (or both) is optimal for R. After R's entry to $\left(\frac{1}{2}, \frac{1}{2}\right)$, L's unique best response is to enter at $\left(\frac{1}{2}, \frac{1}{2}\right)$, and in particular entering at $\left(\frac{2}{3}, 0\right)$ and entering at $\left(0, \frac{2}{3}\right)$ are both suboptimal.
2. Conditional on the event under which L will not enter, both entering at $\left(\frac{1}{2}, \frac{1}{2}\right)$ and not entering are the best for R.

Note that the advantage for R to enter at $\left(\frac{1}{2}, \frac{1}{2}\right)$ is to change L's policy from $\left(\frac{2}{3}, 0\right)$ or $\left(0, \frac{2}{3}\right)$ to $\left(\frac{1}{2}, \frac{1}{2}\right)$ (R's ideal policy). However, such an advantage is only valid when L enters after R enters. Since L will win for sure in such a case, we will prove that, for sufficiently small policy preference $\varepsilon > 0$, it is uniquely optimal for R not to enter:

Lemma 18 *Suppose $t_L^1 < t_L^2$. Fix $-t < -t_L^1$ and L's strategy such that he enters at $\left(\frac{2}{3}, 0\right)$ or $\left(0, \frac{2}{3}\right)$ for all times in $\left(-t, -t_L^1\right)$, and enters at $\left(\frac{1}{2}, \frac{1}{2}\right)$ for all times in $(-t_L^1, 0]$. Then, conditional on any history at time $-t$ at which no candidate has entered and R receives an opportunity, not entering is R's unique best response.*

Proof. Fix time $-t < -t_L^1$. Since R not entering at all in the time interval $[-t, -t_L^1)$ is one of the feasible continuation strategies, it suffices to show that, for each $-t$, this strategy is strictly better for R than her entering at $-t$. Consider the following two cases:

1. L obtains an opportunity in the time interval $(-t, -t_L^1)$. Conditional on this event, if R enters at $x \in X$ at time $-t$, then L enters at $y(L, x)$ and wins for sure. Hence, assuming that R enters, the optimal policy for her to enter is $\left(\frac{1}{2}, \frac{1}{2}\right)$ and it gives R a payoff of $\varepsilon\frac{1}{2}$. Meanwhile, if R does not enter until $-t_L^1$, then R obtains

$$\underbrace{\left(1 - e^{-\lambda_R t_L^1}\right)}_{R \text{ can enter by the deadline after } t_L^1} \cdot \left(1 + \varepsilon\frac{1}{3}\right) + e^{-\lambda_R t_L^1}\varepsilon \cdot 0 \geq 1 - e^{-\lambda_R t_L^1}.$$

Since (35) implies that $t_L^1 = \frac{\ln \frac{1}{2}}{-\lambda_R} = \frac{\ln 2}{\lambda_R}$ and (33) implies $\varepsilon < 1$, straightforward algebra shows that not entering is uniquely optimal for R at $-t$.

2. L does not obtain an opportunity in the time interval $(-t, -t_L^1)$. Conditional on this event, since R's unique best response is not to enter at time $-t_L^1$ (note that, conditional on the event that L does not obtain an opportunity in $(-t, -t_L^1)$, R wants to enter at $-t$ if and only if she wants to enter at $-t_L^1$), it is uniquely optimal for R not to enter at $-t$.

Therefore, conditional on both events, it is uniquely optimal for R not to enter at time $-t$. ■

Let $\bar{\sigma}_L$ be candidate L's strategy such that, if R has not entered, L enters at $\left(\frac{2}{3}, 0\right)$ or $\left(0, \frac{2}{3}\right)$ for each $-t \in \left(-\infty, -t_L^1\right)$ and at $\left(\frac{1}{2}, \frac{1}{2}\right)$ for each $-t \in (-t_L^1, 0]$ (and L chooses a static best response once R enters); and let $\bar{\sigma}_R$ be candidate R's strategy such that R never enters if L has not entered (and R chooses a static best response once L enters). By $t_1^L < t_2^L$ and Lemma 18, there exists $\varepsilon > 0$ such that $\bar{\sigma}_i$ is optimal for each $-t \geq -t_L^1 - \varepsilon$ and $i \in \{L, R\}$.

For $t < t_L^1$, suppose that the candidates take $\bar{\sigma}$ for each time $-\tau$ with $\tau < t$. Given that R never enters after $-t$, given $t < t_L^1$, we must have $X_t^L = \left\{\left(\frac{2}{3}, 0\right), \left(0, \frac{2}{3}\right)\right\}$. Note that the probability that L wins by entering at $\left(\frac{2}{3}, 0\right)$ or $\left(0, \frac{2}{3}\right)$—equivalently, the probability that R cannot enter after L enters—is decreasing in t and converges to 0 as $t \to \infty$. Hence, the payoff of entering converges to $-\varepsilon\frac{1}{3}$. By (33), for sufficiently large t, there exists $\tau' \in [0, t]$ such that L can obtain a payoff greater than $-\varepsilon\frac{1}{3}$ by instead not entering until $-\tau'$. Hence, given the bound of ε we imposed and continuity of the continuation payoff in time given $\bar{\sigma}$, there exists the smallest t such that L is indifferent between entering and not entering at $-t$. Let t_L^3 be such t.

By continuity of the continuation payoff in t and the continuous-time backward induction, $\bar{\sigma}_i$ is optimal for any $-t > -t_L^3$ in any PBE. Hence, we are left to show that no candidate enters at $-t < -t_L^3$. Let σ^* be a pair of strategies such that neither L nor R enters at $-t < -t_L^3$ and both of them take $\bar{\sigma}$ for any $-t > -t_L^3$. One can show that R chooses a best response in the same way as in Lemma 18 given the continuation play σ^*. L's incentive can be checked as follows: Let v_L^3 be L's payoff of entering at time $-t_L^3$ given the continuation play $\bar{\sigma}$. Entering at $-t < -t_L^3$ gives him a payoff strictly lower than v_L^3 since the probability that R can enter after L enters increases monotonically in t. Not entering until $-t_L^3$ guarantees a payoff of v_L^3 since L is indifferent between entering and not entering at $-t_L^3$ given the continuation play σ^*. Hence, by continuity of the continuation payoffs in time and the continuous-time backward induction, both candidates take σ^* in any PBE.

Finally, we examine the conditions under which we have $t_L^1 < t_L^2$ and $t_L^1 > t_L^2$, respectively. Note that (35) implies that $t_L^1 = \frac{\ln \frac{1}{2}}{-\lambda_R} = \frac{\ln 2}{\lambda_R}$. Since the left-hand side of (36) is negative for $t \in \left(0, t_L^2\right)$, and positive for $t > t_L^2$, we have $t_L^1 < t_L^2$ if and only if the left-hand side of (36) is negative for $t = t_L^1$. Substituting $t = t_L^1 = \frac{\ln 2}{\lambda_R}$, the left-hand side of (36) is equal to

$$\frac{\lambda_R e^{-\lambda_R \frac{\ln 2}{\lambda_R}} - \lambda_L e^{-\lambda_L \frac{\ln 2}{\lambda_R}}}{\lambda_L - \lambda_R} = \frac{\frac{1}{2} - \frac{\lambda_L}{\lambda_R}\left(\frac{1}{2}\right)^{\frac{\lambda_L}{\lambda_R}}}{\frac{\lambda_L}{\lambda_R} - 1}.$$

Letting $l = \frac{\lambda_L}{\lambda_R}$, this is equal to $\frac{\frac{1}{2} - l\left(\frac{1}{2}\right)^l}{l-1}$. Taking the derivative of the numerator with respect to l yields

$$-\left(\frac{1}{2}\right)^l + l\left(\frac{1}{2}\right)^l \ln 2 = \left(\frac{1}{2}\right)^l (1 - l \ln 2).$$

Hence, the numerator is decreasing for $l \leq \frac{1}{\ln 2}$ and increasing for $l \geq \frac{1}{\ln 2}$.

Note that the numerator is zero at $l = 1 < \frac{1}{\ln 2}$ and at $l = 2 > \frac{1}{\ln 2}$. Hence, $\frac{1}{2} - l\left(\frac{1}{2}\right)^l$ is positive for $l < 1$, 0 for $l = 1$, negative for $l \in (1, 2)$, 0 for $l = 2$, and positive for $l > 2$. Together with the denominator (and using l'Hopital rule at $l = 1$), we have

$$\frac{\frac{1}{2} - l\left(\frac{1}{2}\right)^l}{l-1} \begin{cases} < 0 & \text{for } l \in (0,2) \\ = 0 & \text{for } l = 2 \\ > 0 & \text{for } l > 2 \end{cases}.$$

Therefore, $t_L^1 < t_L^2$ if and only if $\frac{\lambda_L}{\lambda_R} < 2$. In a similar vein, one can show that $t_L^1 > t_L^2$ if and only if $\frac{\lambda_L}{\lambda_R} > 2$.

K.2 Persuasion-Cost Election Campaign

In the policy-motivated election campaign in Section 3.3, L enters at suboptimal policies $\left(\frac{2}{3}, 0\right)$ or $\left(0, \frac{2}{3}\right)$ since, when R enters after L, this suboptimal policy will lead R to enter at a more favorable policy for L. Such a consideration does not occur if L does not care about what policy R picks when R wins. In such a case, the equilibrium dynamics are simpler than in the model in Section 3.3, while we can still conduct comparative statics with respect to the distribution of voters and the ideal points of the candidates more easily, keeping the advantage of the policy-motivated model over the purely office-motivated model as in Section 3.2.

Let X be an arbitrary policy space that is a full-dimensional compact subset of $\mathbb{R}^n$ for some n, and recall that $|\cdot|$ denotes the Euclidian distance. A unit mass of voters are distributed over X according to the distribution $\mu(x)$ over X. The voter located at x has utility of $-|x-y|$ from policy y.

There are two candidates L and R, and we let $\mathcal{X}_i = \{X\} \cup \left(\bigcup_{x \in X} \{\{x\}\}\right)$ for each $i = L, R$. Given a profile of policies $(x_L, x_R) \in X \times X$, we define candidate i's vote share $S_i(x_L, x_R)$ and probability of i's winning $P_i(x_L, x_R)$ as in Section 3.2. The definition of $P_i(X_i, X_j)$ when $X_i = X$ or $X_j = X$ holds is given later. We assume that $(X, \mu) \notin \mathcal{M}$.

The ideal policies of candidates L and R are x_L^* and x_R^*, respectively. The ideal policies are common knowledge among voters and candidates. The utility for candidate i is equal to

$$\begin{cases} \mathbb{I}_{i \text{ wins}} - \varepsilon \left|x_i^* - x\right| & \text{if } X_i = \{x\} \subseteq X \\ \mathbb{I}_{i \text{ wins}} & \text{if } X_i = X \end{cases},$$

where $\varepsilon > 0$. That is, the candidate incurs a cost $|x_i^* - x|$ associated with the policy to which she commits, regardless of whether she wins the election. For example, if the voters believe that x_i^* is i's ideal policy, committing to x far from x_i^* requires the cost of persuading the voters. Without specifying the policy—with $X_i = X$—, in contrast, she does not have to pay such a cost. We assume that

$$\varepsilon < \frac{1}{\max_{i \in \{R,L\}, x \in X} |x_i^* - x|}. \tag{37}$$

This condition implies that, the minimum (with respect to $x \in X$) of the payoffs from entering at some x and winning exceeds the payoff from not entering and losing. The denominator of the right-hand side of (37) is strictly positive because X is a full-dimensional subset of $\mathbb{R}^n$ and it is finite because X is compact.

Suppose that the voters believe that the candidates will implement their ideal policies once they get elected without specifying a policy. That is, we assume $S_i(X, x_j) = S_i(x_i^*, x_j)$, $S_i(x_i, X) = S_i(x_i, x_j^*)$, $S_i(X, X) = S_i(x_i^*, x_j^*)$, and the probability of winning P_i is accordingly defined when X is chosen by at least one candidate. They vote for the candidate whose policy implementation gives them the higher expected payoff. The candidate who attracts more votes will win the election. Given this, we assume that $P_R(x_R^*, x_L^*) = 1$, that is, R will win if neither candidate specifies their policies.[26] The payoff function v_i for each $i = L, R$ is specified accordingly. As in the policy-motivated election campaign in Section 3.3, we assume that the tie is broken in favor of the last candidate to specify the policy if the candidates enter at different times.[27]

Call this game a *persuasion-cost election campaign*. It is characterized by a tuple $(X, \mu, \varepsilon, T, \lambda_L, \lambda_R)$.

Let X^* be the set of policies with which L attracts weakly more votes than R if R does not specify a policy:

$$X^* = \left\{ \hat{x} : \int_x 1_{\{|x - x_R^*| \geq |x - \hat{x}|\}} \mu(x)\, dx \geq \frac{1}{2} \right\}.$$

In addition, given $x \in X$, let $X^*(x)$ be the set of policies such that R attracts weakly more votes than L given that L enters at x:

$$X^*(x) = \left\{ \hat{x} : \int_{\tilde{x}} 1_{\{|\tilde{x} - x| \leq |\tilde{x} - \hat{x}|\}} \mu(\tilde{x})\, d\tilde{x} \geq \frac{1}{2} \right\}.$$

[26] The case in which $P_R(x_L^*, x_R^*) = 0$ can be analyzed in a symmetric manner, so its analysis is omitted.

[27] We assume such a tie-breaking rule because $(X, \mu) \notin \mathcal{M}$ and thus there is no best response once the opponent enters. As in footnote 39 of the main text, the assumption corresponds to taking a limit of unique PBEs in the models with discrete policy spaces.

Given X^* and $X^*(x)$, we can characterize PBE:

Proposition 17 *The persuasion-cost election campaign with $(X, \mu, \varepsilon, T, \lambda_L, \lambda_R)$ has a PBE. Moreover, there exists $t_L^* < \infty$ such that for any PBE, the following hold:*

1. *L enters at $x \in \arg\min_{x \in X^*} |x_L^* - x|$ for $-t > -t_L^*$, while he does not enter for $-t < -t_L^*$.*

2. *R never enters unless L enters. Once L enters at x, R enters as soon as possible at $x' \in \arg\min_{x' \in X^*(x)} |x_R^* - x'|$.*

Candidate R does not have an incentive to enter before L enters since (i) R can win without entering if L cannot obtain an opportunity and (ii) R will lose by entering if L can obtain an opportunity afterward. Given this strategy of R, since L cannot win without entering, he enters if the deadline is near. If the deadline is far, then the probability that R can enter afterward is very large. Hence, entering gives L the payoff close to 0 (or negative if he pays the persuasion cost). Therefore, L does not enter when the deadline is far.

Once we specify x_R^*, x_L^*, and μ, it is straightforward to derive the distribution of the announced policies at the deadline. Thus, we can conduct the comparative statics about observable variables.[28]

K.2.1 Proof of Proposition 17

Consider a PBE. Given (37), there exists $\bar{t} > 0$ such that for all time $-t \in (\bar{t}, 0]$, L enters at some policy with which he can win. In addition, for each $-t$, if R has already entered, L takes a static best response.

Since R can win without incurring the persuasion cost if L does not enter, we can show that, for each $-t$, R does not enter:

Lemma 19 *Fix candidate L's strategy in which he takes a static best response after R enters. Then, conditional on any history at time $-t$ at which no candidate has entered and R receives an opportunity, not entering is R's unique best response.*

Proof. Since R's not entering until L enters is one of the feasible continuation strategies, it suffices to show that this strategy, denoted by $\bar{\sigma}^R$, is strictly better for R than her entering at $-t$ for each $t \geq 0$.

[28] The policy to which candidates enter is generically unique in Proposition 17.

Fix time $-t$ and a history at time $-t$ such that no candidate has entered. Consider the following two cases:

1. L obtains an opportunity in the time interval $(-t, 0]$. Fix time $\bar{t} > 0$ such that L enters for each $[-\bar{t}, 0]$ if no candidate enters. Conditional on this event, let p be the probability that L obtains an opportunity at some $-\tilde{t} \in [-\bar{t}, 0)$, and then R has an opportunity in some $-\hat{t} \in (-\tilde{t}, 0]$.

   Conditional on this event, entering at x gives R a payoff of $-\varepsilon |x - x_R| \leq 0$ while $\bar{\sigma}^R$ gives R a payoff no less than $p(1 - \max_{x \in X} |x_R^* - x|) > 0$ (strict inequality follows from (37)) since (i) if L has an opportunity at $-\tilde{t} \in [-\bar{t}, 0)$, then either L will have entered by $-\tilde{t}$ or he enters at $-\tilde{t}$, and (ii) if R has an opportunity at some $-\hat{t} \in (-\tilde{t}, 0]$, then she wins for sure by $\bar{\sigma}^R$.

2. L does not obtain an opportunity in the time interval $(-t, 0]$. Conditional on this event, $\bar{\sigma}^R$ gives R a payoff of 1, which is her largest feasible payoff.

Since $\bar{\sigma}^R$ is optimal conditional on each of these two events and the incentive is strict in the first case, it is uniquely optimal for R not to enter given the conditions in the statement of the lemma. ■

After L's entry, candidate R enters at the policy x' with which R can win with the lowest persuasion cost:

$$x' \in \arg \min_{x' \in X^*(x)} |x_R - x'| .$$

Given this reaction of R, L's payoff of entering at x at time $-t$ is $e^{-\lambda_R t} - \varepsilon |x_L - x|$. Hence, if he enters, then he enters at the policy with which L can win with the lowest persuasion cost assuming that R will not enter. His payoff of entering at $-t$ is, therefore,

$$e^{-\lambda_R t} - \min_{x \in X^*} \varepsilon |x_L - x| .$$

In contrast, his payoff of not entering at $-t$, given that he will enter as soon as possible in the interval $(-t, 0]$, is

$$\begin{aligned}
& \int_0^t \lambda_L e^{-\lambda_L \tau} \left(e^{-\lambda_R (t-\tau)} - \min_{x \in X^*} \varepsilon \left| x_L - x \right| \right) d\tau \\
= \quad & \frac{e^{-\lambda_L t} - \lambda_L e^{-\lambda_R t}}{\lambda_L - \lambda_R} - (1 - e^{-\lambda_L t}) \min_{x \in X^*} \varepsilon \left| x_L - x \right|.
\end{aligned}$$

Let

$$t_L^* = \frac{\log \frac{\lambda_L - (\lambda_L - \lambda_R) \min_{x \in X^*} \varepsilon |x_L - x|}{\lambda_R}}{\lambda_L - \lambda_R} \in (0, \infty)$$

be the smallest t such that L is indifferent between entering and not entering. By the continuous-time backward induction, for $(-t_L^*, 0]$, L enters at $x \in \arg\min_{x \in X^*} \varepsilon \left| x_L - x \right|$. We are left to show that L does not enter at any time $-t < -t_L^*$. Let σ_L^* be L's strategy such that, at any time $-t$, if R has not entered before $-t$, (i) L does not enter if $t > t_L^*$ and (ii) he enters at some $x \in \arg\min_{x \in X^*} \varepsilon \left| x_L - x \right|$ if $t < t_L^*$.

Consider the following two cases:

1. R obtains an opportunity in the time interval $(-t, -t_L^*)$. Conditional on this event, if L enters at x at time $-t$, then L's payoff is $-\varepsilon \left| x_L - x \right|$, while σ_L^* gives him a payoff of $e^{-\lambda_R t_L} - \min_{x \in X^*} \varepsilon \left| x_L - x \right|$ since no candidate will enter before $-t_L^*$ and L is indifferent between entering and not entering at $-t_L^*$. Hence it is uniquely optimal not to enter at $-t$.

2. R does not obtain an opportunity in the time interval $(-t, -t_L^*)$. Conditional on this event, L is indifferent between entering and not entering since he is indifferent between entering and not entering at $-t_L^*$.

Hence, it is uniquely optimal not to enter at $-t$ with $t > t_L^*$.

Overall, we have identified the equilibrium dynamics described in the statement of the proposition.

L Proofs for Section 3.4

Our convention throughout the paper is that, when we write $f_i(x, y)$ for a function f_i where $i \in \{A, B\}$, x is associated with candidate i and y is associated with candidate j. However, in this section of the appendix, we have that x is associated with candidate A and y is associated with candidate B. We use this alternative convention to avoid confusion about which candidate spends how much money to the campaign.

L.1 Proof of Part 1 of Proposition 6

Fix a PBE σ and an associated belief β. For any history h_i at time $-t$ at which i does not have an opportunity at $-t$, let $v_{i,t}(h_i)$ be i's continuation payoff given strategy profile σ and belief β. Let $H_{i,t}(x_A, x_B)$ be the set of i's histories such that the current time is $-t$ and the minimum spending amounts that are currently available are x_A for candidate A and x_B for candidate B. Since candidates A and B are symmetric, it is sufficient to consider the incentives in the histories in $H_{A,t}(0,0)$, $H_{A,t}(0,L)$, $H_{B,t}(0,L)$, $H_{A,t}(0,H)$, $H_{B,t}(0,H)$, $H_{A,t}(L,L)$, $H_{A,t}(L,H)$, and $H_{B,t}(L,H)$ (there is no choice to be made in histories in $H_i(H,H)$ for each $i = A, B$).

The simplest case is that a candidate has already spent H. In this case, the opponent spends H as soon as possible.

Lemma 20 *For any t and $h_A \in H_{A,t}(0,H) \cup H_{A,t}(L,H)$, candidate A spends H as soon as possible under σ, and the following hold.*

$$v_{A,t}(h_A) = \begin{cases} v_{A,t}(H,H) := \frac{\alpha}{2} & \text{if } h_A \in H_{A,t}(H,H) \\ v_{A,t}(L,H) := e^{-\lambda t}\left[\alpha \frac{L}{H+L} + (1-\alpha)(H-L)\right] + \left(1 - e^{-\lambda t}\right)\frac{\alpha}{2} & \text{if } h_A \in H_{A,t}(L,H) \\ v_{A,t}(0,H) := e^{-\lambda t}(1-\alpha) H + \left(1 - e^{-\lambda t}\right)\frac{\alpha}{2} & \text{if } h_A \in H_{A,t}(0,H) \end{cases}.$$

$$v_{B,t}(h_B) = \begin{cases} v_{B,t}(H,H) := \frac{\alpha}{2} & \text{if } h_B \in H_{B,t}(H,H) \\ v_{B,t}(L,H) := e^{-\lambda t}\alpha \frac{H}{H+L} + \left(1 - e^{-\lambda t}\right)\frac{\alpha}{2} & \text{if } h_B \in H_{B,t}(L,H) \\ v_{B,t}(0,H) := e^{-\lambda t}\alpha + \left(1 - e^{-\lambda t}\right)\frac{\alpha}{2} & \text{if } h_B \in H_{B,t}(0,H) \end{cases}.$$

Proof. This follows from the fact that $\{H\}$ is a unique static best response against $\{H\}$. ■

The next simple case is that both candidates have spent L, that is, the current profile of policy sets (X_A, X_B) satisfies $X_A = X_B = \{L, H\}$. Each candidate spends H as soon as possible in this case as well:

Lemma 21 *For any time $-t$ and $h_i \in H_{i,t}(L,L)$, each candidate i spends H as soon as possible under σ. Moreover, for each $i = A, B$, we have*

$$v_{i,t}(h_i) = v_{i,t}(L,L) := \frac{\alpha}{2} + e^{-\lambda t}\left(1 - \alpha \frac{1}{2(H+L)}\right)(H-L)$$

if $h_i \in H_{i,t}(L,L)$.

Proof. At histories in $H_{i,t}(L,L)$ for each $i = A, B$, the available policy sets for each candidate are $\{L,H\}$ and $\{H\}$. The payoff matrix of the relevant policy sets (that is, the payoffs of taking these policy sets at the deadline) is as follows.

	$\{L,H\}$	$\{H\}$
$\{L,H\}$	$\alpha\frac{1}{2} + (1-\alpha)(H-L), \alpha\frac{1}{2} + (1-\alpha)(H-L)$	$\alpha\frac{L}{H+L} + (1-\alpha)(H-L), \alpha\frac{H}{H+L}$
$\{H\}$	$\alpha\frac{H}{H+L}, \alpha\frac{L}{H+L} + (1-\alpha)(H-L)$	$\alpha\frac{1}{2}, \alpha\frac{1}{2}$

Note that H is a strictly dominant policy in this normal-form game. Hence, by part 3 of Theorem 4, spending H is uniquely optimal.

Given Lemma 20, for each $h_A \in H_{A,t}(L,L)$, we have

$$v_{A,t}(h_A) = \int_0^t 2\lambda e^{-2\lambda t} \frac{v_{A,t}(H,L) + v_{A,t}(L,H)}{2} + \underbrace{e^{-2\lambda t}}_{\text{No candidate has an opportunity until the deadline}} \times \left(\frac{\alpha}{2} + \alpha(H-L)\right).$$

Straightforward algebra shows that

$$v_{A,t}(h_A) = \frac{\alpha}{2} + e^{-\lambda t}\left(1 - \alpha \frac{1}{2(H+L)}\right)(H-L).$$

By symmetry, we have

$$v_{B,t}(h_B) = \frac{\alpha}{2} + e^{-\lambda t}\left(1 - \alpha \frac{1}{2(H+L)}\right)(H-L).$$

■

We are left to analyze histories in $H_{A,t}(0,L)$, $H_{B,t}(0,L)$, $H_{A,t}(0,0)$, and $H_{B,t}(0,0)$. Suppose now that $H_{A,t}(0,L)$ is reached. Note that $\{H\}$ strictly dominates $\{L,H\}$ for each $h_A \in H_{A,t}(0,L)$ for candidate A since $v_{A,t}(L,L) < v_{A,t}(H,L)$ for each t given Lemma 21. Hence, the remaining questions are (i) given $h_A \in H_{A,t}(0,L)$, whether candidate A wants to spend 0 or H; and (ii) given $h_B \in H_{B,t}(0,L)$, whether candidate B wants to stay at L or spend H.

At the deadline, the payoff matrix for the relevant policy sets is as follows.

	$\{L,H\}$	$\{H\}$
$\{0,L,H\}$	$(1-\alpha)H, \alpha+(1-\alpha)(H-L)$	$(1-\alpha)H, \alpha$
$\{H\}$	$\alpha\frac{H}{H+L}, \alpha\frac{L}{H+L}+(1-\alpha)(H-L)$	$\frac{\alpha}{2}, \frac{\alpha}{2}$

Note that candidate A's $\{H\}$ is a unique best response to candidate B's $\{L,H\}$, and candidate B's $\{L,H\}$ is a unique best response to candidate A's $\{0,L,H\}$. Hence, by continuity of the continuation payoffs in time, there exists $\bar{t}>0$ such that for all $t \in [0,\bar{t})$, in any PBE, candidate A takes $\{H\}$ while candidate B stays at $\{L,H\}$ at histories $h_A \in H_{A,t}(0,L)$ and $h_B \in H_{B,t}(0,L)$, respectively.

Let σ' be a strategy profile such that candidate A takes $\{H\}$ while candidate B stays at $\{L,H\}$ at histories $h_A \in H_{A,t}(0,L)$ and $h_B \in H_{B,t}(0,L)$ for any $t \geq 0$, respectively, and each candidate i follows σ_i if $h_i \notin H_{i,t}(0,L)$. Given the original β, let $\bar{v}_{i,t}(h_i)$ be candidate i's payoff at history h_i under σ' and β:

$$\begin{aligned}
\bar{v}_{A,t}(h_A) = v_{A,t}(0,L) &:= \int_0^t \lambda e^{-\lambda t} v_{A,\tau}(H,L)\,d\tau + e^{-\lambda t}(1-\alpha)H \\
&= e^{-\lambda t}\left(H + \alpha\left(e^{\lambda t}\frac{1}{2} - \frac{1}{2} - H + \frac{H-L}{2(H+L)}\lambda t\right)\right), \\
\bar{v}_{B,t}(h_B) = v_{B,t}(0,L) &:= \int_0^t \lambda e^{-\lambda t} v_{B,\tau}(H,L)\,d\tau + e^{-\lambda t}\left[\alpha + (1-\alpha)(H-L)\right] \\
&= e^{-\lambda t}\left((H-L)(1+\lambda t) + \alpha\begin{pmatrix} \frac{H}{2(H+L)}\left(1+e^{\lambda t}-\lambda t\right) - \frac{H^2}{(H+L)}(1+\lambda t) \\ +\frac{L}{2(H+L)}\left(e^{\lambda t} + (1+2L)(1+\lambda t)\right)\end{pmatrix}\right). \qquad (38)
\end{aligned}$$

Since there exists $\bar{t}>0$ such that σ and σ' coincide for all $t \in [0,\bar{t})$, candidate A takes $\{H\}$ and candidate B takes $\{L,H\}$ at time $-t$ under σ if

$$v_{A,t}(0,L) < v_{A,t}(H,L) \text{ and } v_{B,t}(0,L) > v_{B,t}(0,H).$$

By straightforward algebra, we have

$$v_{A,t}(0,L) \leq v_{A,t}(H,L) \Leftrightarrow t \leq t^* := \frac{1}{\lambda} \tag{39}$$

and

$$v_{B,t}(0,L) \geq v_{B,t}(0,H) \Leftrightarrow t \leq t^{**} := \begin{cases} \frac{1}{\frac{\alpha}{2(1-\alpha)(H+L)}-1}\frac{1}{\lambda} & \text{if } \frac{\alpha}{2(1-\alpha)(H+L)} - 1 > 0 \\ +\infty & \text{otherwise} \end{cases}. \tag{40}$$

In both (39) and (40), the first inequality holds with equality if and only if the second holds with equality. Intuitively, near the deadline, since $\{H\}$ is a static best response against $\{L, H\}$, candidate A chooses $\{H\}$ as soon as possible. Since $\{H\}$ is a static best response to $\{H\}$ and candidate A enters at H regardless of candidate B staying at $\{L, H\}$ or choosing $\{H\}$, candidate B wants to spend H if the deadline is far ($-t < -t^{**}$) and hence it is sufficiently likely that candidate A enters at H.

By (2), we have $t^{**} < t^*$. Hence, for $t \leq t^{**}$, we have $v_{A,t}(h_A) = \bar{v}_{A,t}(h_A)$ and $v_{B,t}(h_B) = \bar{v}_{B,t}(h_B)$, and under σ, candidate A takes $\{H\}$ as a unique best response at any history $h_A \in H_{A,t}(0,L)$ for all time $-t \leq -t^{**}$ while candidate B stays at $\{L, H\}$ as a unique best response at any history $h_B \in H_{B,t}(0,L)$ for all time $-t < -t^{**}$. By continuity of the continuation payoff in time, the following result holds:

Lemma 22 *Under σ, there exists $\varepsilon > 0$ such that, for each $-t \in (-(t^{**} + \varepsilon), 0]$, candidate A takes $\{H\}$ at any $h_A \in H_{A,t}(0,L)$.*

Proof. The result follows from (39), $t^{**} < t^*$, and the continuity of the continuation payoff in time. ■

Before analyzing the candidates' incentives at $h_A \in H_{A,t}(0,L)$ and $h_B \in H_{B,t}(0,L)$ such that $-t < -t^{**}$, we consider their incentives when no candidate has spent anything during $-t > -t^{**}$.

Lemma 23 *Under σ, for any $-t \in (-t^{**}, 0]$, candidates A and B take $\{L, H\}$ at any $h_A \in H_{A,t}(0,0)$ and $h_B \in H_{B,t}(0,0)$, respectively, and we have*

$$v_{A,t}(h_A) = v_{A,t}(0,0) := \frac{1}{2}\alpha + (1-\alpha)(H + (H-L)\lambda t)e^{-\lambda t},$$
$$v_{B,t}(h_B) = v_{B,t}(0,0) := \frac{1}{2}\alpha + (1-\alpha)(H + (H-L)\lambda t)e^{-\lambda t}.$$

Proof. At the deadline, taking $\{L, H\}$ is a static best response against $\{0, L, H\}$. Hence, by continuity of the continuation payoffs in time, there exists $\bar{t} > 0$ such that for all $t \in [0, \bar{t})$, candidates A and B take $\{L, H\}$ at histories $h_A \in H_{A,t}(0,0)$ and $h_B \in H_{B,t}(0,0)$, respectively.

Let σ'' be a strategy profile such that candidates A and B take $\{L, H\}$ at histories $h_A \in H_{A,t}(0,0)$ and $h_B \in H_{B,t}(0,0)$, respectively, and follow σ if $h_i \notin H_{i,t}(0,0)$. Given the original β, let $\hat{v}_{i,t}(h_i)$ be candidate i's payoff at history h_i under σ'' and β:

$$\begin{aligned}\hat{v}_{i,t}(h_i) &= \int_0^t 2\lambda e^{-2\lambda\tau} \frac{v_{A,\tau}(L,0) + v_{A,\tau}(0,L)}{2} d\tau + e^{-2\lambda t}\left(\frac{\alpha}{2} + (1-\alpha) H\right) \\ &= \frac{1}{2}\alpha + (1-\alpha)(H + (H-L)\lambda t) e^{-\lambda t}.\end{aligned}$$

Since there exists $\bar{t} > 0$ such that σ and σ'' coincide for $t \in [0, \bar{t})$, both candidates A and B take $\{L, H\}$ at histories $h_A \in H_{A,t}(0,0)$ and $h_B \in H_{B,t}(0,0)$ in σ if

$$\begin{aligned} v_{A,t}(L,0) &> \max\{v_{A,t}(H,0), v_{A,t}(0,0)\}, \text{ and} \\ v_{B,t}(0,L) &> \max\{v_{B,t}(0,H), v_{B,t}(0,0)\}.\end{aligned}$$

By symmetry, we focus on the case $v_{B,t}(0,L) > \max\{v_{B,t}(0,H), v_{B,t}(0,0)\}$. By continuity of the continuation payoff in time, it suffices to show that $v_{B,t}(0,L) > \max\{v_{B,t}(0,H), v_{B,t}(0,0)\}$ for each $t < t^{**}$.

Since $v_{B,t'}(0,L) > v_{B,t'}(0,H)$ for $t' < t^{**}$ by (40), it suffices to show that, for each $t \le t^{**}$,

$$v_{B,t}(0,L) > v_{B,t}(0,0) \Leftrightarrow (H+L)(\alpha - 2L(1-\alpha)) - (H-L)\alpha t\lambda > 0. \tag{41}$$

The right hand side of (41) holds because, for each $t \le t^{**}$,

$$\begin{aligned} &(H+L)(\alpha - 2L(1-\alpha)) - (H-L)\alpha t\lambda \\ &\ge (H+L)(\alpha - 2L(1-\alpha)) - (H-L)\alpha t^*\lambda \quad (\text{since } t \le t^{**} < t^*) \\ &= (H+L)(\alpha - 2L(1-\alpha)) - (H-L)\alpha \\ &> 0\end{aligned}$$

because $\alpha \geq \frac{H+L}{H+L+\frac{1}{2}} \geq \frac{H+L}{H+L+1}$. ■

A similar proof shows that neither candidate A nor B stays at $\{0, L, H\}$ at any $h_A \in H_{A,t}(0,0)$ and $h_B \in H_{B,t}(0,0)$, respectively.

Lemma 24 *Under σ, there exists $\varepsilon' > 0$ such that, for any $-t \in (-(t^{**} + \varepsilon'), 0]$, no candidate chooses $\{0, L, H\}$ at any $h_A \in H_{A,t}(0,0)$ and $h_B \in H_{B,t}(0,0)$, respectively.*

Proof. By continuity of the continuation payoff in time and symmetry, it suffices to show that $\max\{v_{B,t^{**}}(0,L), v_{B,t^{**}}(0,H)\} > v_{B,t^{**}}(0,0)$. In particular, it is sufficient to have $v_{B,t^{**}}(0,L) > v_{B,t^{**}}(0,0)$. This inequality follows from (41). ■

Given this lemma, we have the following result:

Lemma 25 *Under σ, there exists $\varepsilon'' > 0$ such that, for each $-t \in (-(t^{**} + \varepsilon''), -t^{**})$, both candidates A and B take $\{H\}$ at any $h_A \in H_{A,t}(0,0) \cup H_{A,t}(L,0)$ and $h_B \in H_{B,t}(0,0) \cup H_{B,t}(0,L)$, respectively.*

Proof. Fix $\varepsilon > 0$ such that Lemma 22 holds, and fix $\varepsilon' > 0$ such that Lemma 24 holds. Take $\varepsilon'' \in (0, \min\{\varepsilon, \varepsilon'\})$, and suppose candidate i obtains an opportunity at $-t \in (-(t^{**} + \varepsilon''), -t^{**})$. Let x_i be i's optimal policy at $-t$. By Lemma 24, we have $x_i \in \{\{L, H\}, \{H\}\}$. Consider the following three scenarios that can happen after $-t$: For each $i \in \{A, B\}$,

1. The probability that the two candidates have opportunities in the time interval $(-t, -t^{**})$ is equal to $\left(1 - e^{-\lambda\varepsilon'}\right)^2$.

2. The probability that only candidate j has an opportunity in the time interval $(-t, -t^{**})$ is equal to $e^{-\lambda\varepsilon'}\left(1 - e^{-\lambda\varepsilon'}\right)$. Then, since $x_i \in \{\{L, H\}, \{H\}\}$, Lemmas 20 and 22 (and symmetry) imply that candidate j takes $\{H\}$. Hence, $h_A \in H_{A,t^{**}}(x_{i,t^{**}}, x_{j,t^{**}})$ and $h_B \in H_{B,t^{**}}(x_{i,t^{**}}, x_{j,t^{**}})$ with $(x_{i,t^{**}}, x_{j,t^{**}}) = (x_i, H)$ will be realized at $-t^{**}$. Candidate i strictly prefers $\{H\}$ to $\{L, H\}$ by Lemma 20:

$$v_{i,t^{**}}(H, H) - v_{i,t^{**}}(L, H) > 0.$$

   Note that this inequality holds independently from the choice of ε'.

3. Candidate j does not have an opportunity in the time interval $(-t, -t^{**})$. Conditional on this event, candidate i prefers $\{L, H\}$ to $\{H\}$ at $-t$ if and only if she prefers $\{L, H\}$ to $\{H\}$ at $-t^{**}$. Hence, both $\{H\}$ and $\{L, H\}$ are optimal by the definition of t^{**}.

Since the likelihood ratio of case 2 against case 1 goes to ∞ as $\varepsilon' \to 0$, there exists $\varepsilon' \in (0, \varepsilon)$ such that for each $-t \in (-(t^{**} + \varepsilon'), -t^{**})$, $\{H\}$ is the unique optimal policy set at $-t$. ■

Fix $t > t^{**}$ arbitrarily, and fix the continuation play such that both candidates A and B strictly prefer $\{H\}$ to $\{L, H\}$ and $\{0, L, H\}$ at any $h_A \in H_{A,t'}(0,0) \cup H_{A,t'}(L,0)$ and $h_B \in H_{B,t'}(0,0) \cup H_{B,t'}(0,L)$ for any $-t' \in (-t, -t^{**})$. Then both candidates A and B take $\{H\}$ at any $h_A \in H_{A,t}(0,0) \cup H_{A,t}(L,0)$ and $h_B \in H_{B,t}(0,0) \cup H_{B,t}(0,L)$. To see why, consider the following four scenarios for each $i \in \{A, B\}$:

1. If the two candidates have opportunities in the time interval $(-t, -t^{**})$, then regardless of the strategy at $-t$, $h_A \in H_{A,t^{**}}(H,H)$ and $h_B \in H_{B,t^{**}}(H,H)$ will be realized. In this case, candidate A's spending at $-t$ does not affect i's payoff.

2. If only candidate j has an opportunity in the time interval $(-t, -t^{**})$, then $h_A \in H_{A,t^{**}}(x_{i,t^{**}}, x_{j,t^{**}})$ and $h_B \in H_{B,t^{**}}(x_{i,t^{**}}, x_{j,t^{**}})$ with $(x_{i,t^{**}}, x_{j,t^{**}}) = (x_i, H)$ will be realized at $-t^{**}$, where x_i is i's spending at $-t$. In this case, candidate i strictly prefers $\{H\}$ to $\{L, H\}$ and $\{0, L, H\}$.

3. If only candidate i has an opportunity in the time interval $(-t, -t^{**})$, then $h_A \in H_{A,t^{**}}(x_{i,t^{**}}, x_{j,t^{**}})$ and $h_B \in H_{B,t^{**}}(x_{i,t^{**}}, x_{j,t^{**}})$ with $(x_{i,t^{**}}, x_{j,t^{**}}) = (H, x_j)$ will be realized, where x_j is j's spending at $-t$. In this case, candidate i's spending at $-t$ does not affect i's payoff.

4. If no candidate has an opportunity in the time interval $(-t, -t^{**})$, then $h_A \in H_{A,t^{**}}(x_{i,t^{**}}, x_{j,t^{**}})$ and $h_B \in H_{B,t^{**}}(x_{i,t^{**}}, x_{j,t^{**}})$ with $(x_{i,t^{**}}, x_{j,t^{**}}) = (x_i, x_j)$ will be realized, where (x_i, x_j) is the spending profile at $-t$. In this case, both $\{H\}$ and $\{L, H\}$ are optimal by the definition of t^{**}.

In total, spending H is uniquely optimal, as desired.

Finally, together with Lemma 25, under σ, both candidates A and B strictly prefer $\{H\}$ to $\{L, H\}$ and $\{0, L, H\}$ at any $h_A \in H_{A,t}(0,0) \cup H_{A,t}(L,0)$ and $h_B \in H_{B,t}(0,0) \cup H_{B,t}(0,L)$, respectively, for any $-t \in (-\infty, -t^{**})$.

L.2 Proof of Part 2 of Proposition 6

Fix a PBE σ and an associated belief β. As in part 1 of Proposition 6, it is sufficient to consider the incentives in $H_{A,t}(0,0)$, $H_{A,t}(0,L)$, $H_{B,t}(0,L)$, $H_{A,t}(0,H)$, $H_{B,t}(0,H)$, $H_{A,t}(L,L)$, $H_{A,t}(L,H)$, and $H_{B,t}(L,H)$. Lemmas 20 and 21 still hold. In particular, for $h_A \in H_{A,t}(0,L)$, candidate A always prefers $\{H\}$ to $\{L,H\}$.

Moreover, by the same calculation as (39) and (40), we derive the following: Let σ' be a strategy profile such that candidate A takes $\{H\}$ while candidate B stays at $\{L,H\}$ at histories $h_A \in H_{A,t}(0,L)$ and $h_B \in H_{B,t}(0,L)$ for any $t \geq 0$, respectively. Given the original β, the candidates' payoffs under σ' and β satisfy $\bar{v}_{A,t}(h_A) = v_{A,t}(0,L)$ and $\bar{v}_{B,t}(h_B) = v_{B,t}(0,L)$.

Since there exists $\bar{t} > 0$ such that σ and σ' coincide for $t \in [0,\bar{t})$ as in the proof of part 1, candidate A takes $\{H\}$ and candidate B takes $\{L,H\}$ in σ if

$$v_{A,t}(0,L) < v_{A,t}(H,L) \text{ and } v_{B,t}(0,L) > v_{B,t}(0,H).$$

Again, we have

$$v_{A,t}(0,L) \leq v_{A,t}(H,L) \Leftrightarrow t \leq t^*, \text{ and}$$

$$v_{B,t}(0,L) \geq v_{B,t}(0,H) \Leftrightarrow t \leq t^{**}. \tag{42}$$

In both of these equivalence relationships, the first inequality holds with equality if and only if the second holds with equality.

By (3), we have $t^{**} > t^*$. Hence, under σ, candidate A takes $\{H\}$ as a unique best response at any history $h_A \in H_{A,t}(0,L)$ for all time $-t < -t^*$ while candidate B stays at $\{L,H\}$ as a unique best response at any history $h_B \in H_{B,t}(0,L)$ for all time $-t \leq -t^*$.

For $-t < -t^*$, we show that candidate A stays at $\{0,L,H\}$ for $h_A \in H_{A,t}(0,L)$ and candidate B stays at $\{L,H\}$ for $h_B \in H_{B,t}(0,L)$. First, we show that this is true for $-t$ close to $-t^*$:

Lemma 26 *Under σ, there exists $\varepsilon > 0$ such that, for any $-t \in (-(t^*+\varepsilon), -t^*)$, candidate A chooses $\{0,L,H\}$ for each $h_A \in H_{A,t}(0,L)$ and candidate B chooses $\{L,H\}$ for $h_B \in H_{B,t}(0,L)$.*

Proof. Since $t^{**} > t^*$ and $v_{B,t}(0,L) > v_{B,t}(0,H)$ if $t < t^{**}$, by the continuity of the continuation payoff in time, there exists $\varepsilon_B > 0$ such that, for $h_B \in H_{B,t}(0,L)$ for any $-t \in (-(t^*+\varepsilon_B), -t^*)$, candidate B strictly prefers staying at $\{L,H\}$ at $-t$.

Given this incentive, we consider candidate A's incentive for each $h_A \in H_{A,t}(0,L)$ at $-t \in (-(t^* + \varepsilon_B), -t^*)$. There are the following four cases:

1. Suppose that each of the two candidates receives an opportunity in the time interval $[-t, -t^*)$. In this case, if candidate A always stays at $\{0, L, H\}$ at $-t$, then $h_A \in H_{A,t^*}(0,L)$ will be realized at $-t^*$ because candidate B stays at $\{L, H\}$. Otherwise, $h_A \in H_{A,t^*}(H,H)$ is realized. Since $v_{A,t^*}(0,L) = v_{A,t^*}(H,L) > v_{A,t^*}(H,H)$, she strictly prefers staying at $\{0, L, H\}$.

2. Suppose only candidate B has an opportunity in the time interval $[-t, -t^*)$. In this case, if candidate A stays at $\{0, L, H\}$ at $-t$, $h_A \in H_{A,t^*}(0,L)$ will be realized at $-t^*$ because candidate B stays at $\{L, H\}$. Otherwise, $h_A \in H_{A,t^*}(H,H)$ is realized. Again, she strictly prefers staying at $\{0, L, H\}$.

3. Suppose only candidate A has an opportunity in the time interval $[-t, -t^*)$. In this case, if candidate A always stays at $\{0, L, H\}$, then $h_A \in H_{A,t^*}(0,L)$ will be realized at $-t^*$ because candidate B stays at $\{L, H\}$. Otherwise, $h_A \in H_{A,t^*}(H,L)$ is realized. Since $v_{A,t^*}(0,L) = v_{A,t^*}(H,L)$, she weakly prefers staying at $\{0, L, H\}$ all the time.

4. Suppose no candidate receives an opportunity in the time interval $[-t, -t^*)$. In this case, if candidate A stays at $\{0, L, H\}$ at $-t$, then $h_A \in H_{A,t^*}(0,L)$ will be realized at $-t^*$. Otherwise, $h_A \in H_{A,t^*}(H,L)$ is realized. Since $v_{A,t^*}(0,L) = v_{A,t^*}(H,L)$, she weakly prefers staying at $\{0, L, H\}$.

In total, candidate A strictly prefers to choose $\{0, L, H\}$ for each $h_A \in H_{A,t}(0,L)$ at $-t \in (-(t^* + \varepsilon_B), -t^*)$. ■

Given this lemma, we can show the following result:

Lemma 27 *Under σ, for each $-t < -t^*$, candidate A stays at $\{0, L, H\}$ for $h_A \in H_{A,t}(0,L)$ and candidate B stays at $\{L, H\}$ for $h_B \in H_{B,t}(0,L)$. Hence, for $-t < -t^*$, we have*

$$v_{A,t}(h_A) = v_{A,t^*}(0,L) \text{ for } h_A \in H_{A,t}(0,L);$$

$$v_{B,t}(h_B) = v_{B,t^*}(0,L) \text{ for } h_B \in H_{B,t}(0,L).$$

Together with (38), defining

$$\tilde{v}_{i,t}(0,L) := \begin{cases} v_{i,t}(0,L) & \text{for } t \in [0,t^*] \\ v_{i,t^*}(0,L) & \text{for } t \in [t^*,\infty) \end{cases},$$

we have

$$v_{i,t}(h_i) = \tilde{v}_{i,t}(0,L) \text{ for each } i = A, B,\ t \in [0,\infty), \text{ and } h_i \in H_{i,t}(0,L). \tag{43}$$

Proof. Given Lemma 26, it suffices to prove the following claim: For each $-t < -t^*$, given the continuation strategy that candidate A stays at $\{0, L, H\}$ for $h_A \in H_{A,t'}(0,L)$ with $-t' \in (-t,-t^*)$ and candidate B stays at $\{L, H\}$ for $h_B \in H_{B,t'}(0,L)$ with $-t' \in (-t,-t^*)$, candidate A stays at $\{0, L, H\}$ for $h_A \in H_{A,t}(0,L)$ and candidate B stays at $\{L, H\}$ for $h_B \in H_{B,t}(0,L)$. To see why this claim is true, for candidate A, consider the following four scenarios: Since $v_{A,t}(H,L) > v_{A,t}(L,L)$ for each t, she strictly prefers $\{H\}$ to $\{0, L\}$ under the given history. Hence, below we compare $\{0, L, H\}$ and $\{H\}$.

1. Suppose each of the two candidates receives an opportunity in the time interval $(-t,-t^*)$. In this case, if candidate A stays at $\{0, L, H\}$ at $-t$, then $h_A \in H_{A,t^*}(0,L)$ will be realized at $-t^*$. Otherwise, $h_A \in H_{A,t^*}(H,H)$ is realized. Since $v_{A,t^*}(0,L) = v_{A,t^*}(H,L) > v_{A,t^*}(H,H)$, she strictly prefers staying at $\{0, L, H\}$.

2. Suppose only candidate B has an opportunity in the time interval $(-t,-t^*)$. In this case, if candidate A stays at $\{0, L, H\}$ at $-t$, $h_A \in H_{A,t^*}(0,L)$ will be realized at $-t^*$. Otherwise, $h_A \in H_{A,t^*}(H,H)$ is realized. Again, she strictly prefers staying at $\{0, L, H\}$.

3. Suppose only candidate A has an opportunity in the time interval $(-t,-t^*)$. In this case, if candidate A always stays at $\{0, L, H\}$, then $h_A \in H_{A,t^*}(0,L)$ will be realized at $-t^*$. Otherwise, $h_A \in H_{A,t^*}(H,L)$ is realized. Since $v_{A,t^*}(0,L) = v_{A,t^*}(H,L)$, she weakly prefers staying at $\{0, L, H\}$.

4. Suppose no candidate receives an opportunity in the time interval $(-t,-t^*)$. In this case, if candidate A stays at $\{0, L, H\}$ at $-t$, then $h_A \in H_{A,t^*}(0,L)$ will be realized at $-t^*$. Otherwise, $h_A \in H_{A,t^*}(H,L)$ is realized. Since $v_{A,t^*}(0,L) = v_{A,t^*}(H,L)$, she weakly prefers staying at $\{0, L, H\}$.

In total, staying at $\{0, L, H\}$ is uniquely optimal, as desired.

Similarly, for candidate B, consider the following four scenarios:

1. Suppose each of the two candidates has an opportunity in the time interval $(-t, -t^*)$. In this case, if candidate B stays at $\{L, H\}$ at $-t$, then $h_B \in H_{B,t^*}(0, L)$ will be realized at $-t^*$. Otherwise, $h_B \in H_{B,t^*}(H, H)$ is realized. Since $v_{B,t^*}(0, H) > v_{B,t^*}(H, H)$ and $v_{B,t^*}(0, L) > v_{B,t^*}(0, H)$ at $-t^*$ (recall that $t^* < t^{**}$ and $v_{B,\tau}(0, L) > v_{B,\tau}(0, H)$ if $\tau < t^{**}$, and hence $v_{B,t^*}(0, L) > v_{B,t^*}(0, H)$), she strictly prefers staying at $\{L, H\}$.

2. Suppose only candidate A has an opportunity in the time interval $(-t, -t^*)$. In this case, if candidate B stays at $\{L, H\}$ at $-t$, then $h_B \in H_{B,t^*}(0, L)$ will be realized at $-t^*$. Otherwise, $h_B \in H_{B,t^*}(H, H)$ is realized. Again, she strictly prefers staying at $\{L, H\}$.

3. Suppose only candidate B has an opportunity in the time interval $(-t, -t^*)$. In this case, if candidate B always stays at $\{L, H\}$ at $-t$, then $h_B \in H_{B,t^*}(0, L)$ will be realized at $-t^*$. Otherwise, $h_B \in H_{B,t^*}(0, H)$ is realized. She strictly prefers staying at $\{L, H\}$ due to (42) and $t^* < t^{**}$.

4. Suppose no candidate receives an opportunity in the time interval $(-t, -t^*)$. In this case, if candidate B stays at $\{L, H\}$ at $-t$, then $h_B \in H_{B,t^*}(0, L)$ will be realized at $-t^*$. Otherwise, $h_B \in H_{B,t^*}(0, H)$ is realized. She strictly prefers staying at $\{L, H\}$ due to (42).

In total, staying at $\{L, H\}$ is uniquely optimal. ■

Finally, consider $h_A \in H_{A,t}(0, 0)$ and $h_B \in H_{B,t}(0, 0)$:

Lemma 28 *Under σ, for each $-t$, candidates A and B take $\{L, H\}$ for $h_A \in H_{A,t}(0, 0)$ and $h_B \in H_{B,t}(0, 0)$, respectively.*

Proof. Let σ'' be a strategy profile such that, for each candidate $i = A, B$, if the history for i is h_i, i chooses $\{L, H\}$ if $h_i \in H_{i,t}(0, 0)$, and follow σ if $h_i \notin H_{i,t}(0, 0)$. Given the original β, for each $i = A, B$, the candidates' payoffs under σ'' and β can be written as:

$$v_{i,t}(0,0) := \begin{cases} \frac{1}{2}\alpha + (1-\alpha)(H + (H-L)\lambda t)\, e^{-\lambda t} & \text{for } t \in [0, t^*] \\ \frac{1}{2}\alpha + (1-\alpha)\, e^{-1}\left[2H - L + \left(1 - e^{-2(\lambda t - 1)}\right)(H-L)\right] & \text{for } t \in [t^*, \infty) \end{cases}$$

for each $h_i \in H_i(0,0)$. Here, $v_{i,t}(0,0)$ is calculated as follows: For each $t \leq t^*$,

$$v_{i,t}(0,0) = \int_0^t 2\lambda e^{-2\lambda\tau} \frac{v_{A,\tau}(L,0) + V_{A,\tau}(0,L)}{2} d\tau + e^{-2\lambda t}\left(\frac{\alpha}{2} + (1-\alpha) H\right).$$

For $t > t^*$, by Lemma 27, we have

$$v_{A,t}(0,0) = e^{-2\lambda(t-t^*)} v_{A,t^*}(0,0) + \left(1 - e^{-2\lambda(t-t^*)}\right)\left[\frac{1}{2} v_{A,t^*}(L,0) + \frac{1}{2} v_{A,t^*}(0,L)\right].$$

As in part 1, there exists $\bar{t} > 0$ such that, for each $t \in [0, \bar{t})$, at any histories in $H_{A,t}(0,0)$ and $H_{B,t}(0,0)$, σ and σ'' coincide. Together with Lemma 20 and (43), both candidates A and B take $\{L, H\}$ at histories $h_A \in H_{A,t}(0,0)$ and $h_B \in H_{B,t}(0,0)$ under σ if

$$\begin{aligned} \tilde{v}_{A,t}(L,0) &> \max\{v_{A,t}(H,0), v_{A,t}(0,0)\}, \text{ and} \\ \tilde{v}_{B,t}(0,L) &> \max\{v_{B,t}(0,H), v_{B,t}(0,0)\}. \end{aligned}$$

By symmetry, we focus on $\tilde{v}_{B,t}(0,L) > \max\{v_{B,t}(0,H), v_{B,t}(0,0)\}$. By Lemma 27, (40), and $t^* < t^{**}$, we have $\tilde{v}_{B,t}(0,L) > v_{B,t}(0,H)$ for each $t \geq 0$. Hence we are left to show that, for each $t \geq 0$, $\tilde{v}_{B,t}(0,L) > v_{B,t}(0,0)$.

For $-t \geq -t^*$, since $\tilde{v}_{B,t}(0,L) = v_{B,t}(0,0)$, the same proof as (41) implies that candidate B strictly prefers $\{L, H\}$ to $\{0, L, H\}$ for $h_B \in H_{B,t}(0,0)$.

Moreover, $v_{B,t}(0,0)$ is strictly increasing in t and we have

$$v_{B,t}(0,0) < \lim_{t\to\infty} v_{B,t}(0,0) = \frac{1}{2}\alpha + (1-\alpha) e^{-1}(2H - L).$$

Thus, for each $-t < -t^*$, we have

$$\begin{aligned} &\tilde{v}_{B,t}(0,L) - v_{B,t}(0,0) \\ &> v_{B,t^*}(0,L) - \lim_{t\to\infty} v_{B,t}(0,0) \\ &= \frac{L(2\alpha(L+1) - (H+L)(1-\alpha))}{H+L} e^{-1} \geq 0. \end{aligned}$$

Therefore, candidates B and A (by symmetry) prefer $\{L, H\}$ to $\{0, L, H\}$ for each time $-t$ and for each $h_B \in H_{B,t}(0,0)$ and $h_A \in H_{A,t}(0,0)$, respectively. ■

M Proofs for Section 3.5

M.1 Proof of Proposition 7

[The "if" part] We will show that the following is a PBE: each candidate i takes some strategy σ_i^* that satisfies the properties described in the statement of the proposition. After the opponent has entered, then σ_i^* is optimal given (4). Hence, we focus on the history in which the opponent has not entered. We consider the continuation payoff of entering and not entering at each $-t$ (given that each candidate follows σ^* in the continuation play). By symmetry, we only consider candidate R's incentive.

Under (σ_i^*, σ_j^*), the payoff from entering at $-t$ is

$$\begin{aligned} \max_x \left\{-x^R + \exp(-\lambda t)\, x + (1-\exp(-\lambda t))(-x)\right\} &= \max_x \left\{-x^R - (1-2\exp(-\lambda t))\, x\right\} \\ &= \begin{cases} -x^R & \text{if } t \geq \frac{1}{\lambda}\ln 2 \\ -x^R - (1-2\exp(-\lambda t))\, x^R & \text{if } t \leq \frac{1}{\lambda}\ln 2 \end{cases}. \end{aligned}$$

Note that the set of maximizers is $\{0\}$ if $t < \frac{1}{\lambda}\ln 2$, $[0, x^R]$ if $t = \frac{1}{\lambda}\ln 2$, and $\{x^R\}$ if $t > \frac{1}{\lambda}\ln 2$.

In contrast, the continuation payoff of not entering at $-t$ under (σ_i^*, σ_j^*) is $-x^R$ because the following three cases are exhaustive:

1. If candidate L enters next by $-\frac{1}{\lambda}\ln 2$, then R obtains a payoff of $-x^R$ since L enters at 0.

2. If candidate R enters next by $-\frac{1}{\lambda}\ln 2$, then R obtains a payoff of $-x^R$ since R enters at 0.

3. If no candidate enters by $-\frac{1}{\lambda}\ln 2$, then R obtains a payoff of $-x^R$ since, under σ^*, the candidates take symmetric strategies for $t \in \left(-\frac{1}{\lambda}\ln 2, 0\right]$ and they never enter at a policy not in $[-x^R, x^R]$.[29]

Hence, for each $-t < -\frac{1}{\lambda}\ln 2$, both entering at 0 and not entering are optimal; for $-t = -\frac{1}{\lambda}\ln 2$, entering at any policy in $[0, x^R]$ and not entering are optimal; and for $-t > -\frac{1}{\lambda}\ln 2$, entering at x^i is optimal. Hence, (σ_i^*, σ_j^*) is a PBE.

[29] If the current time $-t$ satisfies $t \leq \frac{1}{\lambda}\ln 2$, then this is the only case that happens.

[The "only if" part] First we show that, under the assumption of $p = 1$, (v_A, v_B) is constant-sum. To see this, note that, under any PBE, at any history, each candidate enters at a policy in $[-x^R, x^R]$. This is because, if candidate i enters outside of this interval when the opponent j has not entered, then i will lose for sure since the median voter will prefer j's ideal policy than i's committed policy, and j's best responses are not to enter and to enter at x^j. Candidate i can do strictly better by not entering, which with some strictly positive probability leads i to implement x^i. Therefore, the implemented policy is included in $[-x^R, x^R]$. Restricting our attention to $x \in [-x^R, x^R]$, the game is constant-sum because the sum of two candidates' payoffs when the implemented policy is x is $-\left|x + x^R\right| - \left|x - x^R\right| = -2x^R$.

Now, Theorem 3 implies that, at any time $-t$ under any PBE, the continuation payoff of not entering under (σ_i^*, σ_j^*) that we computed in the "if" part is the continuation payoff of entering, and $-x^R$ is the continuation payoff of not entering. Hence, in any PBE, for each $-t < -\frac{1}{\lambda}\ln 2$, both entering at 0 and not entering are the only optimal actions; for $-t = -\frac{1}{\lambda}\ln 2$, entering at any policy in $[0, x^R]$ and not entering are the only optimal actions; and for $-t > -\frac{1}{\lambda}\ln 2$, entering at x^i is uniquely optimal. Hence, any PBE satisfies the conditions given in the statement of the proposition.

M.2 Proof of Lemma 3

We prove Lemma 3 as well as compute the payoff $v_i(\tilde{p}, \text{enter})$.

Since entering at $x = 0$ ensures that the winning policy is 0, candidate R, when she enters, enters at $x \geq 0$. In addition, given the best response function in (4), we can show that it is suboptimal for R to enter at x with $\tilde{p}x^R < x$:

Claim 1 *For each $x > \tilde{p}x^R$, we have $x \notin \arg\max_{\tilde{x}} v_t(\tilde{p}, \tilde{x})$.*

Proof. Suppose that R receives an opportunity at time $-t$ and enters at x, while L still has not entered. Note that, conditional on this event, if L does not enter until time 0, the median voter votes for candidate R if and only if it is better to vote for candidate R with known policy commitment x than unknown type L:

$$p_0^L\left(-\left|x^L\right|\right) + \left(1 - p_0^L\right)(-0) \leq -x,$$

where p_0^L is the voters' posterior about L being extreme at $-t = 0$. Since in equilibrium each type of candidate L will enter after candidate R enters, if candidate L does not enter, then the voters know that L did not receive an opportunity. Thus, $p_0^L = p_t^L$. Hence, candidate R does not win with a policy $x \geq 0$ with $\tilde{p}x^R < x$.

Therefore, R's payoff from entering at such x is

$$\begin{aligned} &\underbrace{\left(1-e^{-\lambda t}\right)}_{\Pr(L \text{ has an opportunity})} \times \underbrace{\left[(1-\tilde{p})\left(-x^R\right)+\tilde{p}\left(-\left|BR_L(x)-x^R\right|\right)\right]}_{R\text{'s utility when } L \text{ has an opportunity after } R \text{ enters at } x} \\ +\; &\underbrace{e^{-\lambda t}}_{\Pr(L \text{ does not have an opportunity})} \times \underbrace{\left[\tilde{p}\left(-x^R-x^R\right)+(1-\tilde{p})\left(-x^R\right)\right]}_{\substack{R\text{'s utility when } L \text{ does not enter after } R \text{ enters at } x \\ \text{(note that we consider the case in which } L \text{ wins)}}} . \end{aligned}$$

Note that

$$0 \in \arg\max_x \left[(1-\tilde{p})\left(-x^R\right)+\tilde{p}\left(-\left|BR_L(x)-x^R\right|\right)\right]$$

since, once L has an opportunity after R enters at $x > 0$, he enters at 0 if he is normal while he enters at a best response to x given by (4) if he is extreme. Since

$$(1-\tilde{p})\left(-x^R\right)+\tilde{p}\left(-2x^R\right) < -x^R,$$

R's payoff from entering at 0 is:

$$\begin{aligned} &\underbrace{\left(1-e^{-\lambda t}\right)}_{\Pr(L \text{ has an opportunity})} \times \underbrace{\left[(1-\tilde{p})\left(-x^R\right)+\tilde{p}\left(-x^R\right)\right]}_{R\text{'s utility when } L \text{ has an opportunity after } R \text{ enters at } 0} \\ +\; &\underbrace{e^{-\lambda t}}_{\Pr(L \text{ does not have an opportunity})} \times \underbrace{\left(-x^R\right)}_{\text{By entering at 0, } R \text{ ensure } -x^R\text{. Note that the winning policy is guaranteed to be 0}} \\ >\; &\underbrace{\left(1-e^{-\lambda t}\right)}_{\Pr(L \text{ has an opportunity})} \times \underbrace{\left[(1-\tilde{p})\left(-x^R\right)+\tilde{p}\left(-x^R\right)\right]}_{\substack{\text{An upper bound of } R\text{'s utility when } L \text{ has an opportunity after } R \text{ enters at } x \\ \text{since } L \text{ enters at } BR_L(x)\leq 0}} \\ +\; &\underbrace{e^{-\lambda t}}_{\Pr(L \text{ does not have an opportunity})} \times \underbrace{\left[\tilde{p}\left(-x^R-x^R\right)+(1-\tilde{p})\left(-x^R\right)\right]}_{\substack{R\text{'s utility when } L \text{ does not enter after } R \text{ enters at } x \\ \text{(note that we consider the case in which } L \text{ wins)}}} , \end{aligned}$$

which is the payoff of entering at x. Hence, it is better to enter at 0 rather than x with $\tilde{p}x^R < x$.

■

Given this claim, we focus on R's entering at $x \in \left[0, \tilde{p}x^R\right]$. When candidate R enters at $x \in \left[0, \tilde{p}x^R\right]$ at $-t$, the following two cases can happen:

1. Candidate L receives an opportunity in the time interval $(-t, 0]$.

   (a) If he is extreme, then he will enter at $-x$ and candidate R receives $-x - x^R$. This happens with probability $\tilde{p}\left(1 - e^{-\lambda t}\right)$.

   (b) If he is normal, then he will enter at 0 and candidate R receives $-x^R$. This happens with probability $(1-\tilde{p})\left(1 - e^{-\lambda t}\right)$.

2. Candidate L does not receive any opportunity in the time interval $(-t, 0]$. This happens with probability $e^{-\lambda t}$. In such a case, the median voter votes for candidate R given $x \in \left[0, \tilde{p}x^R\right]$.

Therefore, the payoff from entering at x at time $-t$ is

$$\begin{aligned} &\tilde{p}\left(1 - e^{-\lambda t}\right)\left(-x - x^R\right) + (1-\tilde{p})\left(1 - e^{-\lambda t}\right)\left(-x^R\right) - e^{-\lambda t}\left(x^R - x\right) \\ &= -x^R + \left(e^{-\lambda t} - \tilde{p}\left(1 - e^{-\lambda t}\right)\right)x. \end{aligned} \tag{44}$$

If $e^{-\lambda t} > \tilde{p}\left(1 - e^{-\lambda t}\right)$, then R wants to maximize x, so $x = \tilde{p}x^R$ is uniquely optimal; if $e^{-\lambda t} < \tilde{p}\left(1 - e^{-\lambda t}\right)$, then R wants to minimize x, so $x = 0$ is uniquely optimal; if $e^{-\lambda t} = \tilde{p}\left(1 - e^{-\lambda t}\right)$, then R is indifferent among all $x \in \left[0, \tilde{p}x^R\right]$.

M.3 Proof of Lemma 2

Suppose that normal candidate i has an opportunity at $-t$ when the opponent j has not entered. Since there is a positive probability that j is extreme and he has not received any opportunity, for any strategy of j, there is a positive probability that he is extreme.

If i enters at 0, then she obtains a payoff of 0, which is the maximum feasible payoff of this game. If she enters at $x \neq 0$, then her payoff is strictly less than 0 since, if j is extreme and obtains an opportunity in $-\tau \in (-t, 0]$,then he will enter at a policy that is not 0 and win. If she does not enter, then again her payoff is strictly less than 0 since, if no candidate obtains an opportunity in $-\tau \in (-t, 0]$, then j will win with probability $\frac{1}{2}$ and implement her ideal policy which may not be 0. In total, entering at 0 is the unique optimal strategy for i.

M.4 Proof of Lemma 4

Given $\frac{e^{-\lambda t}}{1-e^{-\lambda t}} \leq p(t)$, Lemma 3 implies that R's payoff from entering is $-x^R$. Hence, we are left to show that R's payoff from not entering is greater than $-x^R$.

Since we focus on symmetric PBE and the implemented policy is in $\left[-x^R, x^R\right]$, if L is extreme, then the extreme candidate R obtains $-x^R$ by symmetry. Hence R's payoff from not entering at $-t$ given $p(t) = \tilde{p}$ is given by

$$\begin{aligned} v_t(\tilde{p}, \text{not}) &= \underbrace{\tilde{p}}_{L \text{ is extreme}} \left(-x^R\right) \\ &\quad + \underbrace{(1-\tilde{p})}_{L \text{ is normal}} \left(\begin{array}{c} \underbrace{\left(1-e^{-\lambda t}\right)}_{L \text{ has an opportunity}} \left(-x^R\right) \\ + \underbrace{e^{-\lambda t}}_{L \text{ does not have an opportunity}} \left(\begin{array}{c} q_t \tilde{v}^R \\ + (1-q_t)\left(-\frac{1}{2}x^R\right) \end{array} \right) \end{array} \right), \end{aligned}$$

where q_t is the conditional probability that R enters at some time in $(-t, 0]$ given that L is normal and does not have an opportunity to enter in the time interval $[-t, 0]$, given the equilibrium strategy σ, and $\tilde{v}^R$ is candidate R's expected payoff conditional on the event that R enters at some timing $(-t, 0]$ and L is normal and does not have an opportunity to enter in the time interval $(-t, 0]$.

Since R enters in $\left[0, x^R\right]$ if she enters, we have $\tilde{v}^R \geq -x^R$. Moreover, $1 - q_t \geq e^{-\lambda t}$ (note that $e^{-\lambda t}$ is the probability that R does not obtain any opportunity in $(-t, 0]$) under any σ, we have $v_t(\tilde{p}, \text{not}) > -x^R$.

M.5 Proof of Proposition 8

For each p and T, fix a symmetric PBE. In the following proof, we consider properties of this symmetric PBE. In fact, there exists a unique strategy profile satisfying such properties, and one can show that it is indeed a PBE, which shows the existence of a PBE.

M.5.1 Equilibrium behavior near the deadline

If the deadline is close, then the probability that candidate L will receive an opportunity is close to zero. By (44), if candidate R enters, then she receives

$$\begin{aligned}&-x^R+\max\left\{\left(e^{-\lambda t}-\tilde{p}\left(1-e^{-\lambda t}\right)\right)\tilde{p}x^R,0\right\}\Big|_{t=0}\\&=-\left(1-\tilde{p}\right)x^R,\end{aligned}$$

where $\tilde{p}$ is R's posterior probability with which L is extreme.

If candidate R does not enter, then since she wins with probability $\frac{1}{2}$ at $-t=0$, she receives

$$\begin{aligned}&\frac{1}{2}\left\{\tilde{p}\left(-x^R-x^R\right)+\left(1-\tilde{p}\right)\left(-x^R\right)\right\}+\frac{1}{2}\times 0\\&=-\left(1-\tilde{p}\right)x^R+\frac{1}{2}\left(1-3\tilde{p}\right)x^R.\end{aligned}$$

Therefore, candidate R enters at x with $x=\tilde{p}x^R$ if her posterior $\tilde{p}$ about L being extreme is greater than $\frac{1}{3}$. She does not enter if the belief is less than $\frac{1}{3}$ at $-t=0$. Intuitively, if the probability that candidate L is extreme is very high, then (i) R's payoff when L wins when L has not entered at time 0 is low since L will pick x^L after the election with a higher probability, and (ii) the median voter votes for R even if R takes a large $x<x^R$. Hence, it is more attractive to enter $x>0$ so that R can win.

For each p and λ, for sufficient large T, we have $\frac{e^{-\lambda t}}{1-e^{-\lambda t}}\leq p\leq p(t)$ for $-t\in\left[-\frac{T}{2},-T\right]$ and $p\left(\frac{T}{2}\right)>\frac{1}{3}$. Since $p(0)\geq p\left(\frac{T}{2}\right)>\frac{1}{3}$, candidate R enters at x with $x=p(t)\,x^R$ for sufficiently small t.

M.5.2 Backward Induction

Suppose that candidates enter in the time interval $(-t,0]$. Then we have $p(\tau)=p(t)$ for each $-\tau\in[-t,0]$. Hence, if $p(t)=\tilde{p}$ and the candidates enter in the time interval $(-t,0]$, then R's payoff from entering at $-t$ is

$$-x^R+\left(e^{-\lambda t}-\tilde{p}\left(1-e^{-\lambda t}\right)\right)\tilde{p}x^R,$$

while the payoff from not entering at $-t$ is

$$\begin{aligned} v_t(\tilde{p}, \text{not}) &= \underbrace{\tilde{p}}_{L \text{ is extreme}} (-x^R) \\ &+ \underbrace{(1-\tilde{p})}_{L \text{ is normal}} \left(\begin{array}{c} \underbrace{\left(1-e^{-\lambda t}\right)}_{L \text{ has an opportunity}} (-x^R) \\ + \underbrace{e^{-\lambda t}}_{L \text{ does not have an opportunity}} \left(\underbrace{\left(1-e^{-\lambda t}\right)}_{R \text{ enters at } \tilde{p}x^R} \left(\tilde{p}x^R - x^R\right) + e^{-\lambda t}\left(-\tfrac{1}{2}x^R\right) \right) \end{array} \right) \\ &= -x^R + (1-\tilde{p})\left(e^{-\lambda t}\tilde{p}x^R - e^{-2\lambda t}\left(1-\tilde{p}+\frac{1}{2}\right)x^R\right). \end{aligned}$$

Candidate R enters if

$$-x^R + \left(e^{-\lambda t} - \tilde{p}\left(1-e^{-\lambda t}\right)\right)\tilde{p}x^R \geq -x^R + (1-\tilde{p})\left(e^{-\lambda t}\tilde{p}x^R - e^{-2\lambda t}\left(1-\tilde{p}+\frac{1}{2}\right)x^R\right)$$

$$\Leftrightarrow \quad e^{-\lambda t} = 2\frac{\tilde{p} - \sqrt{\frac{4\tilde{p}^2+3-5\tilde{p}}{2}}}{5\tilde{p}-2\tilde{p}^2-3}\tilde{p}, \tag{45}$$

where the right-hand side is well defined for any $\tilde{p} \in (0,1)$ because $\frac{4\tilde{p}^2+3-5\tilde{p}}{2} \geq \frac{23}{32}$ and $5\tilde{p}-2\tilde{p}^2-3 < 0$ for $\tilde{p} \in (0,1)$.

Note that, if t satisfies (45), we have $\frac{e^{-\lambda t}}{1-e^{-\lambda t}} - \tilde{p} \geq 0$ (and this holds with equality if and only if $\tilde{p} = 0$).

For each p, fix the smallest t such that candidate R weakly prefers not entering at $-t$ given p, and denote it by $t^*(p)$:

$$e^{-\lambda t^*(p)} = 2\frac{p - \sqrt{\frac{4p^2+3-5p}{2}}}{5p-2p^2-3}p. \tag{46}$$

Then, for each p, $-t < -t^*(p)$ and each h_R^t such that $p(t) = p$ and no candidate has entered at $-t$, candidate R's unique best reply is not to enter. To see why, consider the following two scenarios:

1. If L has an opportunity by $-t^*(p)$, then R's action does not affect R's payoff if L is normal. If L is extreme, R is worse off by entering at px than not entering. In particular, if R enters, then her payoff is $-x^R - px^R$. Suppose next that she does not enter until $-t^*(p)$. Then, if L has not entered by $-t^*(p)$, her payoff is $-x^R$ by symmetry. If he enters, in contrast, her

payoff is $-x^R + (1-e^{-\lambda\tilde{t}})p\left(\tilde{t}\right)x^R - e^{-\lambda\tilde{t}}p\left(\tilde{t}\right)x^R$, where L enters at $-\tilde{t} \in (-t, -t^*(p)]$. Note that

$$\begin{aligned} -x^R + (1-e^{-\lambda\tilde{t}})p\left(\tilde{t}\right)x^R - e^{-\lambda\tilde{t}}p\left(\tilde{t}\right)x^R &\geq -x^R + p\left(\tilde{t}\right)x^R - 2e^{-\lambda\tilde{t}}p\left(\tilde{t}\right)x^R \\ &\geq -x^R + px^R - 2e^{-\lambda t^*(p)}x^R \\ &> -x^R - px^R. \end{aligned}$$

The last line holds for the following reason: Straightforward algebra shows that $\frac{p-\sqrt{\frac{4p^2+3-5p}{2}}}{5p-2p^2-3} < \frac{1}{4}$, so by (46), we have $2e^{-\lambda t^*(p)} < p$. Hence, R is strictly better off not entering.

2. If L does not have an opportunity until $-t^*(p)$, candidate R is indifferent between entering and not entering (conditional on the event that L does not have an opportunity until $-t^*(p)$, the belief that L is extreme is constant in the time interval $[-t, -t^*(p)]$).

Since there is a positive probability that L is extreme and has an opportunity to enter, it is uniquely optimal for R not to enter at $-t$.

M.5.3 Equilibrium Dynamics

Fix $p \in (0,1)$ arbitrarily. There exists $\bar{T}_1 < \infty$ such that for all $T \geq \bar{T}_1$, we have that at any $-t \in \left[-\frac{T}{2}, -T\right]$, $\frac{e^{-\lambda t}}{1-e^{-\lambda t}} \leq p \leq p\left(t\right)$ holds and so candidate R does not enter. For such T, in any symmetric PBE, $p\left(t\right)$ evolves according to $\frac{d}{dt}p\left(t\right) = -\lambda p\left(t\right)\left(1-p\left(t\right)\right)$ for $t \in [\frac{T}{2}, T]$ with the initial condition $p(T) = p$ since the normal type always enters. Define $\bar{p}_T : [0,T] \to [0,1]$ by the differential equation $\frac{d}{dt}\bar{p}_T\left(t\right) = -\lambda\bar{p}_T\left(t\right)\left(1-\bar{p}_T\left(t\right)\right)$ with the initial condition $\bar{p}_T(T) = p \in (0,1)$.

We will show that there exists $\bar{T} < \infty$ such that for all $T \geq \bar{T}$, there exists a unique $t \in [0,T]$ such that

$$e^{-\lambda t} = 2\frac{\bar{p}_T\left(t\right) - \sqrt{\frac{4\bar{p}_T(t)^2+3-5\bar{p}_T(t)}{2}}}{5\bar{p}_T\left(t\right) - 2\bar{p}_T\left(t\right)^2 - 3}\bar{p}_T\left(t\right). \tag{47}$$

It will be useful to define $f(q) := 2\frac{q-\sqrt{\frac{4q^2+3-5q}{2}}}{5q-2q^2-3}q$ for $q \in [0,1)$, where we note that this is well defined for $q = 0$ as well.

Proof of Existence:

We first show that the right-hand side of (47) is strictly decreasing in t. Since $\bar{p}_T(t)$ is strictly decreasing in t, it suffices to show that $f(q)$ is strictly increasing in $q \in [0,1)$. Note that

$$\frac{d}{dq}f(q) = \frac{\sqrt{2}}{2}\frac{6-5q}{\left(2q^2-5q+3\right)^2\sqrt{4q^2-5q+3}}\left(6q^2-5q-2\sqrt{2}q\sqrt{4q^2-5q+3}+3\right). \tag{48}$$

Since straightforward algebra implies that, for each $q \in [0,1)$,

$$6q^2-5q-2\sqrt{2}q\sqrt{4q^2-5q+3}+3 \in (0,3),$$

$f(q)$ is strictly increasing in $q \in [0,1)$.

Second, we show that there exists $\bar{T}_2 < \infty$ such that, for each $T \geq \bar{T}_2$, $e^{-\lambda T} < f(\bar{p}_T(T))$ holds. This follows since the left-hand side is strictly decreasing in T and converges to 0 as T goes to infinity, while the right-hand side is fixed at $f(p)$ because $\bar{p}_T(T) = p$, and $f(p) > 0$ because f is strictly increasing and $f(0) = 0$.

Third, we show that, at $t = 0$, $e^{-\lambda \times 0} = 1 > f(\bar{p}_T(0))$. This follows since, given that $f(q)$ is strictly increasing in $q \in [0,1)$, an upper bound of the right-hand side is calculated as $\lim_{q\to 1} f(q) = \frac{1}{2}$.

These three observations establish the existence of $t \in [0,T]$ satisfying (47) for each $T \geq \max\left\{\bar{T}_1, \bar{T}_2\right\}$.

Proof of Uniqueness:

Given the existence, for each $T > \max\left\{\bar{T}_1, \bar{T}_2\right\}$, let $t^{**}(T)$ be the largest $t \in [0,T]$ satisfying (47). To show that the solution is unique, it suffices to show that there exists $\bar{T} \in \left[\max\left\{\bar{T}_1, \bar{T}_2\right\}, \infty\right)$ such that for all $T \geq \bar{T}$, for each $t < t^{**}(T)$,

$$\frac{d}{dt}e^{-\lambda t} < \frac{d}{dt}f\left(\bar{p}_T(t)\right).$$

On the one hand, since $t^{**}(T) \leq \frac{T}{2}$ for each $T \geq \max\left\{\bar{T}_1, \bar{T}_2\right\}$, for each $\varepsilon > 0$, there exists $\bar{T}_3 \in \left[\max\left\{\bar{T}_1, \bar{T}_2\right\}, \infty\right)$ such that for all $T \geq \bar{T}_3$, $\left|\bar{p}_T\left(t^{**}(T)\right) - 1\right| < \varepsilon$. Hence, there exists $\bar{T}_4 \in \left[\max\left\{\bar{T}_1, \bar{T}_2\right\}, \infty\right)$ such that for all $T \geq \bar{T}_4$,

$$\left| f\left(\bar{p}_T\left(t^{**}(T)\right)\right) - \lim_{q\to 1} f(q) \right| < \varepsilon.$$

Noting that $\lim_{q\to 1} f(q) = \frac{1}{2}$, this implies that for each $T \geq \bar{T}_4$, we have $e^{-\lambda t} \geq \frac{1}{2} - \varepsilon$ for each

$t \leq t^{**}(T)$ that is a solution to (47). Since $\frac{d}{dt}e^{-\lambda t} = -\lambda e^{-\lambda t}$ and $\frac{d^2}{dt^2}e^{-\lambda t} = \lambda^2 e^{-\lambda t}$, for each $T \geq \bar{T}_4$ and each $t < t^{**}(T)$, we have

$$\frac{d}{dt}e^{-\lambda t} \leq \left.\frac{d}{dt}e^{-\lambda t}\right|_{t=t^{**}(T)} = -\lambda e^{-\lambda t^{**}(T)} \leq -\lambda\left(\frac{1}{2} - \varepsilon\right).$$

On the other hand, the right-hand side of (47) is decreasing in t. Its derivative is

$$\frac{d}{dt}f(\bar{p}_T(T)) = \left(\left.\frac{d}{dq}f(q)\right|_{q=\bar{p}_T(t)}\right)\cdot\left(\frac{d}{dt}\bar{p}_T(t)\right) = -\left(\left.\frac{d}{dq}f(q)\right|_{q=\bar{p}_T(t)}\right)\cdot\lambda\bar{p}_T(t)(1-\bar{p}_T(t)).$$

Given (48), straightforward algebra yields $\lim_{q\to 1}\frac{d}{dq}f(q) = \frac{1}{16} < \infty$. Recall that, for each $\varepsilon > 0$, there exists $\bar{T}_3 < \infty$ such that, for each $T \geq \bar{T}_3$ and each $t \leq t^{**}(T)$, we have $\bar{p}_T(t) \geq 1-\varepsilon$. Hence, for each $\varepsilon > 0$, there exists $\bar{T}_5 < \infty$ such that, for each $T \geq \bar{T}_5$, we have $\frac{d}{dt}f(\bar{p}_T(T)) > -\varepsilon$.

In total, there exists $\bar{T} < \infty$ such that, for each $T \geq \bar{T}$ and each $t < t^{**}(T)$, we have

$$\frac{d}{dt}e^{-\lambda t} < \frac{d}{dt}f(\bar{p}_T(T)).$$

Equilibrium Dynamics:

Since the solution to (47) is unique, together with (46), we have the following: For a sufficiently large T such that the solution for (47) is unique, let $t^*(p,\lambda,T)$ be this unique solution. In any symmetric PBE, for each $-t < -t^*(p,\lambda,T)$, extreme candidates do not enter and $p(t) = \bar{p}_T(t)$; and for each $-t \geq -t^*(p,\lambda,T)$, extreme candidate R (or L) enters at $\bar{p}_T(t^*(p,\lambda,T))x^R$ (or $\bar{p}_T(t^*(p,\lambda,T))x^L$) and $p(t) = \bar{p}_T(t^*(p,\lambda,T))$ for each t with $-t \geq -t^*(p,\lambda,T)$.

Finally, note that, for any $p \in (0,1)$, for sufficiently large $T < \infty$, in any symmetric PBE, extreme candidates do not enter for any $-t \in [-T, \frac{T}{2}]$. Hence, $p(\frac{T}{2}) = \bar{p}_T(\frac{T}{2}) \to 1$ as $T \to \infty$. Since $\lim_{q\to 1} f(q) = \frac{1}{2}$, for any $p \in (0,1)$ and $\lambda > 0$, we have

$$\lim_{T\to\infty}|\lambda t^*(p,\lambda,T) - \ln 2| = 0.$$

N Proofs for Appendix I.4

N.1 Proof of Proposition 14

Note that, in any PBE, S enters and receives a payoff of 1 if S can move after W enters. In addition, by the same proof as the one for Proposition 2, there exists $\bar{t} > 0$ such that for all time $-t \in (-\bar{t}, 0]$, $Q_t = (N, E)$. Below, we consider the transition of Q_t in the following two cases: $\beta \geq \gamma$ and $\beta < \gamma$.

Since we assume that the positions that S and W enter do not depend on the timing of entry, Assumption 1 of Section 4.1 is satisfied. In addition, Assumption 2 is satisfied. Also, Assumption 3 holds because $v_i^{BR_j} < \sup_{x_i \in \mathcal{X}_i} v_i(x_i, X)$ for each i. Finally, first-mover disadvantage for $i = W$ holds. Moreover, since S does not enter and W enters near the deadline in any PBE, we are in Case 3 for Theorem 1.

N.1.1 Case 1: $\beta \geq \gamma$

Fix a PBE. For all $-t$, W does not enter after S enters if $\beta > \gamma$. If $\beta = \gamma$, then W is indifferent. The following analysis goes through when $\beta = \gamma$ regardless of W's strategy after S enters.

First, let us consider S's incentive. At time $-t$, if W has not entered, S's payoff is $1 - \beta$ if S enters; if S does not enter, then her payoff is $1 - (1 - \alpha) \int_0^t e^{-\lambda_S \tau} \lambda_W e^{-\lambda_W (t-\tau)} d\tau$, given $Q_\tau = (N, E)$ for all $-\tau \in (-t, 0)$. Hence, $\hat{t}_S$ in the notation of Section 4.1 is characterized by the smallest t satisfying the following equation.

$$1 - \beta = 1 - (1 - \alpha) \int_0^t e^{-\lambda_S \tau} \lambda_W e^{-\lambda_W (t-\tau)} d\tau.$$

Defining

$$f_S(t) := \begin{cases} \frac{1}{1-r}\left(e^{-\lambda_S t} - e^{-\lambda_W t}\right) - \frac{\beta}{1-\alpha} & \text{if } r \neq 1 \\ \lambda_W t e^{-\lambda_W t} - \frac{\beta}{1-\alpha} & \text{if } r = 1 \end{cases}, \tag{49}$$

$\hat{t}_S$ is the smallest positive solution for $f_S(t) = 0$. Recall that we have defined t_S to be such a solution for $f_S(t) = 0$ in Section I.4. Hence, $\hat{t}_S = t_S$. If there is no solution, then we define $t_S = \infty$. Note that the function f_S is continuous, so the smallest positive solution always exists or there is no solution.

Second, let us consider W's incentive. At time $-t$, if S has not entered, W's payoff is

$(1-\alpha)\, e^{-\lambda_S t}$ if W enters; if W does not enter, then his payoff is $(1-\alpha)\int_0^t e^{-\lambda_S \tau} \lambda e^{-\lambda_W (t-\tau)} d\tau$, given $Q_\tau = (N, E)$ for $-\tau \in (-t, 0)$. Hence, $\hat{t}_W$ is characterized by the smallest $t > 0$ satisfying the following equation.

$$(1-\alpha)\, e^{-\lambda_S t} = (1-\alpha) \int_0^t e^{-\lambda_S \tau} \lambda_W e^{-\lambda_W (t-\tau)} d\tau$$

$\Leftrightarrow$

$$f_W(t) = 0, \text{ where } f_W(t) := \begin{cases} \frac{1}{1-r}\left(e^{-\lambda_S t} - e^{-\lambda_W t}\right) - e^{-\lambda_S t} & \text{if } r \neq 1 \\ \lambda_W t e^{-\lambda_W t} - e^{-\lambda_W t} & \text{if } r = 1 \end{cases}.$$

Recall that we define t_W as the smallest solution for $f_W(t) = 0$ in Section I.4:

$$\begin{cases} \frac{1}{1-r}\left(e^{-\lambda_S t_W} - e^{-\lambda_W t_W}\right) - e^{-\lambda_S t_W} = 0 & \text{if } r \neq 1 \\ \lambda_W t e^{-\lambda_W t_W} - e^{-\lambda_W t_W} = 0 & \text{if } r = 1 \end{cases}. \tag{50}$$

Hence, $\hat{t}_W = t_W$. Since $f_W(t)$ is continuous, negative for sufficiently small $t > 0$, and positive for sufficiently large t, the smallest positive t such that $f_W(t) = 0$ exists.

The transition of Q_t depends on the relationship between t_S and t_W.

Case 1(a): $-t_S < -t_W$. This inequality means that W's cutoff $-t_W$ is closer to the deadline than S's cutoff (if any), $-t_S$. Since Assumption 4 is satisfied, part 1 of Proposition 12 implies that, under the fixed PBE, S does not enter and W enters for $-t \in (-t_W, 0]$, and no candidate enters for $-t \in (-\infty, -t_W)$. Since this argument holds for any PBE, we have the following:

- $Q_t = (N, E)$ for $-t \in (-t_W, 0]$.
- $Q_t = (N, N)$ for $-t \in (-\infty, -t_W)$.

Hence, part 1(a) of Proposition 14 holds.

Case (1)(b): $-t_S > -t_W$. This inequality means that S's cutoff $-t_S$ is closer to the deadline than W's cutoff $-t_W$. By part 2 of Proposition 12, we are in Case 2 for Theorem 1 with $t_0 = t_S + \varepsilon$ for small $\varepsilon > 0$.

Since first-mover disadvantage for W holds, there exists $t_W^* < \infty$ such that $\bar{v}_{W,t_W^*}$ (not) $=$

v_{W,t_W^*}(enter).[30] In particular, since

$$\begin{aligned} v_{W,t_W^*}(\text{enter}) &= (1-\alpha)\, e^{-\lambda_S t}, \\ \bar{v}_{W,t_W^*}(\text{not}) &= \int_0^{t-t_S} e^{-(\lambda_S+\lambda_W)\tau} \left(\lambda_S \beta + \lambda_W (1-\alpha)\, e^{-\lambda_S (t-\tau)}\right) d\tau + e^{-(\lambda_S+\lambda_W)(t-t_S)}\beta, \end{aligned}$$

t_W^* is the smallest $t > 0$ satisfying the following equation.

$$g_W(t) := e^{-(\lambda_S+\lambda_W)(t-t_S)} \left(\frac{1}{1+r}\beta - (1-\alpha)\, e^{-\lambda_S t_S}\right) + \frac{r}{1+r}\beta = 0.$$

Here, we use the fact that S is indifferent between entering and not entering at $-t_S$, which implies that her payoff at $-t_S$ is $1-\beta$ if no candidates have entered by $-t_S$, and thus W's payoff is β if no candidates have entered by $-t_S$.

In contrast, S always prefers entering at $-t < -t_S$ for the following reason. Suppose W enters at all times $-t \in (-t^*, 0]$ and does not enter at all times $-t \in (-\infty, -t^*)$.

For $-t \geq -t^*$, since S (weakly) prefers entering at $-t_S$, if W has not entered by $-t_S$, S's payoff at $-t_S$ is no more than $1-\beta$. Even if S enters by $-t_S$, S gets at most $1-\beta$. That is, W can guarantee β if W does not enter until $-t_S$. The fact that W (strictly) prefers entering implies that W's payoff when W can enter before S is more than β. Therefore, S's payoff when W can enter before S is less than $1-\beta$. In contrast, by entering, S can guarantee a payoff of $1-\beta$. Hence, entering is S's strict best response at $-t \geq -t^*$.

Moreover, since (i) W does not enter before S enters for $-t \in (-\infty, -t_W^*)$ and (ii) W does not enter after S enters, entering is S's strict best response at all times $-t$ (even if W does not enter).

Given the above characterization, under the fixed PBE, S does not enter for $-t \in (-t_S, 0]$ and enters for $-t \in (-\infty, -t_S)$, while W enters for $-t \in (-t_W^*, 0]$ and does not enter for $-t \in (-\infty, -t_W^*)$. Since this argument holds for any PBE, we have the following transition of Q_t:

- $Q_t = (N, E)$ for $-t \in (-t_S, 0]$.
- $Q_t = (E, E)$ for $-t \in (-t_W^*, -t_S)$.
- $Q_t = (E, N)$ for $-t \in (-\infty, -t_W^*)$.

[30] This notation of t_W^* is introduced in t_W^* in Section 4.1.

Hence, part 1(b) of Proposition 14 holds.

N.1.2 Case 2: $\gamma > \beta$

Fix a PBE. For all $-t$, W enters after S enters. The continuation payoff profile when only S enters at $-t$ is given by $(1-\beta_t, \beta_t)$ with

$$\beta_t = \beta + \left(1 - e^{-\lambda_W t}\right)(\gamma - \beta). \tag{51}$$

When we replace β with β_t in (49), the analysis for the case with $\beta \geq \gamma$ implies the following: $\hat{t}_S$ in the notation of Section 4.1 is characterized by the smallest solution for $f_S(t) = 0$, where

$$f_S(t) \equiv \begin{cases} \frac{1}{1-r}\left(e^{-\lambda_S t} - e^{-\lambda_W t}\right) - \frac{\beta_t}{1-\alpha} & \text{if } r \neq 1 \\ \lambda_W t e^{-\lambda_W t} - \frac{\beta_t}{1-\alpha} & \text{if } r = 1 \end{cases}. \tag{52}$$

Recall that we define t_S as the smaller positive solution for $f_S(t) = 0$ in Section I.4. If there is no solution, then we define $t_S = \infty$.

In contrast, $\hat{t}_W$ is characterized by the smallest positive solution for $f_W(t) = 0$, where

$$f_W(t) \equiv \begin{cases} \frac{1}{1-r}\left(e^{-\lambda_S t} - e^{-\lambda_W t}\right) - e^{-\lambda_S t} & \text{if } r \neq 1 \\ \lambda_W t e^{-\lambda_W t} - e^{-\lambda_W t} & \text{if } r = 1 \end{cases}.$$

Recall that we define t_W as the smallest solution for $f_W(t) = 0$ in Section I.4:

$$\begin{cases} \frac{1}{1-r}\left(e^{-\lambda_S t_W} - e^{-\lambda_W t_W}\right) - e^{-\lambda_S t_W} = 0 & \text{if } r \neq 1, \\ \lambda_W t e^{-\lambda_W t_W} - e^{-\lambda_W t_W} = 0 & \text{if } r = 1. \end{cases}$$

Since $f_W(t)$ is continuous, negative for sufficiently small t, and positive for sufficiently large t, there exists the smallest t such that $f_W(t) = 0$.

The equilibrium dynamics depend on the relationship between t_S and t_W.

Case 2(a): $-t_S < -t_W$. By the same proof as Case 1(a), we can show that, under the fixed PBE, S does not enter and W enters for $-t \in (-t_W, 0]$, and no candidate enters for $-t \in (-\infty, -t_W)$. Since this argument holds for any PBE, we have the following transition of Q_t:

- $Q_t = (N, E)$ for $-t \in (-t_W, 0]$.
- $Q_t = (N, N)$ for $-t \in (-\infty, -t_W)$.

Hence, part 2(a) of Proposition 14 holds.

Case 2(b): $-t_S > -t_W$. This inequality means that S's cutoff $-t_S$ is closer to the deadline than W's cutoff $-t_W$. By part 2 of Proposition 12, there exists $\bar{\varepsilon} > 0$ such that for all $\varepsilon \in (0, \bar{\varepsilon})$, we are in Case 2 for Theorem 1 with $t_0 = t_S + \varepsilon$. Moreover, since first-mover disadvantage for $i = W$ holds, Proposition 10 pins down the dynamics under the fixed PBE, and since the argument holds for any PBE, we have that the transition of Q_t can be one of the following:

Case 2(b)(i)

- $Q_t = (N, E)$ for $-t \in (-t_S, 0]$.
- $Q_t = (E, E)$ for $-t \in (-t_W^*, -t_S)$.[31]
- $Q_t = (E, N)$ for $-t \in (-\infty, -t_W^*)$.

Case 2(b)(ii)

- $Q_t = (N, E)$ for $-t \in (-t_S, 0]$.
- $Q_t = (E, E)$ for $-t \in (-t_W^*, -t_S)$.
- $Q_t = (E, N)$ for $-t \in (-t_S^{**}, -t_W^*)$.
- $Q_t = (N, N)$ for $-t \in (-\infty, -t_S^{**})$.

Case 2(b)(iii)

- $Q_t = (N, E)$ for $-t \in (-t_S, 0]$.
- $Q_t = (E, E)$ for $-t \in (-t_S^*, -t_S)$.[32]
- $Q_t = (N, E)$ for $-t \in (-t_W^{**}, -t_S^*)$.

[31] This notation of t_W^* is introduced in Section 4.1.
[32] This notation of t_S^* is introduced in Section 4.1.

- $Q_t = (N, N)$ for $-t \in (-\infty, -t_W^{**})$.

We can show that there exists $\bar{r} \leq 1$ such that Case 2(b)(i) is not the case if and only if $r \geq \bar{r}$. To see why, we derive the differential equation that characterizes the transition. Let x_t be W's continuation payoff at time $-t$ when W has entered and S has not entered at $-t$; let y_t be W's continuation payoff at time $-t$ when W has not entered and S has entered at $-t$; and let z_t be W's continuation payoff at time $-t$ when no candidate has entered at $-t$.

Suppose x_t, y_t, and z_t satisfy the following differential equations:

$$\frac{dx_t}{dt} = \lambda_S (0 - x_t), \tag{53}$$

$$\frac{dy_t}{dt} = \lambda_W \max \{\gamma - y_t, 0\}, \tag{54}$$

$$\frac{dz_t}{dt} = \lambda_W \max \{x_t - z_t, 0\} + \lambda_S \min \{y_t - z_t, 0\}, \tag{55}$$

with the following condition:

$$x_0 = 1 - \alpha, y_0 = \beta, z_0 = 0.$$

Since this system of ordinary differential equations satisfies Lipschitz continuity, there exists a solution. Such a solution is equilibrium payoffs for the following reasons: Equation (53) means that whenever S can enter after W enters, W's payoff is 0. Equation (54) means that when W can enter after S enters, W enters if and only if his payoff for entering, γ, is greater than the payoff for not entering, y_t. In addition, the first term of (55) means that when W can enter, W enters if and only if his payoff for entering, x_t, is greater than the payoff for not entering, z_t. The second term of (55) means that when S can enter, S enters if and only if her payoff for entering, $1 - y_t$, is greater than her payoff for not entering, $1 - z_t$ (that is, y_t is smaller than z_t). Since we have shown the uniqueness of the value function in Proposition 10, the solution for the system of (53), (54), and (55) is the unique equilibrium payoffs.

To show that there exists $\bar{r} \leq 1$ such that $r \geq \bar{r}$ if and only if there exists $t_S^{**} \in (t_S, \infty)$ such that S does not enter for all $-t \in (-\infty, -t_S^{**})$, we prove the following three claims:

1. [$\bar{r} \leq 1$] For $r \geq 1$, there exists $t_S^{**} \in (t_S, \infty)$ such that S does not enter for all $-t \in (-\infty, -t_S^{**})$.

2. [cutoff from below] If there does not exist $t_S^{**} \in (t_S, \infty)$ such that S does not enter for all

$-t \in (-\infty, -t_S^{**})$ for (λ_S, λ_W), then such t_S^{**} does not exist for (λ_S', λ_W) with $\lambda_S' < \lambda_S$.

3. [cutoff from above] If there exists $t_S^{**} \in (t_S, \infty)$ such that S does not enter for all $-t \in (-\infty, -t_S^{**})$ for (λ_S, λ_W), then such t_S^{**} exists for (λ_S', λ_W) with $\lambda_S' > \lambda_S$.

[Proof of "$\bar{r} \leq 1$"] To analyze the conditions under which there exists $t_S^{**} \in (t_S, \infty)$ such that S does not enter for all $-t \in (-\infty, -t_S^{**})$, let us consider a sufficiently large $t \geq t_W^{**}$. Since $x_t \geq z_t$ if and only if $t \leq t_W^{**}$, as long as $y_t \leq z_t$, we have $\frac{dz_t}{dt} = \lambda_S (y_t - z_t)$. Since $\frac{dz_t}{dt} + \lambda_S z_t = \lambda_S y_t$, we have

$$e^{\lambda_S t} z_t = C + \int_a^t e^{\lambda_S \tau} \lambda_s y_\tau d\tau, \tag{56}$$

where a is the supremum of τ with $x_\tau \geq z_\tau$, and C is determined by the condition $x_a = z_a$. As we have shown above, a is finite, and so is C.

To show that we have $z_t < y_t$ for sufficiently large t for each (λ_S, λ_W) with $\lambda_S \geq \lambda_W$, we consider the following two cases: $r > 1$ and $r = 1$. Suppose first that $r > 1$. The second term of (56) can be explicitly written as follows:

$$\begin{aligned} \int_a^t e^{\lambda_S \tau} \lambda_s y_\tau d\tau &= \int e^{\lambda_S \tau} \lambda_s \left(\beta + \left(1 - e^{-\lambda_W \tau} \right) (\gamma - \beta) \right) d\tau \\ &= \frac{r}{1-r} (\gamma - \beta) e^{(\lambda_S - \lambda_W) t} + \gamma e^{\lambda_S t} \\ &\quad - \frac{r}{1-r} (\gamma - \beta) e^{(\lambda_S - \lambda_W) a} - \gamma e^{\lambda_S a}. \end{aligned}$$

Hence, the payoff z_t is characterized as follows:

$$z_t = C_a e^{-\lambda_S t} + \gamma + \frac{r}{1-r} (\gamma - \beta) e^{-\lambda_W t}, \tag{57}$$

with

$$C_a = C - \frac{r}{1-r} (\gamma - \beta) e^{(\lambda_S - \lambda_W) a} - \gamma e^{\lambda_S a}.$$

In contrast, the payoff y_t is characterized as follows:

$$y_t = \beta + \left(1 - e^{-\lambda_W t} \right) (\gamma - \beta).$$

Therefore, the difference between z_t and y_t (as long as $y_t \le z_t$) is:

$$z_t - y_t = C_a e^{-\lambda_S t} + \frac{1}{1-r}\left(\gamma - \beta\right) e^{-\lambda_W t}.$$

As a result, whether there exists $t_S^{**} \in (t_S, \infty)$ such that S does not enter for all $-t \in (-\infty, -t_S^{**})$ or not depends on

$$\lim_{t\to\infty}\left(C_a e^{-\lambda_S t} + \frac{1}{1-r}\left(\gamma - \beta\right) e^{-\lambda_W t}\right). \tag{58}$$

If $r > 1$, there exists $\bar{t} < \infty$ such that for all $t > \bar{t}$, the second term of (58) dominates. Since $r > 1$ for this case, there exists $\hat{t} < \infty$ such that for all $t > \hat{t}$, (58) is negative. That is, there exists $-t_S^{**}$ such that S does not enter for $-t \in (-\infty, -t_S^{**})$.

We now consider the case with $r = 1$. In this case, we can write $\lambda_S = \lambda_W = \lambda$. On the one hand, the second term of (56) can be explicitly written as follows:

$$\int_a^t e^{\lambda_S \tau} \lambda_s y_\tau d\tau = -\lambda\left(t-a\right)\left(\gamma-\beta\right) + \gamma\left(e^{\lambda t} - e^{\lambda a}\right).$$

Hence, the payoff z_t is characterized as

$$z_t = \gamma + e^{-\lambda t}\left(C - \lambda\left(t-a\right)\left(\gamma-\beta\right) - \gamma e^{-\lambda a}\right).$$

On the other hand, again, the payoff y_t is characterized as

$$y_t = \gamma + e^{-\lambda t}\left(\beta - \gamma\right).$$

Therefore, the difference between z_t and y_t (as long as $y_t \le z_t$) is:

$$z_t - y_t = e^{-\lambda t}\left(C - \lambda\left(t-a\right)\left(\gamma-\beta\right) - \gamma e^{-\lambda a} - \left(\beta - \gamma\right)\right).$$

There exists $\bar{t} < \infty$ such that for all $t > \bar{t}$, the term $-\lambda\left(t-a\right)\left(\gamma-\beta\right)$ dominates the other terms in the parentheses, and so $z_t - y_t < 0$. That is, there exists $-t_S^{**}$ such that S does not enter for $-t \in (-\infty, -t_S^{**})$, as stated in part 2(b)ii of Proposition 14.

[Proof of "cutoff from below"] We show that, if there does not exist $t_S^{**} \in (t_S, \infty)$ such that S does not enter for all $-t \in (-\infty, -t_S^{**})$ for (λ_S, λ_W), then such t_S^{**} does not exist for (λ_S', λ_W) with

$\lambda'_S < \lambda_S$.

To show this monotonicity, we first arbitrarily fix λ_W. Note that y_t is independent of λ_S. Let $x_t(\lambda_S)$ and $z_t(\lambda_S)$ be the values of x_t and z_t respectively, given λ_S for the fixed λ_W. There exists $\bar{t} > 0$ such that for all $t \in (0, \bar{t})$, $z_t(\lambda_S) < z_t(\lambda'_S)$. Define $t^* \equiv \inf_t \{z_t(\lambda_S) \geq z_t(\lambda'_S), t > 0\} \in \mathbb{R}_{++} \cup \{+\infty\}$.

If $t^* = +\infty$, then we have $z_t(\lambda_S) \leq z_t(\lambda'_S)$ for all $\lambda'_S < \lambda_S$. Since y_t is independent of λ_S, the proof is complete in this case. Hence, we concentrate on the case with $t^* < \infty$ and will derive a contradiction.

At $-t^*$, it must be the case that $z_{t^*}(\lambda_S) = z_{t^*}(\lambda'_S)$ and $\dot{z}_{t^*}(\lambda_S) \geq \dot{z}_{t^*}(\lambda'_S)$.[33] From $z_{t^*}(\lambda_S) = z_{t^*}(\lambda'_S)$, we have

$$
\begin{aligned}
\dot{z}_{t^*}(\lambda_S) &= \lambda_W \max\{x_{t^*}(\lambda_S) - z_{t^*}(\lambda_S), 0\} + \lambda_S \min\{y_{t^*} - z_{t^*}(\lambda_S), 0\}, \\
\dot{z}_{t^*}(\lambda'_S) &= \lambda_W \max\left\{x_{t^*}(\lambda'_S) - z_{t^*}(\lambda'_S), 0\right\} + \lambda'_S \min\left\{y_{t^*} - z_{t^*}(\lambda'_S), 0\right\} \\
&= \lambda_W \max\left\{x_{t^*}(\lambda'_S) - z_{t^*}(\lambda_S), 0\right\} + \lambda'_S \min\{y_{t^*} - z_{t^*}(\lambda_S), 0\}.
\end{aligned}
$$

Note that, by definition, we have $x_{t^*}(\lambda'_S) > x_{t^*}(\lambda_S)$. Given this inequality, the following two cases are possible:

1. If $x_{t^*}(\lambda_S) \geq z_{t^*}(\lambda_S)$, then we have $\lambda_W \max\{x_{t^*}(\lambda_S) - z_{t^*}(\lambda_S), 0\} < \lambda_W \max\{x_{t^*}(\lambda'_S) - z_{t^*}(\lambda_S), 0\}$. In addition, we have $\lambda_S \min\{y_{t^*} - z_{t^*}(\lambda_S), 0\} \leq \lambda'_S \min\{y_{t^*} - z_{t^*}(\lambda_S), 0\}$. Hence, we have $\dot{z}_{t^*}(\lambda'_S) > \dot{z}_{t^*}(\lambda_S)$. This is a contradiction.

2. If $x_{t^*}(\lambda_S) < z_{t^*}(\lambda_S)$, then we consider the following two subcases:

   (a) If $y_{t^*} > z_{t^*}(\lambda_S)$, then we have

$$
\lambda_S \min\{y_{t^*} - z_{t^*}(\lambda_S), 0\} < \lambda'_S \min\{y_{t^*} - z_{t^*}(\lambda_S), 0\}.
$$

[33] The first equality follows from the continuity of z_t with respect to t. The second inequality follows from the first equality and the definition of the derivative: For sufficiently small $\varepsilon > 0$,

$$
\begin{aligned}
\dot{z}_{t^*}(\lambda_S) &\approx \frac{z_{t^*}(\lambda_S) - z_{t^*-\varepsilon}(\lambda_S)}{\varepsilon}, \\
\dot{z}_{t^*}(\lambda'_S) &\approx \frac{z_{t^*}(\lambda'_S) - z_{t^*-\varepsilon}(\lambda'_S)}{\varepsilon} = \frac{z_{t^*}(\lambda_S) - z_{t^*-\varepsilon}(\lambda'_S)}{\varepsilon}.
\end{aligned}
$$

Since $z_{t^*-\varepsilon}(\lambda'_S) > z_{t^*-\varepsilon}(\lambda_S)$, it follows that $\dot{z}_{t^*}(\lambda_S) \geq \dot{z}_{t^*}(\lambda'_S)$.

Since we have

$$\lambda_W \max\{x_{t^*}(\lambda_S) - z_{t^*}(\lambda_S), 0\} \leq \lambda_W \max\left\{x_{t^*}(\lambda'_S) - z_{t^*}(\lambda_S), 0\right\},$$

we have $\dot{z}_{t^*}(\lambda'_S) > \dot{z}_{t^*}(\lambda_S)$. This is a contradiction.

(b) If $y_{t^*} \leq z_{t^*}(\lambda_S)$, then we have

$$\lambda_W \max\{x_{t^*}(\lambda_S) - z_{t^*}(\lambda_S), 0\} = 0$$

and

$$\lambda_S \min\{y_{t^*} - z_{t^*}(\lambda_S), 0\} = 0.$$

Therefore, $\dot{z}_{t^*}(\lambda_S) = 0$. For $t > t^*$, since $x_t(\lambda_S)$ is decreasing in t and y_t is increasing in t, we have $\dot{z}_t(\lambda_S) = 0$. Together with $y_{t^*} \leq z_{t^*}(\lambda_S)$, we have $y_t < z_t(\lambda_S)$ so S does not enter for $-t < -t^*$. This contradicts the assumption that there does not exist $t_S^{**} \in (t_S, \infty)$ such that S does not enter for all $-t \in (-\infty, -t_S^{**})$ for λ_S.

[Proof of "cutoff from above"] We prove that, if there exists $\bar{T} < \infty$ such that S does not enter for any $t > \bar{T}$ for a pair (λ_S, λ_W), then for any pair (λ'_S, λ_W) with $\lambda'_S > \lambda_S$, there exists $\bar{T}' < \infty$ such that S does not enter for any $t > \bar{T}'$.

This proof is symmetric to the one for "cutoff from below." We first arbitrarily fix λ_W. Again, y_t is independent of λ_S. Let $x_t(\lambda_S)$ and $z_t(\lambda_S)$ be the value of x_t and z_t, respectively, given λ_S for the fixed λ_W. There exists $\bar{t} > 0$ such that for all $t \in (0, \bar{t})$, $z_t(\lambda_S) > z_t(\lambda'_S)$. Define $t^* \equiv \inf_t \{z_t(\lambda_S) \leq z_t(\lambda'_S), t > 0\} \in \mathbb{R}_{++} \cup \{+\infty\}$.

If $t^* = +\infty$, then we have $z_t(\lambda_S) \geq z_t(\lambda'_S)$ for all $\lambda'_S < \lambda_S$. Since y_t is independent of λ_S, the proof is complete in this case. Hence, we concentrate on the case with $t^* < \infty$.

At $-t^*$, it must be the case that $z_{t^*}(\lambda_S) = z_{t^*}(\lambda'_S)$ and $\dot{z}_{t^*}(\lambda_S) \leq \dot{z}_{t^*}(\lambda'_S)$ by an argument analogous to footnote 33 in the Online Appendix. From $z_{t^*}(\lambda_S) = z_{t^*}(\lambda'_S)$, we have

$$\begin{aligned}
\dot{z}_{t^*}(\lambda_S) &= \lambda_W \max\{x_{t^*}(\lambda_S) - z_{t^*}(\lambda_S), 0\} + \lambda_S \min\{y_{t^*} - z_{t^*}(\lambda_S), 0\} \\
\dot{z}_{t^*}(\lambda'_S) &= \lambda_W \max\left\{x_{t^*}(\lambda'_S) - z_{t^*}(\lambda'_S), 0\right\} + \lambda'_S \min\left\{y_{t^*} - z_{t^*}(\lambda'_S), 0\right\} \\
&= \lambda_W \max\left\{x_{t^*}(\lambda'_S) - z_{t^*}(\lambda_S), 0\right\} + \lambda'_S \min\{y_{t^*} - z_{t^*}(\lambda_S), 0\}.
\end{aligned}$$

Note that, by definition, we have $x_{t^*}(\lambda'_S) < x_{t^*}(\lambda_S)$. Given this inequality, the following two cases are possible:

1. If $x_{t^*}(\lambda_S) > z_{t^*}(\lambda_S)$, then we have $\lambda_W \max\{x_{t^*}(\lambda_S) - z_{t^*}(\lambda_S), 0\} > \lambda_W \max\{x_{t^*}(\lambda'_S) - z_{t^*}(\lambda_S), 0\}$. In addition, we have $\lambda_S \min\{y_{t^*} - z_{t^*}(\lambda_S), 0\} \geq \lambda'_S \min\{y_{t^*} - z_{t^*}(\lambda_S), 0\}$. Hence, we have $\dot{z}_{t^*}(\lambda'_S) < \dot{z}_{t^*}(\lambda_S)$. This is a contradiction.

2. If $x_{t^*}(\lambda_S) \leq z_{t^*}(\lambda_S)$, then we have

$$\lambda_W \max\left\{x_{t^*}(\lambda'_S) - z_{t^*}(\lambda_S), 0\right\} \leq \lambda_W \max\{x_{t^*}(\lambda_S) - z_{t^*}(\lambda_S), 0\} = 0$$

   and so

$$\lambda_W \max\left\{x_{t^*}(\lambda'_S) - z_{t^*}(\lambda_S), 0\right\} = \lambda_W \max\{x_{t^*}(\lambda_S) - z_{t^*}(\lambda_S), 0\} = 0.$$

   We consider the following subcases:

   (a) If $y_{t^*} > z_{t^*}(\lambda_S)$, then we have

$$\lambda_S \min\{y_{t^*} - z_{t^*}(\lambda_S), 0\} > \lambda'_S \min\{y_{t^*} - z_{t^*}(\lambda_S), 0\}.$$

   Hence, we have $\dot{z}_{t^*}(\lambda'_S) < \dot{z}_{t^*}(\lambda_S)$. This is a contradiction.

   (b) If $y_{t^*} \leq z_{t^*}(\lambda_S)$, then we have

$$\lambda_W \max\left\{x_{t^*}(\lambda'_S) - z_{t^*}(\lambda_S), 0\right\} = 0$$

   and

$$\lambda'_S \min\left\{y_{t^*} - z_{t^*}(\lambda'_S), 0\right\} = \lambda'_S \min\{y_{t^*} - z_{t^*}(\lambda_S), 0\} = 0.$$

   Therefore, $\dot{z}_{t^*}(\lambda'_S) = 0$. For $t > t^*$, since $x_t(\lambda'_S)$ is decreasing in t and y_t is increasing in t, we have $\dot{z}_t(\lambda'_S) = 0$. Hence, S does not enter for $-t < -t^*$ with λ'_S, as desired.

In the proof above, all the time-cutoffs described are finite and independent of T, as stated in part 3 of Proposition 14.

N.2 Proof of Remark 7

Before proving Proposition 15, we prove Remark 7. It suffices to show that $\phi < 0$ implies $t_S = \infty$.

By definition, we can write

$$f_S(t) = \begin{cases} \frac{1}{1-r}\left(e^{-\lambda_S t} - e^{-\lambda_W t}\right) - \frac{\beta + \left(1-e^{-\lambda_W t}\right)\max\{(\gamma-\beta),0\}}{1-\alpha} & \text{if } r \neq 1 \\ \lambda_W t e^{-\lambda_W t} - \frac{\beta + \left(1-e^{-\lambda_W t}\right)\max\{(\gamma-\beta),0\}}{1-\alpha} & \text{if } r = 1 \end{cases}.$$

If $\frac{1}{r} - \frac{1-r}{r}\frac{\max\{(\gamma-\beta),0\}}{1-\alpha} \leq 0$, then

$$\begin{aligned} f_S(t) &= \frac{1}{1-r}\left(e^{-\lambda_S t} - e^{-\lambda_W t}\right) - \frac{\beta + \left(1-e^{-\lambda_W t}\right)\max\{(\gamma-\beta),0\}}{1-\alpha} \\ &= \frac{1}{1-r}e^{-\lambda_S t} + \left(\frac{\max\{(\gamma-\beta),0\}}{1-\alpha} - \frac{1}{1-r}\right)e^{-\lambda_W t} - \frac{\beta + \{\max(\gamma-\beta),0\}}{1-\alpha} \end{aligned}$$

is always decreasing in t. Since $f_S(0) = -\frac{\beta}{1-\alpha}$, we have $f_S(t) < 0$ for all t. Therefore, we have $t_S = \infty$ as desired.

Hence, for the rest of the proof, we focus on the case in which $\frac{1}{r} - \frac{1-r}{r}\frac{\max\{(\gamma-\beta),0\}}{1-\alpha} > 0$. Then, the first- and second- order conditions for $f_S(t)$ imply that $f_S(t)$ is single-peaked at

$$t^{\text{peak}} = \begin{cases} \frac{\log\left(\frac{1}{r} - \frac{1-r}{r}\frac{\max\{(\gamma-\beta),0\}}{1-\alpha}\right)}{\lambda_W - \lambda_S} & \text{if } r \neq 1 \\ \frac{1}{\lambda_W}\left(1 - \frac{\max\{(\gamma-\beta),0\}}{1-\alpha}\right) & \text{if } r = 1 \end{cases}.$$

For $r \neq 1$, since

$$\begin{aligned} f_S(t) &= \frac{1}{1-r}\left(e^{-\lambda_S t} - e^{-\lambda_W t}\right) - \frac{\beta + \left(1-e^{-\lambda_W t}\right)\max\{(\gamma-\beta),0\}}{1-\alpha} \\ &= \left(\frac{1}{1-r}\left(re^{-\lambda_S t} - e^{-\lambda_W t}\right) + \frac{e^{-\lambda_W t}\max\{(\gamma-\beta),0\}}{1-\alpha}\right) \\ &\quad - \frac{1}{1-r}(r-1)e^{-\lambda_S t} - \frac{\beta + \max\{(\gamma-\beta),0\}}{1-\alpha} \\ &= -\frac{1}{\lambda_W}f_S'(t) + e^{-\lambda_S t} - \frac{\max\{\beta,\gamma\}}{1-\alpha}, \end{aligned}$$

substituting $f_S'(t^{\text{peak}}) = 0$ and $t^{\text{peak}} = \frac{\log\left(\frac{1}{r} - \frac{1-r}{r}\frac{\gamma-\beta}{1-\alpha}\right)}{\lambda_W - \lambda_S}$ into $f_S(t)$ yields

$$\begin{aligned}
f_S(t^{\text{peak}}) &= e^{-\lambda_S \frac{\log\left(\frac{1}{r} - \frac{1-r}{r}\frac{\max\{(\gamma-\beta),0\}}{1-\alpha}\right)}{\lambda_W - \lambda_S}} - \frac{\max\{\beta,\gamma\}}{1-\alpha} \\
&= e^{-r \frac{\log\left(\frac{1}{r} - \frac{1-r}{r}\frac{\max\{(\gamma-\beta),0\}}{1-\alpha}\right)}{1-r}} - \frac{\max\{\beta,\gamma\}}{1-\alpha} \\
&= e^{\log\left(\frac{1}{r} - \frac{1-r}{r}\frac{\max\{(\gamma-\beta),0\}}{1-\alpha}\right)^{\frac{r}{r-1}}} - \frac{\max\{\beta,\gamma\}}{1-\alpha} \\
&= \left(\frac{1}{r} - \frac{1-r}{r}\frac{\max\{(\gamma-\beta),0\}}{1-\alpha}\right)^{\frac{r}{r-1}} - \frac{\max\{\beta,\gamma\}}{1-\alpha} \\
&= \phi.
\end{aligned}$$

Therefore, if $\phi < 0$, then there is no solution for $f_S(t) = 0$ and so $t_S = \infty$, as desired.

For $r = 1$, since

$$\begin{aligned}
f_S(t) &= \lambda_W t e^{-\lambda_W t} - \frac{\beta + \left(1 - e^{-\lambda_W t}\right)\max\{(\gamma-\beta),0\}}{1-\alpha} \\
&= -\frac{1}{\lambda_W} f_S'(t) + e^{-\lambda_W t} - \frac{\max\{\beta,\gamma\}}{1-\alpha},
\end{aligned}$$

substituting $f_S'(t^{\text{peak}}) = 0$ and $t^{\text{peak}} = \frac{1}{\lambda}\left(1 - \frac{\max\{(\gamma-\beta),0\}}{1-\alpha}\right)$ into $f_S(t)$ yields

$$\begin{aligned}
f_S(t^{\text{peak}}) &= e^{-\lambda_W \frac{1}{\lambda_W}\left(1 - \frac{\max\{(\gamma-\beta),0\}}{1-\alpha}\right)} - \frac{\max\{\beta,\gamma\}}{1-\alpha} \\
&= e^{\frac{\max\{(\gamma-\beta),0\}}{1-\alpha} - 1} - \frac{\max\{\beta,\gamma\}}{1-\alpha} \\
&= \phi.
\end{aligned}$$

Therefore, if $\phi < 0$, then there is no solution for $f_S(t) = 0$ and so $t_S = \infty$ holds, as desired.

N.3 Proof of Proposition 15

Recall that t_S is the smallest positive solution for $f_S(t) = 0$ where

$$f_S(t) \equiv \begin{cases} \frac{1}{1-r}\left(e^{-\lambda_S t} - e^{-\lambda_W t}\right) - \frac{\beta + \left(1 - e^{-\lambda_W t}\right)\max\{(\gamma-\beta),0\}}{1-\alpha} & \text{if } r \neq 1 \\ \lambda_W t e^{-\lambda_W t} - \frac{\beta}{1-\alpha} & \text{if } r = 1 \end{cases}, \tag{59}$$

while t_W is the smallest solution for $f_W(t) = 0$ where

$$f_W(t) \equiv \begin{cases} \frac{1}{1-r}\left(e^{-\lambda_S t} - e^{-\lambda_W t}\right) - e^{-\lambda_S t} & \text{if } r \neq 1 \\ \lambda_W t e^{-\lambda_W t} - e^{-\lambda_W t} & \text{if } r = 1 \end{cases}. \tag{60}$$

We prove each part of the proposition in what follows.

N.3.1 Proof of Part 1 of Proposition 15

When we change r, without loss, we keep λ_W fixed and vary λ_S. First, note that, for sufficiently large r, ϕ is negative:

$$\begin{aligned} & \lim_{r\to\infty}\left(\frac{1}{r} - \frac{1-r}{r}\max\left\{\frac{\gamma-\beta}{1-\alpha}, 0\right\}\right)^{\frac{r}{r-1}} - \frac{\max\{\beta,\gamma\}}{1-\alpha} \\ = \; & -\max\left(\frac{\gamma-\beta}{\alpha-1}, 0\right) - \frac{\max\{\beta,\gamma\}}{1-\alpha} < 0. \end{aligned}$$

Hence, for sufficiently large r, we have $-t_W > -t_S$.

Second, since

$$\begin{aligned} & \lim_{r\to 0}\left(\frac{1}{1-r}\left(e^{-\lambda_S t} - e^{-\lambda_W t}\right) - \frac{\beta + \left(1 - e^{-\lambda_W t}\right)\max\{(\gamma-\beta), 0\}}{1-\alpha}\right) \\ = \; & \lim_{r\to 0}\frac{1}{1-r}\left(e^{-r\lambda_W t} - e^{-\lambda_W t}\right) - \frac{\beta + \left(1 - e^{-\lambda_W t}\right)\max\{(\gamma-\beta), 0\}}{1-\alpha} \\ = \; & 1 - e^{-\lambda_W t} - \frac{\beta + \left(1 - e^{-\lambda_W t}\right)\max\{(\gamma-\beta), 0\}}{1-\alpha} \end{aligned}$$

and

$$\begin{aligned} & \lim_{r\to 0}\left(\frac{1}{1-r}\left(e^{-\lambda_S t} - e^{-\lambda_W t}\right) - e^{-\lambda_S t}\right) \\ = \; & \lim_{r\to 0}\left(\frac{1}{1-r}\left(e^{-r\lambda_W t} - e^{-\lambda_W t}\right) - e^{-r\lambda_W t}\right) \\ = \; & -e^{-t\lambda_W} \end{aligned}$$

hold for each t, $\lim_{r\to 0} t_S < \infty$ and $\lim_{r\to 0} t_W = \infty$. Thus, for sufficiently small r, we have $-t_W < -t_S$.

Therefore, we are left to show that

$$\frac{\partial\left(-t_W\right)}{\partial r}-\frac{\partial\left(-t_S\right)}{\partial r}>0.$$

To this end, in (59) and (60), when λ_S goes up with λ_W fixed, the first terms in $f_S(t)$ and $f_W(t)$ move in the same way while the second terms ($-\frac{\beta+\left(1-e^{-\lambda_W t}\right)\max\{(\gamma-\beta),0\}}{1-\alpha}$ in $f_S\left(t\right)$ and $-e^{-\lambda_S t}$ in $f_W(t)$) become larger only in $f_W(t)$. Hence, we have $\frac{\partial(-t_W)}{\partial r}-\frac{\partial(-t_S)}{\partial r}>0$, as desired.

N.3.2 Proof of Part 2 of Proposition 15

First, note that, for sufficiently large α, ϕ is negative for the following reason: If $r\neq 1$, since

$$\lim_{\alpha\to 1}\left(\frac{1}{r}-\frac{1-r}{r}\max\left\{\frac{\gamma-\beta}{1-\alpha},0\right\}\right)<0,$$

$\phi<0$ in the limit as $\alpha\to 1$. If $r=1$, since

$$\lim_{\alpha\to 1}\left(e^{-\max\left\{\frac{\gamma-\beta}{1-\alpha},0\right\}}-\frac{\max\{\beta,\gamma\}}{1-\alpha}\right)<0,$$

$\phi<0$ in the limit as $\alpha\to 1$. Hence, for sufficiently large α, we have $-t_W>-t_S$.

Therefore, we are left to show that

$$\frac{\partial\left(-t_W\right)}{\partial \alpha}-\frac{\partial\left(-t_S\right)}{\partial \alpha}>0.$$

In (59) and (60), when α goes up, $f_W(t)$ is unchanged. Hence, we are left to show that $\frac{\partial(-t_S)}{\partial\alpha}<0$, that is, the smallest positive t such that $\frac{1}{1-r}\left(e^{-\lambda_S t}-e^{-\lambda_W t}\right)=\frac{\beta+\left(1-e^{-\lambda_W t}\right)\max\{(\gamma-\beta),0\}}{1-\alpha}$ (or $\frac{1}{1-r}\left(e^{-\lambda_S t}-e^{-\lambda_W t}\right)=\frac{\beta}{1-\alpha}$ if $r=1$) increases. Notice that $\frac{1}{1-r}\left(e^{-\lambda_S t}-e^{-\lambda_W t}\right)$ is single-peaked at $\frac{\log\lambda_W-\log\lambda_S}{\lambda_W-\lambda_S}$. Since $\frac{\beta+\left(1-e^{-\lambda_W t}\right)\max\{(\gamma-\beta),0\}}{1-\alpha}$ and $\frac{\beta}{1-\alpha}$ become larger, the smallest positive t such that $\frac{1}{1-r}\left(e^{-\lambda_S t}-e^{-\lambda_W t}\right)=\frac{\beta+\left(1-e^{-\lambda_W t}\right)\max\{(\gamma-\beta),0\}}{1-\alpha}$ (or $\frac{1}{1-r}\left(e^{-\lambda_S t}-e^{-\lambda_W t}\right)=\frac{\beta}{1-\alpha}$ if $r=1$) increases.

N.3.3 Proof of Part 3 of Proposition 15

First, note that, for sufficiently large β, ϕ is negative: If $r \neq 1$, since

$$\begin{aligned}
\lim_{\beta \to 1} \left(\left(\frac{1}{r} - \frac{1-r}{r} \max \left\{ \frac{\gamma - \beta}{1-\alpha}, 0 \right\} \right)^{\frac{r}{r-1}} - \frac{\max\{\beta, \gamma\}}{1-\alpha} \right) &= \left(\frac{1}{r} \right)^{\frac{r}{r-1}} - \frac{1}{1-\alpha} \\
&\leq \max_r \left(\frac{1}{r} \right)^{\frac{r}{r-1}} - \frac{1}{1-\alpha} \\
&= 1 - \frac{1}{1-\alpha} < 0,
\end{aligned}$$

$\phi < 0$ in the limit as $\beta \to 1$. If $r = 1$, since

$$\lim_{\beta \to 1} \left(e^{-\max\left\{\frac{\gamma-\beta}{1-\alpha}, 0\right\}} - \frac{\max\{\beta, \gamma\}}{1-\alpha} \right) = 1 - \frac{1}{1-\alpha} < 0,$$

$\phi < 0$ in the limit as $\beta \to 1$. Hence, for sufficiently large β, we have $-t_W > -t_S$.

Therefore, we are left to show that

$$\frac{\partial(-t_W)}{\partial \beta} - \frac{\partial(-t_S)}{\partial \beta} < 0.$$

In (59) and (60), when β goes up, the second terms $-\frac{\beta + \left(1 - e^{-\lambda_W t}\right) \max\{(\gamma - \beta), 0\}}{1-\alpha}$ and $-\frac{\beta}{1-\alpha}$ in $f_S(t)$ become smaller while $f_W(t)$ is unchanged. Hence, by the same proof as in the case where α increases, $-t_W - (-t_S)$ increases. Hence, we have $\frac{\partial(-t_W)}{\partial \beta} - \frac{\partial(-t_S)}{\partial \beta} > 0$, as desired.

N.3.4 Proof of Part 4 of Proposition 15

It suffices to show that

$$\begin{aligned}
\frac{\partial(-t_W)}{\partial \gamma} - \frac{\partial(-t_S)}{\partial \gamma} &= 0 \text{ if } \gamma \leq \beta, \\
\frac{\partial(-t_W)}{\partial \gamma} - \frac{\partial(-t_S)}{\partial \gamma} &> 0 \text{ if } \gamma > \beta.
\end{aligned}$$

In (59) and (60), if $\beta \geq \gamma$, then neither $f_S(t)$ nor $f_W(t)$ depends on γ. Hence, we have $\frac{\partial(-t_W)}{\partial \gamma} - \frac{\partial(-t_S)}{\partial \gamma} = 0$. Hence, let us focus on the case $\gamma > \beta$. If γ goes up in (59) and (60), then the second term $-\frac{\beta + \left(1 - e^{-\lambda_W t}\right) \max\{(\gamma - \beta), 0\}}{1-\alpha}$ in $f_S(t)$ becomes smaller while $f_W(t)$ is unchanged. Hence, by the same proof as in part 3 of Proposition 15, $-t_W - (-t_S)$ increases, which means $\frac{\partial(-t_W)}{\partial\gamma} - \frac{\partial(-t_S)}{\partial\gamma} > 0$, as desired.

Contents