A Model of Two Party Representative Democracy: Endogenous Party Formation *

Katsuya Kobayashi † Hideo Konishi‡

May 4, 2011

Abstract

This paper presents a two party representative democratic model in which voters choose their parties in order to influence the choice of party representative. After two candidates are selected as the medians of the parties' support groups, Nature plays to determine the candidates' competence. Based on the candidates' political positions and competence, voters vote for more preferable candidate without being tied by their party choice. We show that there exists a nontrivial equilibrium under some conditions, and show that dependent on voter's distribution over their political positions, the equilibrium party line and the ex ante probability of a party's candidate wins are biased. In particular, we show that if a party has a strong subgroup with extreme positions, then the party tends to alienate the moderate group from the party, and the probability of winning in the final election is reduced.

*This project started when Kobayashi was on sabbatical from Hosei University, visiting Boston College in 2009-2010. Kobayashi thanks Hosei University for their financial support, and Boston College for their hospitality. We thank Hülya Eraslan, Antoine Loeper, and the participants of seminars at Hosei and Kobe, of the public economics workshop at Keio and the contract theory workshops at U. Tokyo for their helpful comments.

†Katsuya Kobayashi: Faculty of Economics, Hosei University, Tokyo, Japan E-mail katsuyak@hosei.ac.jp

‡Hideo Konishi: Department of Economics, Boston College, USA. E-mail hideo.konishi@bc.edu

1 Introduction

In representative democratic system, parties usually have support groups from a wide range of policy spectrum. In many cases, some support group of a party has strong voice in determining their representative candidates and policies. Especially, if the issue is to choose a small number of candidates representing their parties (such as presidential election, senate race, and governors' race etc. in the US), then the core groups of parties seem to have strong voices in particular. For example, in the US, it is often viewed that strong religious conservatives of the Republican Party play an important role in selecting their representative candidates.[1] In this paper, we will theoretically investigate how the party line between the liberal and the conservative and how the winning probability of a party's candidate are affected in the two-party political system if the party is subject to a strong subgroup's influence.

Our main idea is as follows. When an extreme subgroup in the political spectrum of a party's support group exercises strong influences in selecting their representative candidates, moderate (potential) party-supporters are alienated, and they may reconsider which party they should support since each candidate have the winning possibility derived from the uncertainty in the election. If they join the other party with more diverse support groups, then the moderates may be able to play a more significant role in choosing the party candidate. As a result, the party line shifts accordingly, so that the diverse party selects a more moderate candidate while the party supported by an extreme group selects a more extreme candidate. Since all citizens can cast their ballot for the most preferable candidate at voting stage without any commitment, if a candidate is close to the median voter's political position, she has a high probability to be a winner. As a result, given the diverse party selects more moderate candidate, its winning probability is raised.

Towards this goal, we build a simple model of a two party representative democracy with a one-dimensional policy space and atomless voters, and analyze how the distribution of voters' political positions determines the party line and the final outcome of election.[2] We assume that voters are strategic in choosing their parties foreseeing their influence on the choice of candidates, and we assume away all other strategic behaviors by voters and candidates. Each candidate is selected to represent a party by the party supporters (voters) based on her political position, and she cannot misrepresent her position. We also concentrate on the case of two parties only.

We add to the model two ingredients that are not commonly employed in the literature. First, we need to endogeneize the party line which should be based on voters' own choice,

[1] Recent Tea Party movement in the Republican Party shows how a subgroup of a party can influence the selection of a party candidate.

[2] There are several researches on party formations assuming ad hoc distribution of voters in stead of a two-party in the citizen candidate literature (Bordignon & Tabellini 2009 and Riviere 1999). On the other hand, one of our research issues is to show how the distribution of voters effects endogenous party formation under a two-party political system.

although each voter's party choice cannot have an influence in selecting a candidate since voters are atomless. As a result, unilateral deviations cannot affect parties' candidate selection processes, so any partition of voters can be a Nash equilibrium. In order to avoid this difficulty, we consider a sequence of small coalitional deviations converging to zero measure, and define a "political equilibrium" as a partition of voters from which no (convergent) sequence of coalitional deviations are profitable in the limit. [3]

Second, there has to be uncertainty in voting outcome. After two candidates are selected by the two parties, debates and political campaigns are taken place in the actual election. Here uncertainty kicks in. As is often seen in the real world, such shocks can change the voting outcome. Some voters may prefer a candidate from the opposite party after seeing their debate and campaign performance, or they may as well vote for her without sticking in their party's candidate. Note that this sort of shock is regarded as a common shock to all voters: not idiosyncratic unlike in the standard probabilistic voting model (Coughlin 1992 and Osborne 1995). [4] That is, after the shock, which party's candidate the median voter will support dictates the result of the voting outcome.

The above two ingredients complicate the analysis quite a bit, so that we simplify other parts of our model. First, we assume away each part's candidate's strategic policy selection by using the citizen-candidate model (Osborne and Slivinski 1996, and Besley and Coate 1997, 2003). That is, we assume that a selected candidate will implement her most favorite policy once she is elected. Second, we greatly simplify each party's candidate selection mechanism: we simply assume that each party's citizen-candidate is the median voter of the support group of the party, which is the majority preferred candidate. Although this is a strong assumption, it appears to be a necessary simplification given that the party line is determined endogenously. Besley and Coate (2003) adopts this assumption in their two jurisdiction model as well. Third, we naturally focus on sorting allocations, that is a party line decides voters into two parties. In our model, voters' party choice problem is quite complicated, and in some occasions, some extreme voters may want to join the opposite party strategically to influence their candidate selections. We will show the conditions to be away from such strategic behavior, and will focus on the sorting allocations. A sorting equilibrium which contains a sorting allocation at the party choice stage is also supported by introducing some psychological cost to voter's utility function. Finally, we conventionally assume that in the final voting stage voters vote sincerely although each of them is atomless.

Our game goes as follows. In stage 1, voters choose their parties, and a candidate is selected

[3]This definition of equilibrium has superficial similarity to "ϵ-club's deviations" of Osborne and Tourky (2008). However, as we will discuss in Section 1.1, the uses of small coalitional deviations in these two equilibrium concepts are very different from each other.

[4]For example, we can recall the loss of the incumbent George Allen, a Republican, in the 2006 Virginia senator race and the victory by Scott Brown, a Republican, in the 2010 Massachusetts senator race to replace late Edward M. "Ted" Kennedy who had been a Democratic senator more than 40 years. These shocks were clearly not idiosyncratic: the shocks can be quite dramatic and devastating.

mechanically (a median of the party). In stage 2, Nature plays and two candidates' relative attractiveness is determined randomly. In stage 3, voters cast their ballot for a candidate who is more preferable on the two candidates' positions and the relative attractiveness of the candidates. Then the winner in the election implements her most favorite policy going around every citizen. Our equilibrium concept, political equilibrium, is a subgame perfect equilibrium except for considering small coalitional deviations in stead of each player's deviation on the party choice stage. More precisely, a political equilibrium is an allocation that is immune to any sequences of coalitional deviations which are positive measure coalitional deviations with the limit being a zero measure. With atomless voters, Nash behavior does not make sense since they have no impact, while if we allow for large coalitional deviations, there may not be an equilibrium. This definition of equilibrium is motivated by the difficulties of coordinating voters' behaviors. We will characterize our political equilibrium, then provide an existence theorem. It turns out that our equilibrium has a number of intriguing properties.

After providing the characterization and existence theorems of political equilibrium, we investigate how the party line is affected by the distribution of voters over policy space: in particular, we show the relation of median voter's position and the equilibrium party line. Other things equal, if a party's support group (in terms of their policy spectrum) becomes more extreme, then the party loses its support, making their candidate more extreme and the opponent party's candidate more moderate. The probability of the former party's candidate winning is reduced by this. Thus, our model can describe the phenomenon that having a strong extreme subgroup of a party alienates moderate potential supporters of the party.

In the next subsection, we provide a brief literature review. In section 2, we present our model. In section 3, we define political equilibrium, and investigate its properties. Using them, we provide some insights on how the party line is affected by the distribution of voters over their political positions. In section 4, we conclude with brief discussion on how relaxing our assumptions will affect our results.

1.1 A Brief Literature Review

There is a numerous number of papers on voting and party formation. Here, we will discuss a small number of papers on party formation that are most related to our paper. One of the issue is what is meant by work "party." On the one hand, Baron (1993) and Jackson and Moselle (2002) extend the well-known simple non cooperative bargaining model by including policy space in two different ways. In their model, legislators form a coalition in order to strengthen their bargaining powers by acting together (by compromising proposing policy), and the coalition is called a party. On the other hand, Feddersen (1992) construct a model in which voters choose political positions, and call a group of voters who choose the same political position a party. Our model is closest to Feddersen (1992), since voters are assumed to be strategic players in his model as well as ours. However, there are a number of differences, too. Feddersen's is deterministic, there can be an arbitrary number of parties, and a multidimensional policy space is allowed. In contrast, uncertainty plays an important role in our model, while we restrict our attention to the two-party case on a single-issue-space. In our model, a party's political position (candidate's position) is determined by aggregating the party supporters' political positions (via median voter's policy).

Our equilibrium concept superficially resembles a part of Osborne and Tourky's (2007), a small club Nash equilibrium. They consider deviations by positive measures of atomless voters to determine each subgame's voting outcome. However, this is just a response to the insensitivity of electoral outcome to a single voter's action, given a continuum of voters. In contrast, in our model, a small coalition's deviation has more meaning on it. By a small coalitional deviation, party candidates' political positions change and voters' utilities are affected by that. [5] We take the sizes of coalitions to measure zero in the limit, and use "differentiation" of utility functions to characterize voters' behavior. Thus, in our paper, small coalitions play an essential role unlike in Osborne and Tourky (2008).

2 The model

2.1 The overview of the model and the game

There is a one-dimensional policy space, and a continuum of atomless citizen-voters is distributed over the interval $[0, 1]$. There are two parties L (left) and R (right). The party name itself does not matter, but for just convenience, we call the party that has more supporters from left side (right side) L (R). As we will mention in the following, these parties are formed by the citizen-voters (hereafter voters). Each party selects a candidate who represents the party, and each voter casts a vote for the most favorite candidate in the two candidates. Following the citizen-candidate models by Osbourne and Slivinski (1996) and Besley and Coate (1997), we assume that the winner becomes the policy maker, who implements her preferred policy which is based on her own political position, which means that the policy maker elected by voters has all authorities, and assume that candidates' political positions are common knowledge and that candidates cannot commit their ex-ante policies. We also assume that candidates' "political competence" for being a policy maker, which means how she has the ability for electoral campaigns, the charisma as a policy maker, persuasive addresses and so on [6] is a random variable that is initially unknown to the voters, which later realize after candidate starts campaign. Note that this random shock is not an idiosyncratic shock across voters, but it is common to all voters, which can affect the voting outcome. [7] Once candidates'

[5] In Osborne and Tourky (2008) (and in Feddersen 1992), voting is assumed to be costly since they are also interested in voter turnout. In order to avoid trivial equilibrium with no voter turnout, "small club" Nash equilibrium is at least needed to meaningful Nash equilibrium.

[6] To be concrete, these are debate performance, campaign gaffes and scandals, which can affect the voting outcome.

[7] It is well known that each candidate takes the median voter's position if there is no uncertainty following the median voter theorem. On the other hand, when "the candidates are uncertain of the distribution of

political competence have been realized, voters' voting behavior depends on the candidates' political competence as well as on their political positions. Thus, some voters can prefer the candidate of the opposite party to that of her party in the voting stage. Since there are not any restrictions on voting, she does not necessarily vote for the candidate of her party.

We consider the following dynamic two-party representative election game. In stage 1, voters choose their parties, in stage 2, in each party, a member of the party is chosen as a representative who will choose a policy once elected, in stage 3, Nature plays and the competence (attractiveness for voters) of each candidate realizes, and in stage 4, all voters vote freely over two candidates. Basically, these stages are analyzed in backward order. However, following Besley and Coate (2003), we greatly simplify stage 2: a median of the support group of each party is selected as the candidate who represents the party. We solve this game by the backward induction, so that the equilibrium is basically the subgame perfect equilibrium. However we need to modify the equilibrium a little bit in stage 1 since each voter is atomless. We introduce a equilibrium notion that is immune to any small coalitional deviations taking the limit of them as we mentioned in the previous section. Regarding small size deviations, Osborne and Tourky (2008) also use a similar deviation named "ϵ-club", and define the "small clubs Nash equilibrium". [8] This small clubs is not a coalition to weakly improve each member's payoff in the club, but a club to improve the sum of the members' payoff in the club by their deviating. While Osborne and Tourky do not consider redistribution among the members in the club, our model is more rigid in terms of the redistribution of each member's payoff in the coalition by considering coalitional deviations.

2.2 Citizens

Each citizen (voter) cares about the policy chosen by the elected representative and cares about her competence that is the ability well enough to implement her policy as the representative. Each voter is atomless and has a type, θ, which is distributed continuously on $[0, 1]$ with density function $g(\theta)$. [9] Type θ voters have the following vNM utility function

$$u(p_k; \theta, \epsilon_k) = -|p_k - \theta| + \epsilon_k,$$

where $p_k \in [0, 1]$ and $\epsilon_k \in \mathbb{R}$ denote the policy implemented by the elected representative $k \in C$ as only a policy maker and a realization of a random variable that describes her competence, respectively. C denotes a candidate set composed of candidates selected from each party.

The random variable ϵ_k follows probability density function f_k with zero expectation ($E(\epsilon_k) = 0$) and symmetric distribution with respect to 0. A positive realization ϵ_k shows

citizens' ideal points", they may take different positions (Osborne 1995). In this paper, by introducing this political competence, a simple uncertainty, the political phenomenon we pointed out in the previous section can be explained.

[8] They use the ϵ-clubs deviations not in the party formation stage but in the voting stage. In their model, candidates and voters are separated, which is not the citizen candidate model.

[9] Thus, even if all voters of type θ form one group, it is still atomless.

that the candidate is competent, while negative realization describes her incompetence. In addition to the competence, we also introduce a psychological cost into the utility. Thus, when the policy implemented by the elected representative in the candidate set C is p_k, the utility of a type θ voter choosing party $i \in \{L, R\}$ is

$$U(i, p_k; \theta, \epsilon_k) = -|p_k - \theta| + \epsilon_k - \Phi(\theta, i),$$

where $\Phi(\theta, i)$ describes a psychological cost from the participation in a party that is not supported by the voters being close to θ's political position. These voters nearby θ are called the neighbors, hereafter. In more detail, type θ's neighbors is defined as the group of voters whose types are within some political distance $d > 0$ from type θ. Each voter is concerned about not only the policy and the competence of each candidate, but also her neighborhood within this political distance d when they choose a party. From stand point analysis, this psychological cost shows up only in some special occasion in order to eliminate unlikely behavior by voters. [10] The definition is slightly involved. Let $g_L : [0,1] \to \mathbb{R}_+$ and $g_R : [0,1] \to \mathbb{R}_+$ that are membership densities of Party L and R respectively be such that for all $\theta \in [0,1]$, $g_L(\theta) + g_R(\theta) = g(\theta)$ holds. [11]

Definition 1 ***Psychological cost:*** *There is a fraction $\bar{\lambda} < 1$ such that for all $\theta \in [0,1]$ and all $i, j \in \{L, R\}$ with $i \neq j$,*

$$\Phi(\theta, i) = \Phi \quad \text{if} \quad \bar{\lambda} \leq \frac{\int_{\theta-d}^{\theta+d} g_j(\theta')d\theta'}{\int_{\theta-d}^{\theta+d} g(\theta')d\theta'} \leq 1,$$

and $\Phi(\theta, i) = 0$, otherwise.

We will set $\bar{\lambda}$ sufficiently high so that only in an extreme isolation, voters feel uncomfortable, and will set $d > 0$ sufficiently small so that voters are only concerned about the closer neighbors. We will explain the role of this psychological cost in more detail in the latter section. Thus, this psychological cost can be ignored if we do not refer to it.

2.3 Citizen-Candidates as a Median Voter of Each Party

In this section, we will explain how each candidate is selected in each party. It depends on the structure of party in this model that who becomes candidate. We will call each party's membership distribution "allocation", hereafter. In general, each party's support group can be overlapped: i.e., for some interval $(\theta_1, \theta_2) \subset [0,1]$, $g_L(\theta) > 0$ and $g_R(\theta) > 0$ for all $\theta \in (\theta_1, \theta_2)$,

[10] However this psychological cost is plausible. Actually, in no matter what communities, if we move from one group to another opposing group, we will be criticized by the previous companions and not feel so good.

[11] Although a measure theoretical definition may be cleaner, we chose this definition since it might be more intuitive.

which is voters of some same types are choosing different parties and which is called "semi-pooling allocation" in this paper. On the other hand, we define "sorting allocation" in the bellow:

Definition 2 ***Sorting allocation*** *is an allocation with a threshold* $\tilde{\theta} \in [0,1]$ *which partitions* $[0,1]$ *into two intervals:* $L = [0, \tilde{\theta})$ *and* $R = (\tilde{\theta}, 1]$ *such that* $g_L(\theta) = g(\theta)$ *and* $g_R(\theta) = 0$ *for all* $\theta \in [0, \tilde{\theta})$*, and* $g_L(\theta) = 0$ *and* $g_R(\theta) = g(\theta)$ *for all* $\theta \in (\tilde{\theta}, 1]$ [12]

Throughout the paper, we will denote these sorting allocations with threshold value $\tilde{\theta}$ *by* $g_L^{\tilde{\theta}}$ *and* $g_R^{\tilde{\theta}}$ *, respectively.* We will focus mainly our attention on a sorting allocation in the latter sections. We will also consider a general case including a semi-pooling allocation in the latter section.

Distribution functions of supporters of parties L and R can be written as $G_L(\theta) = \int_0^\theta g_L(\theta')d\theta'$ and $G_R(\theta) = \int_0^\theta g_R(\theta')d\theta'$, respectively. Thus, we can write $G(\theta) = G_L(\theta) + G_R(\theta)$. We assume that each candidate is elected by the members of each party and is of these majority preferred type, namely a median voter chosen as only a party representative, following Besley and Coate (2003) as we said earlier. Let $x(g_L)$ and $y(g_R)$ be such that

$$\int_0^{x(g_L)} g_L(\theta)d\theta = \int_{x(g_L)}^1 g_L(\theta)d\theta \iff G_L(x(g_L)) = G_L(1) - G_L(x(g_L))$$

and

$$\int_0^{y(g_R)} g_R(\theta)d\theta = \int_{y(g_R)}^1 g_R(\theta)d\theta \iff G_R(y(g_R)) = G_R(1) - G_R(y(g_R)),$$

respectively. They denote candidates of parties L and R, respectively. [13] Obviously, each candidate x and y is depending on each distribution of her supporters (each party's distribution), respectively. On the basis of these characteristics, we will see the candidates in a sorting allocation with threshold type $\tilde{\theta}$. From the definition of sorting allocation, x is determined by $G(x) = G(\tilde{\theta}) - G(x)$ and y is determined by $G(y) - G(\tilde{\theta}) = 1 - G(y)$. In a sorting allocation, each candidate is also depending on the threshold $\tilde{\theta}$. Thus, we will also denote each candidate as a function of $\tilde{\theta}$: i.e. $x = x(\tilde{\theta})$ and $y = y(\tilde{\theta})$ in the following sections when we focus a change of the threshold $\tilde{\theta}$.

2.4 Realization of Competence of a Candidate

After candidates x and y are selected, their competence ϵ_x and ϵ_y realize. In principle, we assume that both ϵ_x and ϵ_y are independent distributed random variables following f_x and f_y,

[12]Although the voters of type $\tilde{\theta}$ is not choosing any parties in this notation, the results in the following sections are not changed even if they are choosing either L or R party since only one type voters, also including $\tilde{\theta}$, are atomless. However we only assume that there are atomless voters of type 0 (1) in L (R) even if $\tilde{\theta} = 0$ ($\tilde{\theta} = 1$) for the convenience.

[13]Strictly speaking, candidate of party L (R) is a voter whose type is x (y). But we abuse notations to call her x (y), since there is no possibility of confusion.

respectively. But for expository simplicity, we will assume that only party L candidate x has random variable ϵ, y has no shock. [14]

2.5 Voting

First note that voters' voting behavior is not bound by the parties they belong to. There is absolutely no commitment: voters just see the candidates and their competence, and decide who to vote for. *We assume that all voters sincerely vote for a candidate.* Let us consider a type θ voter. We define a function of type θ's relative evaluation of y to x that is the difference of type θ's utility from policies chosen by each candidate as $h(x,y;\theta) \equiv -|y-\theta| + |x-\theta|$ when $\epsilon = 0$, or

$$h(x,y;\theta) = \begin{cases} -(y-\theta)+(x-\theta) = x-y \leq 0 & \text{if } \theta \leq x \\ -(y-\theta)+(\theta-x) = 2\theta - x - y & \text{if } x < \theta < y \\ -(\theta-y)+(\theta-x) = y-x \geq 0 & \text{if } y \leq \theta \end{cases}.$$

Clearly, $h(\theta)$ is a weakly increasing function of θ. Especially, the relative evaluation of y increases by two unit as $\theta \in (x,y)$ becomes one unit larger. Let θ_{med} be the median voter type: i.e., $\int_0^{\theta_{med}} g(\theta)d\theta = \int_{\theta_{med}}^1 g(\theta)d\theta$. Then the competence ϵ which makes the median voters indifferent between both candidates is depending on x and y Thus, we will denote the competence as a function of x and y: i.e.

$$\epsilon(x,y) \equiv h(x,y;\theta_{med}) = -|y-\theta_{med}| + |x-\theta_{med}| = 2\theta_{med} - x - y, \tag{1}$$

We assume that a candidate which wins a plurality of vote, more ballots that the other, at voting stage becomes the elected representative, then we have the following lemma (the proof is in Appendix):

Lemma 1 *If $\epsilon > \epsilon(x,y)$, then x is the winner. If $\epsilon < \epsilon(x,y)$, then y is the winner.*

Since ϵ is a random variable drawn from a probability distribution with density function f, once x and y are determined, $F(\epsilon(x,y))$ and $1-F(\epsilon(x,y))$ are the winning probabilities of candidates x and y from this lemma. Taking the probabilities and the political positions of both Candidate's into account, voters choose their parties.

Before the next section, we have the below corollary of the related the above lemma:

Corollary 1 *If $\theta_{med} < \frac{x+y}{2}$, then $F(\epsilon(x,y)) < 1-F(\epsilon(x,y))$. If $\theta_{med} > \frac{x+y}{2}$, then $F(\epsilon(x,y)) > 1-F(\epsilon(x,y))$. If $\theta_{med} = \frac{x+y}{2}$, then $F(\epsilon(x,y)) = 1-F(\epsilon(x,y))$.*

[14] We can calculate that each voter's expected utility in the case where candidate x and y have competence ϵ_x and ϵ_y, respectively is the same as that in the case where only a candidate x has a competence ϵ by assuming that both of ϵ_x and ϵ_y are independent and that $f(\epsilon) = \int_{-\infty}^{+\infty} f_x(\epsilon+\epsilon_y)f_y(\epsilon_y)d\epsilon_y$.

This corollary means that the winning probabilities of x and y depend on the distance between the median voter's point and each candidate's position. In more detail, the nearer candidate from the median type has the larger winning probability than the further candidate. See Figure 1.

2.6 Party Choice by Voters

In stage 1, all voters choose either party L or party R. We assume that there is no option of joining no party. Each citizen chooses party $i \in \{L, R\}$ in order to influence on the choice of the party's candidate as the representative of the party. *Note that since all voters are atomless, each voter's party choice has absolutely no impact on intra-party selection of candidate.* The expected utility of voter of type θ when two candidates are x and y is

$$\begin{aligned} Eu(x,y;\theta) &= \int_{-\infty}^{\epsilon(x,y)} f(\epsilon)(-|y-\theta|)d\epsilon + \int_{\epsilon(x,y)}^{+\infty} f(\epsilon)(-|x-\theta|+\epsilon)d\epsilon \\ &= \int_{-\infty}^{\epsilon(x,y)} f(\epsilon)d\epsilon(-|y-\theta|) + \int_{\epsilon(x,y)}^{+\infty} f(\epsilon)d\epsilon(-|x-\theta|) + \int_{\epsilon(x,y)}^{+\infty} \epsilon f(\epsilon)d\epsilon \\ &= \underbrace{F(\epsilon(x,y))}_{\text{prob. } y \text{ winning}} \times \underbrace{(-|y-\theta|)}_{\text{utility from } y \text{ winning}} + \underbrace{(1-F(\epsilon(x,y)))}_{\text{prob. } x \text{ winning}} \times \underbrace{(-|x-\theta|)}_{\text{utility from } x \text{ winning}} \\ &\quad + \underbrace{\int_{\epsilon(x,y)}^{+\infty} \epsilon f(\epsilon)d\epsilon}_{\text{ave. of } \epsilon \text{ when } x \text{ wins}} . \end{aligned} \tag{2}$$

Thus, we denote the expected utility in stage 1 when each voter choose a party as EU;

$$EU(i,x,y;\theta) = Eu(x,y;\theta) + \Phi(\theta,i).$$

This implies that any partition of voters can constitute a Nash equilibrium since each voter is atomless. Thus, in this paper, we allow coalitional deviations by small coalitions. The equilibrium concept we employ is described in the following section.

3 Political Equilibrium

3.1 The Definition of Political Equilibrium

As we mentioned in the previous section, since each voter is atomless, any partition can be a Nash equilibrium at the party choice stage, so that we adopt a equilibrium which is immune to any small coalition composed of voters at the stage as defining in the bellow. To put it more concretely, we consider a sequence of deviations by coalitions of which size converging to zero. As we will show below, it turns out that such an equilibrium allocation is immune to deviation of small coalitions with positive measure. Formally,

Definition 3 *A* ***political equilibrium*** *is defined as a population distributions of parties L and R, $g_L : [0,1] \to \mathbb{R}_+$ and $g_R : [0,1] \to \mathbb{R}_+$, such that there is no sequence of coalitions described by functions $\{\gamma_k\}_{k=1}^{\infty}$ such that*

1. $\int_0^1 \gamma_k(\theta)d\theta > 0$ *holds for all* $k = 1, 2, ...,$
2. $\lim_{k\to\infty} \int_0^1 \gamma_k(\theta)d\theta = 0,$
3. *for all* $k = 1, 2, ...,$ *we have either*
   (a) $\gamma_k(\theta) \leq g_L(\theta)$ *and* $EU(x(g_L - \gamma_k), y(g_R + \gamma_k); \theta) > EU(x(g_L), y(g_R); \theta)$ *for all* θ *with* $\gamma_k(\theta) > 0$*, or*
   (b) $\gamma_k(\theta) \leq g_R(\theta)$ *and* $EU(x(g_R + \gamma_k), y(g_R - \gamma_k); \theta) > EU(x(g_L), y(g_R); \theta)$ *for all* θ *with* $\gamma_k(\theta) > 0$.

Since our purpose in this paper is to show how voters form each party and the characterization, the equilibrium definition focus on the voters' party choice stage by already assuming in the previous section candidates as a median voter of each party and voters' sincerely voting. The definition is that there is no sequence of strictly-improving coalition deviations (by switching parties) of which measures go to zero such as k goes to infinity. Conditions (a) and (b) correspond to deviations from parties L and R, respectively.

3.2 Examining Large Coalitions

It is tricky to analyze a majority voting equilibrium in a model with a continuum of voters. On the one hand, as we discussed in the previous section, no single voter can influence a party's candidate selection by joining the party as well. On the other hand, if an arbitrary large coalition of voters switch their parties, it destabilizes party structure very easily as is seen in the following example.

Example 1 *Assume that g is uniform $g(\theta) = 1$ for all $\theta \in [0,1]$, and that f is very widely spread (for example, $f(\epsilon) = \frac{1}{2a}$ for all $\epsilon \in [-a, a]$ with very large number a. In this case, whoever are the two candidates x and y, their winning probabilities are always almost $\frac{1}{2}$ and $\frac{1}{2}$ since $f(\epsilon)$ is sufficiently small, whatever $\epsilon(x, y)$ is. Now, since everything is totally symmetric, consider the allocation $g_L(\theta) = g(\theta)$ for all $\theta < \frac{1}{2}$ and $g_R(\theta) = g(\theta)$ for all $\theta > \frac{1}{2}$. In this case, $x = \frac{1}{4}$ and $y = \frac{3}{4}$. Can this be immune to a large scale deviation? We denote a coalitional deviation as γ. Consider a deviation from party R to L: $\gamma(\theta) = g(\theta)$ for all $\theta \in (\frac{1}{2}, \frac{3}{4} - \delta) \cup (\frac{3}{4} + \delta, 1]$ where $\delta > 0$ is an arbitrary small positive number. That is, after the deviation, almost all voters except for a small R-party group around $\theta = \frac{3}{4}$. This makes $x' \simeq \frac{1}{2} = \theta_{med}$, and $y' = \frac{3}{4}$ still holds. Given a wide-spread f, still winning probabilities of x' and y' are almost $\frac{1}{2}$ and $\frac{1}{2}$. Then, deviators in γ has closer candidate from L who anyway wins with probability $\frac{1}{2}$, so they are all better off.* □

Although this example appears to be an extreme one, but in this model, large coalitional deviations tend to destabilize party structure easily. In order to overcome this problem, we consider deviations by small coalitions with positive measures.

3.3 Deviation Incentives by Small Coalitions

We will focus on sorting equilibria. A sorting allocation is described by the threshold value $\tilde{\theta}$. Let us partition the space of voter types into three intervals: $[0, x)$, (x, y), and $(y, 1]$. [15] Recall that in a sorting allocation, x and y are determined by $\tilde{\theta}$, so that *each candidate's position will be denoted by* $x(\tilde{\theta})$ *and* $y(\tilde{\theta})$*, respectively.* We will consider a size δ coalition's deviation from each interval (we can deal with size δ coalitions across the intervals easily by combining the cases) in the bellow.

Let us start with a coalitional deviation with size $\delta > 0$ that belongs to the interval (x, y), moving from R to L. [16] In this case, the coalitional deviation reduces the population of party R and increases that of party L by δ. In order to avoid confusions, we denote δ in this case by $\delta_{(x,y)}^{R \to L} > 0$. We can easily construct such a deviation. Consider $\gamma_{(x,y)} : [0, 1] \to \mathbb{R}_+$ such that $\int_0^1 \gamma_{(x,y)}(\theta) d\theta = \int_x^y \gamma_{(x,y)}(\theta) d\theta = \delta_{(x,y)}^{R \to L}$ and $\gamma_{(x,y)}(\theta) \leq g_R(\theta)$ for all $\theta \in (x, y)$. After the deviation by $\delta_{(x,y)}^{R \to L}$, party L's population distribution is $g_L^{\tilde{\theta}} + \gamma_{(x,y)}$, while party R's population distribution is $g_R^{\tilde{\theta}} - \gamma_{(x,y)}$. That is, the new median voter type x' of party L is determined by

$$G(x') = G(\tilde{\theta}) + \delta_{(x,y)}^{R \to L} - G(x'),$$

and y' of party R is by

$$G(y') - G(\tilde{\theta}) - \delta_{(x,y)}^{R \to L} = G(1) - G(y').$$

Since we are considering a small coalitional deviation, we will take $\delta_{(x,y)}^{R \to L} \to 0$. By totally differentiating them, [17] we have

$$g(x)dx = d\delta_{(x,y)}^{R \to L} - g(x)dx,$$

or

$$\frac{dx}{d\delta_{(x,y)}^{R \to L}} = \frac{1}{2g(x)},$$

and similarly we have,

$$\frac{dy}{d\delta_{(x,y)}^{R \to L}} = \frac{1}{2g(y)}.$$

[15] The borders x and y are measure zero, so we ignore them.

[16] If a coalitional deviation in the interval (x, y) involves groups who switch parties $R \to L$ and $L \to R$, then the effect of the deviation is simply reduced by canceling them out. So, we can concentrate on one-sided move: either $R \to L$ or $L \to R$.

[17] In this case, x' and y' are functions of the calitional deviation size $\delta_{(x,y)}^{R \to L}$, so that taking the difference of this equation between before and after deviation $G(x') - G(x) = \delta_{(x,y)}^{R \to L} - (G(x') - G(x))$, dividing it by $\delta_{(x,y)}^{R \to L}$ and taking $\delta_{(x,y)}^{R \to L}$ to zero, we can obtain the same result.

These derivatives describe that by the small coalitional deviation $\delta^{R\to L}_{(x,y)} \to 0$, both x and y move to the right. Thus, type θ's expected payoff is affected by such a deviation through changes in x and y. Since we are checking the incentive of the coalition member to join the deviation, we consider voters of R in $(\tilde{\theta}, y)$. Thus, for $\theta \in (\tilde{\theta}, y)$ we have

$$Eu(x, y; \theta) = -F(\epsilon(x, y))(y - \theta) - (1 - F(\epsilon(x, y)))(\theta - x) + \int_{\epsilon(x,y)}^{+\infty} \epsilon f(\epsilon) d\epsilon.$$

Note that $\theta_{med} \in [x, y]$ holds. Suppose that $\theta_{med} < x < y$. Then, since x and y are the medians of parties L and R, we reach a contradiction. The case where $x < y < \theta_{med}$ is also the same logic. Thus, $\theta_{med} \in [x, y]$ must hold. This implies $\epsilon(x, y) = 2\theta_{med} - x - y$, and the impact of the coalitional deviation from the interval $(\tilde{\theta}, y)$ is written as

$$\begin{aligned}
&\frac{dEu(x, y; \theta)}{d\delta^{R\to L}_{(x,y)}} \\
&= \frac{1}{2}\Bigg[\underbrace{-\frac{F(\epsilon(x, y))}{g(y)} + \frac{1 - F(\epsilon(x, y))}{g(x)}}_{\text{changes in candidates' positions}} \underbrace{- (2\theta - x - y - \epsilon(x, y)) f(\epsilon(x, y)) \left(\frac{1}{g(x)} + \frac{1}{g(y)}\right)}_{\text{changes in winning probabilities}}\Bigg] \\
&= \frac{1}{2}\left[-\frac{F(\epsilon(x, y))}{g(y)} + \frac{1 - F(\epsilon(x, y))}{g(x)} - (h(x, y; \theta) - h(x, y; \theta_{med})) f(\epsilon(x, y)) \left(\frac{1}{g(x)} + \frac{1}{g(y)}\right)\right].
\end{aligned} \tag{3}$$

The first terms in the brackets of (3) are changes of the expected utility which both candidates moving to right by the deviation bring. The second term in the brackets is a change of the expected utility which is brought by the change of winning probability of y, namely $f(\epsilon(x, y)) \frac{d\epsilon(x,y)}{d\delta^{R\to L}_{(x,y)}}$. Especially, $h(x, y; \theta) - h(x, y; \theta_{med})$ in the second term means the difference of the relative evaluations of y between θ and θ_{med}, so that a change of the winning probability of y brings a change of the expected utility by the difference of relative evaluation of y to type θ. The difference becomes $h(x, y; \theta) - h(x, y; \theta_{med}) = 2(\theta - \theta_{med})$, that is two times value of the distance between θ and θ_{med}.

Note that θ shows up only in the latter term (the effect due to the changes in winning probabilities), which is an decreasing function in θ. Thus, clearly, the second term in the brackets of (3) is decreasing in θ. Suppose that $\frac{dEU(x,y;\tilde{\theta})}{d\delta^{R\to L}_{(x,y)}} = 0$ with the threshold $\tilde{\theta}$ that divides into L and R ; i.e. $\tilde{\theta}$ is satisfied with

$$-\frac{F(\epsilon(x, y))}{g(y)} + \frac{1 - F(\epsilon(x, y))}{g(x)} = 2(\tilde{\theta} - \theta_{med}) f(\epsilon(x, y)) \left(\frac{1}{g(x)} + \frac{1}{g(y)}\right).$$

(It will be shown in section 3.5 that (3) becomes zero with $\tilde{\theta}$.) Then, for all $\theta < \tilde{\theta}$, we have $\frac{dEU(\theta)}{d\delta^{R\to L}_{(x,y)}} > 0$, while for all $\theta > \tilde{\theta}$, we have $\frac{dEU(\theta)}{d\delta^{R\to L}_{(x,y)}} < 0$. This implies that coalitions do not want to move from R to L if they are composed by the types in $(\tilde{\theta}, y)$, while some small coalitions composed by types in $(x, \tilde{\theta})$ want to move from R to L if there are some voters belonging to R in $(x, \tilde{\theta})$. However, since we are considering a sorting allocation at present, all voters of

$\theta \in [0, \tilde{\theta})$ are in L, so that there are no small coalitions wants to move from R to L. From the analysis above, it is easy to see that if we consider a coalitional deviation with size $\delta \to 0$ that belongs to the interval (x, y), moving from L to R, then the analysis is symmetrically reversed. This argument shows that sorting is consistent with voters' incentives.

In order to be a political equilibrium of a sorting allocation, all voters in $[0, x]$ need to join L, and all voters in $[y, 1]$ need to join R. With the assumption of psychological cost, no small coalition would deviate in these regions.

3.4 Consideration to Psychological Costs

Do we really need the assumption of psychological cost? The answer is "yes". In order to confirm this fact, we will consider small coalitional deviations in the interval $(y, 1]$ switching from R to L. Thus, for $\theta \in (y, 1]$ we have

$$Eu(x, y; \theta) = -F(\epsilon(x, y))(\theta - y) - (1 - F(\epsilon(x, y)))(\theta - x) + \int_{\epsilon(x,y)}^{+\infty} \epsilon f(\epsilon) d\epsilon.$$

Similar calculations as above show that the impact of the coalitional deviation from the interval $(\tilde{\theta}, y)$, switching from R to L, is written as $\frac{dx}{d\delta^{R\to L}_{(y,1]}} = \frac{1}{2g(x)}$, $\frac{dy}{d\delta^{R\to L}_{(y,1]}} = -\frac{1}{2g(y)}$ and

$$\begin{aligned}
&\frac{dEu(x, y; \theta)}{d\delta^{R\to L}_{(y,1]}} \\
= \; & \frac{1}{2}\Bigg[\underbrace{-\frac{F(\epsilon(x, y))}{g(y)} + \frac{1 - F(\epsilon(x, y))}{g(x)}}_{\text{changes in candidates' positions}} + \underbrace{(y - x - \epsilon(x, y)) f(\epsilon(x, y)) \left(\frac{1}{g(y)} - \frac{1}{g(x)}\right)}_{\text{changes in winning probabilities}}\Bigg] \\
= \; & \frac{1}{2}\left[-\frac{F(\epsilon(x, y))}{g(y)} + \frac{1 - F(\epsilon(x, y))}{g(x)} + 2(y - \theta_{med}) f(\epsilon(x, y)) \left(\frac{1}{g(y)} - \frac{1}{g(x)}\right)\right]. \qquad (4)
\end{aligned}$$

Note that $y - x - \epsilon(x, y) = h(x, y; \theta) - h(x, y; \theta_{med}) = 2(y - \theta_{med})$ for $\theta \in (y, 1]$, and that the above formula does not contain θ. This means that for all voters of type $\theta > y$, the relative evaluation of y is common, and then the incentive to deviate is also common if psychological cost Φ could be ignored. On the other hand, we will also consider the deviations $\delta^{L\to R}_{[0,x)} \to 0$. By similar calculations, we obtain

$$\begin{aligned}
&\frac{dEu(x, y; \theta)}{d\delta^{L\to R}_{[0,x)}} \\
= \; & -\frac{1}{2}\left[-\frac{F(\epsilon(x, y))}{g(y)} + \frac{1 - F(\epsilon(x, y))}{g(x)} + 2(\theta_{med} - x) f(\epsilon(x, y)) \left(\frac{1}{g(y)} - \frac{1}{g(x)}\right)\right]. \qquad (5)
\end{aligned}$$

Note that first terms in the brackets of both (4) and (5) are common. Thus if the distances between the median and each candidate are same, (5) is negative when (4) is positive, and vice versa. This fact shows that when voters in $(y, 1]$ have an incentive to deviate from R to L, those in $[0, x)$ have no incentive. In other word, voters in either $(y, 1]$ or $[0, x)$ have always an incentive to deviate to the other party when each candidate is on the same distance from the median. In addition to this, even if each candidate is not on the same distance position from the median, both of candidates may have an incentive for deviation since the signs of the brackets in both (4) and (5) are not always negative. Thus, we need psychological cost to eliminate such incentives to upset the intuitive class of equilibria; a sorting political equilibrium. In the next subsection, we will show the necessary and sufficient conditions for a sorting political equilibrium by considering the expected utility EU containing the psychological cost.

3.5 Existence of Sorting Political Equilibrium

In this section, we will consider whether there is a political equilibrium with a sorting allocation or not. More concretely, we will show the conditions of the existence of a sorting political equilibrium. Let us begin with considering how type $\tilde{\theta}$'s expected utility changes when the threshold $\tilde{\theta}$ slightly slides to right. Note that (3) is a monotonically decreasing function in θ, in a sorting allocation with threshold $\tilde{\theta}$. Differentiating (2) by $\tilde{\theta}$, we have

$$\begin{aligned}\frac{dEU(x(\tilde{\theta}), y(\tilde{\theta}); \tilde{\theta})}{d\tilde{\theta}} &= \frac{g(\tilde{\theta})}{2}\Bigg[-\frac{F(\epsilon(x(\tilde{\theta}), y(\tilde{\theta})))}{g(y(\tilde{\theta}))} + \frac{1 - F(\epsilon(x(\tilde{\theta}), y(\tilde{\theta})))}{g(x(\tilde{\theta})} \\ &\quad -2(\tilde{\theta} - \theta_{med})f(\epsilon(x(\tilde{\theta}), y(\tilde{\theta})))\left(\frac{1}{g(x(\tilde{\theta}))} + \frac{1}{g(y(\tilde{\theta}))}\right)\Bigg]\end{aligned}$$

Note that the contents of the brackets in this expression is the same as that of (3) with $\theta = \tilde{\theta}$. Focusing on the contents of brackets in this expression or (3), we define a function $\varphi : [0, 1] \to \mathbb{R}$:

$$\begin{aligned}\varphi(\tilde{\theta}) &\equiv -\frac{F(\epsilon(x(\tilde{\theta}), y(\tilde{\theta})))}{g(y(\tilde{\theta}))} + \frac{1 - F(\epsilon(x(\tilde{\theta}), y(\tilde{\theta}))))}{g(x(\tilde{\theta}))} \\ &\quad -2(\tilde{\theta} - \theta_{med})f(\epsilon(x(\tilde{\theta}), y(\tilde{\theta})))\left(\frac{1}{g(y(\tilde{\theta}))} + \frac{1}{g(x(\tilde{\theta}))}\right). \qquad (6)\end{aligned}$$

By using this function, we will show that the sorting allocations that divide into party L and party R with a threshold $\tilde{\theta}$ become political equilibria, namely sorting political equilibria, in the following lemmas and propositions. First, in the below lemma, we have the difference of the expected utility of type θ when the threshold changes from $\tilde{\theta}$ to $\tilde{\theta} + \Delta$ or to $\tilde{\theta} - \Delta$.

Lemma 2 *Consider sorting allocations described by $\tilde{\theta}$ and $\tilde{\theta} + \Delta$ such that $\Delta > 0$ and that Δ is sufficiently small. Then, we have*

$$\begin{aligned}&EU(x(\tilde{\theta} + \Delta), y(\tilde{\theta} + \Delta); \theta) - EU(x(\tilde{\theta}), y(\tilde{\theta}); \theta) \\ &= \int_{\tilde{\theta}}^{\tilde{\theta}+\Delta} \frac{g(\theta')}{2}\left[\varphi(\theta') - 2(\theta - \theta')f(\epsilon(x(\theta'), y(\theta')))\left(\frac{1}{g(x(\theta'))} + \frac{1}{g(y(\theta'))}\right)\right] d\theta'.\end{aligned}$$

As a consequence, $EU(x(\tilde{\theta} + \Delta), y(\tilde{\theta} + \Delta); \theta) - EU(x(\tilde{\theta}), y(\tilde{\theta}); \theta)$ is decreasing in θ for all

$\theta \in (\tilde{\theta}, y(\tilde{\theta}))$. *Similarly, consider sorting allocations described by* $\tilde{\theta}$ *and* $\tilde{\theta} - \Delta$. *Then, we have*

$$\begin{aligned} & EU(x(\tilde{\theta}-\Delta), y(\tilde{\theta}-\Delta); \theta) - EU(x(\tilde{\theta}), y(\tilde{\theta}); \theta) \\ = & -\int_{\tilde{\theta}-\Delta}^{\tilde{\theta}} \frac{g(\theta')}{2} \left[\varphi(\theta') - 2(\theta - \theta') f(\epsilon(x(\theta'), y(\theta'))) \left(\frac{1}{g(x(\theta'))} + \frac{1}{g(y(\theta'))} \right) \right] d\theta'. \end{aligned}$$

As a consequence, $EU(x(\tilde{\theta}-\Delta), y(\tilde{\theta}-\Delta); \theta) - EU(x(\tilde{\theta}), y(\tilde{\theta}); \theta)$ *is increasing in* θ *for all* $\theta \in (\tilde{\theta}, y(\tilde{\theta}))$.

Remark. The formula has a clear interpretation, though it is rather messy. The first term $\frac{g(\theta')}{2}\varphi(\theta')$ shows the change of the expected utility of type θ' by moving the threshold value around $\theta' \in [\tilde{\theta}, \tilde{\theta}+\Delta]$, in other words, the first term is the change of the border type's expected utility. Thus, integrating it from $\tilde{\theta}$ to $\tilde{\theta}+\Delta$, voter's type that is evaluating is also moving from $\tilde{\theta}$ to $\tilde{\theta}+\Delta$. This is not what we need, so that we have to correct this by adjusting the evaluator's type to θ at each θ'. This is done in the second term. Note that

$$\frac{dF(\epsilon(x(\theta'), y(\theta')))}{d\theta'} = \frac{g(\theta')}{2} f(\epsilon(x(\theta'), y(\theta'))) \left(\frac{1}{g(x(\theta'))} + \frac{1}{g(y(\theta'))} \right).$$

Since $F(\epsilon(x(\theta'), y(\theta')))$ is the probability of y winning, this means the change of the expected utility of the border type θ' by changing the probability of y winning when the threshold slightly moves to right. (Note that increasing y's winning probability means decreasing x's winning probability, and vice versa, so that this expression also means the change of the probability of x losing at the same time.) In addition to this, the utility which each type's voters obtain from candidate y and x is different, in other words, each type evaluates candidate y and x differently. Thus, when evaluator changes from θ' to θ, this difference of evaluation has to be adjusted. [18] More concretely, since the difference of utility of y and x is $-|y-\theta| + |x-\theta| - \epsilon(x,y) = 2(\theta - \theta_{med})$, the difference of this between θ' and θ is $2(\theta-\theta_{med}) - 2(\theta'-\theta_{med}) = 2(\theta-\theta')$. As a result, for this adjustment, we have to subtract the term of $2(\theta-\theta')$ multiplied by the change of the probability $\frac{g(\theta')}{2} f(\epsilon(x(\theta'), y(\theta')))(\frac{1}{g(x(\theta'))} + \frac{1}{g(y(\theta'))})$ from $\frac{g(\theta')}{2}\varphi(\theta')$. Thus, voters of the edge type $\tilde{\theta}+\Delta$ in the coalition of $[\tilde{\theta}, \tilde{\theta}+\Delta]$ have the smallest utility in the coalition members by the deviation since the formula in lemma 2 is a decreasing function of θ.

Next, the bellow lemma shows that there is a coalitional deviation around $\tilde{\theta}$ which has the same effect as small coalitions deviating from $(\tilde{\theta}, y)$ or $(x, \tilde{\theta})$ to the other party.

Lemma 3 *Consider an improving coalitional deviation* γ *from a sorting allocation with* $\tilde{\theta}$ *such that* $Supp(\gamma) \subset [\tilde{\theta}, y(\tilde{\theta})]$. *Then, there is another improving coalitional deviation* γ_R^{Δ} *such that (i)* $\gamma_R^{\Delta}(\theta) = g(\theta)$ *for all* $\theta \in (\tilde{\theta}, \tilde{\theta}+\Delta)$ *and* $\gamma_R^{\Delta}(\theta) = 0$, *otherwise; and (ii)* $\int_{\tilde{\theta}}^{\tilde{\theta}+\Delta} \gamma_R^{\Delta}(\theta) d\theta = \int_{\tilde{\theta}}^{y(\tilde{\theta})} \gamma(\theta) d\theta$. *Similarly, consider an improving coalitional deviation* γ *from a sorting allocation*

[18] On the other hand, the change of the expected utility of the border type θ' by changing each candidate's position is the same as that of the evaluator type θ since the linear utility is assumed.

with $\tilde{\theta}$ such that $Supp(\gamma) \subset [x(\tilde{\theta}+\Delta), \tilde{\theta}]$. Then, there is another improving coalitional deviation γ_L^Δ such that (i) $\gamma_L^\Delta(\theta) = g(\theta)$ for all $\theta \in (\tilde{\theta} - \Delta, \tilde{\theta})$ and $\gamma_L^\Delta(\theta) = 0$, otherwise; and (ii) $\int_{\tilde{\theta}-\Delta}^{\tilde{\theta}} \gamma_L^\Delta(\theta) d\theta = \int_{x(\tilde{\theta})}^{\tilde{\theta}} \gamma(\theta) d\theta$.

The following lemma is a direct consequence of the above two lemmas.

Lemma 4 *Consider a sorting allocation with threshold $\tilde{\theta}$. This allocation is immune to a coalitional deviation γ with $Supp(\gamma) \subset (\tilde{\theta}, y(\tilde{\theta}))$ if and only if $EU(x(\tilde{\theta}+\Delta), y(\tilde{\theta}+\Delta); \tilde{\theta}+\Delta) \leq EU(x(\tilde{\theta}), y(\tilde{\theta}); \tilde{\theta}+\Delta)$ holds for Δ defined by γ^Δ. Similarly, this allocation is immune to a coalitional deviation γ with $Supp(\gamma) \subset (x(\tilde{\theta}+\Delta), \tilde{\theta})$ if and only if $EU(x(\tilde{\theta}-\Delta), y(\tilde{\theta}-\Delta); \tilde{\theta}-\Delta) \leq EU(x(\tilde{\theta}), y(\tilde{\theta}); \tilde{\theta}-\Delta)$ holds for Δ defined by γ^Δ.*

As a result, in order to check whether a sorting allocation is a political equilibrium, this lemma tells us to confirm whether the type that is the furthest from $\tilde{\theta}$ in every coalition has an incentive for taking part in the coalition. More precisely, if there exists $\bar{\Delta} > 0$ such that (i) $EU(x(\tilde{\theta}+\Delta), y(\tilde{\theta}+\Delta); \tilde{\theta}+\Delta) \leq EU(x(\tilde{\theta}), y(\tilde{\theta}); \tilde{\theta}+\Delta)$ is held for all $\Delta \in (0, \bar{\Delta})$, and (ii) $EU(x(\tilde{\theta}-\Delta), y(\tilde{\theta}-\Delta); \tilde{\theta}-\Delta) \leq EU(x(\tilde{\theta}), y(\tilde{\theta}); \tilde{\theta}-\Delta)$ is held for all $\Delta \in (0, \bar{\Delta})$, then a sorting allocation with threshold $\tilde{\theta}$ is a political equilibrium.

We will simplify the above conditions by using φ function. First, we provide a simple necessary condition to be a political equilibrium.

Lemma 5 *Suppose that $\varphi(\theta)$ is continuous. Then, a sorting allocation with threshold $\tilde{\theta}$ is a political equilibrium only if $\varphi(\tilde{\theta}) = 0$.*

Thus, we will assume $\varphi(\tilde{\theta}) = 0$ in order to characterize political equilibrium. By applying the first-order Taylor expansion, we can approximate the utility change of the critical coalition member's utility in the bellow lemma when a coalition γ_R^Δ deviates.

Lemma 6 *Suppose that $\varphi(\tilde{\theta}) = 0$ and that f and g are differentiable functions. Then, for sufficiently small $\Delta > 0$, $EU(x(\tilde{\theta}+\Delta), y(\tilde{\theta}+\Delta); \tilde{\theta}+\Delta) - EU(x(\tilde{\theta}), y(\tilde{\theta}); \tilde{\theta}+\Delta)$ is approximated as*

$$
\begin{aligned}
& EU(x(\tilde{\theta}+\Delta), y(\tilde{\theta}+\Delta); \tilde{\theta}+\Delta) - EU(x(\tilde{\theta}), y(\tilde{\theta}); \tilde{\theta}+\Delta) \\
= & \int_{\tilde{\theta}}^{\tilde{\theta}+\Delta} \frac{g(\theta')}{2} \left[\varphi(\theta') - 2(\tilde{\theta}+\Delta-\theta') f(\epsilon(x(\theta''))) \left(\frac{1}{g(x(\theta'))} + \frac{1}{g(y(\theta'))} \right) \right] d\theta' \\
\simeq & \frac{\Delta^2 g(\tilde{\theta})}{2} \left[\frac{\varphi'(\tilde{\theta})}{2} - f(\epsilon(x(\tilde{\theta}), y(\tilde{\theta}))) \left(\frac{1}{g(x(\tilde{\theta}))} + \frac{1}{g(y(\tilde{\theta}))} \right) \right].
\end{aligned}
$$

The following proposition follows directly from the above lemma.

Proposition 1 *Suppose that f and g are differentiable. A sorting allocation with threshold $\tilde{\theta}$ is a political equilibrium if (i) $\varphi(\tilde{\theta}) = 0$ and (ii) $\frac{\varphi'(\tilde{\theta})}{2} - f(\epsilon(x(\tilde{\theta}), y(\tilde{\theta}))) \left(\frac{1}{g(x(\tilde{\theta}))} + \frac{1}{g(y(\tilde{\theta}))} \right) < 0$. On the other hand, a sorting allocation with threshold $\tilde{\theta}$ is a political equilibrium only if (i) $\varphi(\tilde{\theta}) = 0$ and (ii') $\frac{\varphi'(\tilde{\theta})}{2} - f(\epsilon(x(\tilde{\theta}), y(\tilde{\theta}))) \left(\frac{1}{g(x(\tilde{\theta}))} + \frac{1}{g(y(\tilde{\theta}))} \right) \leq 0$.*

Proposition 1 says that $\varphi(\tilde{\theta}) = 0$ is not sufficient but one of necessary conditions and that we also need a slope condition of $\varphi(\tilde{\theta})$ so that a sorting allocation becomes a political equilibrium.

From the above proposition, we can find an easy sufficient condition for a sorting allocation being a political equilibrium.

Corollary 2 *Suppose that f and g are differentiable. A sorting allocation with threshold $\tilde{\theta}$ is a political equilibrium if (i) $\varphi(\tilde{\theta}) = 0$ and (ii) $\varphi'(\tilde{\theta}) \leq 0$.*

We can also have the sufficient condition for the existence of a sorting political equilibrium in the bellow theorem.

Theorem 1 *Suppose that f and g are differentiable, $Supp(f) \supset [-1, 1]$, and $g(\theta) > 0$ for all $\theta \in (0, 1)$ and $g(0) = g(1) = 0$. Then, there exists a sorting political equilibrium with an interior threshold $\tilde{\theta} \in (0, 1)$.*

Note that the conditions of the theorem are unnecessarily strong in order to have a clear statement. As it is easily seen from the proof, we only need the existence of $\underline{\theta}$ and $\bar{\theta}$ with $\underline{\theta} < \bar{\theta}$, $\varphi(\underline{\theta}) > 0$ and $\varphi(\bar{\theta}) < 0$.

We can actually show an example that there is a sorting political equilibrium satisfies the sufficient condition, or satisfies the necessary and sufficient condition including $\varphi'(\tilde{\theta}) < 0$. See appendix C. Moreover, the semi-pooling allocation case is also shown in Appendix B.

On the other hand, there are always one party equilibria. The reason is that any voters can not form a small coalitional deviation in one party situation since any isolated coalitional deviations make psychological costs. However, in one party equilibria which are $\tilde{\theta} = 0$ or $\tilde{\theta} = 1$, since the winner is always θ_{med} that is the same as the result of the median voter theorem, this result is different from our analysis interest in this paper.

3.6 The Party Structure in a Sorting Political Equilibrium

Finally, we will see how voters' type distribution is important to determining the party structure in a sorting political equilibrium. First, we start with comparing with the traditional "median voter" case in the bellow proposition after we define the changes in candidates positions in $\varphi(\tilde{\theta})$ as $\xi(\tilde{\theta})$:

$$\xi(\tilde{\theta}) \equiv -\frac{F(\epsilon(x(\tilde{\theta}), y(\tilde{\theta})))}{g(y(\tilde{\theta}))} + \frac{1 - F(\epsilon(x(\tilde{\theta}), y(\tilde{\theta})))}{g(x(\tilde{\theta}))}.$$

Proposition 2 *If $\xi(\tilde{\theta}) \neq 0$ such that $\varphi(\tilde{\theta}) = 0$ and if $Supp(f) \supset [-1, 1]$ and $g(\theta) > 0$ for all $\theta \in (0, 1)$ and $g(0) = g(1) = 0$, then θ_{med} does not become an allocation with a sorting political equilibrium.*

This proposition tells us that a sorting political equilibrium does not always divide the voter into two-party at the median type. Next, we will consider the condition where the median type becomes the threshold of two-party. It turns out that it is not enough to have symmetric g and f, though the additional condition seems often satisfied.

Proposition 3 *Suppose that g and f are symmetric. Moreover,*

$$\frac{g(\theta_{med})}{4g(x)^2}\left(4f(0) - \frac{g'(x)}{g(x)}\right) - \frac{4f(0)}{g(x)} < 0$$

Then, there is a political equilibrium with $\tilde{\theta} = \theta_{med}$.

The above proposition tells us that, in general, $\tilde{\theta}$ and θ_{med} have no reason to coincide with each other. In addition to this, we can obtain the following fact:

Proposition 4 *Assume that g and f are continuous, $Supp(f) \supset [-1, 1]$, and $g(\theta) > 0$ for all $\theta \in (0, 1)$ and $g(0) = g(1) = 0$, then there exists a sorting equilibrium with threshold $\tilde{\theta}^* > (<)\theta_{med}$ if $\xi(\theta_{med}) > (<)0$.*

The condition $\xi(\theta_{med}) > 0$ says that if a small group near θ_{med} switches their party from R to L, then x moves more to the right than y's move to the left. That is, the gain of x coming closer is more than the loss of y moves away from θ_{med}, and such a small group in R prefers to switch their party. Thus, the statement of Proposition 4 is intuitive. However, note that it does not say that there is no other type of equilibria. In order to state the condition that enables us to get a stronger statement, we need some additional concepts. When $\tilde{\theta} = 0$ or $\tilde{\theta} = 1$, there may be one of the parties can be interpreted empty, but let us assume that $x(0) = 0$ and $y(0) = \theta_{med}$, and $x(1) = \theta_{med}$ and $y(1) = 1$ hold, assuming that zero measure parties still elect candidates. We have the following proposition.

Lemma 7 *Assume that g and f are continuous, $Supp(f) \supset [-1, 1]$, and $g(\theta) > 0$ for all $\theta \in (0, 1)$ and $g(0) = g(1) = 0$. Assuming that $x(0) = 0$ and $y(0) = \theta_{med}$, and that $x(1) = \theta_{med}$ and $y(1) = 1$, the expected utility of type θ_{med} is monotonically increasing (decreasing) as the border type $\tilde{\theta}$ is increasing in $[0, \theta_{med}]$ (in $[\theta_{med}, 1]$) if and only if for all $\tilde{\theta} \in [0, \theta_{med}]$ (for all $\tilde{\theta} \in [\theta_{med}, 1]$), $\xi(\tilde{\theta}) \geq (\leq)0$.*

Remark. As $\tilde{\theta}$ increases, $x(\tilde{\theta})$ becomes closer to θ_{med}, while $y(\tilde{\theta})$ moves further away from $\tilde{\theta}$. Condition $\xi(\tilde{\theta}) \geq 0$ means that the expected utility of θ_{med} benefits from the former effect exceeds the latter counter effect. Note that although this condition looks a strong necessary and sufficient condition for the monotonically changing of θ_{med}'s expected utility, it may not need such a strong condition actually for the below reason. Note that the additional term $-2(\tilde{\theta} - \theta_{med})f(\epsilon(x, y))\left(\frac{1}{g(x)} + \frac{1}{g(y)}\right)$ in $\varphi(\tilde{\theta})$ tends to be large in absolute values at $\tilde{\theta} = 0$ and 1 (the first parenthesis is large and the third parenthesis tends to be large if $g(\theta)$ is smaller around $\theta = 0$ and 1). Thus, even if $\xi(\tilde{\theta})$ is not staying either positive in $[0, \theta_{med}]$ or negative in $[\theta_{med}, 1]$, in other wards, the conditions that the above lemma is demanding are not satisfied, the following proposition may still hold since Either $\varphi(\tilde{\theta}) > 0$ in $[0, \theta_{med}]$ or $\varphi(\tilde{\theta}) < 0$ in $[\theta_{med}, 1]$ is enough to tell the same messages as the below.

Proposition 5 *Assume that g and f are continuous, $Supp(f) \supset [-1,1]$, and $g(\theta) > 0$ for all $\theta \in (0,1)$ and $g(0) = g(1) = 0$. If the expected utility of θ_{med} is monotonically increasing (decreasing) as the border type $\tilde{\theta}$ is getting larger in $(0, \theta_{med}]$ (in $[\theta_{med}, 1)$), then $\tilde{\theta}^* \geq (\leq) \theta_{med}$.*

Now, we can use the above facts to discuss how the distribution of voters g affects the equilibrium party lines. Using the above remark, we suppose that when $\xi(\theta_{med}) \geq 0$ (≤ 0), $\varphi(\theta) > 0$ holds for all $\theta \in [0, \theta_{med})$ ($\varphi(\theta) < 0$ holds for all $\theta \in (\theta_{med}, 1]$). Using

$$\xi(\tilde{\theta}) = -\frac{F(\epsilon(x(\tilde{\theta}), y(\tilde{\theta})))}{g(y(\tilde{\theta}))} + \frac{1 - F(\epsilon(x(\tilde{\theta}), y(\tilde{\theta})))}{g(x(\tilde{\theta}))} = -\frac{\Pr(y(\tilde{\theta}))}{g(y(\tilde{\theta}))} + \frac{\Pr(x(\tilde{\theta}))}{g(x(\tilde{\theta}))}$$

where $\Pr(y(\tilde{\theta}))$ is the winning probability of y and $\Pr(x(\tilde{\theta}))$ is that of x, we can talk about equilibrium predictions by focusing on the sign of $\xi(\theta_{med})$. If $g(y(\theta_{med})) = g(x(\theta_{med}))$, then what matters is $\Pr(y(\theta_{med})) \gtreqless \Pr(x(\theta_{med}))$. If $y(\theta_{med})$ is further away from θ_{med} than $x(\theta_{med})$, then $\Pr(y(\theta_{med})) < \Pr(x(\theta_{med}))$ holds since Corollary 1, and $\theta_{med} < \tilde{\theta}$ such that $\varphi(\tilde{\theta}) = 0$ occurs. This means that if the support group of party R goes more extremely right in the policy spectrum, then moderate R supporters tend to switch from R to L. If $\Pr(y(\theta_{med})) = \Pr(x(\theta_{med}))$ that is $\theta_{med} = \frac{x+y}{2}$ while $g(y(\theta_{med})) < g(x(\theta_{med}))$, then a coalitional move from R to L does not change the position of R's candidate much, while the position of L's candidate is pulled toward the median. Thus, the moderate R supporters tend to switch their parties. That is, other things equal, if a party's support group (in terms of their policy spectrum) becomes more extreme, then the the party loses its support, making their candidate more extreme and the opponent party's candidate more moderate (See Figure 4). The probability of the former party's candidate winning is reduced by this.

4 Conclusion

In this paper, we considered a two-party representative democracy, and investigated how the distribution of voters' policy positions on a one-dimensional issue space affects the party line and the probability of each party's winning. We introduce a common shock which affects each voter's utility instead of standard idiosyncratic shocks in the probabilistic voting model. We also introduce a new equilibrium concept, political equilibrium, that is immune to any small coalitional deviations instead of Nash equilibrium and strong equilibrium. This notion makes the characterization of the equilibrium simple by focusing on a sorting allocation case. However we need the psychological cost to be immune to deviation by extreme voters.

In future research, we may consider two extensions although they may be difficult. First, we may try to generalize the functional form of voters' utility function. To be concrete, we may consider strictly convex utility (Osborne 1995) case. In this case, voters are more sensitive to candidates' positions who are close to their own positions. Convex utility function means that voters who are further away from candidates do not take so much care about them. We redefine the utility function of voter as

$$u(p_k; \theta, \epsilon) = -v(|p_k - \theta|) + \epsilon,$$

where $v'(\cdot) < 0$ and $v''(\cdot) > 0$ and $k \in C$ is a winner. [19] This type of utility function is discussed in Osborne (1995) and Kamada & Kojima (2009) in the literature of probabilistic voting model. With such utility function, one may think that extreme left or right voters — voters to the left (right) of the median of party L (R) — do not have incentives to switch parties, and we may be able to drop the assumption of "psychological costs." It is perhaps true that such convex cost function reduces the incentives of switching parties, but it would not totally resolve the problem, since a voter with an extreme position may be better off by having her party's candidate becoming more moderate with a higher chance of winning even if a voter does not care about the other party's candidate's position. It all depends on the relative magnitudes of two effects: dissatisfaction from her party candidate's position becoming more moderate and satisfaction from the candidate's winning probability increase.

Second, it would be interesting to think about the way to make each party supporters to select their candidate strategically in the original Besley-Coate model (Besley and Coate, 1997). One way is to assume that given the party line, each voter tries to find her ideal candidate for the party (depending on her policy position and her candidate's winning probability). It may be possible for us to drop our simple median voter assumption in order to show the existence of equilibrium. However, the characterization of equilibrium can be hard.

Appendix A: Proofs

Proof of Lemma 1 Since each candidate is a median type of each party, $x \leq \theta_{med} \leq y$. Assume that ϵ makes type $\hat{\theta} \in [x, y]$ being indifferent between x and y. Then, $\forall \theta \in [x, y)$ such that $\theta < \hat{\theta}$, and $\forall \bar{\theta} \in [0, x)$,

$$\begin{aligned}
0 = h(x, y; \hat{\theta}) - \epsilon &= -(y - \hat{\theta}) + (\hat{\theta} - x) - \epsilon = 2\hat{\theta} - x - y - \epsilon \\
&> 2\theta - x - y - \epsilon = h(x, y; \theta) - \epsilon \\
&\geq 2x - x - y - \epsilon \\
&= x - y - \epsilon = -(y - \bar{\theta}) + (x - \bar{\theta}) - \epsilon = h(x, y; \bar{\theta}) - \epsilon.
\end{aligned}$$

This means that all citizen-voters of $\theta < \hat{\theta}$ type vote for x when ϵ.

[19] We can also more extreme form with respect not to care about "far candidate" at all, namely a piecewise linear utility of the following form:

$$u(p_k, \theta, \epsilon) = \begin{cases} -|p_k - \theta| + \epsilon & \text{if } \theta - d \leq p_k \leq \theta + d \\ -d + \epsilon & \text{otherwise} \end{cases}$$

That is, voters care only about within policy distance d from their policy positions.

If $\epsilon > \epsilon(x,y)$, then, from

$$\begin{aligned}0 = h(x,y;\hat{\theta}) - \epsilon &= 2\hat{\theta} - x - y - \epsilon \\ &< 2\hat{\theta} - x - y - \epsilon(x,y) \\ &= 2\hat{\theta} - x - y - h(x,y;\theta_{med}) = 2(\hat{\theta} - \theta_{med}),\end{aligned}$$

we have $\hat{\theta} > \theta_{med}$. Hence, x gets majority and wins when $\epsilon < \epsilon(x,y)$.

Similarly, if $\epsilon < \epsilon(x,y)$, $\hat{\theta} < \theta_{med}$ and every type $\theta > \hat{\theta}$ vote for y, and y wins. □

Proof of of Corollary 1 We consider the $h(x,y;\theta)$ such that $\theta = \frac{x+y}{2}$:

$$h(x,y;\frac{x+y}{2}) = -(y - \frac{x+y}{2}) + (\frac{x+y}{2} - x) = 0$$

This means that type $\frac{x+y}{2}$ is indifferent between x and y when $\epsilon = 0$. Thus, when $\theta_{med} < \theta = \frac{x+y}{2}$, we have the below inequality:

$$0 = h(x,y;\frac{x+y}{2}) > h(x,y;\theta_{med}) = \epsilon(x,y)$$

From the assumptions $E(\epsilon) = 0$ and the symmetry of the distribution of ϵ, $F(\epsilon(x,y)) < \frac{1}{2} < 1 - F(\epsilon(x,y))$ can be obtained. The other cases is shown as well as the above. □

Proof of Lemma 2 Differentiating $EU(x(\tilde{\theta}), y(\tilde{\theta}); \theta)$ with respect to $\tilde{\theta}$, we obtain:

$$\begin{aligned}&\frac{dEU(x(\tilde{\theta}), y(\tilde{\theta}); \theta)}{d\tilde{\theta}} \\ &= \frac{g(\tilde{\theta})}{2}\left[-\frac{F(\epsilon(x,y))}{g(y)} + \frac{1-F(\epsilon(x,y))}{g(x)} - 2(\theta - \theta_{med})f(\epsilon(x,y))\left(\frac{1}{g(x)} + \frac{1}{g(y)}\right)\right] \\ &= \frac{g(\tilde{\theta})}{2}\left[-\frac{F(\epsilon(x,y))}{g(y)} + \frac{1-F(\epsilon(x,y))}{g(x)} - 2(\tilde{\theta} - \theta_{med})f(\epsilon(x,y))\left(\frac{1}{g(x)} + \frac{1}{g(y)}\right)\right. \\ &\qquad \left. - 2(\theta - \tilde{\theta})f(\epsilon(x,y))\left(\frac{1}{g(x)} + \frac{1}{g(y)}\right)\right] \\ &= \frac{g(\tilde{\theta})}{2}\left[\varphi(\tilde{\theta}) - 2(\theta - \tilde{\theta})f(\epsilon(x,y))\left(\frac{1}{g(x)} + \frac{1}{g(y)}\right)\right].\end{aligned}$$

This implies that for small $\Delta > 0$, we have

$$\begin{aligned}& EU(x(\tilde{\theta}+\Delta), y(\tilde{\theta}+\Delta); \theta) \\ =\ & EU(x(\tilde{\theta}), y(\tilde{\theta}); \theta) + \int_{\tilde{\theta}}^{\tilde{\theta}+\Delta} \frac{dEU(x(\theta'), y(\theta'); \theta)}{d\theta'} d\theta' \\ =\ & EU(x(\tilde{\theta}), y(\tilde{\theta}); \theta) \\ & + \int_{\tilde{\theta}}^{\tilde{\theta}+\Delta} \frac{g(\theta')}{2}\left[\varphi(\theta') - 2(\theta - \theta')f(\epsilon(x(\theta'), y(\theta')))\left(\frac{1}{g(x(\theta'))} + \frac{1}{g(y(\theta'))}\right)\right] d\theta'.\end{aligned}$$

θ appears only in the brackets as $-2(\theta - \theta')$, so that the second term in this expression is decreasing in θ. Hence, $EU(x(\tilde{\theta}+\Delta), y(\tilde{\theta}+\Delta); \theta) - EU(x(\tilde{\theta}), y(\tilde{\theta}); \theta)$ is decreasing in θ. The latter half of the statement in lemma 2 can be shown by a symmetric argument. □

Proof of Lemma 3 Note that as long as $Supp(\gamma) \subset [\tilde{\theta}, y(\tilde{\theta})]$, the effects of γ switching party from R to L on x and y are the same as the ones of γ_R^Δ switching party from R to L on x and y. Moreover, from the previous lemma, we know that $EU(x(\tilde{\theta}+\Delta), y(\tilde{\theta}+\Delta); \theta) - EU(x(\tilde{\theta}), y(\tilde{\theta}); \theta)$ is decreasing in θ for $\theta \in (\tilde{\theta}, y(\tilde{\theta}))$. Thus, if there is an incentive to join the coalition for $\theta > \tilde{\theta} + \Delta$; i.e., $EU(x(\tilde{\theta} + \Delta), y(\tilde{\theta} + \Delta); \theta) > EU(x(\tilde{\theta}), y(\tilde{\theta}); \theta)$, then all $\theta' \leq \tilde{\theta} + \Delta$ have incentives to join the deviation. A symmetric argument proves the latter half of the statement. □

Proof of Lemma 5 Suppose that $\varphi(\tilde{\theta}) > 0$. Since φ is continuous, there exists $\tilde{\Delta} > 0$ such that $\varphi(\theta) > 0$ for all $\theta \in [\tilde{\theta}, \tilde{\theta} + \tilde{\Delta}]$. Then, from lemma 2, we have

$$\begin{aligned} & EU(x(\tilde{\theta} + \Delta), y(\tilde{\theta} + \Delta); \tilde{\theta} + \Delta) - EU(x(\tilde{\theta}), y(\tilde{\theta}); \tilde{\theta} + \Delta) \\ = & \int_{\tilde{\theta}}^{\tilde{\theta}+\Delta} \frac{g(\theta')}{2} \left[\varphi(\theta') - 2((\tilde{\theta} + \Delta) - \theta') f(\epsilon(x(\theta'), y(\theta'))) \left(\frac{1}{g(x(\theta'))} + \frac{1}{g(y(\theta'))} \right) \right] d\theta'. \end{aligned}$$

By choosing Δ small enough (smaller than $\tilde{\Delta}$), the absolute value of the second term in the brackets becomes smaller than $\varphi(\theta')$, so that we can find an improving coalitional deviation γ_R^Δ. Similarly, if $\varphi(\tilde{\theta}) < 0$, then there is an improving coalitional deviation γ_L^Δ. Hence a sorting allocation with $\tilde{\theta}$ is a political equilibrium only if $\varphi(\tilde{\theta}) = 0$ □

Proof of Lemma 6 First, we will approximate $EU(x(\tilde{\theta}+\Delta), y(\tilde{\theta}+\Delta); \tilde{\theta}+\Delta) - EU(x(\tilde{\theta}), y(\tilde{\theta}); \tilde{\theta}+\Delta)$ by using the first-order Taylor expansion.

$$\begin{aligned} & EU(x(\tilde{\theta} + \Delta), y(\tilde{\theta} + \Delta); \tilde{\theta} + \Delta) - EU(x(\tilde{\theta}), y(\tilde{\theta}); \tilde{\theta} + \Delta) \\ = & \int_{\tilde{\theta}}^{\tilde{\theta}+\Delta} \frac{g(\theta')}{2} \left[\varphi(\theta') - 2(\tilde{\theta} + \Delta - \theta') f(\epsilon(x(\theta'), y(\theta'))) \left(\frac{1}{g(x(\theta'))} + \frac{1}{g(y(\theta'))} \right) \right] d\theta' \\ = & \frac{1}{2} \int_{\tilde{\theta}}^{\tilde{\theta}+\Delta} \varphi(\theta') g(\theta') d\theta' \\ & + \int_{\tilde{\theta}}^{\tilde{\theta}+\Delta} 2(\tilde{\theta} + \Delta - \theta') f(\epsilon(x(\theta'), y(\theta'))) \left(-\frac{g(\theta')}{2} \right) \left(\frac{1}{g(x(\theta'))} + \frac{1}{g(y(\theta'))} \right) d\theta'. \end{aligned}$$

Noting $\varphi(\tilde{\theta}) = 0$, the first term is approximated as

$$\begin{aligned} \frac{1}{2} \int_{\tilde{\theta}}^{\tilde{\theta}+\Delta} \varphi(\theta') g(\theta') d\theta' & \simeq \frac{1}{2} \int_{\tilde{\theta}}^{\tilde{\theta}+\Delta} (\varphi(\tilde{\theta}) g(\tilde{\theta}) + (\varphi'(\tilde{\theta}) g(\tilde{\theta}) + \varphi(\tilde{\theta}) g'(\tilde{\theta}))(\theta' - \tilde{\theta})) d\theta' \\ & = \frac{1}{2} \int_{\tilde{\theta}}^{\tilde{\theta}+\Delta} \varphi'(\tilde{\theta}) g(\tilde{\theta}) (\theta' - \tilde{\theta}) d\theta' \\ & = \frac{1}{2} \varphi'(\tilde{\theta}) g(\tilde{\theta}) \left[\frac{(\theta' - \tilde{\theta})^2}{2} \right]_{\tilde{\theta}}^{\tilde{\theta}+\Delta} \\ & = \frac{\Delta^2}{4} \varphi'(\tilde{\theta}) g(\tilde{\theta}). \end{aligned}$$

In order to calculate the second term, first note that

$$\frac{d}{d\theta'}F(\epsilon(x(\theta'),y(\theta'))) = f(\epsilon(x(\theta'),y(\theta')))\left(-\frac{g(\theta')}{2}\right)\left(\frac{1}{g(x(\theta'))}+\frac{1}{g(y(\theta'))}\right).$$

Thus, partially integrating the second term, we obtain

$$\begin{aligned} & \int_{\tilde\theta}^{\tilde\theta+\Delta} 2\left(\tilde\theta+\Delta-\theta'\right) f(\epsilon(x(\theta'),y(\theta')))\left(-\frac{g(\theta')}{2}\right)\left(\frac{1}{g(x(\theta'))}+\frac{1}{g(y(\theta'))}\right)d\theta' \\ = & \int_{\tilde\theta}^{\tilde\theta+\Delta} 2\left(\tilde\theta+\Delta-\theta'\right)\frac{d}{d\theta'}F(\epsilon(x(\theta'),y(\theta')))d\theta' \\ = & \underbrace{\left[2\left(\tilde\theta+\Delta-\theta'\right)F(\epsilon(x(\theta'),y(\theta'))\right]_{\tilde\theta}^{\tilde\theta+\Delta}}_{A} + \underbrace{\int_{\tilde\theta}^{\tilde\theta+\Delta} 2F(\epsilon(x(\theta'),y(\theta')))d\theta'}_{B}. \end{aligned}$$

Now, term A is rewritten as

$$\begin{aligned} & \left[2\left(\tilde\theta+\Delta-\theta'\right)F(\epsilon(x(\theta'),y(\theta'))\right]_{\tilde\theta}^{\tilde\theta+\Delta} \\ = & 2\left[\left(\tilde\theta+\Delta-\left(\tilde\theta+\Delta\right)\right)F(\epsilon(x(\tilde\theta+\Delta),y(\tilde\theta+\Delta))-\left(\tilde\theta+\Delta-\tilde\theta\right)F(\epsilon(x(\tilde\theta),y(\tilde\theta))\right] \\ = & -2F(\epsilon(x(\tilde\theta),y(\tilde\theta))\Delta. \end{aligned}$$

Since $F(\epsilon(x(\theta'),y(\theta'))) \simeq F(\epsilon(x(\tilde\theta),y(\tilde\theta)) + f(\epsilon(x(\tilde\theta),y(\tilde\theta))\epsilon'(x(\tilde\theta),y(\tilde\theta))\left(\theta'-\tilde\theta\right)$, by substituting $\epsilon'(x(\tilde\theta),y(\tilde\theta)) = f(\epsilon(x(\tilde\theta),y(\tilde\theta)))(-\frac{g(\tilde\theta)}{2})\left(\frac{1}{g(x(\tilde\theta))}+\frac{1}{g(y(\tilde\theta))}\right)$ into this approximation, term B can be approximated as

$$\begin{aligned} & \int_{\tilde\theta}^{\tilde\theta+\Delta} 2F(\epsilon(x(\theta'),y(\theta')))d\theta' \\ \simeq & 2F(\epsilon(x(\tilde\theta),y(\tilde\theta))\int_{\tilde\theta}^{\tilde\theta+\Delta}d\theta' + 2f(\epsilon(x(\tilde\theta),y(\tilde\theta))\epsilon'(x(\tilde\theta),y(\tilde\theta))\int_{\tilde\theta}^{\tilde\theta+\Delta}(\theta'-\tilde\theta)d\theta' \\ = & 2F(\epsilon(x(\tilde\theta),y(\tilde\theta))\Delta + f(\epsilon(x(\tilde\theta),y(\tilde\theta))\left(-\frac{g(\tilde\theta)}{2}\right)\left(\frac{1}{g(x(\tilde\theta))}+\frac{1}{g(y(\tilde\theta))}\right)\Delta^2. \end{aligned}$$

Thus, the second term is $A+B = f(\epsilon(x(\tilde\theta),y(\tilde\theta))(-\frac{g(\tilde\theta)}{2})\left(\frac{1}{g(x(\tilde\theta))}+\frac{1}{g(y(\tilde\theta))}\right)\Delta^2$. Hence, we have the approximation formula:

$$\begin{aligned} & EU(x(\theta'),y(\theta');\tilde\theta+\Delta) - EU(x(\tilde\theta),y(\tilde\theta);\tilde\theta+\Delta) \\ \simeq & \frac{\Delta^2}{4}\varphi'(\tilde\theta)g(\tilde\theta) + f(\epsilon(x(\tilde\theta),y(\tilde\theta))\left(-\frac{g(\tilde\theta)}{2}\right)\left(\frac{1}{g(x(\tilde\theta))}+\frac{1}{g(y(\tilde\theta))}\right)\Delta^2 \\ = & \frac{\Delta^2 g(\tilde\theta)}{2}\left[\frac{\varphi'(\tilde\theta)}{2} - f(\epsilon(x(\tilde\theta),y(\tilde\theta)))\left(\frac{1}{g(x(\tilde\theta))}+\frac{1}{g(y(\tilde\theta))}\right)\right]. \end{aligned}$$

We have completed the proof. □

Proof of Corollary 2 Since f and g are density function, their values are nonnegative. Thus, from lemma 6, we get the conclusion directly. $\square$

Proof of Theorem 1 When $\tilde{\theta} = 0$, candidates of L and R are $x = 0$ and $y = \theta_{med}$, respectively. Then, we have

$$\varphi(0) = -\frac{F(\theta_{med})}{g(\theta_{med})} + \frac{1 - F(\theta_{med})}{g(0)} + 2\theta_{med} f(\theta_{med}) \left(\frac{1}{g(\theta_{med})} + \frac{1}{g(0)} \right).$$

Since $0 < 1 - F(\theta_{med})$ holds from $Supp(f) \supset [-1, 1]$, $g(0) = 0$ assures that $\varphi(0) = \infty$. This implies that there is a small $\underline{\theta} > 0$ close to 0 with $\varphi(\underline{\theta}) > 0$.

When $\tilde{\theta} = 1$, candidates of L and R are $x = \theta_{med}$ and $y = 1$, respectively. Then, we have

$$\varphi(1) = -\frac{F(\theta_{med} - 1)}{g(1)} + \frac{1 - F(\theta_{med} - 1)}{g(\theta_{med})} - 2(1 - \theta_{med}) f(\theta_{med} - 1) \left(\frac{1}{g(1)} + \frac{1}{g(\theta_{med})} \right).$$

Since $0 < F(\theta_{med}) - 1$ holds from $Supp(f) \supset [-1, 1]$, $g(1) = 0$ assures $\varphi(1) = -\infty$.This implies that there is a large $\bar{\theta} < 1$ close to 1 with $\varphi(\bar{\theta}) < 0$.

Since g and f are continuous in θ, $\varphi(\theta)$ is continuous. Thus, there exists at least a $\tilde{\theta} \in (\underline{\theta}, \bar{\theta})$ such that $\varphi(\tilde{\theta}) = 0$ and $\varphi'(\tilde{\theta}) \leq 0$. $\square$

Proof of Proposition 2 From the necessary and sufficient conditions of the sorting political equilibrium in Theorem 1, if $\tilde{\theta}^*$ is a sorting political equilibrium, then

$$\varphi(\tilde{\theta}^*) = -\frac{F(\epsilon(x, y))}{g(y)} + \frac{1 - F(\epsilon(x, y))}{g(x)} - (\tilde{\theta} - \theta_{med}) f(\epsilon(x, y)) \left(\frac{1}{g(y)} + \frac{1}{g(x)} \right) = 0.$$

In addition, if $Supp(f) \supset [-1, 1]$ and $g(\theta) > 0$ for all $\theta \in (0, 1)$ and $g(0) = g(1) = 0$, then a sorting political equilibrium becomes a interior solution from Proposition 1. Thus, with an equilibrium, $f(\epsilon(x, y))(\frac{1}{g(y)} + \frac{1}{g(x)}) > 0$. From these facts, if $-\frac{F(\epsilon(x,y))}{g(y)} + \frac{1-F(\epsilon(x,y))}{g(x)} \neq 0$ in $\varphi(\tilde{\theta})$, then $\tilde{\theta} - \theta_{med} \neq 0$. Hence θ_{med} does not an allocation of a sorting political equilibrium. $\square$

Proof of Proposition 3 Let $\tilde{\theta} = \theta_{med} = \frac{1}{2}$. Then, by symmetry of g, we have $\theta_{med} - x(\theta_{med}) = y(\theta_{med}) - \theta_{med}$, $g(x(\theta_{med})) = g(y(\theta_{med}))$ and $g'(x) = -g'(y)$. Thus, $\epsilon(x(\theta_{med}), y(\theta_{med})) = h(x, y; \theta_{med}) = 0$ is obtained. Since f is symmetric, $1 - F(0) = F(0) = \frac{1}{2}$. Then,

$$-\frac{F(\epsilon(x, y))}{g(y)} + \frac{1 - F(\epsilon(x, y))}{g(x)} = 0.$$

Thus, when $\tilde{\theta} = \theta_{med}$, $\varphi(\theta_{med}) = 0$. In addition to this, by using the above facts, we have

$$\varphi'(\tilde{\theta}) = \frac{g(\theta_{med})}{2g(x)^2} \left[4f(0) - \frac{g'(x)}{g(x)} \right] - \frac{4f(0)}{g(x)}.$$

Moreover, the necessary and sufficient condition in Theorem 1,

$$\frac{\varphi'(\tilde{\theta})}{2} - f(0) \left(\frac{1}{g(x(\tilde{\theta}))} + \frac{1}{g(y(\tilde{\theta}))} \right) < 0$$

is equivalent to

$$\frac{g(\theta_{med})}{4g(x)^2}\left(4f(0) - \frac{g'(x)}{g(x)}\right) - \frac{4f(0)}{g(x)} < 0.$$

Hence, if this condition is satisfied, there is a political equilibrium with $\tilde{\theta} = \theta_{med}$. $\square$

Proof of Proposition 4 Note that $\xi(\theta_{med}) = \varphi(\theta_{med})$. The assumptions guarantee $\varphi(0) > 0$ and $\varphi(1) < 0$. From Theorem 1 (continuity of φ) and Proposition 1, we know that there exists $\tilde{\theta}$ such that $\varphi(\tilde{\theta}) = 0$ with $\varphi'(\tilde{\theta}) \leq 0$ under the assumption of this proposition. This proves the statement of the proposition. $\square$

Proof of Lemma 7 We consider the change of the expected utility of θ_{med} when $\tilde{\theta}$ is moved to the right slightly. Since

$$\begin{aligned}\frac{dEU(x(\tilde{\theta}), y(\tilde{\theta}); \theta)}{d\tilde{\theta}} &= \frac{g(\tilde{\theta})}{2}[-\frac{F(\epsilon(x(\tilde{\theta}), y(\tilde{\theta})))}{g(y(\tilde{\theta}))} + \frac{1 - F(\epsilon(x(\tilde{\theta}), y(\tilde{\theta})))}{g(x(\tilde{\theta}))} \\ &\quad - 2(\theta - \theta_{med})f(\epsilon(x(\tilde{\theta}), y(\tilde{\theta})))\left(\frac{1}{g(x(\tilde{\theta}))} + \frac{1}{g(y(\tilde{\theta}))}\right)\Bigg],\end{aligned}$$

we have

$$\frac{dEU(x(\tilde{\theta}), y(\tilde{\theta}); \theta_{med})}{d\tilde{\theta}} = \frac{g(\tilde{\theta})}{2}\left[-\frac{F(\epsilon(x(\tilde{\theta}), y(\tilde{\theta})))}{g(y(\tilde{\theta}))} + \frac{1 - F(\epsilon(x(\tilde{\theta}), y(\tilde{\theta})))}{x(\tilde{\theta})}\right] = \frac{g(\tilde{\theta})}{2}\xi(\tilde{\theta}). \quad (7)$$

Clearly, for all $\tilde{\theta} \in [0, \theta_{med}]$ (for all $\tilde{\theta} \in [\theta_{med}, 1]$), $\xi(\tilde{\theta}) \geq (\leq) 0$, if and only if $\frac{dEU(x(\tilde{\theta}), y(\tilde{\theta}); \theta_{med})}{d\tilde{\theta}} \geq (\leq) 0$ for all $\tilde{\theta} \in [0, \theta_{med}]$ (for all $\tilde{\theta} \in [\theta_{med}, 1]$). This means that $EU(x(\tilde{\theta}), y(\tilde{\theta}); \theta_{med})$ is monotonically increasing (decreasing), if and only if $\xi(\tilde{\theta}) \geq (\leq) 0$. $\square$

Proof of Proposition 5 From Lemma 7, the expected utility of θ_{med} is monotonically increasing as the border type $\tilde{\theta}$ is getting larger in $[0, \theta_{med}]$ (in $[\theta_{med}, 1]$) if and only if $\xi(\tilde{\theta}) \geq (\leq) 0$. Recall

$$\begin{aligned}\varphi(\tilde{\theta}) &= -\frac{F(\epsilon(x(\tilde{\theta}), y(\tilde{\theta})))}{g(y(\tilde{\theta}))} + \frac{1 - F(\epsilon(x(\tilde{\theta}), y(\tilde{\theta})))}{g(x(\tilde{\theta}))} - 2(\tilde{\theta} - \theta_{med})f(\epsilon(x(\tilde{\theta}), y(\tilde{\theta})))\left(\frac{1}{g(y(\tilde{\theta}))} + \frac{1}{g(x(\tilde{\theta}))}\right) \\ &= \xi(\tilde{\theta}) - 2(\tilde{\theta} - \theta_{med})f(\epsilon(x(\tilde{\theta}), y(\tilde{\theta})))\left(\frac{1}{g(y(\tilde{\theta}))} + \frac{1}{g(x(\tilde{\theta}))}\right).\end{aligned}$$

Note that the latter term is positive for all $\tilde{\theta} < \theta_{med}$, and is negative for all $\tilde{\theta} > \theta_{med}$. This proves that if $\xi(\tilde{\theta}) \geq 0$ for all $\tilde{\theta} \in [0, \theta_{med}]$, then $\varphi(\tilde{\theta}) > 0$ for all $\tilde{\theta} \in [0, \theta_{med})$. Similarly, if $\xi(\tilde{\theta}) \leq 0$ for all $\tilde{\theta} \in [\theta_{med}, 1]$, then $\varphi(\tilde{\theta}) < 0$ for all $\tilde{\theta} \in (\theta_{med}, 1]$. From Theorem 1 and Proposition 1, we know that there exists $\tilde{\theta}$ such that $\varphi(\tilde{\theta}) = 0$ under the assumption of this proposition. These facts mean that there exist all equilibria in $[\theta_{med}, 1)$ if $\xi(\tilde{\theta}) \geq 0$ for all $\tilde{\theta} \in [0, \theta_{med}]$, and there exist all equilibria in $(0, \theta_{med}]$ if $\xi(\tilde{\theta}) \leq 0$ for all $\tilde{\theta} \in [\theta_{med}, 1]$ This proves the statement of the proposition. $\square$

Appendix B: Mixed Distribution Case

Let us start from the case of having general support functions g_R and g_L: i.e., $g_L(\theta) + g_R(\theta) = g(\theta)$ for all $\theta \in [0, 1]$, and let us partition the space of citizen-voter types into three intervals: $[0, x)$, (x, y), and $(y, 1]$. [20] Recall x and y are determined by $G_L(x) = G_L(1) - G_L(x)$ and $G_R(y) = G_R(1) - G_R(y)$ with $x < y$. We will consider a coalitional deviation with size δ from each interval (we can deal with size δ coalitions across the intervals easily by combining the cases).

Let us start with a coalitional deviation with size δ that belongs to the interval (x, y), moving from L to R. In this case, the coalitional deviation reduces the population of party L and increases that of party R by δ. In order to avoid confusions, we denote δ in this case by $\delta^{L\to R}_{(x,y)}$. That is, the new median citizen-voter type x' of party L is determined by

$$G_L(x') = G_L(1) - \delta^{L\to R}_{(x,y)} - G_L(x'),$$

and y' of party R is by

$$G_R(y') + \delta^{L\to R}_{(x,y)} = G_R(1) - G_R(y').$$

By totally differentiating them, we have

$$g_L(x)dx = -d\delta^{L\to R}_{(x,y)} - g_L(x)dx,$$

or

$$\frac{dx}{d\delta^{L\to R}_{(x,y)}} = -\frac{1}{2g_L(x)},$$

and similarly we have,

$$\frac{dy}{d\delta^{L\to R}_{(x,y)}} = -\frac{1}{2g_R(y)}.$$

These derivatives describe that by the small coalitional deviation, both x and y move to the left. Thus, type θ's expected payoff is affected by such a deviation through changes in x and y. First, in order to check the incentives for joining a coalitional deviation from L to R, we consider a citizen-voter of type $\theta \in (x, y)$. For $\theta \in (x, y)$, we have

$$EU(\theta) = -F(\epsilon(x,y))(y-\theta) - (1 - F(\epsilon(x,y)))(\theta - x) + \int_{\epsilon(x,y)}^{+\infty} \epsilon f(\epsilon)d\epsilon.$$

Note that $\theta_{med} \in [x, y]$ holds. Suppose that $\theta_{med} < x < y$. Then, since x and y are the medians of parties L and R, we reach a contradiction. Thus, $\theta_{med} \in [x, y]$ must hold. This implies $\epsilon(x, y) = 2\theta_{med} - x - y$ from (1), and the impact of the coalitional deviation from L

[20] The borders x and y are measure zero, so we ignore them.

to R in the interval (x, y) is written as

$$\begin{aligned}
&\frac{dEU(\theta)}{d\delta^{L\to R}_{(x,y)}}\\
&= \frac{1}{2}\Bigg[\underbrace{\frac{F(\epsilon(x,y))}{g_R(y)} - \frac{1-F(\epsilon(x,y))}{g_L(x)}}_{\text{changes in candidates' positions}}\\
&+ \underbrace{(2\theta - x - y - \epsilon(x,y))\, f(\epsilon(x,y)) \left(\frac{1}{g_L(x)} + \frac{1}{g_R(y)}\right)}_{\text{changes in winning probabilities}}\Bigg].
\end{aligned}$$

Note that θ shows up only in the latter term (the effect due to the changes in winning probabilities), which is an increasing function in θ. Let $\bar{\theta}$ be such that $\frac{dEU(\bar{\theta})}{d\delta^{L\to R}_{(x,y)}} = 0$. Then, for all $\theta < \bar{\theta}$, we have $\frac{dEU(\theta)}{d\delta^{L\to R}_{(x,y)}} < 0$, while for all $\theta > \bar{\theta}$, we have $\frac{dEU(\theta)}{d\delta^{L\to R}_{(x,y)}} > 0$. This implies that coalitions want to move from L to R if they are composed by the types of $\theta \in (\bar{\theta}, y)$ while no small coalition that is composed by types in $(x, \bar{\theta})$ wants to move from L to R. From the above analysis, it is easy to see that if we consider a coalitional deviation with size δ that belongs to the interval (x, y), moving from R to L, then the analysis is symmetrically reversed, more accurately, coalitions want to move from R to L if they are composed by the types of $\theta \in (x, \bar{\theta})$ while no small coalition that is composed by types in $(\bar{\theta}, y)$ wants to move from R to L. This shows that in the interval (x, y), we have the below lemma:

Lemma 8 *In any political equilibrium, a sorting occurs in the interval (x, y):*

$$\begin{aligned}
g_L(\theta) &= g(\theta) \text{ and } g_R(\theta) = 0 \text{ for all } \theta \in (x, \bar{\theta})\\
g_L(\theta) &= 0 \text{ and } g_R(\theta) = g(\theta) \text{ for all } \theta \in (\bar{\theta}, y),
\end{aligned}$$

where $\bar{\theta}$ satisfies $\frac{dEU(\bar{\theta})}{d\delta^{L\to R}_{(x,y)}} = 0$ (if there exists such $\bar{\theta} \in (x, y)$: otherwise, all support either L or R).

Furthermore, if there is not $\bar{\theta} \in (x, y)$ such that $\varphi(\bar{\theta}) = 0$, then either $\frac{dEU(\theta)}{d\delta^{L\to R}_{(x,y)}} > 0$ or $\frac{dEU(\theta)}{d\delta^{R\to L}_{(x,y)}} > 0$ hold. In the former case, citizen-voters in L want to move to R, and in the latter case, they in R want to move to L by the similar logic. These facts show that there is no political equilibrium with semi-pooling allocation in the interval (x, y). In other wards, they show that even if there exists a semi-pooling political equilibrium, it must have a sorting allocation in the interval (x, y) at least.

Next, we consider coalitions that belong to the interval $[0, x)$ and $(y, 1]$, respectively. We assume that there is positive measure of both L and R at least either in $[0, x)$ or in $(y, 1]$, and that the allocation in (x, y) is a sorting allocation from Lemma 8. Then, we consider small coalitions in the interval $[0, x)$ switching from L to R. Similar calculations as above show that the new median citizen-voter type x' of party L is determined by

$$G_L(x') - \delta^{L\to R}_{[0,x)} = G_L(1) - G_L(x'),$$

and y' of party R is by

$$G_R(y') + \delta^{L\to R}_{[0,x)} = G_R(1) - G_R(y').$$

By totally differentiating them, we have

$$\frac{dx}{d\delta^{L\to R}_{[0,x)}} = \frac{1}{2g_L(x)},$$

and similarly we have,

$$\frac{dy}{d\delta^{L\to R}_{[0,x)}} = -\frac{1}{2g_R(y)}.$$

These derivatives describe that by the small coalitional deviation, x moves to the right and y move to the left. Thus, type θ's expected payoff is affected by such a deviation through changes in x and y. Since we are checking the incentive of the coalition member to join the deviation, we have $\theta < x$. Thus, for $\theta \in [0, x)$ we have

$$EU(\theta) = -F(\epsilon(x,y))(y-\theta) - (1 - F(\epsilon(x,y)))\,(x-\theta) + \int_{\epsilon(x,y)}^{+\infty} \epsilon f(\epsilon) d\epsilon.$$

Since $\theta_{med} \in [x, y]$, $\epsilon(x, y) = 2\theta - x - y$. Thus, the impact of the coalitional deviation from L to R in the interval $[0, x)$ is written as

$$\begin{aligned}&\frac{dEU(\theta)}{d\delta^{L\to R}_{[0,x)}}\\ =\ & \frac{1}{2}\Bigg[\underbrace{\frac{F(\epsilon(x,y))}{g_R(y)} - \frac{1-F(\epsilon(x,y))}{g_L(x)}}_{\text{changes in candidates' positions}} + \underbrace{(y - x - \epsilon(x,y))\, f(\epsilon(x,y)) \left(\frac{1}{g_L(x)} - \frac{1}{g_R(y)}\right)}_{\text{changes in winning probabilities}}\Bigg]. \quad (8)\end{aligned}$$

Note that the above formula does not contain θ. This means that for all $\theta < x$, the incentive to deviate is common, namely not depending on type. Moreover, we have

$$\frac{dEU(\theta)}{d\delta^{L\to R}_{[0,x)}} = -\frac{dEU(\theta)}{d\delta^{R\to L}_{[0,x)}}.$$

This implies that if both L and R has positive measure in $[0, x)$, then there are incentives for moving from R to L unless $\frac{dEU(\theta)}{d\delta^{L\to R}_{[0,x)}} = 0$ happen to hold. In addition, L needs to have some positive measure in the interval $[0, x)$ at a political equilibrium under the two-party. Thus, there is a semi-pooling equilibrium with positive measure of R in the interval $[0, x)$ only if (8) equals zero.

Finally, consider small coalition in the interval $(y, 1]$ switching from L to R, similarly. The new median citizen-voter type x' of party L is determined by

$$G_L(x') + \delta^{L\to R}_{(y,1]} = G_L(1) - G_L(x') - \delta^{L\to R}_{(y,1]},$$

and y' of party R is by

$$G_R(y') = G_R(1) - G_R(y') + \delta^{L\to R}_{(y,1]}.$$

By totally differentiating them, we have

$$\frac{dx}{d\delta^{L\to R}_{(y,1]}} = -\frac{1}{2g_L(x)},$$

and similarly we have,

$$\frac{dy}{d\delta^{L\to R}_{(y,1]}} = \frac{1}{2g_R(y)}.$$

These derivatives are exactly the opposite of the case of $\delta^{L\to R}_{[0,x)}$ described above. Thus, for $\theta \in (y,1]$ we have

$$EU(\theta) = -F(\epsilon(x,y))(\theta - y) - (1 - F(\epsilon(x,y)))(\theta - x) + \int_{\epsilon(x,y)}^{+\infty} \epsilon f(\epsilon) d\epsilon.$$

Thus, the impact of the "L to R" coalitional deviation from the interval $(y,1]$ is written as

$$\begin{aligned} &\frac{dEU(\theta)}{d\delta^{L\to R}_{(y,1]}} \\ = \;& \frac{1}{2}\Bigg[\underbrace{\frac{F(\epsilon(x,y))}{g_R(y)} - \frac{1-F(\epsilon(x,y))}{g_L(x)}}_{\text{changes in candidates' positions}} + \underbrace{(y - x - \epsilon(x,y))\, f(\epsilon(x,y)) \left(\frac{-1}{g_L(x)} + \frac{1}{g_R(y)}\right)}_{\text{changes in winning probabilities}}\Bigg]. \quad (9) \end{aligned}$$

Thus, the first terms of $\frac{dEU(\theta)}{d\delta^{L\to R}_{[0,x)}}$ and $\frac{dEU(\theta)}{d\delta^{L\to R}_{(y,1]}}$ (and $\frac{dEU(\theta)}{d\delta^{L\to R}_{(x,y)}}$) are all common, while the second terms of $\frac{dEU(\theta)}{d\delta^{L\to R}_{[0,x)}}$ and $\frac{dEU(\theta)}{d\delta^{L\to R}_{(y,1]}}$ have opposite signs with the same absolute value. The above formula does not contain θ, and for all $y < \theta$, the incentive to deviate is common. Moreover, we have

$$\frac{dEU(\theta)}{d\delta^{R\to L}_{(y,1]}} = -\frac{dEU(\theta)}{d\delta^{L\to R}_{(y,1]}}.$$

This implies that if both L and R has positive measure in $(y,1]$, then there are incentives for moving from L to R unless $\frac{dEU(\theta)}{d\delta^{L\to R}_{[0,x)}} = 0$ happen to hold. In addition, R needs to have some positive measure in the interval $(y,1]$ at a political equilibrium under the two-party. Thus, there is a semi-pooling equilibrium with positive measure of L in the interval $(y,1]$ only if (9) equals zero.

From the above discussion, we recognized that semi-pooling equilibria are only if $\frac{dEU(\theta)}{d\delta^{L\to R}_{[0,x)}} = 0$ or $\frac{dEU(\theta)}{d\delta^{L\to R}_{[0,x)}} = 0$. However, considering the following coalitional deviation, we can understand that there might be often no semi-pooling equilibria.

Now, we consider a sufficiently small coalitional deviation of the following components. One component is sufficiently small $\delta^{L\to R}_{[0,x)} > 0$ in the interval $[0,x)$ switching from L to R such that there are both parties' citizen-voters of the same types as those of citizen-voters participating in the coalition, namely $g_L(\theta) > 0$ and $g_R(\theta) > 0$ for all θ participating in the coalition. [21] The other component is the same size component, $\delta^{L\to R}_{(x,y)} = \delta^{L\to R}_{[0,x)}$ in the interval

[21] Precisely, it is sufficient that there are sufficiently many citizen-voters of the other party R within the distance d from each member of the coalition, so that the psychological cost is not occurred. Recall Definition 1.

$(\bar{\theta}-\eta, \bar{\theta})$ switching from L to R such that $\eta > 0$ and sufficiently small. We assume $\bar{\theta} > \theta_{med}$. This coalitional deviation does not affect x, namely x is not moved by this coalitional deviation since the deviation size on x's left is the same as that on x's right. On the other hand, the coalitional deviation move y to left. Then, the expected utilities of all members participating in the coalition strictly improve. The effect of this coalitional deviation is shown as the following:

$$G_L(x') - (\delta^{L\to R}_{[0,x)} - \delta^{L\to R}_{(x,y)}) = G_L(1) - G_L(x')$$

Since $\delta^{L\to R}_{(x,y)} = \delta^{L\to R}_{[0,x)}$, we have $G_L(x') = G_L(1) - G_L(x')$. By totally differentiating this, we have

$$2g_L(x)dx = 0 \tag{10}$$

Thus, $dx = 0$, namely x does not change. On the other hand, y' of party R is by

$$G_R(y') + (\delta^{L\to R}_{[0,x)} + \delta^{L\to R}_{(x,y)}) = G_R(1) - G_R(y').$$

By totally differentiating this,

$$d\delta^{L\to R}_{[0,x)} + d\delta^{L\to R}_{(x,y)} = -2g_R(y)dy \tag{11}$$

These total differentials describe that by the small coalitional deviation, as we mentioned above, x does not move and y moves to the left. Thus, type θ's expected payoff is affected by such a deviation through changes in only y. We are checking the incentive of the coalition member to join the deviation. First, for $\theta \in [0, x)$, we have again

$$EU(\theta) = -F(\epsilon(x,y))(y-\theta) - (1-F(\epsilon(x,y)))(x-\theta) + \int_{\epsilon(x,y)}^{+\infty} \epsilon f(\epsilon)d\epsilon.$$

Since $\theta_{med} \in [x, y]$, $\epsilon(x,y) = 2\theta - x - y$. Thus, by taking total differential, and by substituting (10) and (11), the impact of the coalitional deviation is written as

$$dEU(\theta) = \frac{1}{2}\left[\frac{F(\epsilon(x,y))}{g_R(y)} + 2(\theta_{med} - x)f(\epsilon(x,y))\frac{1}{g_R(y)}\right](d\delta^{L\to R}_{[0,x)} + d\delta^{L\to R}_{(x,y)}) > 0$$

Second, similarly, for $\theta \in (\bar{\theta}-\eta, \bar{\theta})$ such that η is positive and $\bar{\theta}-\eta > \theta_{med}$, we have

$$EU(\theta) = -F(\epsilon(x,y))(y-\theta) - (1-F(\epsilon(x,y)))(\theta-x) + \int_{\epsilon(x,y)}^{+\infty} \epsilon f(\epsilon)d\epsilon,$$

and we have

$$dEU(\theta) = \frac{1}{2}\left[\frac{F(\epsilon(x,y))}{g_R(y)} + 2(\theta - \theta_{med})f(\epsilon(x,y))\frac{1}{g_L(x)}\right](d\delta^{L\to R}_{[0,x)} + d\delta^{L\to R}_{(x,y)}) \tag{12}$$

Since $\bar{\theta} > \theta_{med}$ is assumed, (12) is always positive by the other component made in $(\theta_{med}, \bar{\theta})$. Thus, there is a coalitional deviation in the semi-pooling allocation. In other wards, there are no semi-pooling equilibria in this case. On the other hand, if $\bar{\theta} \leq \theta_{med}$, (12) is not always positive since, while the first term in the brackets of (12) is positive, the second term in that might be negative. The second term is the disutility that type θ dislike the decrease of the x's winning probability since all $\theta \in (x, \bar{\theta})$ prefer x to y when $\bar{\theta} \leq \theta_{med}$. Nevertheless, if $\bar{\theta}$ is nearby θ_{med}, then θ participating in the coalition is also nearby θ_{med}, so that the effect of the second term become small. At this time, (9) become positive, and the coalitional deviation occurs.

Appendix C: An Example

In order to illustrate how political equilibrium looks like, including the case where some citizen-voters are distributed extremely in one side, we provide the following example. This example shows that there may be a sorting equilibrium with threshold $\tilde{\theta} \in (0,1)$ such that $\varphi(\tilde{\theta}) = 0$, $\varphi'(\tilde{\theta}) > 0$, and that there may not be no sorting equilibrium with two parties.

Example 2 *Consider the case where $g(\theta)$ is a step function and $f(\epsilon)$ is uniform.*

$$g(\theta) = \begin{cases} \frac{1}{2\theta_{med}} & \text{if } \theta \leq \theta_{med} \\ \frac{1}{2(1-\theta_{med})} & \text{if } \theta > \theta_{med} \end{cases}$$

$$f(\epsilon) = \begin{cases} \frac{1}{2a} & \text{if } \epsilon \in [-a, a] \\ 0 & \text{otherwise} \end{cases}$$

where $a > \frac{1}{2}$ to assure that none of probabilities of x and y wins becomes zero $(1 - F(x) > 0$ and $F(y) > 0)$. In addition, since $g(\theta)$ is symmetrical with respect to $\theta_{med} < \frac{1}{2}$ and $\theta_{med} > \frac{1}{2}$, without loss of generality, we also assume $\theta_{med} \leq \frac{1}{2}$.

With this example, we can explicitly calculate φ function and under the population distribution. Noting that each candidate satisfies $x \leq \theta_{med} \leq y$ under any sorting political equilibria, the types of both candidates become

$$x(\tilde{\theta}) = \frac{\tilde{\theta}}{2} \quad \text{and} \quad y(\tilde{\theta}) = \theta_{med} + \frac{\tilde{\theta}}{2\theta_{med}} - \frac{\tilde{\theta}}{2}.$$

On the basis of these, we need to consider two cases where (i) $\tilde{\theta} \leq \theta_{med}$ and (ii) $\tilde{\theta} > \theta_{med}$.

(i) $\tilde{\theta} \leq \theta_{med}$

First, we calculate the winning probability of each candidate. Since $\epsilon(x, y) = 2\theta_{med} - x - y$ from (1), we have

$$\epsilon(x(\tilde{\theta}), y(\tilde{\theta})) = \theta_{med} - \frac{\tilde{\theta}}{2\theta_{med}}.$$

Thus, the winning probabilities of x and y are

$$1 - F(\epsilon(x(\tilde{\theta}), y(\tilde{\theta}))) = \frac{1}{2a}\left(a - \theta_{med} + \frac{\tilde{\theta}}{2\theta_{med}}\right),$$

and

$$F(\epsilon(x(\tilde{\theta}), y(\tilde{\theta}))) = \frac{1}{2a}\left(\theta_{med} + a - \frac{\tilde{\theta}}{2\theta_{med}}\right),$$

respectively. Substituting them into $\varphi(\tilde{\theta})$*, we obtain*

$$\begin{aligned}\varphi(\tilde{\theta}) &= -2(1-\theta_{med})\frac{1}{2a}\left(\theta_{med} + a - \frac{\tilde{\theta}}{2\theta_{med}}\right) + 2\theta_{med}\frac{1}{2a}\left(a - \theta_{med} + \frac{\tilde{\theta}}{2\theta_{med}}\right) \\ &\quad -2(\tilde{\theta} - \theta_{med})\frac{1}{2a}(2\theta_{med} + 2(1-\theta_{med}))\end{aligned}$$

If $\tilde{\theta}$ *is a threshold of a sorting political equilibrium, it satisfies* $\varphi(\tilde{\theta}) = 0$ *from Lemma 5, namely, at a sorting political equilibrium,*

$$\tilde{\theta} = \frac{2\theta_{med}((2a+1)\theta_{med} - a)}{4\theta_{med} - 1}$$

holds.

The sufficient condition of the sorting political equilibrium in Corollary 2 is

$$\begin{aligned}\varphi'(\tilde{\theta}) &= \frac{1}{a}\left[-(1-\theta_{med})\left(-\frac{1}{2\theta_{med}}\right) + \theta_{med}\frac{1}{2\theta_{med}} - 2\right] \\ &= \frac{1}{a}\left(\frac{1}{2\theta_{med}} - 2\right) \leq 0\end{aligned}$$

which is equivalent to $\theta_{med} \geq 1/4$*.*

The necessary and sufficient condition of the sorting political equilibrium in Proposition 1 is

$$\begin{aligned}&\frac{\varphi'(\tilde{\theta})}{2} - f(\epsilon(x(\tilde{\theta}), y(\tilde{\theta})))\left(\frac{1}{g(x(\tilde{\theta}))} + \frac{1}{g(y(\tilde{\theta}))}\right) \\ &= \frac{1}{2a}\left(\frac{1}{2\theta_{med}} - 2\right)\frac{1}{4\theta_{med}} + \frac{1}{2a}\left(-\frac{1}{2\theta_{med}}\right) \\ &= \frac{1 - 8\theta_{med}}{16a\theta_{med}^2} < 0,\end{aligned}$$

which is equivalent to $\theta_{med} > 1/8$*.*

In addition, we have to obtain the conditions of the interior solution, $\tilde{\theta} \in [0, \theta_{med}]$ *such that* $\varphi(\tilde{\theta}) = 0$*. The conditions of* $\tilde{\theta} \geq 0$ *are* $a \leq \frac{\theta_{med}}{1-2\theta_{med}}$ *when* $\theta_{med} > \frac{1}{4}$ *and* $a \geq \frac{\theta_{med}}{1-2\theta_{med}}$ *when* $\theta_{med} < \frac{1}{4}$*. However, the letter condition is not binding since we are actually assuming* $a \geq \frac{1}{2}$*, so that* $\forall\theta_{med} \in (\frac{1}{8}, \frac{1}{4})$*,* $\frac{1}{2} > \frac{\theta_{med}}{1-2\theta_{med}}$*. On the other hand, the condition of* $\tilde{\theta} \leq \theta_{med}$ *is* $a \geq \frac{1}{2}$ *and* $\theta_{med} \leq \frac{1}{2}$*. These conditions are the same as the assumption.*

If $a \geq \frac{\theta_{med}}{1-2\theta_{med}}$ *in* $\theta_{med} > \frac{1}{4}$*, then* $\varphi(\tilde{\theta}) < 0$ *for all* $\tilde{\theta} \in [0, \theta_{med}]$*. This means that, in this parameter range, there is no political equilibrium such that* $\tilde{\theta} \leq \theta_{med}$ *in* $[0, \theta_{med}]$*. Furthermore, when* $\theta_{med} < \frac{1}{8}$*, there is also no political equilibrium such that* $\tilde{\theta} \leq \theta_{med}$ *since the necessary and sufficient condition is not satisfied. See the figure 3.*

(ii) $\tilde{\theta} \geq \theta_{med}$

In this case, (1) becomes

$$\epsilon(x(\tilde{\theta}), y(\tilde{\theta})) = \frac{-2\theta_{med}^2 + 4\theta_{med} - 1 - \tilde{\theta}}{2(1-\theta_{med})}.$$

Thus, the winning probabilities of x and y become

$$1 - F(\epsilon(x,y)) = \frac{1}{4a(1-\theta_{med})}[2\theta_{med}^2 - (2a+4)\theta_{med} + 2a + 1 + \tilde{\theta}],$$

and

$$F(\epsilon(x,y)) = \frac{1}{4a(1-\theta_{med})}[-2\theta_{med}^2 + (4-2a)\theta_{med} + 2a - 1 - \tilde{\theta}],$$

respectively. Substituting them into $\varphi(\tilde{\theta})$, we obtain

$$\begin{aligned}\varphi(\tilde{\theta}) &= -2(1-\theta_{med})\frac{1}{4a(1-\theta_{med})}[-2\theta_{med}^2 + (4-2a)\theta_{med} + 2a - 1 - \tilde{\theta}] \\ &\quad +2\theta_{med}\frac{1}{4a(1-\theta_{med})}[2a + 1 - (2a+4)\theta_{med} + 2\theta_{med}^2 + \tilde{\theta}] \\ &\quad -2(\tilde{\theta} - \theta_{med})\frac{1}{2a}(2\theta_{med} + 2(1-\theta_{med}))\end{aligned}$$

If $\tilde{\theta}$ is a threshold of a sorting political equilibrium, it satisfies $\varphi(\tilde{\theta}) = 0$ as well as the case (i), namely,

$$\tilde{\theta} = \frac{-(2+4a)\theta_{med}^2 + 6a\theta_{med} - 2a + 1}{3 - 4\theta_{med}}$$

holds.

In addition, the sufficient condition in Corollary 2 is

$$\begin{aligned}\varphi'(\tilde{\theta}) &= -\frac{1}{2a} + \frac{\theta_{med}}{2a(1-\theta_{med})} - \frac{2}{a} \\ &= \frac{-3 + 4\theta_{med}}{2a(1-\theta)} \leq 0\end{aligned}$$

which is equivalent to $\theta_{med} \leq 3/4$. The necessary and sufficient condition in Proposition 1 is

$$\begin{aligned}&\frac{\varphi'(\tilde{\theta})}{2} - f(\epsilon(x(\tilde{\theta}), y(\tilde{\theta})))\left(\frac{1}{g(x(\tilde{\theta}))} + \frac{1}{g(y(\tilde{\theta}))}\right) \\ &= \frac{1}{2}\frac{-3 + 4\theta_{med}}{2a(1-\theta_{med})} + \frac{2}{2a} \\ &= \frac{-7 + 8\theta_{med}}{4a(1-\theta_{med})} < 0\end{aligned}$$

which is equivalent to $\theta_{med} < 7/8$. These conditions are symmetric with case (i), and not biding since $\theta_{med} \leq \frac{1}{2}$ has already been assumed.

As well as the case (i), the conditions of interior solution, $\tilde{\theta} \in (\theta_{med}, 1]$ such that $\varphi(\tilde{\theta}) = 0$ has to be obtained. First, calculating the conditions of $\tilde{\theta} > \theta_{med}$, we have those of $\tilde{\theta} > \theta_{med}$ are $a < \frac{1}{2}$ when $\theta_{med} \leq \frac{1}{2}$ and $a > \frac{1}{2}$ when $\theta_{med} > \frac{1}{2}$. However, neither of them satisfies the assumptions which are $a \geq \frac{1}{2}$ and $\theta_{med} \leq \frac{1}{2}$. (Just in case, the condition of $\tilde{\theta} \leq 1$ is

$a \geq \frac{1-\theta_{med}}{2\theta_{med}-1}$. This condition is always satisfied since the right side is negative.) From the above, there is no political equilibrium such that $\tilde{\theta} > \theta_{med}$ in $(\theta_{med}, 1]$ under the assumptions.

As a results, under those given assumptions, we found out that there is a sorting equilibrium which satisfies the sufficient condition or satisfies the necessary and sufficient condition with $\varphi(\tilde{\theta}) > 0$. □

References

[1] Baron, D.P., (1993), "Government Formation and Endogenous Parties," *American Political Science Review* 87, pp. 34-47.

[2] Besley, T., and S. Coate, (1997), "An Economic Model of Representative Democracy," *Quartery Journal of Economics* 112, pp. 85-114.

[3] Besley, T., and S. Coate, (2003), "Centralized versus Decentralized Provision of Local Public Economy Approach," *Journal of Public Economics* 87, pp. 2611-2637.

[4] Bordignon & G. Tabellini (2009), "Moderating Political Extremism: Single Round vs Runoff Elections under Plurality Rule," IGIER, Bocconi University, working paper No. 348.

[5] Coughlin, P.J., *Probabilistic Voting Theory*, Cambridge University Press (New York), 1992.

[6] Feddersen, T.J. (1992), "A Voting Model Implying Duverger's Law and Positive Tournout," *American Journal of Political Science* 36 - 4, pp. 938 - 962.

[7] Jackson, M.O., and B. Moselle, "Coalition and Party Formation in a Legislative Voting Game," *Journal of Economic Theory* 103, pp. 49-87.

[8] Kamada, Y., and F. Kojima, (2009), "Voter Preferences, Polarization, and Electoral Policies", mimeo.

[9] Osborne, M.J., (1995), "Spatial Models of Political Competition Under Plurality Rule: A Survey of Some Explanations of the Number of Candidates and the Positions They Take," *Canadian Journal of Economics* 26 - 2, pp. 261-301.

[10] Osborne, M.J., and A. Slivinski, (1996) "A Model of Political Competition with Citizen Candidates," *Quarterly Journal of Economics* 111, pp. 65-96.

[11] Osborne, M.J., and R. Tourky (2008), "Party Formation in Single-Issue Politics," *Journal of the Europian Economic Association* 6 - 5, pp. 974 - 1005.

[12] Riviere, A. (1999), "Citizen Candidacy, Party Formation and Duveger's Law," mimeo.

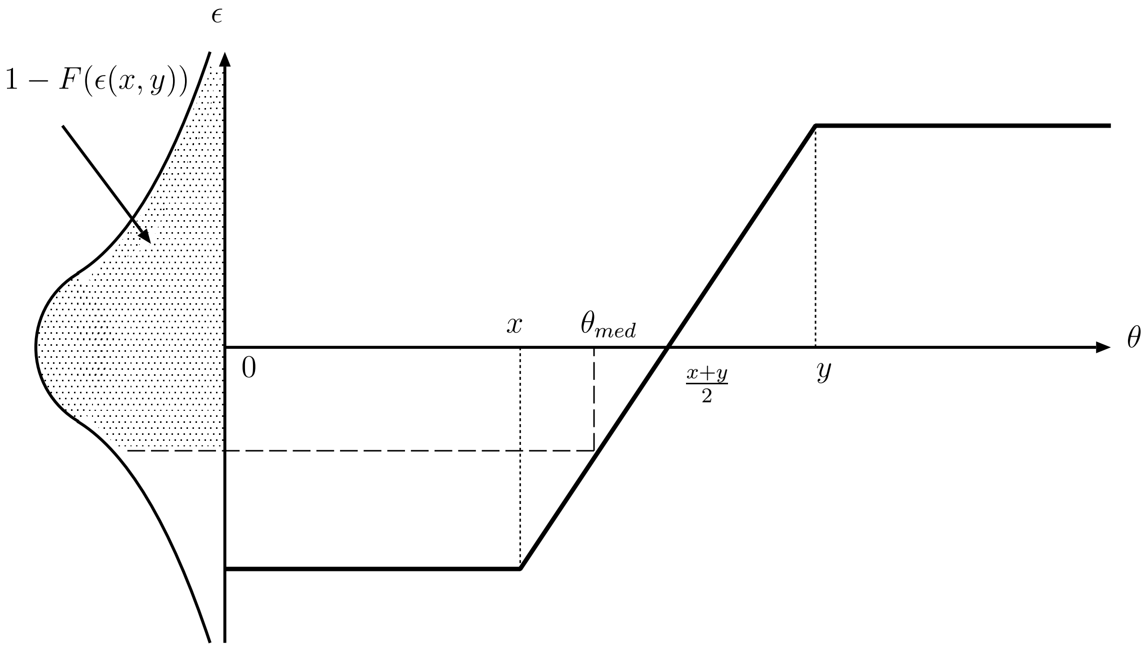

Figure 1

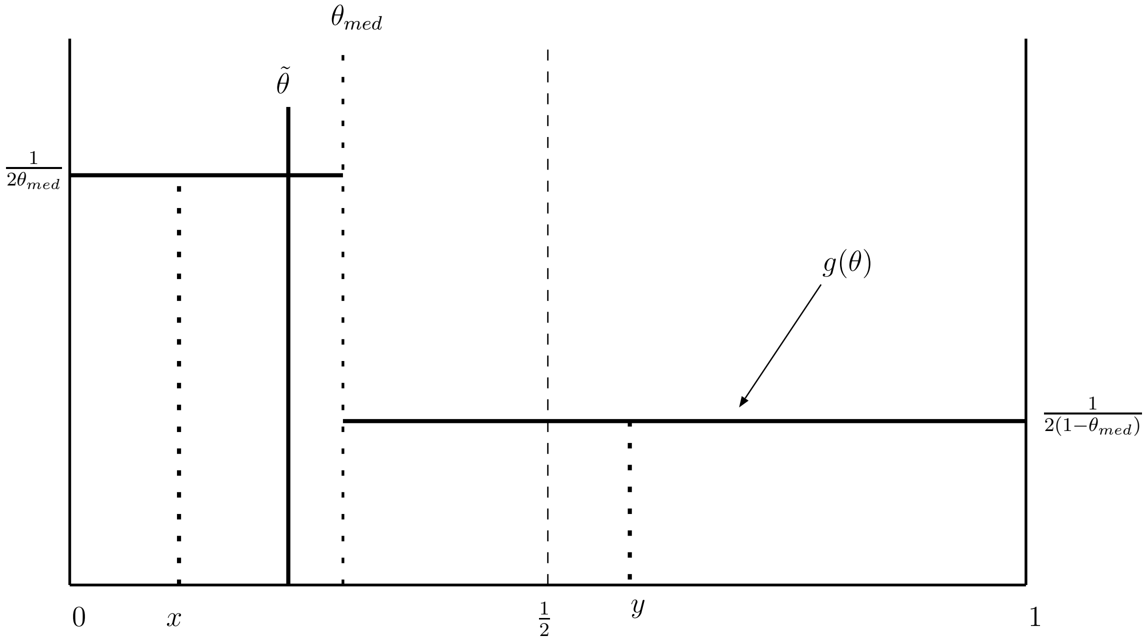

Figure 3: **Example 2**

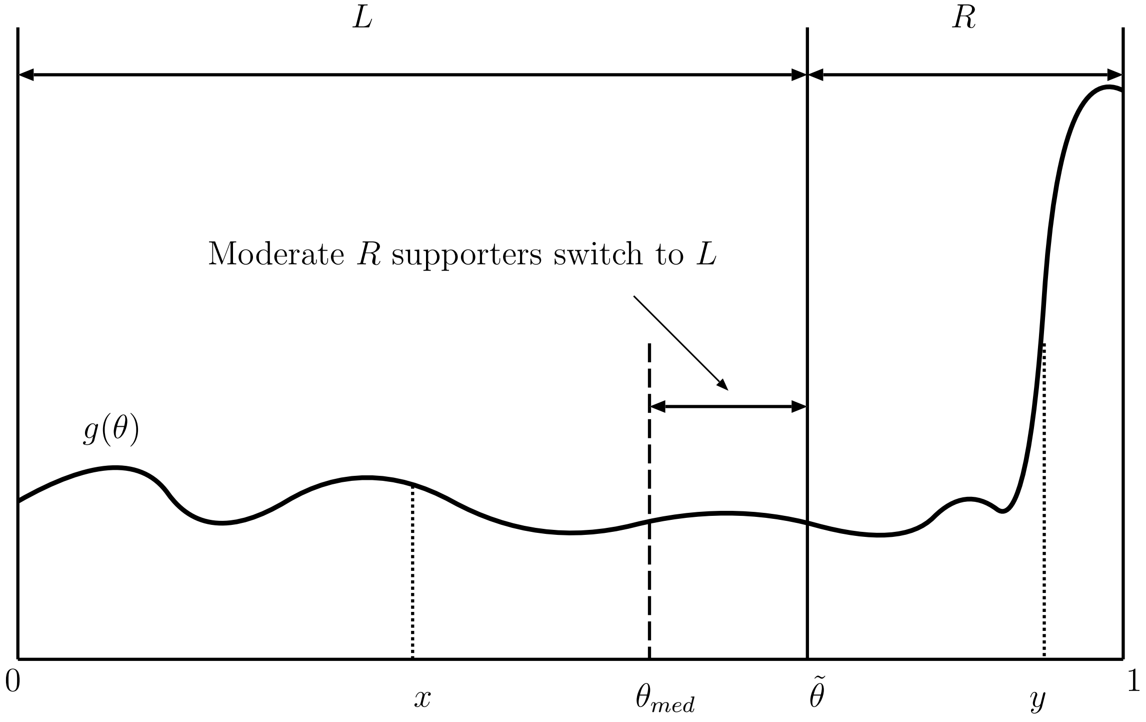
L
R
Moderate R supporters switch to L
$g(\theta)$
0
x
θ_{med}
$\tilde{\theta}$
y
1

Figure 4