早稲田大学現代政治経済研究所

Campaign Promises as an Imperfect Signal: How does an Extreme Candidate Win against a Moderate Candidate?

Yasushi Asako

No.E1411

Working Paper Series

Waseda Institute of Political Economy

Waseda University

169-8050 Tokyo,Japan

Campaign Promises as an Imperfect Signal: How does an Extreme Candidate Win against a Moderate Candidate?

Yasushi Asako*

Waseda University

September 2014

Abstract

This study develops a political competition model in which campaign platforms are partially binding. A candidate who implements a policy that differs from his/her platform must pay a cost of betrayal, which increases with the size of the discrepancy. I also assume that voters are uncertain about candidates' policy preferences. If voters believe that a candidate is likely to be extreme, there exists a semi-separating equilibrium: an extreme candidate imitates a moderate candidate, with some probability, and approaches the median policy with the remaining probability. Although an extreme candidate will implement a more extreme policy than will a moderate candidate, regardless of imitation or approach, partial pooling ensures that voters prefer an extreme candidate who does not pretend to be moderate over an uncertain candidate who announces an extreme platform. As a result, a moderate candidate never has a higher probability of winning than does an extreme candidate.

Keywords: electoral competition, voting, campaign promise, signaling game

JEL Classification Numbers: C72, D72, D82

---

*School of Political Science and Economics, Waseda University, 1-6-1 Nishi-Waseda, Shinjuku-ku, Tokyo 169-8050, JAPAN. E-mail: yasushi.asako@aoni.waseda.jp. The author benefited from the comments of Scott Gehlbach, Marzena Rostek, and William Sandholm, as well as from the comments of Swati Dhingra, Steven Durlauf, Yoichi Hizen, Ching-Yang Lin, John Morrow, Daniel Quint, Ryoji Sawa, Ricardo Serrano-Padial, Marek Weretka, and Mian Zhu. Financial support from the Nakajima Foundation and JSPS KAKENHI (Grant Number 26780178) is gratefully acknowledged.

1 Introduction

Most past studies on the two-candidate model of political competition show a politician's convergence with the ideal policy of the median voter. However, observed patterns in real-world elections are frequently polarized. One simple reason for this is polarized politicians. That is, politicians' policy preferences are extreme, which means these policies are further from the median voter's ideal policy. Some candidates renege on their campaign promises to the electorate and implement an extreme policy after an election. However, most past studies show that more moderate candidates beat more extreme candidates when these two candidates run. If so, winning politicians should not be so polarized. To understand polarization, it is important to answer the following questions. How do candidates with extreme policy preferences beat those whose policy preferences are less extreme? Why do extreme candidates have an incentive to win by compromising their preferences and approaching the median policy in an electoral campaign, but ultimately implementing a polarized, extreme policy after winning?

The striking result of this study is that an extreme candidate has a higher probability of winning than does a moderate candidate, *although the extreme candidate will implement a more extreme policy*. The important reason for the extreme candidate's higher probability of winning is that he has a stronger incentive to prevent an opponent from winning because his ideal policy is further from the opponent's policy than is that of a moderate candidate. My model describes this incentive for an extreme candidate by introducing two reasonable assumptions: *partially binding platforms* and uncertainty about a candidate's preference.

My model with partially binding platforms assumes that a candidate can choose any policy after the election. However, a politician who implements a policy that differs from his/her campaign platform must pay a "cost of betrayal," which increases with the degree of betrayal. I also introduce asymmetric information by assuming that candidate policy preferences are private information. A politician's preferences may change depending on local conditions or the important issues in a given election. In particular, when candidates are not well known, it is difficult to know their preferences. Even in a political competition among parties, a party's preferences may change depending on the leader of the party. Voters are also uncertain of how power will be distributed between factions as a result of an intra-party negotiation after an election. In particular, if several parties form a coalition government, voters will be especially uncertain about implemented policies, since these will depend on the negotiation among the members of the coalition.

The model supposes a two-candidate political competition in a one-dimensional policy space. One candidate's ideal policy is to the left of the median policy, while that of the other candidate is to the right. Candidates are entirely motivated by policy. Each candidate is one of two types —moderate or extreme— and the moderate type's ideal policy is closer to the median policy than that of the extreme type. A candidate knows his/her own type, but voters and the opponent do not. In the remainder of the paper, I refer to an extreme type as "he" and a moderate type as "she."

If voters believe in advance that a candidate is likely to be extreme, an extreme type has a higher probability of winning than a moderate type in a semi-separating equilibrium. In this equilibrium, an extreme type chooses a mixed strategy. With some probability, the extreme type announces the same platform as the moderate type, and with the remaining probability, approaches the median policy, thus revealing his type to voters. We call an extreme type a "pooling extreme type" when he imitates the moderate type, but a "separating extreme type" when he approaches the median policy.

A separating extreme type implements a more moderate policy than a pooling extreme type, but this is still a more extreme policy than that of a moderate type in equilibrium. This is because an extreme type will betray his platform to a greater extent than a moderate type, even though the platform of a separating extreme type is more moderate than that of a moderate type. While voters know the type of a separating extreme type, they remain uncertain about the type of candidate who announces a moderate type's platform, because this may be a moderate or a pooling extreme type candidate. Voters around the median policy wish to avoid electing a pooling extreme type who will implement the most extreme policy. Thus, they forgo the chance to elect a moderate type who will implement the most moderate policy and choose a separating extreme type whose implemented policy lies between those of a moderate and a pooling extreme type. This is also the reason why a separating extreme type can implement a more extreme policy than a moderate type, but still defeat her. As a result, a separating extreme type has a higher probability of winning than a moderate type (and a pooling extreme type) in equilibrium.

On the other hand, if voters believe that a candidate is likely to be a moderate type, a separating extreme type needs to rigorously approach the median policy to win. In this case, an extreme type just imitates a moderate type with certainty (a pooling equilibrium). As a result, a moderate type can never have a higher probability of winning than an extreme type.

In several countries, an extreme party won an election by announcing a moderate platform, such as the Justice and Development Party of Turkey in 2007 and the Liberal Democratic Party of Japan in the early 2000s. According to my model, an extreme candidate or party approaches the median policy and wins because he wants to prevent the opposition from winning. This is an important reason why extremists compromised and won in some countries. I discuss these examples in detail in Appendix B.

1.1 Related Literature

While my study assumes partially binding platforms, most previous studies use one of two polar assumptions about platforms. First, models with *completely binding platforms* assume that a politician cannot implement any policy other than that of the platform.[1] Second, models with *nonbinding platforms* assume that a politician can implement any policy freely, without cost.[2] In other words, a politician implements his/her ideal policy regardless of platform. However, in reality, politicians often betray their platforms, but doing so is likely to be costly. When politicians implement a policy that differs from their platform, their approval rating may fall as a result of criticism from voters and the media.[3] They will need to undertake costly negotiations with Congress, the party may punish them,[4] and the possibility of losing the next election may increase. Therefore, platforms should be partially binding. That is, a politician should decide on a policy based on the platform and the cost of betrayal.

Some previous studies have considered a similar idea to the cost of betrayal. In particular, Banks (1990) and Callander and Wilkie (2007) show that a platform can signal a policy to be implemented. My study differs from these in two important ways. First, in their studies, candidates automatically implement their own ideal policy after an election. However, if

[1] Electoral competition models in the Downsian tradition (Downs, 1957; Wittman, 1973).

[2] For example, this approach is taken in citizen candidate models, such as Besley and Coate (1997), Osborne and Slivinski (1996), and in retrospective voting models such as Barro (1973) and Ferejohn (1986).

[3] Some studies, such as Reinikka and Svensson (2005) and Djankov et al. (2003), show a relationship between the media and the credible commitment of politicians.

[4] Cox and McCubbins (1994), McGillivray (1997), Aldrich (1997), Snyder and Groseclose (2000), McCarty et al. (2001), and Grossman and Helpman (2005, 2008) indicate that there is party discipline.

there is a cost of betrayal, a rational candidate would adjust the implemented policy to reduce this cost after an election. Second, Banks (1990) and Callander and Wilkie (2007) consider that candidates care about policy only when they win; their utility is set to zero when they lose, regardless of the policy their opponent implements. However, *policy-motivated* candidates should care about policy when they lose. With these two assumptions, Banks (1990) and Callander and Wilkie (2007) show that a moderate type may defeat an extreme type when there is asymmetric information about each of the candidate's ideal policies. I relax these assumptions and add more reasonable ones by examining the rational choices when implementing a policy, as well as candidates who care about policy regardless of the election results. Thus, my results are the opposite of those in the aforementioned studies. That is, an extreme type has a higher probability of winning than a moderate type.[5]

These two differences are critical to my results. First, if candidates implement their own ideal policies automatically, an extreme candidate will lose against a moderate type when he reveals his type to voters. Thus, a separating extreme type cannot obtain a higher probability of winning in a semi-separating equilibrium. Second, if a candidate does not care about policy when he/she loses, an extreme type will not have as strong an incentive to prevent an opponent from winning. Therefore, under the assumptions of Banks (1990) and Callander and Wilkie (2007), an extreme type does not have either a method or an incentive to win against a moderate type. I will clarify these points in Section 3.5.1.[6] The other important setting to obtain my result is asymmetric information. If voters know the candidate's type, an extreme type cannot pretend to be a moderate type, so both pooling and

[5] There is another difference between my study and those of Banks (1990) and Callander and Wilkie (2007). The latter studies assume a continuum of candidate types, while I assume just two types of candidate, namely a moderate and an extreme type. If my model used a continuum of types, a semi-separating equilibrium would not exist, but a pooling equilibrium would exist in some circumstances. To have a semi-separating equilibrium, an extreme type's platform should have sufficient distance from that of a moderate (and a pooling extreme) candidate. A continuum makes it difficult to have such a sufficient distance. However, candidates' types do not need to be on a continuum in a real-world election. Usually, there are a few factions or groups on each side. Voters may be uncertain about the candidates' group or faction, or which group has power in a party or in Congress. Such a situation should be analyzed using distinct types rather than on a continuum.

[6] In Callander (2008), candidates care about what happens when they lose. Here, there are two choices—policy and a level of effort—and candidates can fully commit to a policy before an election (completely binding), or decide on a level of effort after winning (nonbinding). Then, Callander (2008) shows that a policy position can signal a future level of effort. Hummel (2010) supposes that there exist costs when policy announcements are different between primary and general elections. Other studies also consider that a completely binding platform is a signal for the functioning of the economy (Schulz, 1996) and the candidate's degree of honesty (Kartik and McAfee, 2007).

semi-separating equilibria do not exist. Indeed, Asako (forthcoming) shows that a moderate type wins against an extreme type when assuming the same settings as this study apart from asymmetric information.

The idea that an implemented policy will lie somewhere between a platform and an ideal policy is not entirely original. Several other studies discuss ideas concerning partially binding platforms with complete information. Austen-Smith and Banks (1989) consider a two-period game based on a retrospective voting model in which, if office-motivated candidates betray their platforms, the probability of winning in the next election decreases. Grossman and Helpman (2005, 2008) develop a legislative model in which office-motivated parties announce platforms before an election, and the victorious legislators, who are policy-motivated, decide the policy. If legislators betray the party platform, the party punishes them. In contrast, my model is based on a prospective and two-candidate competition model, and I consider that candidates who are policy-motivated decide both the platform and policy.[7] In Huang (2010), candidates strategically choose both a platform and an implemented policy, but do not care about policy when they lose. Huang (2010) also supposes sufficiently large benefits from holding office and shows that candidates cluster around or at the median policy. On the other hand, my model supposes that candidates care about policy even after losing, and does not consider (large) benefits from holding office.

Several past studies consider the effects of a candidate's character or personality, as indicated by Stokes (1963), as "valence." These past studies assume that the valence of a candidate is given exogenously, and voters care not only about the policy, but also about valence. They show that a candidate with good valence has a higher probability of winning an election.[8] Therefore, an extreme candidate may win against a moderate candidate if he has good valence, as shown by Kartik and McAfee (2007) and Callander (2008).[9] On the

[7]Austen-Smith and Banks (1989) consider only a decrease in the probability of winning as the cost of betrayal, and Grossman and Helpman (2005, 2008) consider only party discipline as the cost of betrayal. However, as I have indicated, the cost of betrayal also includes other types of costs, such as a decrease in approval ratings or a negotiation cost with Congress. Therefore, I include these in the current term as the cost of betrayal. In addition, Harrington (1993), Aragones et al. (2007), and Ishihara (2014) show that, in a repeated game, nonbinding platforms can be completely binding in equilibrium.

[8]For example, in Ansolabehere and Snyder (2000), Aragones and Palfrey (2002), Groseclose (2001), Kartik and McAfee (2007), and Callander (2008), there are advantaged and disadvantaged candidates. An advantage is given exogenously as a valence, and a voter's utility is affected by both a policy and this valence.

[9]Kartik and McAfee (2007) give good "character" exogenously, and interpret it as integrity. Callander (2008) gives this valence endogenously using policy motivation, but a candidate's policy motivation is given exogenously. In both studies, more extreme policy can be a signal for good valence.

other hand, my study shows that an extreme candidate wins against a moderate candidate without such exogenous valence.[10]

The remainder of the paper is structured as follows. Section 2 presents the model, and Section 3 analyzes the equilibria. Finally, Section 4 concludes the paper.

2 Setting

The policy space is $\Re$. There is a continuum of voters, and their ideal policies are distributed on some interval of $\Re$. The distribution function is continuous and strictly increasing, so there is a unique median voter's ideal policy, x_m. There are two candidates, L and R, and each candidate is one of two types: moderate or extreme. Let x_i^M and x_i^E denote the respective *ideal policies* for the moderate and extreme types, where $i = L$ or R, and $x_L^E < x_L^M < x_m < x_R^M < x_R^E$. The superscripts M and E represent a moderate or extreme type, respectively, and the moderate type's ideal policy is closer to the median policy. Assume $x_m - x_L^t = x_R^t - x_m$ for $t = M$ or E. That is, the ideal policies of the same type are equidistant from the median policy. A candidate knows his/her own type, but voters and the opponent are uncertain about the candidate's type. For both candidates, $p^M \in (0, 1)$ is the prior probability that the candidate is a moderate type. Thus, the prior probability that the candidate is an extreme type is $1 - p^M$.

After the types of candidates are decided, each candidate announces a *platform*, denoted by $z_i^t \in \Re$, where $i = L$ or R and $t = M$ or E. On the basis of these platforms, voters decide on a winner according to a majority voting rule. After an election, the winning candidate chooses an *implemented policy*, denoted by χ_i^t, where $i = L$ or R and $t = M$ or E. The relations among an ideal policy, a platform, and an implemented policy are summarized in Figure 1.

[Figure 1 here]

[10] The other important topic related to campaign platforms is political ambiguity. Here, a candidate announces an ambiguous platform that includes more than one policy choice (Zeckhauser, 1969; Shepsle, 1972; Aragones and Cukierman, 1990; Glazer, 1990; Aragones and Neeman, 2000; Aragones and Postlewaite, 2002; Meirowitz, 2005; Callander and Wilson, 2008). However, it is difficult to obtain an equilibrium with an ambiguous platform. As a result, they use a very simple setting, such as a restricted number of policy choices, degrees of ambiguity, and/or a special form of voters' preferences. Thus, it is important to derive an equilibrium using an ambiguous platform with more standard settings. However, this topic is outside the scope of this study's purpose, so I do not analyze political ambiguity here.

If the implemented policy is different from the candidate's ideal policy, all candidates—both winner and loser—experience disutility. This disutility is represented by $-v(|\chi - x_i^t|)$, where $i = L$ or R, $t = M$ or E, and χ is the policy implemented by the winner. Assume that $v(.)$ satisfies $v(0) = 0$, $v'(0) = 0$, $v'(d) > 0$, and $v''(d) > 0$ when $d > 0$. If the implemented policy is not the same as the platform, the winning candidate needs to pay a cost of betrayal. The function describing the cost of betrayal is $c(|z_i - \chi|)$. Assume that $c(.)$ satisfies $c(0) = 0$, $c'(0) = 0$, $c'(d) > 0$, and $c''(d) > 0$ when $d > 0$. Here, the loser does not pay a cost. Moreover, I assume throughout that $c'(d)/c(d)$, $v'(d)/v(d)$, and $v''(d)/v'(d)$ strictly decrease as d increases.[11] For example, if the function is a monomial, this assumption holds, and many polynomial functions satisfy the assumptions as well. After an election, the winning candidate chooses a policy that maximizes $-v(|\chi - x_i^t|) - c(|z_i - \chi|)$. That is, the winner chooses $\chi_i^t(z_i) = argmax_\chi \ - v(|\chi - x_i^t|) - c(|z_i - \chi|)$.

Upon observing a platform, the utility of voter n when candidate i of type t wins is $-u(|\chi_i^t(z_i) - x_n|)$. Assume that $u(.)$ satisfies $u'(d) > 0$ when $d > 0$. Let $p_i(t|z)$ denote the voters' revised beliefs that candidate i is of type t upon observing platform z. The expected utility of voter n when the winner is candidate i, who promises z_i, is $-p_i(M|z_i)u(|\chi_i^M(z_i) - x_n|) - (1 - p_i(M|z_i))u(|\chi_i^E(z_i) - x_n|)$. Voters vote sincerely, which means that they vote for the most preferred candidate, and weakly dominated strategies are ruled out. Assume that all voters and the opponent have the same beliefs about a candidate's type.

Let $\pi_i^t(z_i^t, z_j^s)$ denote the probability of candidate i of type t winning against opponent j of type s, given z_i^t and z_j^s. Let $F_i^t(.)$ denote the distribution function of the mixed strategy chosen by candidate i of type t. The expected utility of candidate i of type t who promises z_i^t is

$$\begin{aligned} & V_i^t((F_j^M(z_j^M), F_j^E(z_j^E)), z_i^t) \\ = & \sum_{s=M,E} \left[p^s \int_{z_j^s} \pi_i^t(z_i^t, z_j^s) dF_j^s(z_j^s) \right] \left[-v(|\chi_i^t(z_i^t) - x_i^t|) - c(|z_i^t - \chi_i^t(z_i^t)|) \right] \\ - & \sum_{s=M,E} p^s \int_{z_j^s} (1 - \pi_i^t(z_i^t, z_j^s)) v(|\chi_j^s(z_j^s) - x_i^t|) dF_j^s(z_j^s), \qquad (1) \end{aligned}$$

where $i, j = L, R$ and $t = M, E$.[12] The first term indicates when the candidate defeats

[11] This assumption means that the relative marginal cost (disutility) decreases as $|z_i^t - \chi|$ ($|x_i^t - \chi|$) increases, and the Arrow–Pratt measure of absolute risk aversion is decreasing in $|x_i^t - \chi|$.

[12] To simplify, I do not introduce a benefit from holding office (i.e., a positive utility from winning) in a

each type of opponent. The second term indicates when the candidate loses to each type of opponent. In summary, the timing of events is as follows.[13]

1. Nature decides each candidate's type, and a candidate knows his/her own type.
2. The candidates announce their platforms.
3. Voters vote.
4. The winning candidate chooses which policy to implement.

In what follows, I concentrate on a symmetric, pure-strategy, perfect Bayesian equilibrium consisting of strategies and beliefs.

3 Equilibrium

3.1 Policy Implemented by the Winner

Following an election, the winning candidate implements a policy that maximizes his/her utility following a win, $-v(|\chi_i^t(z_i^t) - x_i^t|) - c(|z_i^t - \chi_i^t(z_i^t)|)$.

Lemma 1 *The implemented policy $\chi_i^t(z)$ satisfies $v'(|\chi_i^t(z) - x_i^t|) = c'(|z - \chi_i^t(z)|)$, for $\chi_i^t(z) \in (x_i^t, z)$ when $z > x_i^t$, and $\chi_i^t(z) \in (z, x_i^t)$ when $z < x_i^t$.*

---

candidate's utility. A benefit from holding office would not change my results significantly when the benefit is small. Here, candidates approach the median policy more closely, but the main characteristics of the equilibria do not change. Therefore, I simply assume that the benefit from holding office is zero. However, if the benefits from holding office are great, candidates' implemented policies converge to the median policy regardless of type. This is what Huang (2010) shows by introducing sufficiently high benefits from holding office, which compensate for all disutility resulting from the policy and the cost of betrayal.

[13] I do not consider candidates' decisions to run, and the number of candidates (two) is given here. If an endogenous decision to run is introduced based on a citizen-candidate model (Besley and Coate, 1997; Osborne and Slivinski, 1996), there is a possibility that more than two candidates decide to run. When three or more candidates run, it is well known that there are many equilibria in a political competition, assuming completely binding platforms (e.g., see Adams, Merrill III, and Grofman, 2005). Such problems arise even when analyzing partially binding platforms. If it is simply supposed that there are only two *potential* candidates (with two possible types), both candidates have an incentive to run regardless of their type if the cost to run is sufficiently low, even if one candidate has a lower probability of winning. Indeed, Asako (forthcoming) shows that a candidate whose probability of winning is zero runs in order to change the opponent's (i.e., the winner's) policy with a partially binding platform and complete information. A similar implication can be obtained here. Thus, such endogenous decisions to run are not analyzed here, and I implicitly assume that the cost to run is sufficiently low to induce two candidates to run, but sufficiently high that it deters other citizens whose ideal policy is around the median policy from entering the race.

When the platform differs from the ideal policy, the implemented policy must differ from the platform or the ideal policy since it is decided by the winner (who no longer cares about the loser's platform, but cares about the cost of betrayal) after an election. The implemented policy will lie between the platform and the candidate's ideal policy, as shown in Figure 1. If voters know the candidate's type (ideal policy), they can also know the future implemented policy by observing the platform. However, with asymmetric information, they may not know the candidate's type. Here, the median voter x_m is pivotal. Thus, if the candidate is more attractive to the median voter than is the opponent, this candidate is certain to win.

3.2 Pooling Equilibrium

3.2.1 Definition and Proposition

If both types of opponent announce the same platform z_j (i.e., a pooling strategy), the expected utility of candidate i of type t when the opponent wins is $-p^M v(|\chi_j^M(z_j) - x_i^t|) - (1-p^M)v(|\chi_j^E(z_j) - x_i^t|)$, where $i, j = L, R$, $i \neq j$, and $t = M, E$. The utility of candidate i of type t after he/she wins is $-v(|\chi_i^t(z_i^t) - x_i^t|) - c(|z_i^t - \chi_i^t(z_i^t)|)$. If the utility when candidate i of type t wins is strictly lower than the expected utility when his/her opponent wins ($-p^M v(|\chi_j^M(z_j) - x_i^t|) - (1-p^M)v(|\chi_j^E(z_j) - x_i^t|) > -v(|\chi_i^t(z_i^t) - x_i^t|) - c(|z_i^t - \chi_i^t(z_i^t)|)$), candidate i prefers the opponent winning to winning him/herself. On the other hand, in the inverse case (i.e., $-p^M v(|\chi_j^M(z_j) - x_i^t|) - (1-p^M)v(|\chi_j^E(z_j) - x_i^t|) < -v(|\chi_i^t(z_i^t) - x_i^t|) - c(|z_i^t - \chi_i^t(z_i^t)|)$), candidate i prefers to win.

In a pooling equilibrium, a moderate type chooses a platform z_i^{M*} such that she is indifferent between her winning and the opponent winning. That is, the above two expected utilities are the same for a moderate type:

$$-p^M v(|\chi_j^M(z_j^{M*}) - x_i^M|) - (1-p^M)v(|\chi_j^E(z_j^{M*}) - x_i^M|) = -v(|\chi_i^M(z_i^{M*}) - x_i^M|) - c(|z_i^M - \chi_i^M(z_i^{M*})|), \quad (2)$$

where z_L^{M*} and z_R^{M*} are symmetric ($|x_m - z_L^{M*}| = |x_m - z_R^{M*}|$). An extreme type mimics a moderate type by choosing the same platform.[14]

[14] This section concentrates on a pooling equilibrium in which both types announce z_i^{M*}. The other possible pooling equilibria are discussed in Section 3.4.3.

Definition 1 *In a pooling equilibrium, z_i^{M*} is chosen by a candidate regardless of his/her type.*

Then, a pooling equilibrium exists if the prior belief that a candidate is of a moderate type, p^M, is sufficiently high. The parameter, $\overline{p}$, will be defined later (Equation (6)).

Proposition 1 *Suppose that the off-path beliefs of voters are $p_i(M|z_i) = 0$. If $p^M \geq \overline{p}$, the pooling equilibrium defined in Definition 1 exists.*

Proof See Appendix A.1.

3.2.2 An Extreme Type's Choice

Intuitively, a moderate type is indifferent between her winning and the opponent winning, with z_i^{M*}, so she does not deviate. On the other hand, this subsection shows that an extreme type prefers his winning to the opponent winning, with z_i^{M*}.

Let $z_i^t(z_j)$ denote the "*cut-off*" platform, where the utility when candidate i wins and the expected utility when opponent j wins are the same for type-t candidate i, given the opponent's pooling strategy, z_j. That is, a candidate is indifferent between his/her winning and the opponent winning:

$$-p^M v(|\chi_j^M(z_j)-x_i^t|)-(1-p^M)v(|\chi_j^E(z_j)-x_i^t|) = -v(|\chi_i^t(z_i^t(z_j))-x_i^t|)-c(|z_i^t(z_j)-\chi_i^t(z_i^t(z_j))|). \tag{3}$$

For example, if the opponent chooses z_j^{M*}, z_i^{M*} is the cut-off platform of moderate candidate i ($z_i^M(z_j^{M*}) = z_i^{M*}$). If a candidate approaches the median policy, the disutility after winning and the cost of betrayal increase. Thus, if type-t candidate i announces a platform that is further from his/her ideal policy (i.e., more moderate) than $z_i^t(z_j)$, then his/her utility after winning is lower than the expected utility when the opponent wins, and vice versa.

Note that when I use the term "more moderate platform," this means that "this platform is further from the candidate's ideal policy." In Figure 2, $z_R^E(z_L)$ is further from x_R^E and x_R^M than $z_R^M(z_L)$. Therefore, $z_R^E(z_L)$ is "more moderate" than $z_R^M(z_L)$. Note also that "approaching the median policy" means that a candidate announces a platform such that an implemented policy given this platform approaches the median policy.[15]

[15]However, first, a more moderate platform does not mean a more moderate implemented policy, given

[Figure 2 here]

If an extreme type's $z_i^E(z_j)$ is always more moderate than a moderate type's $z_i^M(z_j)$, given any z_j ($z_L^M(z_R) < z_L^E(z_R)$ and $z_R^E(z_L) < z_R^M(z_L)$), the extreme type prefers his winning to the opponent winning when both candidates announce z_i^{M*} and z_j^{M*}. The following lemma shows that this is always true. Figure 2 demonstrates the lemma in the case of R.

Lemma 2 *Suppose that an opponent announces the same platform z_j regardless of type. Given any p^M, (i) an extreme type's cut-off platform is more moderate than that of a moderate type ($z_L^M(z_R) < z_L^E(z_R)$ and $z_R^E(z_L) < z_R^M(z_L)$), but (ii) an extreme type's implemented policy, given the cut-off platform, is more extreme than that of a moderate type ($\chi_L^E(z_L^E(z_R)) > \chi_L^M(z_L^M(z_R))$ and $\chi_R^M(z_R^M(z_L)) > \chi_R^E(z_R^E(z_L))$).*

Proof See Appendix A.2.

Lemma 2 shows the following two facts.

1. An extreme type has an incentive to announce a more moderate platform than does a moderate type.

2. A moderate type has an incentive to commit to implementing a more moderate implemented policy than does an extreme type.

If a candidate approaches the median policy and wins against the opponent with certainty, this candidate will pay a cost of betrayal with certainty. This marginal cost of approaching the median policy depends on the cost of betrayal, $-c(|z_i^t - \chi_i^t(z_i^t)|)$. On the other hand, this candidate can avoid the opponent's victory and decrease his/her expected disutility from the implemented policy. This marginal benefit depends on the difference in the (expected) disutilities when the candidate wins and when the opponent wins, $p^M v(|\chi_j^M(z_j) - x_i^t|) + (1 - p^M)v(|\chi_j^E(z_j) - x_i^t|) - v(|\chi_i^t(z_i^t) - x_i^t|)$.

Following an election, an extreme type will betray the platform more severely and pay a higher cost of betrayal. However, at the same time, the ideal policy for an extreme type

this platform. As shown in Figure 2, since an extreme type will betray his platform to a greater extent than will a moderate type, the extreme type's implemented policy, $\chi_R^E(z_R^E(z_L))$, is more extreme than that of the moderate type, $\chi_R^M(z_R^M(z_L))$. Second, a more moderate platform may not mean that this platform is closer to the median policy, because there is a possibility that a platform encroaches on the opponent's side of the policy space (i.e., $z_R^t < x_m < z_L^t$). In Figure 2, if both types' platforms encroach on the opponent's side, $z_R^E(z_L)$ is further from the median policy than $z_R^M(z_L)$. See Appendix C for more detail.

is further from the median policy than that of a moderate type, which means that his ideal policy is also further from the opponent's implemented policy. Thus, an extreme type has a higher disutility from the opponent's victory. Therefore, he finds it especially costly if the opponent wins, more so than a moderate type does. As a result, an extreme type has higher marginal benefit and cost than a moderate type. Since an extreme type has a higher marginal cost of betrayal, he does not have an incentive to choose a more moderate *implemented policy* than a moderate type (Lemma 2-(ii)). However, an extreme type has a higher marginal benefit, so he has an incentive to announce a more moderate *platform* than a moderate type (Lemma 2-(i)).

Therefore, when both candidates announce a pooling strategy, z_i^{M*}, an extreme type prefers his winning to the opponent winning. Thus, an extreme type does not have an incentive to lose. However, there is a possibility that an extreme type has an incentive to further approach the median policy and win with certainty, since he prefers his winning to the opponent winning at z_i^{M*}. To check this incentive of an extreme type, the off-path belief of voters must be specified.

3.2.3 Beliefs

Suppose that when a candidate deviates from z_i^{M*}, voters believe with a probability of one that the candidate is an extreme type. That is, I consider the simple off-path beliefs $p_i(M|z_i) = 0$. This simple off-path belief is *partially* based on the idea of the intuitive criterion in Cho and Kreps (1987). In a pooling equilibrium, a moderate type is indifferent to her winning and the opponent winning at z_i^{M*}. Thus, a moderate type never chooses a more moderate platform than z_i^{M*} since her expected payoff decreases from the equilibrium expected-payoff even if voters believe this candidate to be a moderate type as a result of this deviation.[16] As a result, if a platform is announced that is more moderate than z_i^{M*}, this candidate must not be a moderate type. On the other hand, an extreme type has an incentive to announce a platform more extreme than his cut-off platform, given z_j^{M*} (i.e., $z_i^E(z_j^{M*})$). Lemma 2 shows $z_i^E(z_j^{M*})$ is more moderate than z_i^{M*}, so the off-path belief for a

[16] To be precise, her equilibrium expected-payoff is $(1/2)[-p^M v(|\chi_j^M(z_j^{M*}) - x_i^M|) - (1-p^M)v(|\chi_j^E(z_j^{M*}) - x_i^M|) - v(|\chi_i^M(z_i^{M*}) - x_i^M|) - c(|z_i^M - \chi_i^M(z_i^{M*})|)] = -v(|\chi_i^M(z_i^{M*}) - x_i^M|) - c(|z_i^M - \chi_i^M(z_i^{M*})|)$ since (2) holds. If a moderate type deviates by choosing any platform z_i which is more moderate than z_i^{M*}, and she wins with certainty, her expected payoff is $-v(|\chi_i^M(z_i) - x_i^M|) - c(|z_i - \chi_i^M(z_i)|) < -v(|\chi_i^M(z_i^{M*}) - x_i^M|) - c(|z_i^M - \chi_i^M(z_i^{M*})|)$.

platform that lies between z_i^{M*} and $z_i^E(z_j^{M*})$ should be $p_i(M|z_i) = 0$. The intuitive criterion cannot apply to any other off-path beliefs, so I simply assume that $p_i(M|z_i) = 0$ for all off-path strategies.

3.2.4 The Existence of the Pooling Equilibrium

Voters do not know candidates' types in a pooling equilibrium, so their expected utility is the weighted average of the utility between a moderate and an extreme type. On the other hand, if an extreme type deviates by approaching the median policy, voters believe that this candidate's type is extreme because of the off-path belief. If an extreme type deviates to announce a sufficiently moderate platform, this extreme type can win over an uncertain opponent who chooses z_i^{M*}. Here, I denote z_i' such that it satisfies

$$-u(|\chi_i^E(z_i') - x_m|) = -p^M u(|\chi_j^M(z_j^{M*}) - x_m|) - (1 - p^M)u(|\chi_j^E(z_j^{M*}) - x_m|), \qquad (4)$$

The left-hand side is the utility of the median voter when extreme candidate i, who deviates to z_i' wins. The right-hand side is the expected utility of the median voter when candidate j, who announces the pooling platform z_j^{M*} wins. That is, the median voter is indifferent between candidates who announce z_i' and z_j^{M*}. If an extreme candidate announces a platform that is slightly more moderate than z_i', this candidate wins over an uncertain opponent. Figure 3(a) shows z_i' in the case of R when voters have linear utility. Note that because voters are uncertain about the type of the candidate who announces z_i^{M*}, an extreme type who deviates needs to implement a more moderate policy than an extreme type who chooses a pooling platform ($\chi_j^E(z_j^{M*})$), but does not need to implement a more moderate policy than a moderate type ($\chi_j^M(z_j^{M*})$), as shown in Figure 3(a).

[Figure 3 here]

An extreme type does not deviate by announcing a platform that is more moderate than z_i' if

$$\begin{aligned}&-v(|\chi_i^E(z_i') - x_i^E|) - c(|z_i' - \chi_i^E(z_i')|) \leq \frac{1}{2}\big[-p^M v(|\chi_j^M(z_j^{M*}) - x_i^E|) \\ &-(1 - p^M)v(|\chi_j^E(z_j^{M*}) - x_i^E|) - v(|\chi_i^E(z_i^{M*}) - x_i^E|) - c(|z_i^{M*} - \chi_i^E(z_i^{M*})|)\big]. \qquad (5)\end{aligned}$$

The right-hand side is the extreme candidate i's expected utility when he stays in a pooling equilibrium. His expected utility from this deviation is slightly lower than the left-hand side. If (5) holds, this extreme type does not deviate, so a pooling equilibrium where all types announce z_i^{M*} exists. Now, denote

$$\overline{p} \equiv \frac{\left[\begin{array}{c} v(|\chi_j^E(z_j^{M*}) - x_i^E|) + v(|\chi_i^E(z_i^{M*}) - x_i^E|) + c(|z_i^{M*} - \chi_i^E(z_i^{M*})|) \\ -2\left[v(|\chi_i^E(z_i') - x_i^E|) + c(|z_i' - \chi_i^E(z_i')|)\right] \end{array}\right]}{v(|\chi_j^E(z_j^{M*}) - x_i^E|) - v(|\chi_j^M(z_j^{M*}) - x_i^E|)}. \tag{6}$$

Then, condition (5) can be represented by $p^M \geq \overline{p}$, as shown in Proposition 1. This $\overline{p}$ is always positive and less than one.

Corollary 1 $\overline{p} \in (0, 1)$

Proof See Appendix A.3.

If $p^M \geq \overline{p}$, a pooling equilibrium exists. The next section shows that if $p^M < \overline{p}$, a semi-separating equilibrium exists. Thus, the previous corollary implies that for all parameter values and functional forms, if p^M is large enough, a pooling equilibrium always exists; otherwise, a semi-separating equilibrium exists.

The intuitive reasoning is as follows. Suppose that L chooses z_L^{M*} as a pooling equilibrium, and R is an extreme type who originally announces z_R^{M*}. If p^M is high, extreme type R needs to announce a very moderate platform to win with certainty, because the expected utility to the median voter of choosing L is quite high given that there is a strong possibility that L is moderate and will implement a moderate policy. Thus, as in Figure 3(b), there is a significant distance between z_R' and R's ideal policy, so this deviation decreases his expected utility. As a result, a pooling equilibrium exists. However, if p^M is sufficiently low, the expected utility of the median voter who chooses L is quite low. Thus, if R slightly approaches the median policy, the policy he implements improves for the median voter. That is, z_R' is closer to z_R^{M*}, as in Figure 3(c), so the extreme type will deviate from a pooling strategy.[17]

[17] Comparative statics for $\overline{p}$ are not obvious and depend on the parameter values. This is because a change of one parameter affects all variables, such that z_i' (mainly decided by voters' utility functions), z_i^{M*}, $\chi_i^M(z_i^{M*})$ (mainly decided by a moderate type's utility and cost functions), and $\chi_i^E(z_i^{M*})$ (mainly decided by an extreme type's utility and cost functions). For example, if a moderate type's ideal policy becomes more moderate, she will announce a more extreme platform z_i^{M*}, and implement a more moderate policy $\chi_i^M(z_i^{M*})$ in a pooling equilibrium, as per Lemma 2. Thus, z_i' may become more moderate since a moderate type will implement a more moderate policy. However, since z_i^{M*} becomes more extreme, $\chi_i^E(z_i^{M*})$ will also

Note that this pooling equilibrium exists in the broader value of the off-path beliefs. For example, suppose $p_i(M|z_i) = p^M$ if the platform is more extreme than z_i^{M*}. Then a candidate still has no incentive to deviate to a more extreme platform than z_i^{M*}, since he/she will then be certain to lose and the expected utility decreases or is unchanged by this deviation. A pooling equilibrium can exist when the off-path belief, $p_i(M|z_i)$, is lower than p^M for z_i, which is more extreme than z_i^{M*}.

3.3 Semi-separating Equilibrium

If $p^M < \overline{p}$, a pooling equilibrium does not exist since an extreme type will deviate from z_i^{M*} to z_i'. In this case, a moderate type still announces one platform with certainty (a pure strategy). On the other hand, an extreme type chooses a mixed strategy: With some probability, an extreme type announces the same platform as a moderate type (a *pooling extreme type*). With the remaining probability, an extreme type approaches the median policy (a *separating extreme type*). That is, if $p^M < \overline{p}$, a semi-separating equilibrium exists.

3.3.1 Definition and Proposition

In a semi-separating equilibrium, a moderate type announces z_i^*, and an extreme type announces z_i^* with probability $\sigma^M \in (0,1)$, and $\bar{z}_i$ with the remaining probability, $1-\sigma^M$. I call this type of a semi-separating equilibrium a *two-policy semi-separating equilibrium.* This equilibrium is shown in Figure 4(a) and is defined as follows.

Definition 2 *In a two-policy semi-separating equilibrium, a moderate type chooses z_i^*, and an extreme type chooses z_i^* with probability $\sigma^M \in (0,1)$, and $\bar{z}_i$ with probability $1-\sigma^M$.*

[Figure 4 here]

Now, denote z_i^* (and z_j^*) such that it satisfies

$$\begin{aligned} & -\frac{p^M}{p^M+\sigma^M(1-p^M)}v(|\chi_j^M(z_j^*)-x_i^M|) - \frac{\sigma^M(1-p^M)}{p^M+\sigma^M(1-p^M)}v(|\chi_j^E(z_j^*)-x_i^M|) \\ = & -v(|\chi_i^M(z_i^*)-x_i^M|) - c(|z_i^* - \chi_i^M(z_i^*)|)). \end{aligned} \quad (7)$$

be more extreme, which may induce z_i' to be more extreme. Moreover, if voters are sufficiently risk averse, when the distance between $\chi_i^E(z_i^{M*})$ and $\chi_i^M(z_i^{M*})$ increases, z_i' may become more extreme, since it becomes more risky to vote for a candidate of uncertain type who chooses a pooling strategy. Thus, the effect of the position of x_i^M on $\overline{p}$ is unclear, as it depends on voters' utility functions and the utility and cost functions of the moderate and extreme candidates.

where z_i^* and z_j^* are symmetric ($|x_m - z_i^*| = |x_m - z_j^*|$). The left-hand side is the expected utility of moderate candidate i when the opponent promising z_j^* wins. The right-hand side is the utility of moderate candidate i when she wins. That is, a moderate type is indifferent between her winning and the opponent who wins and announces z_j^*.[18]

Next, denote $\bar{z}_i$ such that it satisfies

$$-u(|\chi_i^E(\bar{z}_i)-x_m|) > -\frac{p^M}{p^M+\sigma^M(1-p^M)}u(|\chi_j^M(z_j^*)-x_m|)-\frac{\sigma^M(1-p^M)}{p^M+\sigma^M(1-p^M)}u(|\chi_j^E(z_j^*)-x_m|). \tag{8}$$

That is, the median voter prefers $\bar{z}_i$ to z_j^* when $\bar{z}_i$ is announced by an extreme type. The right-hand side is the expected utility of the median voter when candidate j wins and announces z_j^*. The left-hand side is the expected utility of the median voter when extreme candidate i wins and announces $\bar{z}_i$. Moreover, $\bar{z}_i$ denotes the most extreme platform that satisfies (8).[19]

Then, the following proposition is derived.

Proposition 2 *Suppose that the off-path beliefs of voters are $p_i(M|z_i) = 0$. If $p^M < \overline{p}$, a semi-separating equilibrium exists in which a moderate type chooses z_i^*.*

Proof See Appendix A.4.

3.3.2 Each Player's Choice

The details of the semi-separating equilibrium are given in Appendix A.4, so I provide the intuitive reasoning in this section. Suppose again that when a candidate deviates from the equilibrium platform, voters believe with a probability of one that the candidate is an extreme type. That is, $p_i(M|z_i) = 0$. This off-path belief is also partially based on the idea of the intuitive criterion. Since a moderate type is indifferent between her winning and the opponent winning (and announcing z_j^*), a moderate type never chooses a platform that is more moderate than z_i^*. On the other hand, from Lemma 2, an extreme type prefers his

[18] There exist other semi-separating equilibria, as discussed in Section 3.4.3. The difference between the equilibrium in this subsection and the others is only that a moderate type chooses another platform. Thus, the basic characteristics are the same.

[19] More precisely, because the policy space is continuous, there is no maximal (minimal) value of z_R (z_L) that satisfies (8). Instead, it is possible to define $\bar{z}_i$ such that a platform satisfies (8) with equality, and to assume that if an extreme type announces $\bar{z}_i$, he defeats an opponent who announces z_j^*. It is also possible to suppose that a policy space is discrete with a grid of policies. That is, there are a large number of policy choices, and the distance between sequential policies is ϵ. If ϵ is very close to zero, thus approximating a continuous policy space, then there exists a most extreme platform that satisfies (8). The following results do not change given these settings.

winning to the winning of an opponent who announces z_j^* when he announces z_i^*.[20] Thus, an extreme type has an incentive to choose a more moderate platform (until his cut-off platform, given z_j^*). The intuitive criterion cannot apply to any other off-path beliefs, so I simply assume that $p_i(M|z_i) = 0$ for all off-path strategies.

The following summarizes each player's rational choice.

Voters: From the definition of $\bar{z}_i$ (Inequality (8)), an extreme type who announces $\bar{z}_i$ can win against an opponent who announces z_j^* (and tie with an opponent who announces $\bar{z}_j$). Suppose that candidate R (an extreme type) announces $\bar{z}_R$, while candidate L announces z_L^*. Voters can know that the type of R is extreme, but remain uncertain about the type of L who announces z_L^*, because an extreme type L will still pretend to be moderate and announce z_L^* with probability σ^M. Therefore, to defeat L (i.e., to satisfy (8)), R does not need to implement a more moderate policy than a moderate type L. That is, $x_m - \chi_L^E(z_L^*) > \chi_R^E(\bar{z}_R) - x_m > x_m - \chi_L^M(z_L^*)$. In other words, for the median voter, a moderate type L will implement the best policy ($\chi_L^M(z_L^*)$). However, if L wins, there is the possibility that L is an extreme type who implements the worst policy for the median voter ($\chi_L^E(z_L^*)$). Thus, the median voter forgoes the chance of electing a moderate type L to avoid electing an extreme type L, and chooses the second best candidate, R, who is a separating extreme type.

A Moderate Type: As a result of this off-path belief, a moderate type does not have an incentive to deviate from z_i^*. If a moderate type deviates to a more extreme platform than z_i^* or $z_i \in (\bar{z}_i, z_i^*)$, she will be certain to lose and her expected utility will remain unchanged as she is indifferent between her winning and the opponent winning at z_i^*. A moderate type does not have an incentive to approach the median policy by more than $\bar{z}_i$ since her utility from a winning will then become lower than her utility when the opponent wins from Lemma 2.

An Extreme Type: An extreme type prefers his winning to the opponent winning and announcing z_j^* when he announces z_i^*. Thus, he does not have an incentive to deviate to a more extreme platform than z_i^* or $z_i \in (\bar{z}_i, z_i^*)$ since he will be certain to lose and his expected utility will decrease. When an extreme type announces $\bar{z}_i$, his disutility following a

[20] Replace p^M by $p^M/[p^M + \sigma^M(1-p^M)]$ in (3). This result can then be derived in the same way as Lemma 2.

win and the cost of betrayal are higher, but the probability of winning is greater than when he announces z_i^*. Thus, an extreme type can be indifferent between $\bar{z}_i$ and z_i^* when $p^M < \overline{p}$.

3.3.3 The Existence of the Semi-separating Equilibrium

A two-policy semi-separating equilibrium exists if $p^M < \overline{p}$ and

$$-v(|\chi_i^E(\bar{z}_i) - x_i^E|) - c(|\bar{z}_i - \chi_i^E(\bar{z}_i)|) \leq -v(|\chi_j^E(\bar{z}_j) - x_i^E|). \tag{9}$$

When (9) holds, an extreme type has no incentive to defeat with certainty an extreme opponent who announces $\bar{z}_j$ by approaching the median policy. This is because (9) means that, for extreme candidate i, the utility when extreme opponent j who announces $\bar{z}_j$ wins is higher than the utility when i wins. However, an extreme type with $\bar{z}_i$ does not want to deviate to a more extreme platform because that would mean he would also lose to an opponent with z_j^* and his expected utility would decrease.

If (9) does not hold, an extreme type still has an incentive to converge by more than $\bar{z}_i$ to beat an extreme opponent who announces $\bar{z}_j$. Therefore, a two-policy semi-separating equilibrium does not exist, but a *continuous semi-separating equilibrium* does exist. In a continuous semi-separating equilibrium, an extreme type's mixed strategy includes z_i^* and a connected support, $[\bar{z}_L, \underline{z}_L]$ for L and $[\underline{z}_R, \bar{z}_R]$ for R, as shown in Figure 4(b). An extreme type chooses any platform in this support with probability $1 - \sigma^M$, and has a continuous distribution function, $F(.)$, within the support. More specifically, the distribution is $(1 - \sigma^M)F(.)$. Platform $\bar{z}_i$ is defined in the same way as $\bar{z}_i$ in a two-policy semi-separating equilibrium, so the basic results are the same as those of a two-policy semi-separating equilibrium. That is, a separating extreme type defeats an uncertain type (a moderate type and a pooling extreme type).

A semi-separating equilibrium exists in the broader value of the off-path beliefs. For example, suppose $p_i(M|z_i) = p^M/(p^M + \sigma^M(1 - p^M))$ if the platform is more extreme than z_i^*. Then a candidate still has no incentive to deviate to a more extreme platform than z_i^*, since he/she will be certain to lose and the expected utility decreases or is unchanged by this deviation. Thus, a semi-separating equilibrium can exist when the off-path belief $p_i(M|z_i)$ is lower than $p^M/(p^M + \sigma^M(1 - p^M))$ for z_i, which is more extreme than z_i^*.

3.4 Other Possible Equilibria

The previous sections only discuss a pooling equilibrium in which both types choose z_i^{M*} and a semi-separating equilibrium in which a moderate type chooses z_i^*. This section discusses other equilibria that may exist.

3.4.1 Separating Equilibrium in which a Moderate Type Wins

If a separating equilibrium exists in which a moderate type wins against an extreme type, both types must choose different platforms, and the moderate type must implement a more moderate policy than does the extreme type. However, such a separating equilibrium does not exist.

Proposition 3 *There is no separating equilibrium in which a moderate type wins against an extreme type, regardless of off-path beliefs.*

Proof See Appendix A.5.

This result is true because an extreme type always has an incentive to pretend to be moderate.

If a separating equilibrium exists in which a moderate type wins, the moderate type must prefer her winning to the opponent winning in this equilibrium. However, this means an extreme type also prefers to win over having the opponent win at the moderate type's platform, for the same reasoning given in Lemma 2. Moreover, if an extreme type deviates by pretending to be a moderate type, he will have a higher probability of winning. Thus, an extreme type has an incentive to deviate by choosing the moderate type's platform within this separating strategy.

This result contradicts that of Banks (1990) and Callander and Wilkie (2007) who show that there exists a separating equilibrium in which a moderate type wins.

3.4.2 Separating Equilibrium in which an Extreme Type Wins

A separating equilibrium in which an extreme type wins against a moderate type exists if off-path beliefs are $p_i(M|z_i) = 0$. For example, suppose that an extreme type announces a platform, $\hat{z}_i^E$, such that he is indifferent between his winning and an extreme-type opponent winning. To be precise, $\hat{z}_i^t$ satisfies $-v(|\chi_i^t(\hat{z}_i^t) - x_i^t|) - c(|\hat{z}_i^t - \chi_i^t(\hat{z}_i^t)|) = -v(|\chi_j^t(\hat{z}_j^t) - x_i^t|)$

where $|x_m - \hat{z}_i^t| = |x_m - \hat{z}_j^t|$. Suppose also that a moderate type announces $z_i^{M\prime}$, such that the moderate type will implement a more extreme policy than the extreme type. That is, $|x_m - \chi_i^E(\hat{z}_i^E)| < |x_m - \chi_i^M(z_i^{M\prime})|$. As a result of the above off-path beliefs, although a moderate type approaches the median policy, voters believe that this candidate is an extreme type. To increase the probability of winning, a moderate type needs to approach the median policy in a significant way. This may decrease her expected utility. An extreme type does not deviate either, because he is indifferent between winning and losing. As a result, a separating equilibrium in which an extreme type wins exists.

However, this separating equilibrium is less important, since it exists in a very restricted set of off-path beliefs. For example, suppose that the off-path beliefs are $p_i(M|z_i) = p^M$ if the platform is more extreme than $\hat{z}_i^M$. Then a moderate type has an incentive to choose $\hat{z}_i^M$, where she is indifferent between her winning and her opponent winning. If a moderate type announces $\hat{z}_i^M$, an extreme type does not have an incentive to win against a moderate type by committing to implement $\chi_i^E(z_i^E)$, which is more moderate than $\chi_i^M(\hat{z}_i^M)$, from Lemma 2. Therefore, this separating equilibrium does not exist with such off-path beliefs.

3.4.3 Other Pooling and Semi-separating Equilibria

In the previous sections, I assume off-path beliefs as $p_i(M|z_i) = 0$. Under this assumption, there exist multiple pooling and semi-separating equilibria. First, there could be a pooling equilibrium in which both types announce a platform, say z_i^{M**}, that is more extreme than z_i^{M*} ($z_L^{M**} < z_L^{M*}$ and $z_R^{M*} < z_R^{M**}$). Since the off-path belief is $p_i(M|z_i) = 0$, a moderate type needs to approach the median policy significantly to be sure of winning because voters believe that the candidate is extreme when a he/she deviates from z_i^{M**}, regardless of the real type. Thus, a moderate type may not want to deviate. If an extreme type also has no incentive to deviate, a pooling equilibrium with z_i^{M**} exists. Second, there could be a semi-separating equilibrium in which a moderate type (and a pooling extreme type) announces a platform, say z_i^{**}, that is more extreme than z_i^*. A moderate type needs to approach the median policy significantly to win against an opponent who announces z_j^{**}, because this deviation leads voters to believe that the candidate is extreme. Therefore, a moderate type may have no incentive to deviate from z_i^{**}.

These equilibria (including a separating equilibrium in which an extreme type wins) have several problems. First, a moderate type does not want to approach the median policy because voters will likely *misunderstand* the candidate to be an extreme type. Second, if the off-path beliefs, $p_i(M|z_i)$, exceed zero for some off-path platforms, many of the above equilibria will be eliminated.[21] Thus, these equilibria exist for restricted values of off-path beliefs.

On the other hand, the equilibria analyzed in the previous sections exist in broader values of off-path beliefs than do other equilibria. As I discussed, a pooling equilibrium with z_i^{M*} exists when $p_i(M|z_i) = p^M$ if the platform is more extreme than z_i^{M*}. A semi-separating equilibrium with z_i^* exists when $p_i(M|z_i) = p^M/[p^M + \sigma^M(1-p^M)]$ if the platform is more extreme than z_i^*.

3.5 Differences from Past Papers

3.5.1 Assumptions

As I indicated in the introduction, there are two important differences between my study and those of Banks (1990) and Callander and Wilkie (2007): (1) candidates choose a policy to implement strategically and (2) they care about policy when they lose.

In a semi-separating equilibrium, an extreme type reveals his type by approaching the median policy with some probability since he can obtain a higher probability of winning. However, if candidates implement their own ideal policies automatically, voters only believe that an extreme type will implement his ideal policy, so a separating extreme type cannot increase his probability of winning by revealing his type. That is, a strategic choice of an implemented policy provides a *way* to win for an extreme type.

Suppose that a candidate does not care about policy when he loses. From (1), the expected utility of candidates is

$$\sum_{s=M,E}\left[p^s\int_{z_j^s}\pi_i^t(z_i^t,z_j^s)dF_j^s(z_j^s)\right]\left[-v(|\chi_i^t(z_i^t)-x_i^t|)-c(|z_i^t-\chi_i^t(z_i^t)|)\right].$$

Obviously, candidates will announce their ideal policy as platforms (and his/her expected utility is zero) because if not, they will incur disutility from the policy and a cost of betrayal, and the expected utility becomes negative. Thus, benefits from holding office should be introduced to induce candidates to approach the median policy. However, even with benefits

[21] For example, suppose that off-path beliefs are $p_i(M|z_i) = p^M$ $(p^M/[p^M + \sigma^M(1-p^M)])$ if the platform is more extreme than z_i^{M*} (z_i^*), a moderate type has an incentive to compromise until z_i^{M*} (z_i^*).

from holding office, a semi-separating equilibrium does not exist when candidates do not care about policy after losing. An extreme type has a stronger incentive to prevent the opponent from winning in my model, but to a lesser extent when they do not care about an opponent's policy. Thus, caring about policy after losing provides an extreme type with an *incentive* to win.

3.5.2 Universal Divinity

Banks (1990), Callander and Wilkie (2007), and Huang (2010) employ universal divinity, as introduced by Banks and Sobel (1987). If universal divinity is applied to my model, in short, as Lemma 2 shows, an extreme type always has a greater incentive to announce a more moderate *platform* than does a moderate type. This means that a moderate type always has a stronger incentive to announce a more extreme platform than does an extreme type. Suppose a pooling equilibrium at z_i^{M*}. With universal divinity, if a platform is more moderate than z_i^{M*}, $p_i(M|z_i) = 0$. If not, $p_i(M|z_i) = 1$. With these off-path beliefs, both have an incentive to deviate to a more extreme platform than z_i^{M*}, be thought a moderate type by voters, and win. For the same reasons, a semi-separating equilibrium does not exist, and only a separating equilibrium exists in which an extreme type wins against a moderate type, which is discussed in Section 3.4.2. However, such a separating equilibrium seems peculiar, as I discussed. Moreover, in this separating equilibrium, an extreme type always wins against a moderate type, so my main result does not change.[22]

4 Conclusion

This study examined how an extreme candidate wins against a moderate candidate to provide one reason to have political polarization, based on the model of partially binding platforms with asymmetric information. There are two main equilibria, namely semi-separating and pooling, and voters cannot determine a candidate's political preferences in any equilibrium. In a pooling equilibrium, an extreme candidate imitates a moderate candidate by announcing the same platform as the moderate candidate. In a semi-separating equilibrium, an extreme candidate imitates a moderate candidate with some probability, and with the

[22] Criteria D1 and D2 in Cho and Kreps (1987) have the same result as universal divinity. On the other hand, divinity in Banks and Sobel (1987) cannot reduce to an equilibrium because it supports $p_i(M|z_i) = 0$ for any off-path platform.

remaining probability reveals his own preferences by approaching the median policy. An extreme candidate who reveals his preference type will defeat an unknown candidate who may be moderate or extreme, even though this extreme candidate will implement a more extreme policy than the moderate candidate. This is because voters wish to avoid electing an extreme type who imitates a moderate candidate and will implement the most extreme policy. As a result, an extreme candidate has a higher expected probability of winning than a moderate candidate. The important reason for this result is that an extreme candidate has a stronger incentive to prevent an opponent from winning because his ideal policy is further from the opponent's policy than is the ideal policy of a moderate candidate.

More work is needed to investigate these findings. In particular, this study assumes that candidates are symmetric, and that there are only two types of candidates. However, in reality, these characteristics may differ. First, voters may have asymmetric degrees of uncertainty about candidates. For example, in an election between an incumbent and a challenger, an incumbent's type is usually less uncertain. Second, the cost of betrayal should be derived endogenously. For example, if the candidate is more senior and has an incumbent advantage, he or she may not care about the next election. In this case, such a candidate will not care as much about the cost of betrayal. Third, candidates' ideal positions may not be symmetric, and one candidate's ideal policy may be closer to the median policy than the ideal policies of other candidates.[23]

To show the basic implication of the findings, namely the advantage of an extreme type, this study employs the simplest settings. As shown, partially binding platforms induce different predictions and, therefore, applications of a model of partially binding platforms are interesting subjects for future research. For such applications, some assumptions need to be relaxed or added, but these depend on the nature of the application. As a result, this study concentrates only on the simplest case.

[23] Asako (forthcoming) analyzes candidates who have asymmetric characteristics (positions of ideal policies, costs of betrayal, and policy preferences) with partially binding platforms, but with complete information.

A Proofs

A.1 Proposition 1

First, a candidate has no incentive to deviate to a more extreme platform than z_i^{M*} or $z_i \in [z_i', z_i^{M*})$ because off-path beliefs are $p_i(M|z_i) = 0$. Hence, this candidate will be certain to lose and the expected utility will decrease for an extreme type (from Lemma 2) and will remain unchanged for a moderate type (from the definition of z_i^{M*}). A moderate type does not have an incentive to approach the median policy by more than z_i', because her utility from winning will become lower than her utility if her opponent wins (from the definition of z_i^{M*}). An extreme type also does not have an incentive to deviate by approaching the median policy by more than z_i' when $p^M \geq \bar{p}$, from the discussion in Section 3.2.4. As a result, this is an equilibrium in which both candidates announce z_i^{M*}. $\square$

A.2 Lemma 2

Consider a case of R, without loss of generality. Let $\chi_R^t = \chi_R^t(z_R^t(z_L))$ denote the situation where the utility when type t R wins is the same as the expected utility for type t R when L wins, given z_L. This means that

$$p^M v(x_R^t - \chi_L^M) + (1 - p^M) v(x_R^t - \chi_L^E) - v(x_R^t - \chi_R^t) = c(\chi_R^t - z_R^t(\chi_R^t)), \tag{10}$$

where $z_R^t(\chi_R^t)$ represents the platform in which the candidate implements χ_R^t. First, I prove the second statement.

(ii) Differentiate both sides of (10) with respect to x_R^t, given the opponent's strategies ($\chi_L^M = \chi_L^M(z_L)$ and $\chi_L^E = \chi_L^E(z_L)$). Then,

$$\begin{aligned} & p^M v'(x_R^t - \chi_L^M) + (1 - p^M) v'(x_R^t - \chi_L^E) - v'(x_R^t - \chi_R^t)\left(1 - \frac{\partial \chi_R^t}{\partial x_R^t}\right) \\ = \; & c'(\chi_R^t - z_R^t(\chi_R^t))\left(\frac{\partial \chi_R^t}{\partial x_R^t} - \left(\frac{\partial z_R^t(\chi_R^t)}{\partial x_R^t} + \frac{\partial z_R^t(\chi_R^t)}{\partial \chi_R^t}\frac{\partial \chi_R^t}{\partial x_R^t}\right)\right). \end{aligned} \tag{11}$$

From Lemma 1, $v'(x_R^t - \chi_R^t) = c'(\chi_R^t - z_R^t(\chi_R^t))$. Moreover, fix χ_R^t and differentiate $v'(x_R^t -$

$\chi_R^t) = c'(\chi_R^t - z_R^t(\chi_R^t))$ with respect to x_R^t. Then,

$$\frac{\partial z_R^t(\chi_R^t)}{\partial x_R^t} = -\frac{v''(x_R^t - \chi_R^t)}{c''(\chi_R^t - z_R^t(\chi_R^t))} < 0.$$

Substitute these values into (11), which then becomes

$$\frac{\partial \chi_R^t}{\partial x_R^t} = \frac{\frac{v''(x_R^t - \chi_R^t)c'(\chi_R^t - z_R^t(\chi_R^t))}{c''(\chi_R^t - z_R^t(\chi_R^t))} - (p^M v'(x_R^t - \chi_L^M) + (1-p^M)v'(x_R^t - \chi_L^E) - v'(x_R^t - \chi_R^t))}{v'(x_R^t - \chi_R^t)\frac{\partial z_R^t(\chi_R^t)}{\partial \chi_R^t}}. \tag{12}$$

If (12) is positive, an extreme type will implement a more extreme policy than will a moderate type. In the same way as $\partial z_R^t(\chi_R^t)/\partial x_R^t$ was derived,

$$\frac{\partial z_R^t(\chi_R^t)}{\partial \chi_R^t} = 1 + \frac{v''(x_R^t - \chi_R^t)}{c''(\chi_R^t - z_R^t(\chi_R^t))} > 0,$$

so the denominator of (12) is positive. To prove that (12) is positive, it is sufficient to show that the numerator of (12) is positive. In other words,

$$\frac{p^M v'(x_R^t - \chi_L^M) + (1-p^M)v'(x_R^t - \chi_L^E) - v'(x_R^t - \chi_R^t)}{v''(x_R^t - \chi_R^t)} < \frac{c'(\chi_R^t - z_R^t(\chi_R^t))}{c''(\chi_R^t - z_R^t(\chi_R^t))}. \tag{13}$$

Note that, from (10) and Lemma 1,

$$\frac{p^M v(x_R^t - \chi_L^M) + (1-p^M)v(x_R^t - \chi_L^E) - v(x_R^t - \chi_R^t)}{v'(x_R^t - \chi_R^t)} = \frac{c(\chi_R^t - z_R^t(\chi_R^t))}{c'(\chi_R^t - z_R^t(\chi_R^t))}. \tag{14}$$

Since I assume that $c'(d)/c(d)$ strictly decreases as d increases, $c'(\chi_R^t - z_R^t(\chi_R^t))/c''(\chi_R^t - z_R^t(\chi_R^t)) > c(\chi_R^t - z_R^t(\chi_R^t))/c'(\chi_R^t - z_R^t(\chi_R^t))$. The right-hand side of (13) is greater than the right-hand side of (14). Therefore, if the left-hand side of (13) is less than the left-hand side of (14), (13) holds. This means

$$\begin{aligned} &p^M\left(\frac{v'(x_R^t - \chi_L^M)}{v''(x_R^t - \chi_R^t)} - \frac{v(x_R^t - \chi_L^M)}{v'(x_R^t - \chi_R^t)}\right) + (1-p^M)\left(\frac{v'(x_R^t - \chi_L^E)}{v''(x_R^t - \chi_R^t)} - \frac{v(x_R^t - \chi_L^E)}{v'(x_R^t - \chi_R^t)}\right) \\ < \; &\frac{v'(x_R^t - \chi_R^t)}{v''(x_R^t - \chi_R^t)} - \frac{v(x_R^t - \chi_R^t)}{v'(x_R^t - \chi_R^t)} \end{aligned}$$

Since I assume that $v'(d)/v(d)$ strictly decreases as d increases, the right-hand side is positive. If $\chi_L^E = \chi_L^M = \chi_R^t$, both sides are the same. If χ_L^E and χ_L^M become further from x_R^t than χ_R^t, the left-hand side decreases. The reason is as follows. I differentiate $v'(x_R^t - \chi_L^k)/v''(x_R^t - \chi_R^t) - v(x_R^t - \chi_L^k)/v'(x_R^t - \chi_R^t)$ with respect to $x_R - \chi_L^k$, then $v''(x_R^t - \chi_L^k)/v''(x_R^t - \chi_R^t) - v'(x_R^t - \chi_L^k)/v'(x_R^t - \chi_R^t)$. This value is negative because I assume that $v''(d)/v'(d)$ strictly decreases as d increases. As a result, the left-hand side of (13) is less than the left-hand side of (14), so (13) holds, and (12) is positive.

(i) The first statement is true because

$$\frac{\partial z_R^t(\chi_R^t)}{\partial x_R^t} + \frac{\partial z_R^t(\chi_R^t)}{\partial \chi_R^t}\frac{\partial \chi_R^t}{\partial x_R^t} =$$
$$-\frac{1}{v'(x_R^t - \chi_R^t)}(p^M v'(x_R^t - \chi_L^M) + (1-p^M)v'(x_R^t - \chi_L^E) - v'(x_R^t - \chi_R^t)) < 0. \square$$

A.3 Corollary 1

When $p^M = 0$, $z_i' = z_i^{M*}$, from (4). Thus, the left-hand side of (5) is $-v(|\chi_i^E(z_i^{M*}) - x_i^E|) - c(|z_i^{M*} - \chi_i^E(z_i^{M*})|)$, and is strictly greater than the right-hand side, since $-v(|\chi_j^E(z_j^{M*}) - x_i^E|) < -v(|\chi_i^E(z_i^{M*}) - x_i^E|) - c(|z_i^{M*} - \chi_i^E(z_i^{M*})|)$. That is, (5) does not hold when $p^M = 0$.

From (4), when $p^M = 1$, $z_i' = z_i^E(\chi_i^M(z_i^{M*}))$. Here, if an extreme type announces the platform $z_i^E(\chi_i^M(z_i^{M*}))$, he will implement the policy of the moderate type, $\chi_i^M(z_i^{M*})$. The left-hand side of (5) is $-v(|\chi_i^M(z_i^{M*}) - x_i^E|) - c(|z_i^{M*}(\chi_i^M(z_i^{M*})) - \chi_i^M(z_i^{M*})|)$, which is strictly less than its right-hand side, since $-v(|\chi_j^M(z_j^{M*}) - x_i^E|) > -v(|\chi_i^M(z_i^{M*}) - x_i^E|) - c(|z_i^{M*}(\chi_i^M(z_i^{M*})) - \chi_i^M(z_i^{M*})|)$, from Lemma 2, and $-v(|\chi_i^E(z_i^{M*}) - x_i^E|) - c(|z_i^{M*} - \chi_i^E(z_i^{M*})|) > -v(|\chi_i^M(z_i^{M*}) - x_i^E|) - c(|z_i^{M*}(\chi_i^M(z_i^{M*})) - \chi_i^M(z_i^{M*})|)$. That is, (5) is satisfied when $p^M = 1$.

Both sides of (5) change continuously with p^M, so there always exists $p^M = \overline{p} \in (0,1)$ in which both sides of (5) are equal. $\square$

A.4 Proposition 2

The precise definition of a semi-separating equilibrium is as follows. Denote the expected utility of an extreme type who announces z_i as $V_i^E(z_i)$.

Definition 3 *A continuous semi-separating equilibrium is a collection* $(z_i^*, \sigma^M, F(.), \Pi)$ *and a two-policy semi-separating equilibrium is a collection* $(z_i^*, \sigma^M, \bar{z}_i, \Pi)$*, where* z_i^* *is a platform chosen by a moderate type,* σ^M *is the probability of an extreme type choosing* z_i^* *in a mixed strategy,* $F(.)$ *is a distribution function with the support of* $[\bar{z}_L, \underline{z}_L]$ *for* L *and* $[\underline{z}_R, \bar{z}_R]$ *for*

R, and Π *is a scalar, such that: (a)* $\Pi = V_i^E(z_i) = V_i^E(z_i^*)$*, for all* z_i *in support of* $F(.)$ *in a continuous semi-separating equilibrium; and (b)* $\Pi = V_i^E(z_i^*) = V_i^E(\bar{z}_i)$ *in a two-policy semi-separating equilibrium.*

A.4.1 Define σ^M and Π

First, I discuss a continuous semi-separating equilibrium.[24] When an extreme type announces z_i^*, the expected utility is $V_i^E(z_i^*) = (1/2)\Big[-p^M v(|\chi_j^M(z_j^*) - x_i^E|) - \sigma^M(1-p^M)v(|\chi_j^E(z_j^*) - x_i^E|) - (p^M + \sigma^M(1-p^M))\big[v(|\chi_i^E(z_i^*) - x_i^E|) + c(|z_i^* - \chi_i^E(z_i^*)|)\big]\Big] - (1-\sigma^M)(1-p^M)\int_{\bar{z}_j}^{z_j} v(|\chi_j^E(z_j) - x_i^E|)dF(z_j)$. When an extreme type announces $\bar{z}_i$, the expected utility is $V_i^E(\bar{z}_i) = (p^M + \sigma^M(1-p^M))\big[-v(|\chi_i^E(\bar{z}_i) - x_i^E|) - c(|\bar{z}_i - \chi_i^E(\bar{z}_i)|)\big] - (1-\sigma^M)(1-p^M)\int_{\bar{z}j}^{z_j} v(|\chi_j^E(z_j) - x_i^E|)dF(z_j)$. The value of σ^M is decided at the point at which the extreme type's expected utilities under z_i^* and $\bar{z}_i$ are the same:

$$\begin{aligned} &\frac{1}{2}\big[-\frac{p^M}{p^M+\sigma^M(1-p^M)}v(|\chi_j^M(z_j^*) - x_i^E|) - \frac{\sigma^M(1-p^M)}{p^M+\sigma^M(1-p^M)}v(|\chi_j^E(z_j^*) - x_i^E|) \\ - \;& v(|\chi_i^E(z_i^*) - x_i^E|) - c(|z_i^* - \chi_i^E(z_i^*)|)\big] \\ = \;& -v(|\chi_i^E(\bar{z}_i) - x_i^E|) - c(|\bar{z}_i - \chi_i^E(\bar{z}_i)|). \end{aligned} \tag{15}$$

When $\sigma^M = 1$, the left-hand side of (15) is less than the right-hand side because (5) does not hold. If the left-hand side is greater than the right-hand side when $\sigma^M = 0$, the value of $\sigma^M \in (0,1)$ under which an extreme type is indifferent between z_i^* and $\bar{z}_i$ exists. The following condition means that the left-hand side is greater than the right-hand side of (15) when $\sigma^M = 0$:

$$\begin{aligned} &-\frac{1}{2}\big[v(|\chi_j^M(z_j^*) - x_i^E|) + v(|\chi_i^E(z_i^*) - x_i^E|) + c(|z_i^* - \chi_i^E(z_i^*)|)\big] \\ > \;& -v(|\chi_i^E(\bar{z}_i) - x_i^E|) - c(|\bar{z}_i - \chi_i^E(\bar{z}_i)|). \end{aligned} \tag{16}$$

First, $-v(|\chi_i^E(z_i^*) - x_i^E|) - c(|z_i^* - \chi_i^E(z_i^*)|) > -v(|\chi_i^E(\bar{z}_i) - x_i^E|) - c(|\bar{z}_i - \chi_i^E(\bar{z}_i)|)$ because $\bar{z}_i$ is more moderate than z_i^*. Second, $\bar{z}_i$ is the platform with which an extreme type can defeat

[24] To determine the mixed strategy in a continuous semi-separating equilibrium, I build on techniques introduced by Burdett and Judd (1983). They consider price competition and show that firms randomize prices when there is a possibility that consumers will observe only one price. Just as Burdett and Judd (1983) show that firms are indifferent over a range of prices, I show that an extreme type is indifferent over a range of platforms.

a moderate type who announces z_i^*. When σ^M becomes zero, voters guess that a candidate announcing z_i^* is a moderate type. From the definition of $\bar{z}_i$, an extreme type's implemented policy, $\chi_i^E(\bar{z}_i)$, needs to be more moderate than a moderate type's implemented policy, $\chi_i^M(z_i^*)$. From Lemma 2, a moderate type has a greater incentive to choose a more moderate *implemented policy* than an extreme type, and a moderate type is indifferent to winning or losing at z_i^*. This means that $-v(|\chi_j^M(z_j^*) - x_i^E|) > -v(|\chi_i^E(\bar{z}_i) - x_i^E|) - c(|\bar{z}_i - \chi_i^E(\bar{z}_i)|)$. As a result, (16) holds, so a value of $\sigma^M \in (0,1)$ under which an extreme type is indifferent between $\bar{z}_i$ and z_i^* exists.

A.4.2 The Other Bound of Support for $F(.)$

The distribution function, $F(.)$, satisfies the following lemma.

Lemma 3 *Suppose that a continuous semi-separating equilibrium exists. In such an equilibrium, $F(.)$ is continuous with connected support.*

Proof: If $F(.)$ has a discontinuity at some policy, say z_i' (i.e., $F(z_i'+) > F(z_i'-)$), there is a strictly positive probability that an opponent also chooses z_j' (the probability density function is $f(z_j') > 0$). If this candidate approaches the median policy by an infinitesimal degree, it increases the probability of winning by $f(z_j')/2 > 0$. On the other hand, because this approach is minor, the expected utility changes by slightly less than $(1/2)f(z_j')[-v(|\chi_i^E(z_i') - x_i|) - c(|z_i' - \chi_i^E(z_i')|) - (-v(|\chi_j^E(z_j') - x_i|))]$, and is positive (or negative). This implies that if $F(.)$ has a discontinuity, it cannot be part of a continuous semi-separating equilibrium. Assume that $F(.)$ is constant in some region, $[z_1, z_2]$, in the convex hull of the support. If a candidate chooses z_1, he has an incentive to deviate to z_2 because the probability of winning does not change. However, the implemented policy will approach the candidate's own ideal policy, so the expected utility increases. Thus, the support of $F(.)$ must be connected. □

At $\underline{z}_i$, the expected utility is $V_i^E(\underline{z}_i) = -v(|\chi_i^E(\underline{z}_i) - x_i^E|) - c(|\underline{z}_i - \chi_i^E(\underline{z}_i)|)$ because, from Lemma 3, $F(\underline{z}_L) = 0$, and the probability of winning is one. If (9) does not hold, $V_i^E(\underline{z}_i)$ is higher than $V_i^E(\bar{z}_i)$ when $\underline{z}_i = \bar{z}_i$, so $\bar{z}_i \neq \underline{z}_i$ in equilibrium. Therefore, a continuous semi-separating equilibrium exists. If (9) holds, the extreme bound and the moderate bound are equivalent (a two-policy semi-separating equilibrium). Suppose that (9) does not hold. In equilibrium, $V_i^E(\underline{z}_i)$ and $V_i^E(\bar{z}_i)$ should be the same, so $\underline{z}_i$ and $F(.)$ should satisfy the following equation. Further, suppose R, without loss of generality, then

$$\begin{aligned} & -v(x_R^E-\chi_R^E(\underline{z}_R))-c(\chi_R^E(\underline{z}_R)-\underline{z}_R) \\ = \; & (p^M+\sigma^M(1-p^M))\big[-v(x_R^E-\chi_R^E(\bar{z}_R))-c(\chi_R^E(\bar{z}_R)-\bar{z}_R)\big] \\ & -(1-\sigma^M)(1-p^M)\int_{\bar{z}_L}^{\underline{z}_L} v(x_R^E-\chi_L^E(z_L))dF(z_L). \end{aligned} \tag{17}$$

I assume that the two candidates' positions are symmetric, so when $\underline{z}_R$ decreases, $\underline{z}_L$ increases. Then, $V_R^E(\bar{z}_R)$ increases because $\int_{\bar{z}_L}^{\underline{z}_L} v(x_R^E-\chi_L^E(z_L))dF(z_L)$ decreases while $V_i^E(\underline{z}_i)$ decreases. In addition, $F(.)$ adjusts the value of $\int_{\bar{z}_L}^{\underline{z}_L} v(x_R^E-\chi_L^E(z_L))dF(z_L)$. Thus, there exist combinations of $\bar{z}_i$ and $F_j(.)$ that satisfy (17).

I denote $\hat{z}_i^E$ such that $-v(|\chi_j^E(\hat{z}_j^E)-x_i^E|)=-v(|\chi_i^E(\hat{z}_i^E)-x_i^E|)-c(|\hat{z}_i^E-\chi_i^E(\hat{z}_i^E)|)$. The moderate bound, $\underline{z}_i$, should be more extreme than $\hat{z}_i^E$. If $\underline{z}_i$ is more moderate than $\hat{z}_i^E$, it means $-v(|\chi_j^E(\underline{z}_j)-x_i^E|)>-v(|\chi_i^E(\underline{z}_i)-x_i^E|)-c(|\underline{z}_i-\chi_i^E(\underline{z}_i)|)$. Thus, an extreme type with $\underline{z}_i$ has an incentive to deviate and lose to an extreme opponent with a platform close to $\underline{z}_j$. Any platform in the support of $F(.)$, say z_i', needs to satisfy $-v(|\chi_j^E(z_j')-x_i^E|)>-v(|\chi_i^E(z_i')-x_i^E|)-c(|z_i'-\chi_i^E(z_i')|)$ to avoid deviating to lose. Therefore, $\chi_i^E(\underline{z}_i)$ is more extreme than $\chi_i^M(z_i^*)$, because $\chi_i^E(\hat{z}_i^E)$ is more extreme than $\chi_i^M(z_i^*)$.

A.4.3 Define $F(.)$

Suppose R, without loss of generality. Let $X(z_L')=\int_{z_L'}^{\underline{z}_L} v(x_R^E-\chi_L^E(z_L))dF(z_L)$. For any $z_R'\in(\underline{z_R},\ \bar{z}_R)$, the expected utility should be the same as Π.[25] This means that

$$F_X(z_R')=\frac{\Pi+v(x_R^E-\chi_R^E(z_R'))+c(\chi_R^E(z_R')-z_R')X(z_L')}{(1-\sigma^M)(1-p^M)(v(x_R^E-\chi_R^E(z_R'))+c(\chi_R^E(z_R')-z_R'))}.$$

The distribution function, $F_X(.)$, is defined by the above equation for any platform in support of $F(.)$, given $X(z_L')$. When $F_X(z_R')=0$, it is $\Pi+v(x_R^E-\chi_R^E(z_R'))+c(\chi_R^E(z_R')-z_R'^M)X(z_L')=0$. This equation holds if and only if $z_R'=\underline{z_R}$ and $X(z_L')=0$, such that $\Pi=V_R^E(\underline{z}_R)$. Then, $X(z_L')=0$ if and only if $z_L'=\underline{z}_L$. Therefore, when z_R' and z_L' become $\underline{z}_R$ and $\underline{z}_L$, respectively, $F(z_R')$ becomes zero.

When $F(z_R')=1$, it is $\Pi=(p^M+\sigma^M(1-p^M))(-v(x_R^E-\chi_R^E(z_R'))-c(\chi_R^E(z_R')-z_R'^M)(1-$

[25] When extreme type R chooses $z_R'\in[\underline{z}_R,\bar{z}_R]$, the expected utility is $(1-p^M)F(z_R')\big[-v(x_R^E-\chi_R^E(z_R'))-c(\chi_R^E(z_R')-z_R')\big]-(1-\sigma^M)(1-p^M)\int_{z_L'}^{\underline{z}_L} v(x_R^E-\chi_L^E(z_L))dF(z_L)$.

$p^M)X(z_L')$. This equation holds if and only if $z_R' = \bar{z}_R$ and $X(z_L') = \int_{\underline{z}_L}^{\bar{z}_L} v(x_R^E - \chi_L^E(z_L))dF(z_L)$, such that $\Pi = V_R^E(\bar{z}_R)$. This means that when z_R' and z_L' become $\bar{z}_R$ and $\bar{z}_L$, respectively, $F(z_R')$ becomes one.

When z_L' satisfies $|x_m - z_L'| = |x_m - z_R'|$, that is, $F(.)$ is symmetric for both candidates, the value of $X(z_L')$ increases continuously as z_R' (z_L') becomes more extreme. Therefore, if the platform moves from $\underline{z_R}$ to $\bar{z}_R$, $F(z_R')$ increases from zero to one. Thus, if $F(.)$ is symmetric for both candidates, $F_i(.)$ can be defined for $i = L, R$.

A.4.4 An Extreme Type does not Deviate

An extreme type does not deviate to a more moderate platform than $\underline{z_i}$, as the probability of winning is still one. However, the cost of betrayal and the disutility following a win will increase.

If an extreme type deviates to a platform that is more extreme than z_i^*, or between z_i^* and $\bar{z}_i$, this candidate is certain to lose because voters believe that such a candidate is an extreme type based on the off-path belief. Therefore, the expected utility is:

$$\begin{aligned} & -p^M v(|\chi_j^M(z_j^*) - x_i^E|) - \sigma^M(1-p^M)v(|\chi_j^E(z_j^*) - x_i^E|) \\ - & \quad (1-\sigma^M)(1-p^M)\int_{\bar{z}_j}^{\underline{z}_j} v(|\chi_j^E(z_j) - x_i^E|)dF(z_j). \end{aligned} \tag{18}$$

Subtracting (18) from $V_i^E(z_i^*)$ yields

$$\begin{aligned} & -v(|\chi_i^E(z_i^*) - x_i^E|) - c(|z_i^* - \chi_i^E(z_i^*)|) + \frac{p^M}{p^M + \sigma^M(1-p^M)} v(|\chi_j^M(z_j^*) - x_i^E|) \\ + & \quad \frac{\sigma^M(1-p^M)}{p^M + \sigma^M(1-p^M)} v(|\chi_j^E(z_j^*) - x_i^E|). \end{aligned} \tag{19}$$

A moderate type is indifferent to winning and losing at z_i^*. That is, (7) holds. Thus, from Lemma 2, the value of (19) is positive, and this deviation decreases the expected utility. Note that (10) in the proof of Lemma 2 uses p^M, but the same result holds when p^M is replaced by $p^M/[p^M + \sigma^M(1-p^M)]$.

A.4.5 A Moderate Type does not Deviate

Suppose R, without loss of generality. As a moderate type is indifferent between winning and losing at z_R^*, she is indifferent on whether to deviate to a platform that is more extreme than z_R^* or between z_R^* and $\bar{z}_R$. The second possibility involves deviating to any platform in $z'_R \in [\underline{z}_R, \bar{z}_R]$. For an extreme type, the candidate is indifferent between z_R^* and z'_R. This means that

$$\begin{aligned}
& (p^M + \sigma^M(1-p^M))(v(x_R^E - \chi_R^E(z'_R)) + c(\chi_R^E(z'_R) - z'_R)) \\
- & \frac{1}{2}\Big[p^M v(x_R^E - \chi_L^M(z_L^*)) + \sigma^M(1-p^M)v(x_R^E - \chi_L^E(z_L^*)) \\
+ & (p^M + \sigma^M(1-p^M))(v(x_R^E - \chi_R^E(z_R^*)) + c(\chi_R^E(z_R^*) - z_R^*))\Big] \\
= & (1-\sigma^M)(1-p^M)\int_{\bar{z}_L}^{z'_L} v(x_R^E - \chi_L^E(z_L))dF(z_L) \\
- & (1-\sigma^M)(1-p^M)(1-F(z'_L))\big[v(x_R^E - \chi_R^E(z'_R)) + c(\chi_R^E(z'_R) - z'_R)\big].
\end{aligned} \tag{20}$$

A moderate type has no incentive to deviate to z'_i if

$$\begin{aligned}
& (p^M + \sigma^M(1-p^M))(v(x_R^M - \chi_R^M(z'_R)) + c(\chi_R^M(z'_R) - z'_R)) \\
- & \frac{1}{2}\Big[p^M v(x_R^M - \chi_L^M(z_L^*)) + \sigma^M(1-p^M)v(x_R^M - \chi_L^E(z_L^*)) \\
+ & (p^M + \sigma^M(1-p^M))(v(x_R^M - \chi_R^M(z_R^*)) + c(\chi_R^M(z_R^*) - z_R^*))\Big] \\
> & (1-\sigma^M)(1-p^M)\int_{\bar{z}_L}^{z'_L} v(x_R^M - \chi_L^E(z_L))dF(z_L) \\
- & (1-\sigma^M)(1-p^M)(1-F(z'_L))\big[v(x_R^M - \chi_R^M(z'_R)) + c(\chi_R^M(z'_R) - z'_R)\big].
\end{aligned} \tag{21}$$

I disregard $(1-\sigma^M)(1-p^M)$ and differentiate the right-hand side of the above equations with respect to x_R^t to obtain $\int_{\bar{z}_L}^{z'_L} v'(x_R^t - \chi_L^E(z_L))dF(z_L) - (1-F(z'_R))v'(|\chi_R^t(z'_R) - x_R^t|)$. This is positive because the opponent's implemented policy is further from the ideal policy than z'_R, so the right-hand side of (20) is greater than the right-hand side of (21). From (15), at $z'_R = \bar{z}_R$, the left-hand side of (20) is zero. From (7), the left-hand side of (21) is $(p^M + \sigma^M(1-p^M))(v(x_R^M - \chi_R^M(z'_R)) + c(\chi_R^M(z'_R) - z_R'^M) + \sigma^M(1-p^M))(v(x_R^M - \chi_R^M(z_R^*)) + c(\chi_R^M(z_R^*) - z_R^*))$, so is positive because z'_R is smaller than z_R^*. I differentiate the left-hand side with respect to z'_R. Note that σ^M and z_R^* are already decided, so only z'_R changes. Then, $(p^M + \sigma^M(1-p^M))[-v'(|\chi_R^t(z'_R) - x_R^t|)(\partial\chi_R^t(z'_R)/\partial z'_R) + c'(|z'_R - \chi_R^t(z'_R)|))(\partial\chi_R^t(z'_R)/\partial z'_R) -$

$c'(|z'_R - \chi^t_R(z'_R)|)]$. I ignore $p^M + \sigma^M(1-p^M)$. From Lemma 1, this is negative. That is, $-v'(|\chi^t_R(z'_R) - x^t_R|) < 0$. This implies that if z'_R becomes smaller, then the left-hand sides of both equations increase. The next problem is the degree of increase. Differentiating $-v'(|\chi^t_R(z'_R) - x^t_R|)$ with respect to x^t_i yields

$$-v''(|\chi^t_R(z'_R) - x^t_R|)\left(1 - \frac{\partial \chi^t_R(z'_R)}{\partial x^t_R}\right). \tag{22}$$

I differentiate (15) with respect to x^t_R, then

$$\frac{\partial \chi^t_R(z'_R)}{\partial x^t_R} = \frac{v''(|\chi^t_R(z'_R) - x^t_R|)c'(|z'_R - \chi^t_R(z'_R)|)}{v''(|\chi^t_R(z'_R) - x^t_R|)c'(|z'_R - \chi^t_R(z'_R)|) + c''(|z'_R - \chi^t_R(z'_R)|)v'(|\chi^t_R(z'_R) - x^t_R|)} \in (0,1).$$

Thus, the value of (22) is negative. This implies that if x^t_R is more extreme, the increase of the left-hand side is lower when z'_R becomes smaller. At $z'_R = \bar{z}_R$, the left-hand side of (20) is lower than the right-hand side of (21). If z'_R becomes more moderate, both left-hand sides increase, but an increase in (21) is greater than an increase in (20). As a result, for all z'_R, the left-hand side of (20) is lower than the right-hand side of (21), and (21) is satisfied.

Finally, since a moderate type has no incentive to deviate to $\underline{z}_R$, she does not deviate to any policy that is more moderate than $\underline{z}_R$.

A.4.6 A Two-policy Semi-separating Equilibrium

When (9) holds, a two-policy semi-separating equilibrium exists. When an extreme type chooses z^*_i, the expected utility is $V^E_i(z^*_i) = (1/2)\Big[-p^M v(|\chi^M_j(z^*_j) - x^E_i|) - \sigma^M(1-p^M)v(|\chi^E_j(z^*_j) - x^E_i|) - (p^M + \sigma^M(1-p^M))\big[v(|\chi^E_i(z^*_i) - x^E_i|) + c(|z^*_i - \chi^E_i(z^*_i)|)\big]\Big] - (1-\sigma^M)v(|\chi^E_j(\bar{z}_j) - x^E_i|)$. The expected utility when the candidate chooses $\bar{z}_i$ is $V^E_i(\bar{z}_i) = (p^M + \sigma^M(1-p^M))\big[-v(|\chi^E_i(\bar{z}_i) - x^E_i|) - c(|\bar{z}_i - \chi^E_i(\bar{z}_i)|)\big] - (1/2)(1-\sigma^M)(1-p^M)\big[v(|\chi^E_j(\bar{z}_j) - x^E_i|) + v(|\chi^E_i(\bar{z}_i) - x^E_i|) + c(|\bar{z}_i - \chi^E_i(\bar{z}_i)|)\big]$. When $\sigma^M = 1$, $V^E_i(\bar{z}_i)$ is greater than $V^E_i(z^*_i)$ because we assume that (5) does not hold. Assume $\bar{\sigma}^M$, which satisfies (15). If (9) holds, then $V^E_i(\bar{z}_i)$ is less than $V^E_i(z^*_i)$ at $\bar{\sigma}^M$. When σ^M increases continuously from $\bar{\sigma}^M$, $V^E_i(\bar{z}_i)$ increases and $V^E_i(z^*_i)$ decreases continuously, so there exists a σ^M under which $V^E_i(\bar{z}_i) = V^E_i(z^*_i)$, and such σ^M should be higher than $\bar{\sigma}^M$.

Platform $\bar{z}_i$ should be such that $\chi^E_i(\bar{z}_i)$ is between $\chi^M_i(z^*_i)$ and $\chi^E_i(z^*_i)$ if $p^M > 0$ and $\sigma^M > 0$, because in this region, there exists a policy that voters prefer to the expected implemented policy of a candidate with z_i^*. Thus, $\chi_i^E(\bar{z}_i)$ is more extreme than $\chi_i^M(z_i^*)$.

An extreme type does not deviate for the reason explained in Appendix A.4.4. If an extreme type deviates to a platform that is more moderate than $\bar{z}_i$, the expected utility changes by $(1/2)(1-\sigma^M)(1-p^M)\big[v(|\chi_j^E(\bar{z}_j)-x_i^E|)-v(|\chi_i^E(\bar{z}_i)-x_i^E|)-c(|\bar{z}_i-\chi_i^E(\bar{z}_i)|)\big]$. This is negative because (9) holds.

A moderate type does not deviate to a more extreme policy than $\bar{z}_i$ for the reason explained in Appendix A.4.5. A moderate type does not deviate to $\bar{z}_i$ if

$$\begin{aligned}
&(p^M+\sigma^M(1-p^M))[v(|\chi_i^M(\bar{z}_i)-x_i^M|)+c(|\bar{z}_i-\chi_i^M(\bar{z}_i)|)\\
-\;& v(|\chi_i^M(z_i^*)-x_i^M|)-c(|z_i^*-\chi_i^M(z_i^*)|)]\\
-\;& (1-\sigma^M)(1-p^M)\frac{1}{2}\Big[v(|\chi_j^E(\bar{z}_j)-x_i^M|)\\
-\;& v(|\chi_i^M(\bar{z}_i)-x_i^M|)-c(|\bar{z}_i-\chi_i^M(\bar{z}_i)|)\Big]>0. \qquad (23)
\end{aligned}$$

As $\bar{z}_i$ is more moderate than z_i^*, $v(|\chi_i^M(\bar{z}_i)-x_i^M|)+c(|\bar{z}_i-\chi_i^M(\bar{z}_i)|)-v(|\chi_i^M(z_i^*)-x_i^M|)-c(|z_i^*-\chi_i^M(z_i^*)|)$ is positive. For an extreme type, $v(|\chi_j^E(\bar{z}_j)-x_i^E|)-v(|\chi_i^E(\bar{z}_i)-x_i^E|)-c(|\bar{z}_i-\chi_i^E(\bar{z}_i)|)$ is negative because (9) holds. From Lemma 2, its value for a moderate type is lower than for an extreme type, so $v(|\chi_j^E(\bar{z}_j)-x_i^M|)-v(|\chi_i^M(\bar{z}_i)-x_i^M|)-c(|\bar{z}_i-\chi_i^M(\bar{z}_i)|)$ is also negative for a moderate type. As a result, (23) is satisfied.

A.4.7 Asymmetric Equilibrium

There does not exist an asymmetric equilibrium in which candidates choose asymmetric platforms or different values of σ^M or $F(.)$. First, suppose that the support of $F(.)$ is asymmetric. Then, the probability of winning is constant in some regions of the support for at least one candidate, and it cannot be an equilibrium for the reason explained in Lemma 3. This means that σ^M should also be symmetric. Second, suppose that moderate types' platforms are asymmetric. This means that the probability of winning for a moderate type is also asymmetric. Suppose moderate type R announces a more extreme platform than moderate type L, and so loses to L. In this case, extreme type R has no incentive to imitate moderate type R. The values of σ^M should be symmetric, so it cannot be a semi-separating equilibrium. □

A.5 Proposition 3

In such a separating equilibrium, the utility of candidate i of type t when he/she wins is $-v(|\chi_i^t(z_i^t) - x_i^t|) - c(|z_i^t - \chi_i^t(z_i^t)|)$. Then, the utility of candidate i of type t when a *same-type* opponent (type t) wins is $-v(|\chi_j^t(z_j^t) - x_i^t|)$. I denote $\hat{z}_i^t$ as the cut-off platform under which both of these utilities are the same for a type-t candidate, and are symmetric ($|x_m - \hat{z}_i^t| = |x_m - \hat{z}_j^t|$). Figure A-1 shows the positions of $\hat{z}_R^t$ and $\chi_R^t(\hat{z}_R^M)$.

[Figure A-1 here]

Then, the following lemma holds.

Lemma 4 *The extreme type's cut-off platform is more moderate than that of the moderate type ($\hat{z}_R^M - \hat{z}_L^M > \hat{z}_R^E - \hat{z}_L^E$), but the extreme type's implemented policy given the cut-off platform is more extreme than that of the moderate type, $\chi_i^E(\hat{z}_i^E)$ ($\chi_R^E(\hat{z}_R^E) - \chi_L^E(\hat{z}_L^E) > \chi_R^M(\hat{z}_R^M) - \chi_L^M(\hat{z}_L^M)$).*

Proof: To prove this, I differentiate (10) with respect to $x_R^t - x_L^t$ rather than x_R^t. Equation (10) can be rewritten as $v(x_R^t + \chi_R^t - 2x_m) - v(x_R^t - \chi_R^t) = c(\chi_R^t - z_R^t(\chi_R^t))$, where $x_R^t - \chi_L^t = (x_R^t - x_m) + (\chi_R^t - x_m) = x_R^t + \chi_R^t - 2x_m$ because the platforms are symmetric. Differentiating both sides of (10) with respect to $x_R^t - x_L^t$ is the same as differentiating both sides of the rewritten equation with respect to x_R^t. Then,

$$\frac{\partial \chi_R^t}{\partial x_R^t} = \frac{\frac{v''(x_R^t - \chi_R^t)c'(\chi_R^t - z_R^t(\chi_R^t))}{c''(\chi_R^t - z_R^t(\chi_R^t))} - (v'(x_R^t - \chi_L^t) - v'(x_R^t - \chi_R^t))}{v'(x_R^t - \chi_L^t) + v'(x_R^t - \chi_R^t)\frac{\partial z_R^t(\chi_R^t)}{\partial \chi_R^t}}.$$

For the same reason given in Lemma 2, this is positive. Note that $\partial z_R^t(\chi_R^t)/\partial x_R^t = -v''(x_R^t - \chi_R^t)/c''(\chi_R^t - z_R^t(\chi_R^t)) < 0$, and $\partial z_R^t(\chi_R^t)/\partial \chi_R^t = 1 + (v''(x_R^t - \chi_R^t)/c''(\chi_R^t - z_R^t(\chi_R^t))) > 0$. Moreover,

$$\begin{aligned}
&\frac{\partial z_R^t(\chi_R^t)}{\partial x_R^t} + \frac{\partial z_R^t(\chi_R^t)}{\partial \chi_R^t}\frac{\partial \chi_R^t}{\partial x_R^t} = \\
&-\frac{v''(x_R^t - \chi_R^t)}{c''(\chi_R^t - z_R^t(\chi_R^t))}\left(1 - \frac{c'(\chi_R^t - z_R^t(\chi_R^t))\frac{\partial z_R^t(\chi_R^t)}{\partial \chi_R^t}}{v'(x_R^t - \chi_L^M) + v'(x_R^t - \chi_R^t)\frac{\partial z_R^t(\chi_R^t)}{\partial \chi_R^t}}\right) \\
&-\frac{v'(x_R^t - \chi_L^t) - v'(x_R^t - \chi_R^t)}{\frac{v'(x_R^t - \chi_L^M)}{\partial z_R^t(\chi_R^t)/\partial \chi_R^t} + v'(x_R^t - \chi_R^t)}.
\end{aligned}$$

The second term of the right-hand side is negative. The first term is also negative since $v'(x^t_R - \chi^t_R) = c'(\chi^t_R - z^t_R(\chi^t_R))$ from Lemma 1. Thus, this value is negative. □

Then, a symmetric separating equilibrium in which a moderate type wins does not exist. First, if a separating equilibrium in which a moderate type wins against an extreme type exists, regardless of off-path beliefs, an extreme type should announce $\hat{z}^E_i$. If the utility when an extreme candidate wins is higher than the utility when an extreme opponent wins, the extreme candidate has an incentive to win with certainty against the extreme opponent. This is made possible by approaching the median policy, regardless of off-path beliefs.

Second, a moderate type never announces a more moderate platform than $\hat{z}^E_i$. On such a platform, the utility when this moderate type candidate wins is lower than the utility when a moderate opponent wins. Therefore, the moderate type has an incentive to deviate and lose to the moderate opponent. If this moderate type also has an incentive to deviate and lose against an extreme opponent, she will deviate to lose with certainty. If this moderate type has an incentive to win against an extreme opponent, she will deviate to approach $\hat{z}^E_i$, because she can win against an extreme opponent and lose against a moderate opponent.

Finally, suppose that a moderate type announces a more extreme platform than $\hat{z}^E_i$. If an extreme type deviates to a moderate type's platform, the extreme type can improve his chance of winning and can implement a policy closer to his ideal. As a result, the extreme type can increase his expected utility from this deviation.

An asymmetric separating equilibrium in which a moderate type wins (z^M_R and z^M_L are asymmetric) also does not exist. Suppose that z^M_R and z^M_L are asymmetric, so one moderate candidate defeats a moderate opponent with certainty. Without loss of generality, suppose that moderate type R defeats moderate type L. That is, $\chi^M_R(z^M_R) - x_m < x_m - \chi^M_L(z^M_L)$, and moderate type R defeats extreme type L. Note that an extreme type announces $\hat{z}^E_i$. If z^M_R is more extreme than $\hat{z}^E_R$ ($z^M_R < \hat{z}^E_R$), an extreme type R will deviate to pretend to be a moderate type R. Therefore, assume that z^M_R is more moderate than $\hat{z}^E_R$ ($z^M_R < \hat{z}^E_R$). There are three cases.

The first case is that moderate type L loses to or has the same probability of winning as extreme type R ($\chi^E_R(\hat{z}^E_R) - x_m \leq x_m - \chi^M_L(z^M_L)$). Regardless of off-path beliefs, if moderate type R's platform approaches $\hat{z}^E_R$, she can win against both moderate and extreme types of L, and the disutility following a win and the cost of betrayal decrease as the platform approaches her ideal policy.

The second case is that moderate type L defeats extreme type R ($\chi_R^E(\hat{z}_R^E) - x_m > x_m - \chi_L^M(z_L^M)$) when moderate type L announces a more moderate platform than $\hat{z}_L^M$. From Lemma 2, moderate type R has an incentive to deviate and lose to moderate type L. If moderate type R approaches $\hat{z}_R^E$ by more than z_L^M, she can lose to moderate type L and still win against extreme type L.

The final case is that moderate type L defeats extreme type R ($\chi_R^E(\hat{z}_R^E) - x_m > x_m - \chi_L^M(z_L^M)$) when moderate type L announces a platform that is the same as, or closer to her own ideal policy than $\hat{z}_L^M$. If extreme type L deviates to moderate type L's platform (z_L^M), extreme type L can win against extreme type R with certainty, and so gain a higher probability of winning. With this deviation, an extreme type can implement a policy closer to his ideal policy, so he will deviate. □

B Examples

The important equilibrium is a semi-separating equilibrium. Thus, I discuss the national elections in Turkey, Japan, and the UK as examples of this equilibrium.[26] Since my model is simple, these examples do not exactly match my model. For example, my model considers only two symmetric parties, a plurality voting system, and a single-policy issue. In Japan, Turkey, and the UK, there are more than two parties. Turkey uses proportional representation (PR), and Japan employs parallel voting, including both PR and single-member districts. Additionally, it is rare to have an election with a single-policy issue (the only exception may be the 2005 election in Japan in which the privatization of post offices was the main issue) and symmetric parties.[27]

However, I show here that these examples have at least the following four important characteristics of a semi-separating equilibrium: (1) one party's platform is more moderate than the other party's platform; (2) voters guess that such a party is an extreme type; (3) voters are uncertain whether the opposition(s) is an extreme type; and (4) the party that announces the more moderate platform wins the election.

[26] In these examples, the decisions of "parties" are considered instead of "candidates." I use "candidates" in the model, but this can be replaced by "parties." My model supposes majority rule, so only the winner will be in the government. However, in national elections, the losing party will have some share of the sheets in Congress. If a party gets the majority of sheets, I take this to mean that the party wins over the opposition.

[27] These problems are not only applied to my model, but also to most voting models in the Downsian tradition.

B.1 Turkey

In Turkish politics, there are two large groups, namely Political Islam and secular parties. Broadly speaking, secularists, represented by parties such as the Republican People's Party, support democratic systems and politico-religious separation. Political Islam, represented by the Justice and Development Party (AKP), wants to introduce Islamic doctrines into some policies. In recent years, the AKP and the prime minister, Recep Erdogan, have supported politico-religious separation and promoted the AKP as the party of reform, a party that supports democratic systems, including politico-religious separation. Most citizens support secularism in Turkey, and the AKP's promises were almost the same as those encapsulated in the opponent's policies. Nevertheless, voters realized that the AKP is the extreme Islamic party (Dağt, 2006). On the other hand, in the 2007 Turkey presidential election, the Turkish military, which supports secularism, stated that "the Turkish armed forces have been monitoring the situation with concern." People interpreted this as a threat of a coup, and started to worry that the secular parties would support extreme secular policies, such as using violence against Political Islam. Thus, voters were uncertain about the secular party's type. Finally, the AKP won the 2007 elections.

Therefore, this case has the four characteristics of a semi-separating equilibrium: (1) the AKP compromised greatly by promising politico-religious separation; (2) people recognized the AKP as an extreme party that still supported Political Islam; (3) the type of the opponents became uncertain for voters after the threat of a coup; and (4) the AKP won.

B.2 Japan

In Japan, there are two main parties, the Liberal Democratic Party (LDP), which supports increasing government spending on, for example, public works to sustain rural areas, and the Democratic Party of Japan (DPJ), which supports economic reforms and reducing government debt. In 2001, the LDP chose Junichiro Koizumi as their leader. Koizumi promised to implement economic reforms such as reducing government works and debt, and moreover, promised to "destroy" the (traditional) LDP. After the "great depression" during the 1990s, many Japanese supported implementing economic reforms rather than traditional economic policies, so the LDP's position should be further from the median policy than the DPJ after 1990s. Indeed, voters were afraid that the LDP would not implement the economic reforms

(Mulgan, 2002). The opposition, the DPJ, had no experience in government, so voters remained uncertain about the party. Finally, the LDP of Koizumi won the elections in 2001, 2003, and 2005.

This case also satisfies the four characteristics of a semi-separating equilibrium: (1) the LDP compromised greatly by promising to destroy the traditional LDP; (2) people still believed that the LDP was too conservative; (3) the DPJ's type was uncertain for voters, as it had no experience in government; and (4) the LDP won.

B.3 The UK

In the UK, there are two major parties, the Conservative Party and the Labour Party. Broadly speaking, while the Conservative Party supports the free market, the Labour Party is famous for its support of socialist policies and for being supported by labor unions. The Labour Party had not been in government since 1979 because many citizens did not support socialist policies. In 1994, the Labour Party chose Tony Blair as their leader, and he promised the "Third Way" and free-market policies. Most members of the Labour Party supported Blair, although some, such as members of labor unions, still supported socialist policies. This means that the Labour Party compromised greatly by choosing Blair as their leader. On the other hand, voters were uncertain about the Conservative Party preferences because of infighting between factions. As a result, the Labour Party won the 1997 election (Clarke, 2004).

This case matches the four characteristics of a semi-separating equilibrium, as follows: (1) the Labour Party compromised greatly by promising the Third Way; (2) people believed that the Labour Party supported socialist policies, since many members still supported these policies; (3) the type of the Conservative Party was uncertain for voters because of intra-party conflict; and (4) the Labour Party won.

C Position of the Platforms and a Probabilistic Model

In any equilibrium, implemented policies never encroach on the opponent's side of the policy space (i.e., $\chi^t_L(z_L) \leq x_m \leq \chi^t_R(z_R)$), otherwise candidates can always find a better choice in which an implemented policy remains on their own side.

On the other hand, platforms may encroach on the opponent's side (i.e., $z_R^t < x_m < z_L^t$). This study allows for this situation and does not restrict candidates to announcing platforms only within their own halves of the policy space. However, this problem is not a critical one. My model assumes that candidates know every decision-relevant fact about the median voter. If candidates are uncertain about voter preferences—that is, a probabilistic model is considered—in many cases, the above situation does not hold. That candidates have a greater divergence of policies in a probabilistic model is well known (see Calvert (1985)). That is, platforms do not encroach on the opponent's platform in a probabilistic model when the degree of uncertainty is sufficiently high. However, the purpose of this study is to analyze the effect of partially binding platforms on policy, so for the sake of simplicity, I do not consider such probabilistic models here.[28]

References

Adams, James, Samuel Merrill III, and Bernard Grofman, 2005, *A Unified Theory of Party Competition*, Cambridge: Cambridge University Press.

Aldrich, John, 1997, *Why Parties? The Origin and Transformation of Political Parties in America,* Chicago: University of Chicago Press.

Ansolabehere, Stephen and James Snyder Jr., 2001, "Valence Politics and Equilibrium in Spatial Election Models," *Public Choice* 103, 327–336.

Aragones, Enriqueta and Zvika Neeman, 2000, "Strategic Ambiguity in Electoral Competition," *Journal of Theoretical Politics* 12, 183–204.

Aragones, Enriqueta and Thomas Palfrey, 2002, "Mixed Equilibrium in a Downsian Model with a Favored Candidate," *Journal of Economic Theory* 103, 131–161.

Aragones, Enriqueta, Thomas Palfrey, and Andrew Postlewaite, 2007, "Political Reputations and Campaign Promises," *Journal of the European Economic Association* 5, 846–884.

[28]Sometimes, platforms encroach on the opponent's side in real-world elections. For example, in Japan, there are two main parties: the Liberal Democratic Party (LDP), which supports increases in public works to sustain rural areas, and the Democratic Party of Japan (DPJ), which supports economic reform and a reduction in government debt. In 2001, the LDP prime minister, Junichirou Koizumi, promised to implement radical economic reforms that were also suggested by the DPJ, including a reduction in government works and debt. Thus, Koizumi and the LDP promised DPJ policies (Mulgan, 2002, pp. 56–57). Moreover, in the 2007 Upper House election, the LDP and Prime Minister Shinzou Abe promised to continue Koizumi's economic reforms while the DPJ promised policies to recover and support rural areas ("Abe Stumbles on Japan," *The Economist*, July 30, 2007). This was a complete reversal of the original stances of the parties. My model can explain both cases in which the platforms do or do not encroach on the opposing side.

Aragones, Enriqueta and Andrew Postlewaite, 2002, “Ambiguity in Election Games,” *Review of Economic Design* 7, 233–255.

Asako, Yasushi, forthcoming, “Partially Binding Platforms: Campaign Promises vis-à-vis Cost of Betrayal,” *Japanese Economic Review*.

Austen-Smith, David and Jeffrey Banks, 1989, “Electoral Accountability and Incumbency,” in P. Ordeshook ed., *Models of Strategic Choice in Politics*, Ann Arbor: University of Michigan Press.

Banks, Jeffrey, 1990, “A Model of Electoral Competition with Incomplete Information,” *Journal of Economic Theory* 50, 309–325.

Banks, Jeffrey and Joel Sobel, 1987, “Equilibrium Selection in Signaling Games,” *Econometrica* 55, 647–661.

Barro, Robert, 1973, “The Control of Politicians: An Economic Model,” *Public Choice* 14, 19–42.

Besley, Timothy and Stephan Coate, 1997, “An Economic Model of Representative Democracy,” *Quarterly Journal of Economics* 112, 85–114.

Burdett, Kenneth and Kenneth Judd, 1983, “Equilibrium Price Dispersion,” *Econometrica* 51, 955–970.

Callander, Steven, 2008, “Political Motivations,” *Review of Economic Studies* 75, 671–697.

Callander, Steven and Simon Wilkie, 2007, “Lies, Damned Lies, and Political Campaigns,” *Games and Economic Behavior* 80, 262–286.

Callander, Steven and Catherine Wilson, 2008, “Context-Dependent Voting and Political Ambiguity,” *Journal of Public Economics* 92, 565–581.

Cho, In-Koo and David Kreps, 1987, “Signaling Games and Stable Equilibria,” *Quarterly Journal of Economics* 102, 179–222.

Clarke, Peter, 2004, *Hope and Glory: Britain 1900–2000,* London: Penguin Books.

Cox, Gary and Mathew McCubbins, 1994, *Legislative Leviathan: Party Government in the House,* Los Angeles: University of California Press.

Dağt, İsan, 2006, “The Justice and Development Party: Identity, Politics, and Human Rights Discourse in the Search for Security and Legitimacy,” in M. Yavuz ed., *The Emergence of a New Turkey: Democracy and the AK Party*, Salt Lake City: The University of Utah Press.

Djankov, Simeon, Caralee McLiesh, Tatiana Nenova, and Andrei Shleifer, 2003, "Who Owns the Media?" *Journal of Law and Economics* 46, 341–382.

Downs, Anthony, 1957, *An Economic Theory of Democracy*, New York: Harper and Row.

Ferejohn, John, 1986, "Incumbent Performance and Electoral Control," *Public Choice* 50, 5–26.

Glazer, Amihai, 1990, "The Strategy of Candidate Ambiguity," *American Political Science Review* 84, 237–241.

Groceclose, Timothy, 2001, "A Model of Candidate Location when one Candidate has a Valence Advantage," *American Journal of Political Science* 45, 862–886.

Grossman, Gene and Elhanan Helpman, 2005, "A Protectionist Bias in Majoritarian Politics," *Quarterly Journal of Economics* 120, 1239–1282.

Grossman, Gene and Elhanan Helpman, 2008, "Party Discipline and Pork-Barrel Politics," in E. Helpman ed., *Institutions and Economic Performance*, Cambridge MA: Harvard University Press.

Harrington, Joseph, 1993, "The Impact of Re-election Pressures on the Fulfillment of Campaign Promises," *Games and Economic Behavior* 5, 71–97.

Huang, Haifeng, 2010, "Electoral Competition when Some Candidates Lie and Others Pander," *Journal of Theoretical Politics* 22, 333–358.

Hummel, Patrick, 2010, "Flip-flopping from Primaries to General Elections," *Journal of Public Economics* 94, 1020–1027.

Ishihara, Akifumi, 2014, "Strategic Candidacy via Endogenous Commitment," mimeo.

Kartik, Navin and Preston McAfee, 2007, "Signaling Character in Electoral Competition," *American Economic Review* 97, 852–870.

McCarty, Nolan, Keith Poole, and Howard Rosenthal, 2001, "The Hunt for Party Discipline in Congress," *American Political Science Review* 95, 673–687.

McGillivray, Fiona, 1997, "Party Discipline as a Determinant of the Endogenous Formation of Tariffs," *American Journal of Political Science* 41, 584–607.

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

Mulgan, Aurelia, 2002, *Japan's Failed Revolution: Koizumi and the Politics of Economic Reform*, Canberra: Asia Pacific Press.

Osborne, Martin and Al Slivinski, 1996, "A Model of Political Competition with Citizen

Candidates," *Quarterly Journal of Economics* 111, 65–96.

Persson, Torsten and Guido Tabellini, 2000, *Political Economics: Explaining Economic Policy*, Cambridge and London: The MIT Press.

Reinikka, Ritva and Jakob Svensson, 2005, "Fighting Corruption to Improve Schooling: Evidence from a Newspaper Campaign in Uganda," *Journal of European Economic Association* 3, 259–267.

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

Schultz, Christian, 1996, "Polarization and Inefficient Policies," *Review of Economic Studies* 63, 331–344.

Snyder, James and Timothy Groseclose, 2000, "Estimating Party Influence in Congressional Role-Call Voting," *American Journal of Political Science* 44, 193–211.

Stokes, Donald, 1963, "Spatial Models of Party Competition," *American Political Science Review* 57, 368–377.

Wittman, Donald, 1973, "Parties as Utility Maximizers," *American Political Science Review* 67, 490–498.

Zeckhauser, Richard, 1969, "Majority Rule with Lotteries on Alternatives," *Quarterly Journal of Economics* 83, 696–703.

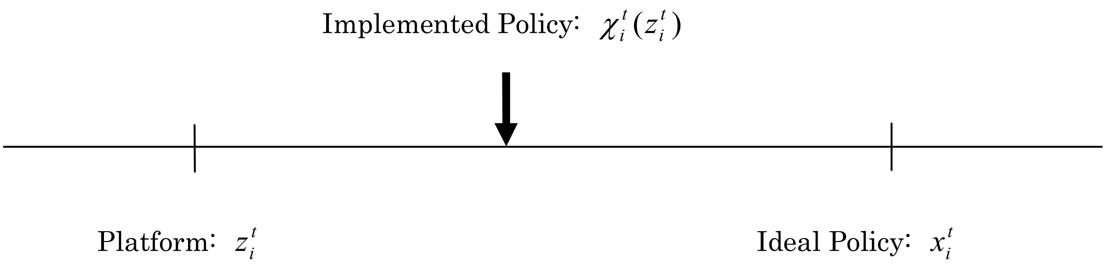


Figure 1: Ideal Policy, Platform and Implemented Policy

Each candidate has an ideal policy, and announces a campaign platform before an election. Given the ideal policy and the platform, the winning candidate decides the implemented policy which will be between the ideal policy and the platform.

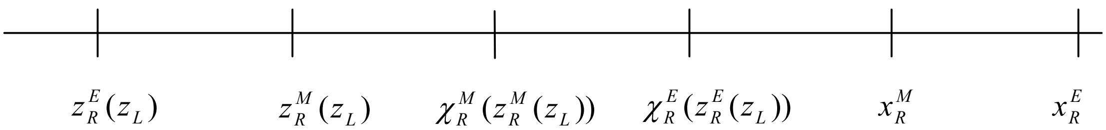
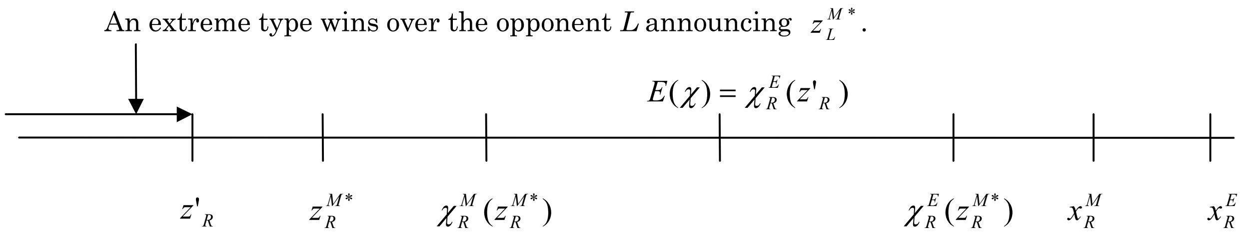
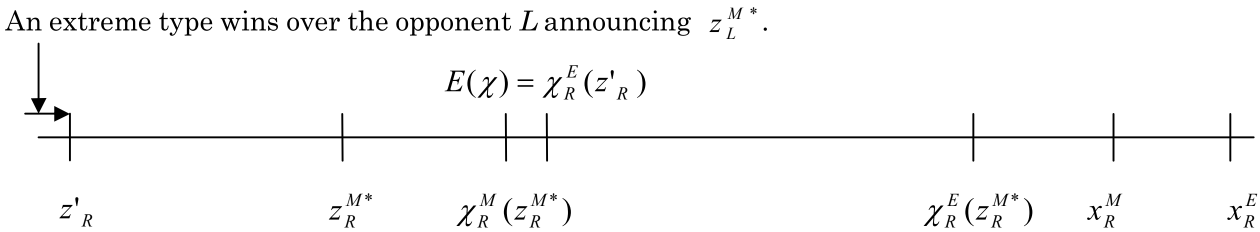


Figure 2: Lemma 2

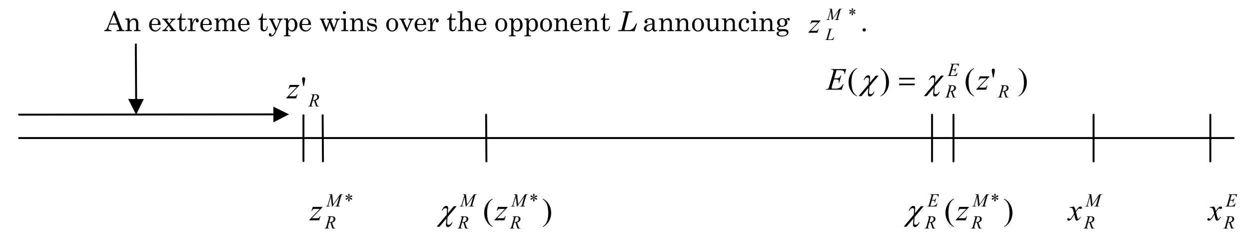
While an extreme type's implemented policy given the cut-off platform $\chi_R^E(z_R^E(z_L))$ is more extreme than a moderate type's one $\chi_R^M(z_R^M(z_L))$, an extreme type's cut-off platform $z_R^E(z_L)$ is more moderate than a moderate type's one $z_R^E(z_L)$.



(a): With $p^M = 1/2$



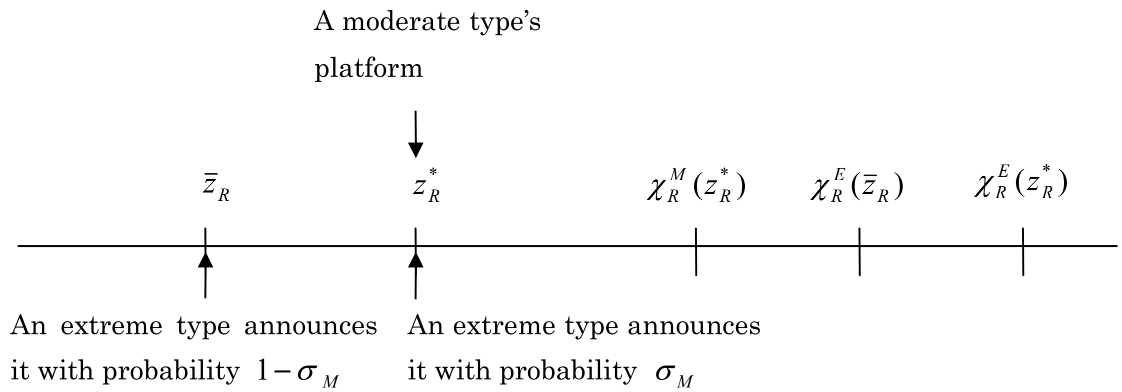
(b): With Sufficiently High p^M



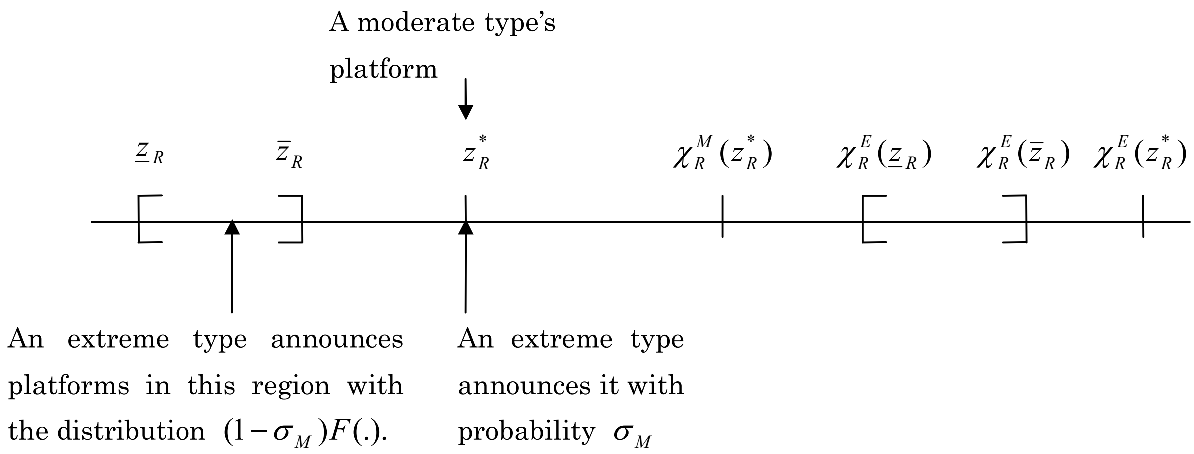
(c): With Sufficiently Low p^M

Figure 3: Pooling Equilibrium

Let $E(\chi) = p^M \chi_R^M(z_R^{M*}) + (1-p^M)\chi_R^E(z_R^{M*})$ denote the expected policy implemented by a candidate announcing z_R^{M*}. Suppose that voters have a linear utility. If an extreme type's platform is more moderate than z'_R, such an extreme type wins over the opponent L announcing z_L^{M*}.



(a) A two-policy semiseparating equilibrium



(b) A continuous semiseparating equilibrium

Figure 4: Semiseparating Equilibrium

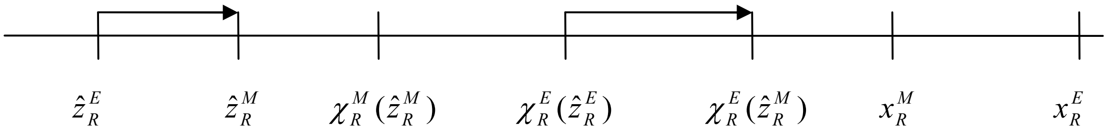


Figure A-1: Separating Equilibrium

An extreme type has an incentive to pretend to be moderate by choosing the moderate type's platform $\hat{z}_R^M$ because the probability of winning increases, and the implemented policy approaches the ideal policy.