How Does Political Uncertainty Affect the Optimal Degree of Policy Divergence?

Aytimur, Emre and Boukouras, Aris and Suen, Richard M. H.

2024

How Does Political Uncertainty Affect the Optimal Degree of Policy Divergence?

R. Emre Aytimur*† Aris Boukouras* Richard M. H. Suen*

September 26, 2024

Abstract

We examine how the optimal degree of policy divergence between two policy platforms in an election is affected by two types of aggregate uncertainty: policy-related and candidate-specific. We show that when the candidate-specific uncertainty is sufficiently large, policy convergence becomes optimal. We also show that when these two types of uncertainty co-exist, only purely office-motivated parties result in policy convergence, in other words, any level of policy motivation of parties results in some policy divergence, making policy motivation undesirable when candidate-specific uncertainty is sufficiently large.

*Department of Economics, Finance and Accounting, School of Business, University of Leicester, University Road, Leicester LE1 7RH, United Kingdom.

† *Corresponding Author.* Email: *rea22@leicester.ac.uk*

1 Introduction

Many experts and the public seem to agree that it is not desirable that political parties champion very similar, if not the same, policies, because voters will have no meaningful choice. In spite of this consensus, the most famous theoretical result on electoral competition, namely the median voter theorem (Downs, 1957a,b), predicts precisely that parties choose the same policy platform, which is the ideal policy of the median voter. Moreover, this is socially optimal as long as the welfare maximizing policy coincides with the median voter's ideal policy. More generally, under some usual assumptions such as symmetric, single-peaked and risk-averse preferences of voters, voters prefer both parties offering the median voter's ideal policy rather than the two parties offering policies symmetric and equally distant from the median voter's ideal policy (such that they each have a chance to win the election).

Bernhardt et al. (2009) propose a potential solution to this puzzle. They argue that, if there is a shock on voter preferences so that the ideal policy positions of voters shift after political parties have made their policy commitments to voters (but before the election takes place), then parties offering different platforms symmetrically situated around the median voter's expected ideal policy are better for *all* voters than parties both offering the median voter's expected ideal policy. This is so because the presence of divergent policy platforms offers to voters hedging against the preference shock. They also show that policy-motivated parties are better in terms of the social welfare for a wide parameter range, because they offer divergent policy platforms, whereas office-motivated parties offer the same policy in equilibrium.

We contribute in this debate by, first, observing that political campaigning often entails more than one source of uncertainty. Indeed, in many recent campaigns in both the US and Europe issues such as the competency of handling the coronavirus pandemic, the moral character of the candidates, and gender or minority-race representation, have been as central as (if not more central than) political platforms on fiscal policies, immigration and foreign policy. While Bernhardt et al. (2009) model explicitly only uncertainty over political preferences, uncertainty over candidate characteristics matters as well. We use the terms candidate-specific uncertainty and preference uncertainty to distinguish between these two broad categories. Accordingly, we generalize Bernhardt et al. (2009) by incorporating both types of uncertainty into a single model.

Second, we find that the inclusion of candidate-specific uncertainty is consequential for the welfare properties of the median-voter theorem. When candidate-specific uncertainty is strong enough, convergent policy platforms are optimal for all voters. Hence, whether convergence of political platforms is socially desirable or not depends on the relative strength of these two uncertainty sources.

Third, and perhaps more importantly, the welfare properties of any political equilibrium depend on the interaction between the two types of uncertainty and the motives of political parties. As stated earlier, Bernhardt et al. (2009) show that when there is only preference uncertainty, then policy-motivated candidates propose platforms closer to the social optimum than office-motivated platforms, provided candidate ideologies are not too extreme. However, when we add candidate-specific uncertainty into their model, the results change. When candidate-specific uncertainty is strong enough, then office-motivated candidates are better positioned to serve the electorates' interests than policy-motivated parties. This is because office-motivated parties generate the first-best political equilibrium, i.e. policy platforms which are convergent to the median voter's expected ideal policy, whereas policy-motivated parties generate policy divergence, which is suboptimal in this case.

Finally, when policy convergence is optimal in our model, then parties with any positive degree of mixed motives between office and policy generate suboptimal political equilibria. This is because, with two sources of uncertainty, parties with mixed-motives always produce some degree of policy divergence no matter how small. Note that this is contrast to the Bernhardt et al. (2009) model, where a finite level of office motivation is sufficient to generate full policy convergence of equilibrium platforms to the median voter's preferred policy. In our model this is impossible, because candidate-specific uncertainty incentivizes parties with mixed motives to deviate from full convergence. Intuitively, they trade-off an infinitesimally small probability of losing the election for a positive payoff due to policy motivations.

As a general message, we see that whether policy convergence and office-motivated parties are desirable depends on the situation, and more specifically on the relative importance of different types of uncertainties. For instance, when a population considers a new area in policy such as big tech regulation, policy preference uncertainty is likely to be high, and this would imply that voters would be better off with different choices coming from policy-motivated parties. On the other hand, when there is high uncertainty with respect to candidate characteristics, which is more likely with the emergence of new candidates in a party structure, office motivation becomes valuable for voters since it induces policy convergence.

Ever since Downs's seminal work (Downs, 1957a,b), candidates' position choice is a central topic in political economy. While the classical median voter framework identifies reasons for platform convergence, many subsequent electoral competition models develop different reasons for policy divergence, including policy motivation (Wittman, 1983; Calvert, 1985; Londregan and Romer, 1993; Osborne and Slivinski, 1996; Besley and Coate, 1997; Martinelli, 2001; Gul and Pesendorfer, 2009); entry deterrence (Palfrey, 1984; Callander, 2005); agency problems (Van Weelden, 2013); incomplete information among voters or candidates (Castanheira, 2003; Bernhardt et al., 2007; Callander, 2008); and differential candidate valence (Bernhardt and Ingerman, 1985; Groseclose, 2001; Krasa and Polborn, 2010, 2012; Bierbrauer and Boyer, 2013).

There are two main differences between this paper and the above literature. First, most of the above papers focus on positive analysis only. The welfare implications of the deviations from the median voter's ideal policy for the rest of the electorate are not examined. In contrast, following Bernhardt et al. (2009), we identify conditions so that the deviation may be welfare enhancing or welfare reducing *for all voters*. Second, we show that, even if a benevolent planner were able to carefully calibrate the degree of office motivation or policy motivation of the candidates (but can not completely remove it), there are cases where the first-best policy can not be implemented as long as parties have any level of policy motivation, however small. That is, political competition is an imperfect instrument for adopting the socially optimal policy in our paper.

Naturally, there are several sources of uncertainty that are relevant for political competition. Voters may exhibit idiosyncratic shocks to their partisanship (Hinich et al., 1972; Lindbeck and Weibull, 1993; Banks and Duggan, 2005), they may be influenced by common shocks to their policy preferences (Bernhardt et al., 2009; Roemer and Roemer, 2009), or unforeseen events may force them to re-evaluate the valence of the candidates (Groseclose, 2001; Adams et al., 2005; Schofield, 2007; Adams and Merrill III, 2009). Callander (2011) proposes an additional source, uncertainty over how policies map into outcomes and advocates for policy experimentation playing a role in platform divergence. All these sources are important determinants of voting decisions and political outcomes. Our analysis focuses on two of them because we want to make a simple point. The welfare properties of a political equilibrium depend not only on which type of uncertainty dominates, but also on the motives of political parties.

Moreover, we are not aware of any other paper that consider both preference and candidate-specific uncertainty simultaneously in the same model. In general, one of these two uncertainties is assumed mainly for analytical convenience. We demonstrate that the inclusion of both types of uncertainty may change the welfare properties of the political equilibrium and especially so when party motives are allowed to vary. Ashworth and De Mesquita (2009) consider a different problem than ours first with preference uncertainty, and second with candidate-specific uncertainty and obtain drastically different results. Therefore, they emphasize "*the overlooked substantive importance of common modeling assumptions*".

We start with a simple example in the next section before we present the general model in Section 3 and our result on optimal policy divergence in Section 4. The following Section 5 studies the political equilibrium with political parties with different policy and office motivations in order to see which types of parties are better fitted for voter welfare. Finally, Section 6 briefly concludes.

2 A simple example

Before we introduce the full details of the model, we illustrate the core idea behind our results through a simple example[1]. We first revisit the intuition presented in Bernhardt et al. (2009), where policy divergence is beneficial for all voters when the only source of uncertainty is preference uncertainty. We then introduce an additional source of uncertainty, namely the candidate-specific uncertainty, to demonstrate how this can lead to the conclusion that policy convergence might be optimal for voters.

Suppose that voters have quadratic loss preferences over policy, with the median voter's ideal policy set at zero. An aggregate shock $\mu \in \{-1, +1\}$, occurring with equal probability for both outcomes, shifts the ideal policy of every voter by $+\mu$. Consequently, a voter v with an initial ideal policy δ_v will have an ideal policy of $\delta_v + \mu$ once the shock is realized. The aggregate shock μ is realised before the election takes place but after candidates have committed to their political platforms. The question is whether voters, from an ex ante perspective, that is, before knowing the realized value of the shock, prefer convergent or divergent political platforms.

We compare the expected utility of a voter between a scenario where both candidates converge at zero and another scenario featuring symmetric divergent platforms: the left-wing candidate L locates at $-a$, and the right-wing candidate R locates at $+a > 0$. In the case of divergent platforms, the left-wing party secures the median voter's support and wins the election if and only if $\mu = -1$. A voter with initial ideology δ_v has an expected utility before the realization of the shock given by:

$$-\frac{1}{2}(\delta_v + 1 - a)^2 - \frac{1}{2}(\delta_v - 1 - (-a))^2 = -(a-1)^2 - \delta_v^2. \tag{1}$$

If the parties converge at zero, this voter's ex ante expected utility becomes:

$$-\frac{1}{2}(\delta_v + 1 - 0)^2 - \frac{1}{2}(\delta_v - 1 - 0)^2 = -1 - \delta_v^2. \tag{2}$$

The expression in (1) exceeds that in (2) for any $a \in (0, 2)$, indicating that the voter benefits more from divergent platforms as long as the divergence is not too strong. This is because platform divergence provides voters with insurance against the risk associated with shifts in their preferences. This improvement applies across all voters, irrespective of their preferences. If $a > 2$, the polarization of policies outweighs the advantage of having a policy choice.

This paper offers a refinement of this insight by showing that when candidate-specific uncertainty, such as candidate valence, is introduced, alongside policy-related uncertainty, the previous findings may reverse. To illustrate, in the earlier example, assume that,

[1] We are very grateful to an anonymous referee for constructing and suggesting this simple example.

in addition to the shock μ, candidates L and R generate net payoffs of $\gamma/2$ and $-\gamma/2$, respectively, to all voters regardless of ideology. We assume $\gamma \in \{-\kappa, \kappa\}$, with $\kappa > 0$ and equal probability for both realizations. We can interpret γ as candidate valence. Therefore, L (respectively R) has higher valence when $\gamma = \kappa$ (respectively $\gamma = -\kappa$). All other aspects of the model remain unchanged.

The median voter, with an ideal policy of zero, remains decisive in this model. If the median voter favors one candidate on both policy and valence dimensions, they will vote for that candidate. Consider the scenario where valence favors L (i.e., $\gamma = \kappa$), but the common policy shock favors R (i.e., $\mu = 1$). In this situation, the median voter prefers the valence-advantaged candidate L if and only if:

$$-(a-1)^2 - \kappa/2 \leq -(-a-1)^2 + \kappa/2,$$

which simplifies to:

$$4a \leq \kappa.$$

Similarly, if $\gamma = -\kappa$ but the common policy shock favors L (i.e., $\mu = -1$), the median voter prefers the valence-advantaged candidate R if and only if $4a \leq \kappa$.

Therefore, if $4a \leq \kappa$, or equivalently $a \leq \frac{\kappa}{4}$, valence plays a sufficiently crucial role such that the median voter prefers the valence-advantaged candidate regardless of the realization of μ. Conversely, when $a > \frac{\kappa}{4}$, the candidates are polarized enough for the median voter to base their decision on policy.

Initially, consider the case where $a \leq \frac{\kappa}{4}$. In this scenario, a voter with ideal point δ_v receives an ex ante expected payoff of:

$$\frac{\kappa}{2} - \frac{1}{2}\left[\frac{1}{2}(\delta_v + 1 - a)^2 + \frac{1}{2}(\delta_v - 1 - a)^2\right] - \frac{1}{2}\left[\frac{1}{2}(\delta_v + 1 - (-a))^2 + \frac{1}{2}(\delta_v - 1 - (-a))^2\right].$$

This expression can be explained as follows: Since the median voter selects on the basis of valence only, the voter always receives $\kappa/2 > 0$ for valence. The rest of the expression arises because each candidate is elected with probability 1/2, irrespective of μ and the policy benefit continues to depend on μ's realization. Simplifying this expression yields

$$\frac{\kappa}{2} - \delta_v^2 - a^2 - 1,$$

which decreases in a for all voters. The key insight is that when the median voter focuses solely on valence, divergent policy platforms merely introduce risk, and the higher level of policy divergence maintains the mean of the policy lottery but increases variance, reducing the expected utility of risk-averse voters.

In the case of policy convergence, the median voter trivially chooses based on valence and her ex ante expected utility is

$$\frac{\kappa}{2} - \delta_v^2 - 1, \tag{3}$$

which is higher than the payoff from policy divergence since convergence eliminates the undesired policy lottery.

Now, consider the case where $a > \frac{\kappa}{4}$. A voter's expected payoff is

$$\frac{1}{2}\left(-(\delta_v + 1 - a)^2 + \frac{1}{2}\frac{\kappa}{2} + \frac{1}{2}(-\frac{\kappa}{2})\right) + \frac{1}{2}\left(-(\delta_v - 1 - (-a))^2 + \frac{1}{2}\frac{\kappa}{2} + \frac{1}{2}(-\frac{\kappa}{2})\right)$$

With probability 1/2, $\mu = 1$ and R is elected, delivering $\kappa/2$ or $-\kappa/2$ to the voter with equal probabilities. With probability 1/2, $\mu = -1$ and L is elected, similarly delivering $\kappa/2$ or $-\kappa/2$ with equal probabilities. This simplifies to:

$$-\delta_v^2 - (a-1)^2. \tag{4}$$

Notice that this is maximized at $a = 1$. When the median voter chooses based on policy, the best case is when the policy matches $\mu \in \{-1, +1\}$.

The payoff from convergence, as given in (3), is bigger than the payoff from any platform divergence $a > \frac{\kappa}{4}$, as given in (4) if and only if

$$\frac{\kappa}{2} + (a-2)a > 0.$$

Whenever $\kappa > 2$, this condition holds for any $a > \frac{\kappa}{4}$. When γ exceeds a certain threshold, the policy platforms for which the median voter votes based on policy become sufficiently polarized, leading voters to prefer policy convergence. Remember that even in the absence of candidate-specific uncertainty, there is a limit of divergence ($a = 2$) beyond which voters prefer policy convergence.

In summary, when $\kappa > 2$, platform convergence is preferable to any degree of symmetric policy divergence. In other words, the conclusion that voters benefit from policy divergence is reversed when candidate-specific uncertainty reaches a sufficient level relative to policy-related uncertainty.

The above example is specific in its assumption of binary distributions for both candidate-specific and policy-related uncertainties. Nonetheless, it highlights the central trade-off between candidate-specific and policy-preference uncertainty. When these uncertainties follow continuous distributions, the preference for policy convergence or divergence will vary depending on the realizations of μ and γ, necessitating an evaluation of their expectations. The following section introduces a more general model that confirms the intuition that policy convergence is optimal when candidate-specific uncertainty is significantly large compared to policy-related uncertainty.

3 The Model

There is a one-dimensional policy space, a continuum of voters, and two candidates. Each voter, indexed by v, has single-peaked preferences over policies, which are symmetric around his bliss point δ_v. Voters' bliss points are distributed on the interval $[\underline{\delta}, \overline{\delta}]$ and the median voter has $\delta_v = 0$.

The two candidates L and R propose, respectively, $a_L \leq 0$ and $a_R \geq 0$ as policy choices to the electorate. We focus on cases where the two candidates position themselves symmetrically around the median. Thus, we denote by $a \geq 0$ the positioning of the right-wing candidate and by $-a \leq 0$ the positioning of the left-wing candidate. When $a = 0$, the positions of the two candidates converge with the median voter's most preferred policy. This would be indeed the equilibrium if there was no uncertainty about voter preferences, as shown by the celebrated *median voter theorem*.

This standard model is enriched by two sources of uncertainty. The first one is with regards to policy preferences as in Bernhardt et al. (2009). This is captured by an aggregate policy preference shock $\mu \in \mathbb{R}$ which shifts the bliss point of all voters either to the left or to the right by the same distance. Therefore, for any realized value μ, the bliss point of voter v is given by $\delta_v + \mu$. The probability density function of μ is denoted by $f(\cdot)$, which has a single peak at zero and is symmetric around it. The value of μ is drawn and revealed after the candidates made their pledges in the political campaign, but before the election takes place. Hence, it is meant to capture all sources of uncertainty that can shift the political preferences of the electorate en masse during the election campaign. For example, a terrorist attack may enhance the preoccupation of voters with security considerations and push the electorate to become more conservative. Alternatively, a financial crisis may make lax fiscal policies more salient and push the electorate to left-leaning policies.

The second source of uncertainty is with regards to the voters' perception of the candidates themselves and is orthogonal to the policy preference shock. Events such as a corruption scandal or the candidates' performance in broadcasted debates may change the voters' perception on their competence or credibility in one direction or another. We assume that candidates L and R generate net payoffs of $\sigma_\gamma \gamma/2$ and $-\sigma_\gamma \gamma/2$, respectively, to all voters regardless of ideology. $\sigma_\gamma > 0$ is a constant and $\gamma \in \mathbb{R}$ is a random variable with mean zero and unit variance, so the variance of $\sigma_\gamma \gamma$ is σ_γ^2. A positive value of γ means that L gives an extra $\sigma_\gamma \gamma$ units of utility to all voters relative to R. The opposite is true if γ is negative. The probability density function of γ is denoted by $h(\cdot)$, which is strictly positive-valued and differentiable with a single peak at zero and symmetric around it. The corresponding cumulative distribution function is denoted by $H(\cdot)$.

Conditional on the realization of μ, voter v's utility from policy a is given by

$$u(\delta_v + \mu - a) = -(\delta_v + \mu - a)^2 . \tag{5}$$

Quadratic utility function is chosen so as to convey the main message of this paper in a clear and straightforward manner. As we will see in Proposition 1, under this type of preferences, there exists a unique threshold value of σ_γ above which all voters would strictly prefer convergent symmetric policy platforms and below which all voters would strictly prefer divergent platforms.[2] The rationale behind this result will be explained fully later.

The timing of the voting game is as follows: The candidates make their political pledges first, then γ and μ materialize. After the voters observe these values, they cast their ballot sincerely and without abstention. If R wins, voter v's utility is $u(\delta_v + \mu - a) - \sigma_\gamma\gamma/2$, and if L wins, his utility is $u(\delta_v + \mu + a) + \sigma_\gamma\gamma/2$. The candidate that garners the majority of votes wins and his policy is implemented.

4 Optimal Policy Divergence

In this section, we ask the following normative question: Which pair of symmetric policy platforms $(-a, a)$ would maximize voters' ex ante welfare? Then in the next section, we will analyze the political equilibrium and discuss whether they correspond to the socially optimal platforms we study in this section. In the equilibrium analysis, we focus on symmetric political equilibria in which candidates choose their policy platforms without knowing the realization of μ and γ. To make the comparison consistent, we confine our attention to ex ante (i.e., before the uncertainties are realized) socially optimal symmetric policy platforms in this section.

The main result of Bernhardt et al. (2009) is that some degree of policy divergence is socially optimal as all voters prefer it to policy convergence at the median voter's ex ante ideal policy. If we call the optimal policy platforms $(-a^*, a^*)$, then Bernhardt et al. (2009, Proposition 2) states that it is socially optimal to have $a^* > 0$. As explained in Section 2, aggregate uncertainty over political preferences generates risk for voters and the divergence of political platforms hedges against it. This hedging mechanism, however, may not work in the presence of candidate-specific uncertainty.[3] Our first main result in this paper qualifies this statement using the model presented in Section 3. Specifically, we show that when the degree of candidate-specific uncertainty exceeds a certain threshold, then

[2] In the working paper version of this paper, we consider a more general utility function $u(x) \equiv w[d(x)]$, where $w : \mathbb{R}_+ \to \mathbb{R}$ is a twice differentiable loss function and $d : \mathbb{R} \to \mathbb{R}_+$ is a twice differentiable distance function. It is shown that under this general specification, policy convergence is preferred by all voters when σ_γ is sufficiently large. The working paper version is available from the authors' personal website.

[3] This type of uncertainty is absent in the model of Bernhardt et al. (2009). Their model is thus equivalent to ours when $\gamma = 0$ with certainty.

policy convergence is preferred by all voters.

We start the analysis with the following observations: First consider the case when both types of uncertainty exist. After the value of μ and γ materialize, voter v will vote for R if and only if

$$u(\delta_v + \mu - a) - \sigma_\gamma \gamma/2 > u(\delta_v + \mu + a) + \sigma_\gamma \gamma/2.$$

Under quadratic preferences, this can be simplified to become

$$\gamma < \frac{4a\left(\delta_v + \mu\right)}{\sigma_\gamma}. \tag{6}$$

The median voter is decisive in the election, therefore R wins if and only if (6) holds for $\delta_v = 0$, i.e.,

$$\widetilde{w} \equiv \frac{4a\mu}{\sigma_\gamma} > \gamma. \tag{7}$$

Therefore, the expected utility of voter v before the resolution of uncertainty is given by:

$$\begin{aligned} E\left[U_v(a)\right] &= \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{\widetilde{w}} \left[u(\delta_v + \mu - a) - \sigma_\gamma \frac{\gamma}{2} \right] h(\gamma) d\gamma + \int_{\widetilde{w}}^{\infty} \left[u(\delta_v + \mu + a) + \sigma_\gamma \frac{\gamma}{2} \right] h(\gamma) d\gamma \right\} f(\mu) d\mu \\ &= \int_{-\infty}^{\infty} u(\delta_v + \mu - a) H(\widetilde{w}) + u(\delta_v + \mu + a)\left[1 - H\left(\widetilde{w}\right)\right] f\left(\mu\right) d\mu \\ &\quad + \sigma_\gamma \int_{-\infty}^{\infty} \left[\int_{\widetilde{w}}^{\infty} \frac{\gamma}{2} h(\gamma) d\gamma - \int_{-\infty}^{\widetilde{w}} \frac{\gamma}{2} h(\gamma) d\gamma \right] f\left(\mu\right) d\mu. \end{aligned} \tag{8}$$

In the above expression, $H(\widetilde{w})$ and $1 - H\left(\widetilde{w}\right)$ denote, respectively, the winning probability of R and L in the presence of candidate-specific uncertainty.

If this type of uncertainty is absent, so that $\gamma = 0$ with certainty, then R wins if and only if $\mu > 0$. In this case, the ex ante expected utility of voter v is given by

$$E\left[\widetilde{U}_v(a)\right] = \int_0^{\infty} u(\delta_v + \mu - a) f(\mu) d\mu + \int_{-\infty}^{0} u(\delta_v + \mu + a) f(\mu) d\mu. \tag{9}$$

The first integral is the expected utility when R wins, while the second one gives the expected utility when L wins. This can be expressed more succinctly as

$$E\left[\widetilde{U}_v(a)\right] = \int_{-\infty}^{\infty} u\left[\chi_v\left(\mu; a\right)\right] f(\mu) d\mu,$$

where

$$\chi_v\left(\mu; a\right) = \begin{cases} \delta_v + \mu - a & \text{if } \mu > 0 \ (R \text{ wins}) \\ \delta_v + \mu + a & \text{if } \mu < 0 \ (L \text{ wins}). \end{cases}$$

For any realized value μ, $\chi_v\left(\mu; a\right)$ is the difference between voter v's ideal policy, $\delta_v + \mu$, and the final implemented policy, which is a if R wins and $-a$ if L wins. It is obvious that, for any $a > 0$, $\mu - a < \mu$ if $\mu > 0$ and $\mu + a > \mu$ if $\mu < 0$. Thus, if the extent of platform divergence is moderate or small, then a pair of symmetric and divergent policy platform has the effect of narrowing the difference between the voters' ideal policy and the final policy, regardless of which candidate wins the election. A graphical illustration of this is shown in Figures 1(a) and 1(b) which plot the probability densities of $\chi_v(\mu; a)$. For $a = 0$, illustrated in Figure 1(a), this distribution is just a horizontal translation of the distribution of μ, i.e., $f(\mu + \delta_v)$. The case of a small positive value of a is illustrated in Figure 1(b). The black left-hand side and the green right-hand side correspond, respectively, to $\mu < 0$ (when L wins) and $\mu > 0$ (when R wins). Moderately divergent platforms have the effect of condensing the distribution and reducing the risk faced by any risk-averse voter. Hence, they strictly prefer this type of platforms to convergent ones, i.e.,

$$E\left[\widetilde{U}_v(a)\right] > E\left[\widetilde{U}_v(0)\right] = \int_{-\infty}^{\infty} u(\delta_v + \mu) f(\mu) d\mu.$$

This explains the hedging mechanism behind Bernhardt et al. (2009, Proposition 2). This mechanism, however, has a limit. If a is "too large", i.e., when there is "too much" polarization, then it is possible that the two halves will drift too far apart [as in Figure 1(c)], and increases the risks faced by the voters instead.

When comparing between (8) and (9), we can see two major changes once candidate-specific uncertainty is introduced. First, instead of having either $u(\delta_v + \mu - a)$ or $u(\delta_v + \mu + a)$ as payoffs under a given pair of μ and a, the voter will now have a weighted average of the two. The weights are determined by the candidates' winning probabilities which are also dependent on μ and a. Second, the candidate-specific factor γ will directly contribute to the voter's expected utility. This is captured by the last term in (8). It follows that changing the policy platforms $(-a, a)$ will have more than one effect on the voters' ex ante welfare. To gain further insights, consider the marginal effect of policy divergence on the voters' expected utility, which is

$$\frac{dE[U_v(a)]}{da} = \int_{-\infty}^{\infty} \left\{-u'(\delta_v + \mu - a) H(\widetilde{w}) + u'(\delta_v + \mu + a)\left[1 - H(\widetilde{w})\right]\right\} f(\mu) d\mu$$
$$+ \int_{-\infty}^{\infty} \left\{u(\delta_v + \mu - a) - u(\delta_v + \mu + a) - \left[u(a - \mu) - u(a + \mu)\right]\right\} h(\widetilde{w}) \frac{d\widetilde{w}}{da} f(\mu) d\mu. \quad (10)$$

The above expression captures two types of effects. First, an increase in a will have a direct impact on the utilities derived from the proposed policies, i.e., $u(\delta_v + \mu - a)$ and $u(\delta_v + \mu + a)$. The average effect across all possible realizations of μ and γ is represented by the first integral in (10). Second, when a increases, the difference between the candidates' policy platforms increases, and this leads to an increased likelihood that policy preferences determine the winner. This effect is captured by the second integral in (10).

When σ_γ is small, candidate-specific uncertainty has only a negligible impact on election outcome and voters' welfare. Thus, voters can benefit from a mild degree of policy divergence as in Bernhardt et al. (2009). But when σ_γ is sufficiently large, $\widetilde{w}$ will be close

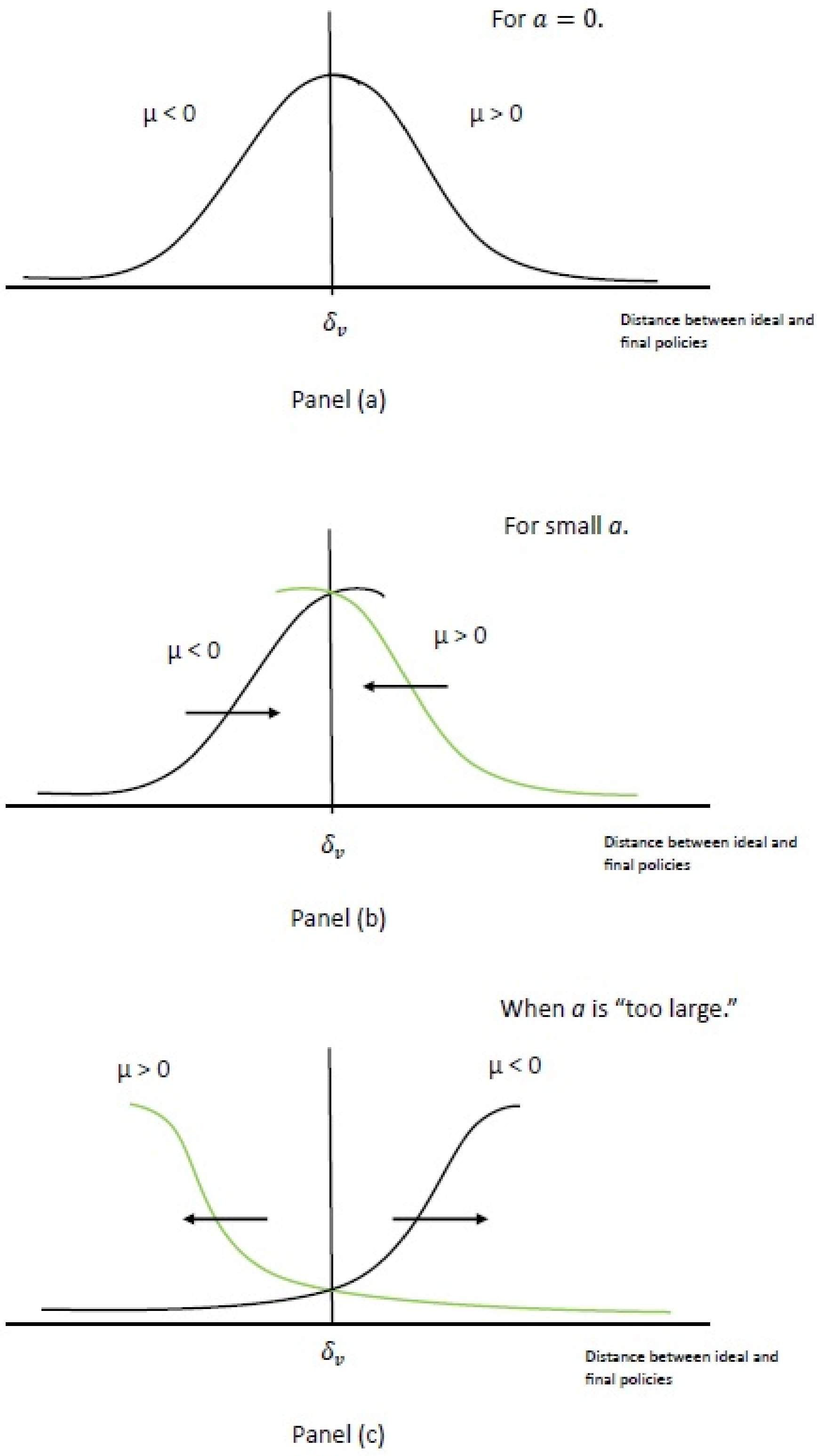

Figure 1: Illustration of the benefit of policy divergence in the absence of candidate-specific uncertainty

to zero except in cases where a or μ is very large, as indicated by equation (7). Therefore, when a is not excessively large, the two candidates will have close to equal probability of winning even when they propose different policies. This is because candidate-specific uncertainty becomes the dominant factor, except in the unlikely event of extremely high realizations of μ. In other words, when σ_γ is sufficiently large, the connection between election outcome and μ weakens sufficiently, leading to the breakdown of the previously discussed hedging mechanism. Finally, in cases where a is very large, policies may still significantly influence the election outcome, but excessive polarization is detrimental to voter welfare, even in the absence of candidate-specific uncertainty. Proposition 1 formally establishes our result. In particular, it is shown that even a small value of a is undesirable for voters if candidate-specific uncertainty is large, that is, when high values of $\sigma_\gamma \gamma$ in absolute value are likely.

Proposition 1. *Under a quadratic policy utility function, all voters have the same preference ordering over policy pairs $(-a, a)$. Define $\sigma_{\min} \equiv 16h(0) \int_0^\infty \mu^2 f(\mu)\, d\mu$.*

1. *If $\sigma_\gamma \geq \sigma_{\min}$, then all voters would strictly prefer convergent symmetric policy platforms $(-a, a) = (0, 0)$ to any other symmetric platforms with $a > 0$. i.e., $E[U_v(0)] > E[U_v(a)]$, for all $a > 0$ and for all v.*

2. *If $\sigma_\gamma < \sigma_{\min}$, then there exists a unique $a^* > 0$ such that all voters would strictly prefer the divergent symmetric policy platforms $(-a^*, a^*)$ to any other symmetric platforms, i.e., $E[U_v(a^*)] > E[U_v(a)]$, for all $a \geq 0$, $a \neq a^*$, and for all v. Furthermore, a^* decreases in σ_γ, i.e. $da^*/d\sigma_\gamma < 0$.*

Under quadratic utility, the derivative in (10) can be shown to be independent of δ_v. Thus, if the median voter prefers policy convergence, then so are the other voters. Bernhardt et al. (2009) also have this result with a quadratic loss function and the policy preference shock μ. Our result shows that this stays true with the additional uncertainty over γ.

An explicit formula for $\sigma_{\min}$ can be obtained if $h(\cdot)$ is the standard normal density function, i.e.,

$$h(\gamma) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\gamma^2\right),$$

and $f(\cdot)$ is the density function of $N\left(0, \sigma_\mu^2\right)$, i.e.,

$$f(\mu) = \frac{1}{\sqrt{2\pi}\sigma_\mu} \exp\left[-\frac{1}{2}\left(\frac{\mu}{\sigma_\mu}\right)^2\right].$$

Then the threshold value $\sigma_{\min}$ can be expressed as[4]

$$\sigma_{\min} = \frac{8}{\pi}\frac{1}{\sigma_\mu}\int_0^\infty \mu^2 \exp\left[-\frac{1}{2}\left(\frac{\mu}{\sigma_\mu}\right)^2\right] d\mu = 4\sqrt{\frac{2}{\pi}}\sigma_\mu^2.$$

This result states that, when the utility function is quadratic and the two shocks are independent normal random variables, the threshold value of σ_γ is directly proportional to σ_μ^2. Thus, whether policy convergence is socially optimal depends on the relative riskiness of the two shocks. In particular, policy convergence is socially optimal if the risk associated with the candidate-specific shock is sufficiently higher than that associated with the policy preference shock. Otherwise, policy divergence is beneficial to all voters. In such cases, the optimal degree of policy divergence decreases as σ_γ increases.

5 Political Equilibrium

We assume that the ideal policies of the two candidates L and R are $-\psi$ and ψ, respectively, with $\psi > 0$. For their payoff functions, we consider two possibilities. First, we assume that they are both policy- and office-motivated; second, we assume that they are purely office-motivated. The timing of the game is as follows: First, L and R announce simultaneously their electoral platforms $a_L \leq 0$ and $a_R \geq 0$, respectively. Then, the uncertainty on μ and γ is resolved, the election takes place, and the winner's electoral platform is implemented. In what follows, we focus on symmetric equilibria, i.e. $a_R = -a_L = a$.

When the candidates are both policy- and office-motivated, given $a_L = -a$, R's payoff function is given by

$$U_R(a_R; -a) = \Pr(\text{R wins})[u(a_R - \psi) + b] + \Pr(\text{L wins})u(-a - \psi)$$

where $u(\cdot)$ is the same quadratic utility function as the voters', $b \geq 0$ represents the office rents in case of victory, and R's winning probability is given by

$$\Pr(\text{R wins}) = \int_{-\infty}^{\infty} H\left(\frac{1}{\sigma_\gamma}(u(a_R - \mu) - u(-a - \mu))\right) f(\mu) d\mu.$$

Finally, $\Pr(\text{L wins}) = 1 - \Pr(\text{R wins})$. An increase in b reflects an increase in the office-motivation component, and conversely, $b = 0$ represents the pure policy-motivation case. When parties are purely office-motivated, their payoff is simply b if they win the election,

[4]The second equality uses the formula

$$\int_0^\infty \mu^2 \exp\left(-A\mu^2\right) d\mu = \frac{1}{4}\sqrt{\frac{\pi}{A^3}},$$

for any $A > 0$, which can be found in any table of integrals. See, for instance, Dwight (1947) Equation 861.7.

and 0 otherwise.

As summarized in the proposition below, we show that policy convergence is an equilibrium if and only if parties are purely office-motivated. When parties are both office- and policy-motivated (or purely policy-motivated), the trade-off of a party is to choose a more moderate policy to increase its chance of winning versus to choose a policy closer to its ideal policy to increase its payoff in case of victory. Bernhardt et al. (2009) shows that if b is large enough, policy convergence is an equilibrium since moving away from the median voter's ideal policy does not pay off since it decreases the winning probability, which is especially damaging when b is high. As opposed to Bernhardt et al. (2009), policy convergence is not an equilibrium in our model even for arbitrarily large values of b, in other words, even when policy motivation is arbitrarily small relative to office motivation. This is because the introduction of candidate-specific uncertainty leads to the result that moving away from the median voter's ideal policy marginally does not reduce the winning probability, i.e. $d\text{Pr}(\text{R wins})/da_R = 0$ at $a_R = a_L = 0$ as shown in the proof of Proposition 2. Therefore parties face no trade-off when moving marginally towards their ideal policies. Intuitively, when, say, party R moves marginally to the right from policy 0, it increases its winning probability for positive realizations of μ but decreases it by the same magnitude for negative realizations of μ. These two effects exactly cancel out in expectation due to the symmetric distribution of μ.

Proposition 2. *When the candidates are not purely office-motivated, in any symmetric equilibrium (when it exists) $(-a^*, a^*)$, $a^* \neq 0$. When the candidates are purely office-motivated, there exists a unique symmetric equilibrium $(-a^*, a^*)$, $a^* = 0$.*

In the Supplementary Material, we show that, when candidates are not purely office-motivated, there exists a unique symmetric equilibrium if μ and γ are normally distributed[5].

Taken together, Propositions 1 and 2 imply that when candidate-specific uncertainty is sufficiently large, purely office-motivated candidates is first-best for voter welfare and that candidates with any level of policy motivation prevents the realization of the first-best outcome.

6 Conclusion

In this paper we show how the relative strength between preference uncertainty and candidate-specific uncertainty determine the optimality of policy convergence (divergence). Our main result is more than a mere generalization of Bernhardt et al. (2009), as it points to how the motivation of political candidates interacts with the source of uncertainty to determine the welfare properties of the canonical political equilibrium. While in Bernhardt

[5]This is a sufficient, rather than a necessary, condition for the existence of a symmetric equilibrium.

et al. (2009) policy divergence is optimal and so some degree of policy motivation is necessary to achieve the first-best political equilibrium, in our model any positive degree of policy motivation is undesirable if candidate-specific uncertainty is strong enough. This is because candidate-specific uncertainty erodes the expected benefit of platform diversification by reducing the probability that the candidate with the policy closest to the median voter wins the election. Since policy motivation, however small, induces some degree of divergence, it generates suboptimal political equilibria.

Appendix

Proof of Proposition 1

The proof focuses on the utility maximization problem for an arbitrary voter v, i.e.,

$$\max_{a\geq 0} E\left[U_v\left(a\right)\right]. \tag{11}$$

We first show that if $H\left(0\right) = 1/2$ (which is true as H is symmetric around zero), then for any δ_v, the first-order condition of this maximization problem is satisfied at $a = 0$, i.e.,

$$\left.\frac{dE\left[U_v\left(a\right)\right]}{da}\right|_{a=0} = 0.$$

Fix $\mu \in \mathbb{R}$. Then voter v's expected utility is given by

$$\begin{aligned}
&\int_{-\infty}^{\widetilde{w}} \left[u\left(\delta_v + \mu - a\right) - \sigma_\gamma \frac{\gamma}{2}\right] h\left(\gamma\right) d\gamma + \int_{\widetilde{w}}^{\infty} \left[u\left(\delta_v + \mu + a\right) + \sigma_\gamma \frac{\gamma}{2}\right] h\left(\gamma\right) d\gamma \\
&= u\left(\delta_v + \mu - a\right) H\left(\widetilde{w}\right) + u\left(\delta_v + \mu + a\right)\left[1 - H\left(\widetilde{w}\right)\right] + \sigma_\gamma \left[\int_{\widetilde{w}}^{\infty} \frac{\gamma}{2} h(\gamma) d\gamma - \int_{-\infty}^{\widetilde{w}} \frac{\gamma}{2} h(\gamma) d\gamma\right] \\
&= \left[u\left(\delta_v + \mu - a\right) - u\left(\delta_v + \mu + a\right)\right] H\left(\widetilde{w}\right) + u\left(\delta_v + \mu + a\right) + \sigma_\gamma \left[\int_{\widetilde{w}}^{\infty} \frac{\gamma}{2} h(\gamma) d\gamma - \int_{-\infty}^{\widetilde{w}} \frac{\gamma}{2} h(\gamma) d\gamma\right].
\end{aligned} \tag{12}$$

Define the following auxiliary functions:

$$\Phi\left(a; \mu, \delta_v\right) \equiv \left[u\left(\delta_v + \mu - a\right) - u\left(\delta_v + \mu + a\right)\right] H\left(\widetilde{w}\right),$$

$$\Psi\left(a; \mu, \delta_v\right) \equiv u\left(\delta_v + \mu + a\right) + \sigma_\gamma \left[\int_{\widetilde{w}}^{\infty} \frac{\gamma}{2} h(\gamma) d\gamma - \int_{-\infty}^{\widetilde{w}} \frac{\gamma}{2} h(\gamma) d\gamma\right].$$

Then the expected utility is (12) is the sum of $\Phi\left(a; \mu, \delta_v\right)$ and $\Psi\left(a; \mu, \delta_v\right)$, for any given $\mu \in \mathbb{R}$.

Given the quadratic utility function in (5),

$$\widetilde{w} = \frac{1}{\sigma_\gamma}\left[\left(\mu + a\right)^2 - \left(\mu - a\right)^2\right] = \frac{4\mu}{\sigma_\gamma} a,$$

and the auxiliary functions $\Phi\left(a; \mu, \delta_v\right)$ and $\Psi\left(a; \mu, \delta_v\right)$ can be simplified to become

$$\Phi\left(a; \mu, \delta_v\right) = 4\left(\delta_v + \mu\right) a H\left(\widetilde{w}\right),$$

$$\Psi\left(a; \mu, \delta_v\right) = -\left(\delta_v + \mu + a\right)^2 + \sigma_\gamma \left[\int_{\widetilde{w}}^{\infty} \frac{\gamma}{2} h(\gamma) d\gamma - \int_{-\infty}^{\widetilde{w}} \frac{\gamma}{2} h(\gamma) d\gamma\right].$$

Straightforward differentiation yields

$$\frac{d}{da}\Phi\left(a;\mu,\delta_v\right)=4\left(\delta_v+\mu\right)\left[H\left(\widetilde{w}\right)+\widetilde{w}h\left(\widetilde{w}\right)\right].$$

$$\frac{d}{da}\Psi\left(a;\mu,\delta_v\right)=-2\left(\delta_v+\mu+a\right)-4\mu\widetilde{w}h\left(\widetilde{w}\right).$$

It follows that

$$\begin{aligned}\frac{dE\left[U_v\left(a\right)\right]}{da} &= \int_{-\infty}^{\infty}\left[\frac{d}{da}\Phi\left(a;\mu,\delta_v\right)+\frac{d}{da}\Psi\left(a;\mu,\delta_v\right)\right]f\left(\mu\right)d\mu \\ &= 4\int_{-\infty}^{\infty}\left[\left(\delta_v+\mu\right)H\left(\widetilde{w}\right)+\delta_v\widetilde{w}h\left(\widetilde{w}\right)\right]f\left(\mu\right)d\mu-2\left(a+\delta_v\right) \\ &= 4\left\{\int_{-\infty}^{\infty}\mu H\left(\widetilde{w}\right)f\left(\mu\right)d\mu+\delta_v\int_{-\infty}^{\infty}H\left(\widetilde{w}\right)f\left(\mu\right)d\mu+\delta_v\int_{-\infty}^{\infty}\widetilde{w}h\left(\widetilde{w}\right)f\left(\mu\right)d\mu\right\} \\ & -2\left(a+\delta_v\right). \end{aligned} \tag{13}$$

The second line uses the assumption that $E\left(\mu\right)=0$. We will evaluate each of the three integrals inside the curly brackets. The first one can be expressed as[6]

$$\int_{-\infty}^{\infty}\mu H\left[\widetilde{w}\left(\mu\right)\right]f\left(\mu\right)d\mu=\int_{0}^{\infty}\mu H\left[\widetilde{w}\left(\mu\right)\right]f\left(\mu\right)d\mu+\int_{-\infty}^{0}\mu H\left[\widetilde{w}\left(\mu\right)\right]f\left(\mu\right)d\mu.$$

Note that in general $\widetilde{w}\left(\mu\right)$ is an odd function, i.e., $\widetilde{w}\left(\mu\right)=-\widetilde{w}\left(-\mu\right)$. Define $z=-\mu$, then

$$\begin{aligned}\int_{-\infty}^{0}\mu H\left[\widetilde{w}\left(\mu\right)\right]f\left(\mu\right)d\mu &= \int_{-\infty}^{0}\left(-z\right)H\left[\widetilde{w}\left(-z\right)\right]f\left(-z\right)d\left(-z\right) \\ &= \int_{-\infty}^{0}zH\left[-\widetilde{w}\left(z\right)\right]f\left(z\right)dz \\ &= -\int_{0}^{\infty}\mu H\left[-\widetilde{w}\left(\mu\right)\right]f\left(\mu\right)d\mu.\end{aligned}$$

The second line uses the assumption that $f\left(\cdot\right)$ is an even function. Hence,

$$\begin{aligned}\int_{-\infty}^{\infty}\mu H\left[\widetilde{w}\left(\mu\right)\right]f\left(\mu\right)d\mu &= \int_{0}^{\infty}\mu H\left[\widetilde{w}\left(\mu\right)\right]f\left(\mu\right)d\mu-\int_{0}^{\infty}\mu H\left[-\widetilde{w}\left(\mu\right)\right]f\left(\mu\right)d\mu \\ &= \int_{0}^{\infty}\mu\left\{H\left[\widetilde{w}\left(\mu\right)\right]-H\left[-\widetilde{w}\left(\mu\right)\right]\right\}f\left(\mu\right)d\mu \\ &= \int_{0}^{\infty}\mu\left\{2H\left[\widetilde{w}\left(\mu\right)\right]-1\right\}f\left(\mu\right)d\mu \\ &= 2\int_{0}^{\infty}\mu H\left[\widetilde{w}\left(\mu\right)\right]f\left(\mu\right)d\mu-\int_{0}^{\infty}\mu f\left(\mu\right)d\mu, \end{aligned} \tag{14}$$

The third lines uses the assumption that $H\left(\cdot\right)$ is a symmetric distribution around zero, so that $H\left(\gamma\right)+H\left(-\gamma\right)=1$.

[6]Here we use the notation $\widetilde{w}\left(\mu\right)$ to showcase the dependence of $\widetilde{w}$ on μ.

Next consider the second integral inside the brackets in (13), which is

$$\int_{-\infty}^{\infty} H\left[\widetilde{w}\left(\mu\right)\right] f\left(\mu\right) d\mu = \int_{0}^{\infty} H\left[\widetilde{w}\left(\mu\right)\right] f\left(\mu\right) d\mu + \int_{-\infty}^{0} H\left[\widetilde{w}\left(\mu\right)\right] f\left(\mu\right) d\mu.$$

Define $z = -\mu$, then

$$\int_{-\infty}^{0} H\left[\widetilde{w}\left(\mu\right)\right] f\left(\mu\right) d\mu = \int_{-\infty}^{0} H\left[\widetilde{w}\left(-z\right)\right] f\left(-z\right) d\left(-z\right) = \int_{0}^{\infty} H\left[-\widetilde{w}\left(z\right)\right] f\left(z\right) dz.$$

Hence,

$$\int_{-\infty}^{\infty} H\left[\widetilde{w}\left(\mu\right)\right] f\left(\mu\right) d\mu = \int_{0}^{\infty} \left\{H\left[\widetilde{w}\left(\mu\right)\right] + H\left[-\widetilde{w}\left(\mu\right)\right]\right\} f\left(\mu\right) d\mu = \frac{1}{2}. \qquad (15)$$

Finally consider the integral $\int_{-\infty}^{\infty} \widetilde{w}\left(\mu\right) h\left[\widetilde{w}\left(\mu\right)\right] f\left(\mu\right) d\mu$. Note that $\xi\left(\mu\right) \equiv \widetilde{w}\left(\mu\right) h\left[\widetilde{w}\left(\mu\right)\right]$ is an odd function, i.e.,

$$\xi\left(-\mu\right) = \widetilde{w}\left(-\mu\right) h\left[\widetilde{w}\left(-\mu\right)\right] = -\widetilde{w}\left(\mu\right) h\left[-\widetilde{w}\left(\mu\right)\right] = -\xi\left(\mu\right).$$

For any odd function $\xi\left(\mu\right)$,

$$\begin{aligned}
\int_{-\infty}^{\infty} \xi\left(\mu\right) f\left(\mu\right) d\mu &= \int_{0}^{\infty} \xi\left(\mu\right) f\left(\mu\right) d\mu + \int_{-\infty}^{0} \xi\left(\mu\right) f\left(\mu\right) d\mu \\
&= \int_{0}^{\infty} \xi\left(\mu\right) f\left(\mu\right) d\mu + \int_{-\infty}^{0} \xi\left(-z\right) f\left(-z\right) d\left(-z\right) \\
&= \int_{0}^{\infty} \xi\left(\mu\right) f\left(\mu\right) d\mu - \int_{0}^{\infty} \xi\left(z\right) f\left(z\right) dz = 0.
\end{aligned}$$

Hence,

$$\int_{-\infty}^{\infty} \widetilde{w}\left(\mu\right) h\left[\widetilde{w}\left(\mu\right)\right] f\left(\mu\right) d\mu = 0. \qquad (16)$$

Substituting (14)-(16) into (13) gives

$$\begin{aligned}
\frac{dE\left[U_v\left(a\right)\right]}{da} &= 4\left\{2\int_{0}^{\infty} \mu H\left[\widetilde{w}\left(\mu\right)\right] f\left(\mu\right) d\mu - \int_{0}^{\infty} \mu f\left(\mu\right) d\mu + \frac{\delta_v}{2}\right\} - 2\left(a + \delta_v\right) \\
&= 8\int_{0}^{\infty} \mu H\left[\widetilde{w}\left(\mu\right)\right] f\left(\mu\right) d\mu - 4\int_{0}^{\infty} \mu f\left(\mu\right) d\mu - 2a.
\end{aligned}$$

This shows that the first-order condition for any turning points (either maximum or minimum) is independent of δ_v. This means all voters have the same preference ordering over policy pairs $\left(-a, a\right)$.

When evaluated at $a = 0$, the above expression can be reduced to

$$\left.\frac{dE\left[U_v\left(a\right)\right]}{da}\right|_{a=0} = 8\int_{0}^{\infty} \mu H\left(0\right) f\left(\mu\right) d\mu - 4\int_{0}^{\infty} \mu f\left(\mu\right) d\mu, \qquad (17)$$

which is zero when $H\left(0\right) = 1/2$. This shows that $a = 0$ always satisfies the first-order condition.

Next, we consider the second-order and third-order derivative of $E[U_v(a)]$ with respect to a, which are

$$\begin{aligned}\frac{d^2E[U_v(a)]}{da^2} &= 8\int_0^\infty \mu h[\widetilde{w}(\mu)]\left(\frac{4\mu}{\sigma_\gamma}\right) f(\mu)\,d\mu - 2 \\ &= \frac{32}{\sigma_\gamma}\int_0^\infty \mu^2 h\left(\frac{4\mu}{\sigma_\gamma}a\right) f(\mu)\,d\mu - 2.\end{aligned}$$

$$\frac{d^3E[U_v(a)]}{da^3} = \frac{128}{\sigma_\gamma^2}\int_0^\infty \mu^3 h'\left(\frac{4\mu}{\sigma_\gamma}a\right) f(\mu)\,d\mu < 0. \tag{18}$$

The last inequality uses the assumption that $h(\cdot)$ is symmetric and single-peaked at zero, hence $h'(\gamma) < 0$ for $\gamma > 0$. Equation (18) implies that $dE[U_v(a)]/da$ is itself a strictly concave function in a.

Suppose

$$\left.\frac{d^2E[U_v(a)]}{da^2}\right|_{a=0} = 2\left[\frac{16}{\sigma_\gamma}h(0)\int_0^\infty \mu^2 f(\mu)\,d\mu - 1\right] \leq 0,$$

which is equivalent to $\sigma_\gamma \geq \sigma_{\min} \equiv 16h(0)\int_0^\infty \mu^2 f(\mu)\,d\mu$. Then (18) implies that

$$\frac{d^2E[U_v(a)]}{da^2} < 0, \qquad \text{for all } a > 0.$$

This, together with (17), implies

$$\frac{dE[U_v(a)]}{da} < 0, \qquad \text{for all } a > 0.$$

These show that $E[U_v(a)]$ and $dE[U_v(a)]/da$ are both decreasing concave functions in a for all $a > 0$. A graphical illustration of $dE[U_v(a)]/da$ is shown in Figure 2(a). Hence, $a = 0$ is the unique global maximizer for all types of voters.

Next, consider the case in which

$$\left.\frac{d^2E[U_v(a)]}{da^2}\right|_{a=0} = 2\left[\frac{16}{\sigma_\gamma}h(0)\int_0^\infty \mu^2 f(\mu)\,d\mu - 1\right] > 0,$$

or equivalently $\sigma_\gamma < \sigma_{\min}$. This condition implies that $dE[U_v(a)]/da$ is strictly increasing at $a = 0$. Since $H[\widetilde{w}(\mu)] \leq 1$ for all μ,

$$\begin{aligned}\frac{dE[U_v(a)]}{da} &= 4\left\{2\int_0^\infty \mu H[\widetilde{w}(\mu)] f(\mu)\,d\mu - \int_0^\infty \mu f(\mu)\,d\mu\right\} - 2a \\ &\leq 2\left[2\int_0^\infty \mu f(\mu)\,d\mu - a\right].\end{aligned}$$

$$\Rightarrow \frac{dE[U_v(a)]}{da} < 0 \qquad \text{for any } a > 2\int_0^\infty \mu f(\mu)\,d\mu > 0.$$

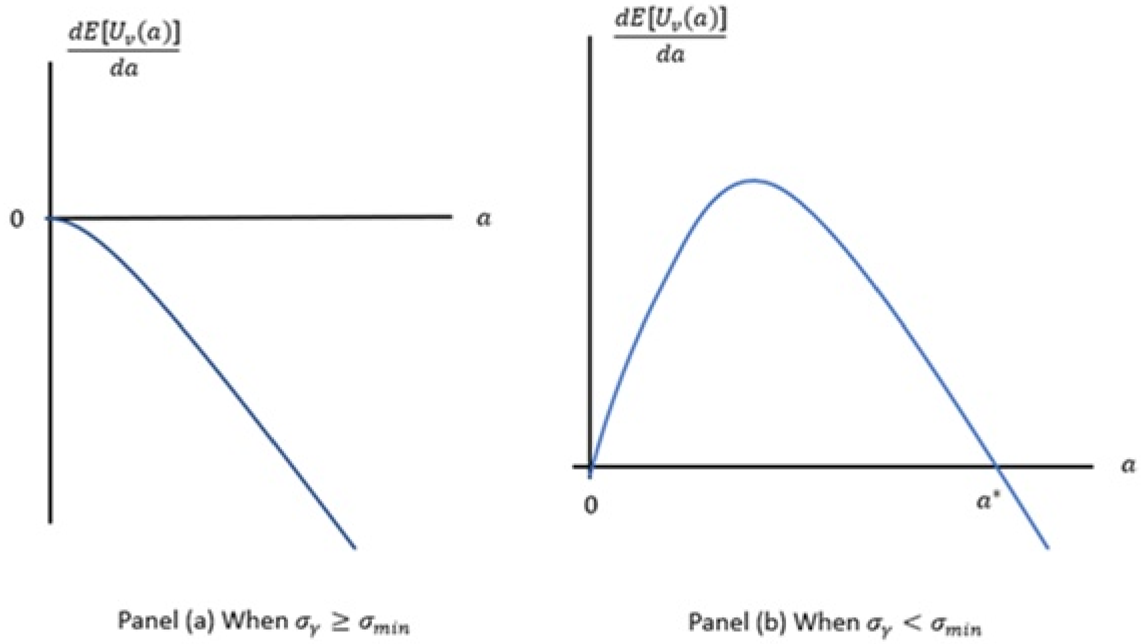

Figure 2: Illustration of Proposition 1

In sum, we have learned three things about $dE[U_v(a)]/da$ when $\sigma_\gamma < \sigma_{\min}$. First, by (18) it is a strictly concave function in $a \geq 0$. Second, it is equal to zero at $a = 0$ but is strictly increasing at this point. Third, it will eventually turn negative when $a > 2\int_0^\infty \mu f(\mu)\, d\mu$. By the intermediate value theorem, there exists a unique value $a^* > 0$, which is strictly less than $2\int_0^\infty \mu f(\mu)\, d\mu$, such that

$$\left.\frac{dE[U_v(a)]}{da}\right|_{a=a^*} = 0 \qquad \text{and} \qquad \left.\frac{d^2E[U_v(a)]}{da^2}\right|_{a=a^*} < 0.$$

Hence, $a^* > 0$ is the unique global maximizer for all types of voters. A graphically illustration of this result is shown in Figure 2(b).

Applying the implicit function theorem on the first-order condition, we obtain

$$\frac{da^*}{d\sigma_\gamma} = -\frac{\frac{d^2E[U_v(a)]}{dad\sigma_\gamma}}{\frac{d^2E[U_v(a)]}{da^2}}.$$

Since the denominator is negative when $\sigma_\gamma < \sigma_{\min}$, $\frac{da^*}{d\sigma_\gamma}$ has the same sign as

$$\frac{d^2E[U_v(a)]}{dad\sigma_\gamma} = 8\int_0^\infty \mu h[\widetilde{w}(\mu)] \frac{d\widetilde{w}(\mu)}{d\sigma_\gamma} f(\mu)\, d\mu,$$

which is negative since $d\widetilde{w}(\mu)/d\sigma_\gamma$ is negative for all $\mu > 0$. This completes the proof of Proposition 1.

Proof of Proposition 2

First consider the case when the candidates are both policy- and office-motivated. Taking $a_L = -a \leq 0$ as given, R's problem is to choose $a_R \geq 0$ so as to maximize his expected utility

$$\begin{aligned} U_R(a_R; -a) &= \Pr(\text{R wins}) \left[-(a_R - \psi)^2 + b \right] - [1 - \Pr(\text{R wins})] (-a - \psi)^2 \\ &= \Pr(\text{R wins}) \left[(-a - \psi)^2 - (a_R - \psi)^2 + b \right] - (-a - \psi)^2 , \end{aligned}$$

where

$$\Pr(\text{R wins}) = \int_{-\infty}^{\infty} H\left(\frac{1}{\sigma_\gamma} \left[(-a - \mu)^2 - (a_R - \mu)^2 \right] \right) f(\mu) d\mu.$$

The first-order condition of this problem is given by

$$\begin{aligned} \frac{dU_R(a_R; -a)}{da_R} &= \left[(-a - \psi)^2 - (a_R - \psi)^2 + b \right] \\ &\times \left(-\frac{2}{\sigma_\gamma} \right) \int_{-\infty}^{\infty} h\left(\frac{1}{\sigma_\gamma} \left[(-a - \mu)^2 - (a_R - \mu)^2 \right] \right) (a_R - \mu) f(\mu) d\mu \\ &- 2(a_R - \psi) \int_{-\infty}^{\infty} H\left(\frac{1}{\sigma_\gamma} \left[(-a - \mu)^2 - (a_R - \mu)^2 \right] \right) f(\mu) d\mu \\ &\leq 0, \end{aligned}$$

with equality holds if $a_R > 0$.

We now show that $a_L = a_R = 0$ cannot be an equilibrium. Setting $a = 0$ and $a_R = 0$ in the above derivative gives

$$\left. \frac{dU_R(a_R; 0)}{da_R} \right|_{a_R = 0} = \frac{2h(0)\, b}{\sigma_\gamma} \int_{-\infty}^{\infty} \mu f(\mu) d\mu + 2\psi H(0) \int_{-\infty}^{\infty} f(\mu) d\mu = \psi > 0.$$

The second equality uses the facts that the first term is 0 (i.e. $d\Pr(\text{R wins})/da = 0$ at $a_L = a_R = 0$) since $E(\mu) = 0$ and that $H(0) = 1/2$. This shows that choosing $a_R = 0$ is not R's best response to $a_L = 0$, hence, $a_L = a_R = 0$ cannot be an equilibrium. This in turn implies that any symmetric equilibrium, if exists, must involve policy divergence, i.e., $a_R = -a_L = a^* > 0$.

Assume now that the candidates are purely office-motivated. Their payoff is b if they win the election, and 0 otherwise. Given $a_L = -a$, R's maximizes

$$\widetilde{U}_R(a_R; -a) = b \int_{-\infty}^{\infty} H\left(\frac{1}{\sigma_\gamma} \left[(-a - \mu)^2 - (a_R - \mu)^2 \right] \right) f(\mu) d\mu.$$

Likewise, given $a_R = a$, candidate L maximizes

$$\widetilde{U}_L(a_L; a) = b\left[1 - \int_{-\infty}^{\infty} H\left\{\frac{1}{\sigma_\gamma}\left[(a_L - \mu)^2 - (a-\mu)^2\right]\right\} f(\mu)\, d\mu\right].$$

The two choice problems are symmetric. Hence it suffices to focus on the first-order condition of R's problem, which is

$$\frac{d\widetilde{U}_R(a_R; -a)}{da_R} = -\frac{2b}{\sigma_\gamma}\int_{-\infty}^{\infty} h\left(\frac{1}{\sigma_\gamma}\left[(-a-\mu)^2 - (a_R - \mu)^2\right]\right)(a_R - \mu) f(\mu) d\mu \leq 0,$$

with equality holds if $a_R > 0$. When evaluated at any symmetric policy platforms, i.e., when $a_R = -a_L = a \geq 0$, the above derivative can be simplified to become

$$\begin{aligned} \left.\frac{d\widetilde{U}_R(a_R; -a)}{da_R}\right|_{a_R = a} &= -\frac{2b}{\sigma_\gamma}\int_{-\infty}^{\infty} h\left(\frac{4\mu a}{\sigma_\gamma}\right)(a-\mu) f(\mu) d\mu \\ &= -\frac{2ba}{\sigma_\gamma}\int_{-\infty}^{\infty} h\left(\frac{4\mu a}{\sigma_\gamma}\right) f(\mu) d\mu < 0. \end{aligned} \tag{19}$$

The second line follows from the fact that $\xi(\mu) \equiv h\left(\frac{4\mu a}{\sigma_\gamma}\right)\mu f(\mu)$ is an odd function. Thus, using the same line of argument as in (16), we can get

$$\int_{-\infty}^{\infty} h\left(\frac{4\mu a}{\sigma_\gamma}\right)\mu f(\mu) d\mu = 0.$$

The condition in (19) shows that $a_R = a$ is not R's best response to $a_L = -a$, except when $a = 0$. This rules out any symmetric equilibrium with $a_R = -a_L = a^* > 0$.

To prove that $(a_L, a_R) = (0, 0)$ is indeed an equilibrium, we need to show that $a_R = 0$ is R's best response to $a_L = 0$, i.e.,

$$\widetilde{U}_R(0; 0) = \frac{b}{2} > \widetilde{U}_R(a_R; 0) = b\int_{-\infty}^{\infty} H\left(\frac{1}{\sigma_\gamma}\left[\mu^2 - (a_R - \mu)^2\right]\right) f(\mu) d\mu, \tag{20}$$

for all $a_R > 0$, and likewise, $a_L = 0$ is L's best response to $a_R = 0$, i.e.,

$$\widetilde{U}_L(0; 0) = \frac{b}{2} > \widetilde{U}_L(a_L; 0) = b\left[1 - \int_{-\infty}^{\infty} H\left\{\frac{1}{\sigma_\gamma}\left[(a_L - \mu)^2 - \mu^2\right]\right\} f(\mu)\, d\mu\right], \tag{21}$$

for all $a_L < 0$.

Fix $\mu \in \mathbb{R}$. By the strict convexity of the square function,

$$\mu^2 = \left[\frac{1}{2}(\mu - a_R) + \frac{1}{2}(\mu + a_R)\right]^2 < \frac{1}{2}(\mu - a_R)^2 + \frac{1}{2}(\mu + a_R)^2$$

$$\Rightarrow \mu^2 - (\mu - a_R)^2 < -\left[\mu^2 - (\mu + a_R)^2\right],$$

for any $a_R > 0$. Since $H(\cdot)$ is strictly increasing, we can get

$$H\left(\frac{1}{\sigma_\gamma}\left[\mu^2 - (a_R - \mu)^2\right]\right) < H\left(-\frac{1}{\sigma_\gamma}\left[\mu^2 - (a_R + \mu)^2\right]\right). \tag{22}$$

Rewrite R's expected utility $\widetilde{U}_R(a_R; 0)$ as

$$\begin{aligned}\widetilde{U}_R(a_R; 0) &= b\int_{-\infty}^{\infty} H\left(\frac{1}{\sigma_\gamma}\left[\mu^2 - (a_R - \mu)^2\right]\right) f(\mu)d\mu \\ &\quad + b\int_{-\infty}^{0} H\left(\frac{1}{\sigma_\gamma}\left[\mu^2 - (a_R - \mu)^2\right]\right) f(\mu)d\mu.\end{aligned} \tag{23}$$

Using the change of variable $\mu = -z$,

$$\int_{-\infty}^{0} H\left(\frac{1}{\sigma_\gamma}\left[\mu^2 - (a_R - \mu)^2\right]\right) f(\mu)d\mu = \int_{0}^{\infty} H\left(\frac{1}{\sigma_\gamma}\left[z^2 - (a_R + z)^2\right]\right) f(z)dz. \tag{24}$$

This uses the assumption that $f(\cdot)$ is an even function. Substituting (24) into (23) gives

$$\begin{aligned}\widetilde{U}_R(a_R; 0) &= b\int_{0}^{\infty} H\left(\frac{1}{\sigma_\gamma}\left[\mu^2 - (a_R - \mu)^2\right]\right) f(\mu)d\mu \\ &\quad + b\int_{0}^{\infty} H\left(\frac{1}{\sigma_\gamma}\left[z^2 - (a_R + z)^2\right]\right) f(z)dz \\ &\quad < b\int_{0}^{\infty} \left\{H\left(-\frac{1}{\sigma_\gamma}\left[\mu^2 - (a_R + \mu)^2\right]\right) + H\left(\frac{1}{\sigma_\gamma}\left[\mu^2 - (a_R + \mu)^2\right]\right)\right\} f(\mu)d\mu.\end{aligned} \tag{25}$$

The inequality follows from (22). Since $H(\cdot)$ is a symmetric distribution around zero, we have $H(\gamma) + H(-\gamma) = 1$, for all $\gamma \in \mathbb{R}$.Hence, for any $a_R > 0$, (25) implies

$$\widetilde{U}_R(a_R; 0) < b\int_0^\infty f(\mu)\, d\mu = \frac{b}{2},$$

which establishes (20). Using the same line of argument, we can show that

$$\int_{-\infty}^{\infty} H\left\{\frac{1}{\sigma_\gamma}\left[(a_L - \mu)^2 - \mu^2\right]\right\} f(\mu)\, d\mu > \frac{1}{2},$$

which then implies (21). Hence $a_R = a_L = 0$ is indeed a symmetric equilibrium. This completes the proof of Proposition 2.

Statements and Declarations:

The authors have no relevant financial or non-financial interests to disclose.

References

Adams, J. and S. Merrill III (2009). Policy-seeking parties in a parliamentary democracy with proportional representation: A valence-uncertainty model. *British Journal of Political Science 39*(3), 539–558.

Adams, J. F., S. Merrill III, and B. Grofman (2005). *A unified theory of party competition: A cross-national analysis integrating spatial and behavioral factors.* Cambridge University Press.

Ashworth, S. and E. B. De Mesquita (2009). Elections with platform and valence competition. *Games and Economic Behavior 67*(1), 191–216.

Banks, J. S. and J. Duggan (2005). Probabilistic voting in the spatial model of elections: The theory of office-motivated candidates. *Social Choice and Strategic Decisions: Essays in Honor of Jeffrey S. Banks*, 15–56.

Bernhardt, D., J. Duggan, and F. Squintani (2007). Electoral competition with privately-informed candidates. *Games and Economic Behavior 58*(1), 1–29.

Bernhardt, D., J. Duggan, and F. Squintani (2009). The case for responsible parties. *American Political Science Review 103*(4), 570–587.

Bernhardt, M. D. and D. E. Ingerman (1985). Candidate reputations and the 'incumbency effect'. *Journal of Public Economics 27*(1), 47–67.

Besley, T. and S. Coate (1997). An economic model of representative democracy. *The Quarterly Journal of Economics 112*(1), 85–114.

Bierbrauer, F. J. and P. C. Boyer (2013). Political competition and Mirrleesian income taxation: A first pass. *Journal of Public Economics 103*, 1–14.

Callander, S. (2005). Electoral competition in heterogeneous districts. *Journal of Political Economy 113*(5), 1116–1145.

Callander, S. (2008). Political motivations. *The Review of Economic Studies 75*(3), 671–697.

Callander, S. (2011). Searching for good policies. *American Political Science Review 105*(4), 643–662.

Calvert, R. L. (1985). Robustness of the multidimensional voting model: Candidate motivations, uncertainty, and convergence. *American Journal of Political Science 29*(1), 69–95.

Castanheira, M. (2003). Why vote for losers? *Journal of the European Economic Association 1*(5), 1207–1238.

Downs, A. (1957a). *An Economic Theory of Democracy.* Harper New York.

Downs, A. (1957b). An economic theory of political action in a democracy. *Journal of Political Economy 65*(2), 135–150.

Dwight, H. B. (1947). Tables of integrals and other mathematical data. *New York: The Macmillan Company.*

Groseclose, T. (2001). A model of candidate location when one candidate has a valence advantage. *American Journal of Political Science 45*(4), 862–886.

Gul, F. and W. Pesendorfer (2009). Partisan politics and election failure with ignorant voters. *Journal of Economic Theory 144*(1), 146–174.

Hinich, M. J., J. O. Ledyard, and P. C. Ordeshook (1972). Nonvoting and the existence of equilibrium under majority rule. *Journal of Economic Theory 4*(2), 144–153.

Krasa, S. and M. Polborn (2010). Competition between specialized candidates. *American Political Science Review 104*(4), 745–765.

Krasa, S. and M. K. Polborn (2012). Political competition between differentiated candidates. *Games and Economic Behavior 76*(1), 249–271.

Lindbeck, A. and J. W. Weibull (1993). A model of political equilibrium in a representative democracy. *Journal of Public Economics 51*(2), 195–209.

Londregan, J. and T. Romer (1993). Polarization, incumbency, and the personal vote. *Political economy: Institutions, competition, and representation*, 355–377.

Martinelli, C. (2001). Elections with privately informed parties and voters. *Public Choice 108*(1-2), 147–167.

Osborne, M. J. and A. Slivinski (1996). A model of political competition with citizen-candidates. *The Quarterly Journal of Economics 111*(1), 65–96.

Palfrey, T. R. (1984). Spatial equilibrium with entry. *The Review of Economic Studies 51*(1), 139–156.

Roemer, J. E. and J. E. Roemer (2009). *Political competition: Theory and applications.* Harvard University Press.

Schofield, N. (2007). The mean voter theorem: necessary and sufficient conditions for convergent equilibrium. *The Review of Economic Studies 74*(3), 965–980.

Van Weelden, R. (2013). Candidates, credibility, and re-election incentives. *Review of Economic Studies 80*(4), 1622–1651.

Wittman, D. (1983). Candidate motivation: A synthesis of alternative theories. *American Political science review 77*(1), 142–157.