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

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

April 8, 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 present the general model in Section 2 and our result on optimal policy divergence in Section 3. Then, we illustrate our result using quadratic utility function 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 The Model

There is a one-dimensional policy space, a set 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 is single-peaked at and symmetric around zero. 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. The term that captures voters' perception of the candidates is $\sigma_\gamma \gamma$, where $\sigma_\gamma > 0$ 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. Intuitively, this means L is viewed as a more favorable candidate than R. The opposite is true if γ is negative. The probability density function of γ is denoted by $h(\cdot)$, which is differentiable, symmetric at and single-peaked around zero. 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

$$u(\delta_v + \mu - a) = w\left[d(\delta_v + \mu - a)\right]. \tag{1}$$

The function $d : \mathbb{R} \to \mathbb{R}_+$ is a distance function that measures the gap between the voter's state-dependent ideal policy and the final policy. It is assumed to be a twice continuously differentiable even function, i.e., $d(z) = d(-z)$, for all $z \in \mathbb{R}$, strictly increasing and convex for $z > 0$, and satisfies $d(z) \geq 0$ for all $z \in \mathbb{R}$. These assumptions imply

$$d'(z) = -d'(-z) > 0 \qquad \text{and} \qquad d''(z) = d''(-z) \geq 0,$$

for all $z > 0$. It follows that $d(\cdot)$ is strictly decreasing and convex for all $z < 0$. The function $w : \mathbb{R}_+ \to \mathbb{R}$ is a loss function, which is twice continuously differentiable, strictly decreasing and concave.

Given these assumptions, the utility function $u(\cdot)$ can be shown to satisfy

$$u^{(n)}(z) = (-1)^n u^{(n)}(-z),$$

for $n \in \{0, 1, 2\}$ and for all $z \in \mathbb{R}$.[1] In particular, the first- and second-order derivatives of (1) with respect to a are given by

$$\frac{d}{da} u(\delta_v + \mu - a) = -w'[d(\delta_v + \mu - a)]\, d'(\delta_v + \mu - a) \lesseqgtr 0 \qquad \text{iff} \qquad a \gtreqless (\mu + \delta_v)$$

$$\frac{d^2}{da^2} u(\delta_v + \mu - a) = \underbrace{w''[d(\delta_v + \mu - a)]}_{(-)} [d'(\delta_v + \mu - a)]^2 + \underbrace{w'[d(\delta_v + \mu - a)]}_{(-)} \underbrace{d''(\delta_v + \mu - a)}_{(+)} < 0.$$

In words, these mean $u(\delta_v + \mu - a)$ is a hump-shaped function in a which peaks at $a = \mu + \delta_v$ and strictly concave on both sides. One example of such $u(\cdot)$ is the quadratic utility function, which is the composition of $w(x) = -x$ and $d(z) = z^2$.

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 $u(\delta_v + \mu - a)$, and if L wins, his utility is $u(\delta_v + \mu + a) + \sigma_\gamma \gamma$. The candidate that garners the majority of votes wins and his policy is implemented.

3 Optimal Policy Divergence

In this section, we ask the following normative question: Which pair of policy platforms $(-a, a)$ would maximize voters' ex ante welfare? Then in the next section, we will solve for 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

[1] The notation $u^{(n)}(\cdot)$ represents the nth order derivative of $u(\cdot)$, with $u^{(0)}(\cdot) \equiv u(\cdot)$.

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.[2] Our first main result is to qualify this statement using the model presented in Section 3. To be more precise, we show that when the degree of candidate-specific uncertainty is sufficiently large, then 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 values of μ and γ materialize, voter v will vote for R if and only if

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

$$\Leftrightarrow \gamma < \frac{1}{\sigma_\gamma}\left[u(\delta_v + \mu - a) - u(\delta_v + \mu + a)\right]. \tag{2}$$

The median voter is decisive in the election, therefore R wins if and only if (2) holds for $\delta_v = 0$, i.e.,

$$\gamma < \frac{1}{\sigma_\gamma}\left[u(\mu - a) - u(\mu + a)\right] \equiv \varpi. \tag{3}$$

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

$$E\left[U_v(a)\right] = \int_{-\infty}^{\infty}\left\{\int_{-\infty}^{\varpi} u(\delta_v + \mu - a)h(\gamma)d\gamma + \int_{\varpi}^{\infty}\left[u(\delta_v + \mu + a) + \sigma_\gamma\gamma\right]h(\gamma)d\gamma\right\} f(\mu)d\mu \tag{4}$$

If the candidate-specific uncertainty is absent, so that $\gamma = 0$ with certainty, then R wins if and only if $\mu > 0$.[3] In this case, the ex ante expected utility of voter v is given by

$$E\left[\widetilde{U}_v(a)\right] = \int_0^{\infty} u(\delta_v + \mu - a)f(\mu)d\mu + \int_{-\infty}^{0} u(\delta_v + \mu + a)f(\mu)d\mu. \tag{5}$$

The first integral in the above expression captures the expected utility when R wins, whereas the second one gives the expected utility when L wins. It is obvious that, for any

[2] 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.

[3] Specifically, R wins the election if and only if he is supported by the median voter, i.e., $\mu - a$ is closer to zero than $\mu + a$. For any $a > 0$, this is true if and only if $\mu > 0$.

$a > 0$, $\mu - a < \mu$ if $\mu > 0$ and $\mu + a > \mu$ if $\mu < 0$. Thus, unless too much divergent, a pair of symmetric and divergent policy platforms has the effect of condensing the arguments of the utility function, i.e. the distance functions, both when R wins and L wins, $\delta_v + \mu - a$ and $\delta_v + \mu + a$ respectively.[4] A graphical illustration of this is shown in Figures 1(a) and 1(b) which illustrate the probability densities of the distance between final and ideal policies. For $a = 0$, illustrated in Figure 1(a), this distribution coincides with the distribution of μ since the final policy is always 0. The case of a small value of a is illustrated in Figure 1(b). The black left-hand side and the green right-hand side correspond respectively to $\mu < 0$ (L wins and the argument is $\delta_v + \mu + a$) and to $\mu > 0$ (R wins and the argument is $\delta_v + \mu - a$). Such a change condenses the distribution and reduces the risk faced by any risk-averse voter. Hence, they will strictly prefer divergent 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.

However, when there is candidate-specific uncertainty, the results change. To illustrate this, a similar figure to Figure 1(b) can be drawn for each possible realization of $\sigma_\gamma \gamma$, and all these figures can be juxtaposed with the weight given by the distribution $h(\gamma)$. Here, we illustrate two figures that correspond to two specific values of $\sigma_\gamma \gamma$ in order to give the intuition. Figure 2(a) assumes small a and a small positive $\sigma_\gamma \gamma$. The difference between Figures 1(b) and 2(a) is that due to its candidate-specific advantage, L wins not only for $\mu < 0$, but also for small positive values of μ, i.e. L wins for $\mu < \tilde{\mu}$ with $\tilde{\mu} > 0$. Again, the black left-hand side and the green right-hand side correspond respectively to L's win and the argument being $\delta_v + \mu + a$) and to R's win and the argument being $\delta_v + \mu - a$). Due to the symmetric distribution of γ, it is equally likely to have the mirror image of this figure where R wins for a large set of values of μ. When these two cases (with same absolute value of $\sigma_\gamma \gamma$, but different signs) are juxtaposed, the resulting distribution will be more spread out than Figure 1(b). Figure 2(b) assumes a higher positive value of $\sigma_\gamma \gamma$, and L wins now for even more positive values μ, i.e. wins for $\mu < \hat{\mu}$ with $\hat{\mu} > \tilde{\mu}$. When Figure 2(b) is juxtaposed with its mirror image, the spread of the distribution will be stronger, compared to the previous case of smaller $\sigma_\gamma \gamma$. Intuitively, even for small a, the risk increases as high realizations of $\sigma_\gamma \gamma$ become predominant. Proposition 1 below shows that even a small value of a is undesirable for voters if candidate-specific uncertainty is large, that is, high values of $\sigma_\gamma \gamma$ in absolute value are likely.

[4] More formally, for any given $a > 0$, define a new random variable χ according to $\chi = \mu - a$ if $\mu \geq 0$ and $\chi = \mu + a$ if $\mu < 0$. If a is not "too large," then χ has the same mean as μ but is less dispersed.

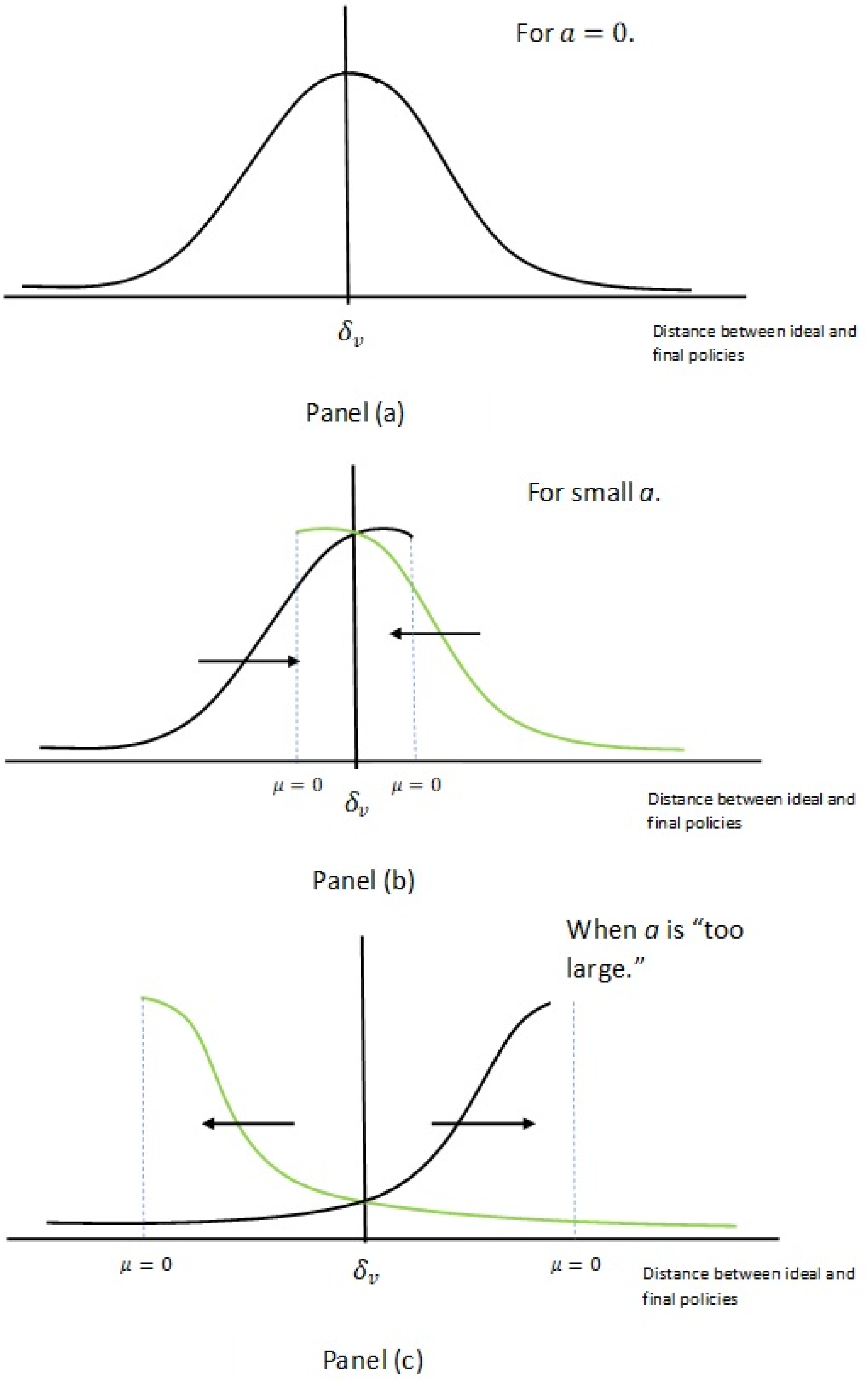

Figure 1: Illustration of the benefit of policy divergence in the absence of candidate-specific uncertainty

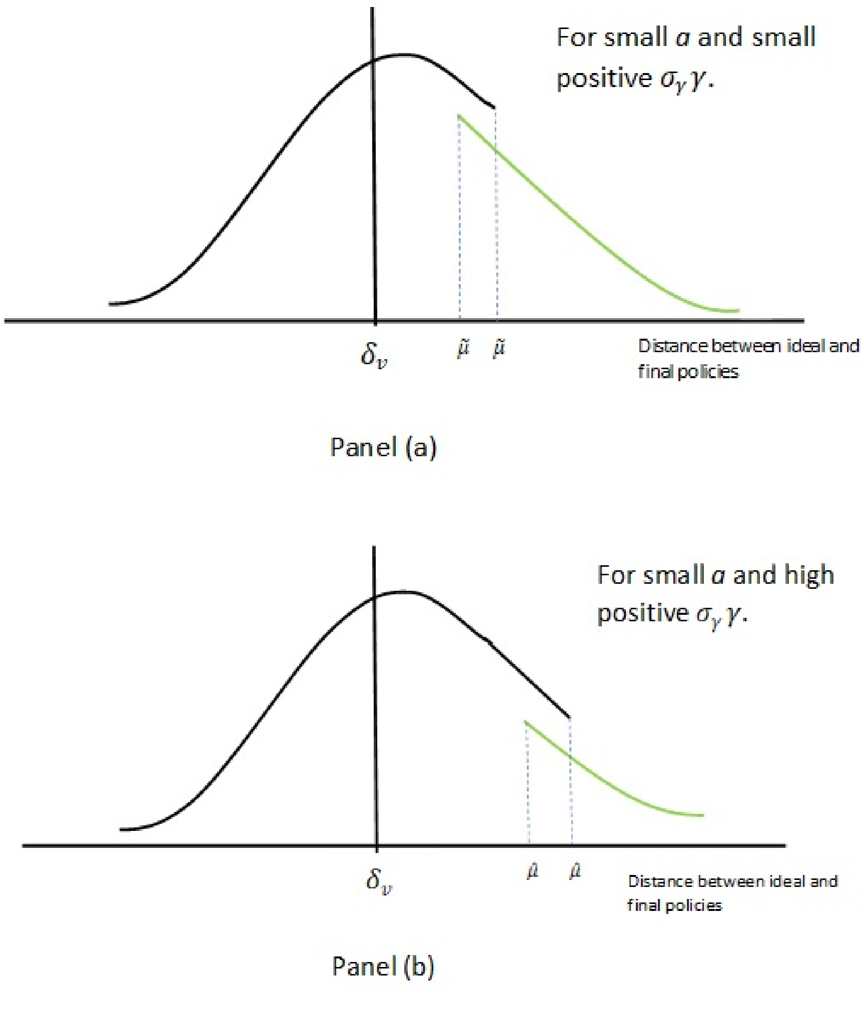

Figure 2: Illustration of the effect of candidate-specific uncertainty

Proposition 1. *For sufficiently large σ_γ, all voters strictly prefer the convergent policy platforms $(-a, a) = (0, 0)$ to any other symmetric policy platforms with $a > 0$.*

In the proof, we show that $(dEU_v(a, -a))/(da) < 0$ for any $a > 0$ and any voter δ_v for sufficiently large σ_γ. To gain further intuition, consider the marginal effect of policy divergence given by:

$$\frac{dEU_v(a, -a; \mu, \gamma)}{da} = \int_{-\infty}^{+\infty} \Big\{ u'(a - \mu - \delta_v) H(\widetilde{w}) + u'(-a - \mu - \delta_v)(1 - H(\widetilde{w}))$$
$$+ h(\widetilde{w}) \frac{d\widetilde{w}}{da} \left[u(a - \mu - \delta_v) - u(-a - \mu - \delta_v) - [u(a - \mu) - u(-a - \mu)] \right] \Big\} f(\mu) d\mu,$$

where the first two terms of the integrand measures the change in policy utility as a result of a larger a when L wins and when R wins, weighted the their winning probabilities. Another effect is that when a increases, the difference between candidates' policy platforms increases, and this leads to an increased likelihood that policy preferences determine the winner. The third term of the integrand captures this effect and is shown to be arbitrarily small when the candidate-specific uncertainty is sufficiently large. Notice that the third term is 0 for the median voter, i.e. for $\delta_v = 0$. This is because when the winner changes, it is exactly because the new winner gives now a marginally higher utility to the median voter than the old winner. We further show in the proof that this term is arbitrarily small for large σ_γ.

As σ_γ increases indefinitely, $\widetilde{w}$ will be close to zero even if $a > 0$. Because of this, the two candidates will have equal probability of winning even if they propose different policies. The intuition behind this can best be explained by considering a two-point distribution of γ. Suppose now there are only two possible values of γ, say γ_L and γ_H, which are equidistant on both sides of zero and happens with equal probability. An increase in the gap $|\gamma_H - \gamma_L|$ will have the same effect as an increase in σ_γ. When this happens, concerns about candidate-specific characteristics (such as competence or credibility) become increasingly important for voters' decisions, while policy preferences become less so. In the extreme case, the realization of γ alone will determine the winner of the election, regardless of the candidates' platforms $(-a, a)$ and the policy preference shock μ.[5] In terms of our notations, this is represented by $\widetilde{w} = 0$ and $H(0) = 1/2$. In this case, the outcome of the election is a pure lottery. Hence, any risk-averse voter would strictly prefer the convergent platforms (which guarantee a certain outcome) to any divergent ones.

[5]The independence between election outcome and μ is the main reason why the aforementioned hedging mechanism breaks down when σ_γ is sufficiently large.

4 Quadratic Utility

In this section, we illustrate the result in Proposition 1 by assuming a quadratic policy utility function, i.e., $u(x) = -x^2$. Using this type of preferences, we are able to derive 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. We are also able to show that the optimal level of policy divergence, when non-zero, decreases as candidate-specific uncertainty increases. These results are formally stated in Proposition 2.

Proposition 2. *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 first-order condition of maximizing $E[U_v(a)]$ with respect to a is 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 the 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[6]

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

This result states that, when the utility function is quadratic and the two shocks are 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 significantly higher than that associated with the policy preference shock. Otherwise, policy divergence is beneficial to all voters.

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

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

where $b \geq 0$ represents the office rents in case of victory, and the probability that R wins 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, and 0 otherwise.

[6] 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.

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. As opposed to Bernhardt et al., 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 policy-unrelated uncertainty leads to the result that the derivative of the winning probability is 0 at $a = 0$, in other words, an infinitely small move from $a = 0$ does not reduce the winning probability, and therefore parties do not choose $a = 0$ as long as they have some policy motivation, however small.

Proposition 3. *When parties are not purely office-motivated, in any symmetric equilibrium (when it exists)* $(-a^*, a^*)$*,* $a^* \neq 0$*. When parties are purely office-motivated, in any symmetric equilibrium* $(-a^*, a^*)$*,* $a^* = 0$*.*

We conclude from this proposition that when candidate-specific uncertainty is sufficiently large, purely office-motivated parties is first-best for voter welfare and that any level of policy motivation for parties prevents the realisation 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 generalisation of Bernhardt et al. (2009), as it points to how party motivation interacts with the source of uncertainty to determine the welfare properties of the canonical political equilibrium. While in Bernhardt 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 party 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,

$$\max_{a\geq 0} E\left[U_v\left(a\right)\right]. \tag{P1}$$

First, it is shown that if $H\left(0\right)=1/2$ (which is true if 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.,

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

Second, it is shown that if σ_γ is sufficiently large, then the second-order condition for a maximum is also satisfied at $a=0$, i.e.,

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

This establishes that $a=0$ is a solution of (P1) when σ_γ is sufficiently large. Finally, it is shown that there is no interior solution (i.e., $a>0$) of (P1) when σ_γ is sufficiently large. Hence, $a=0$ is the unique solution when σ_γ is sufficiently large.

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

$$\begin{aligned}
& \int_{-\infty}^{\widetilde{w}} u\left(\delta_v+\mu-a\right)h\left(\gamma\right)d\gamma+\int_{\widetilde{w}}^{\infty}\left[u\left(\delta_v+\mu+a\right)+\sigma_\gamma\gamma\right]h\left(\gamma\right)d\gamma \\
= \;& u\left(\delta_v+\mu-a\right)H\left(\widetilde{w}\right)+u\left(\delta_v+\mu+a\right)\left[1-H\left(\widetilde{w}\right)\right]+\sigma_\gamma\int_{\widetilde{w}}^{\infty}\gamma h\left(\gamma\right)d\gamma \\
= \;& \left[u\left(\delta_v+\mu-a\right)-u\left(\delta_v+\mu+a\right)\right]H\left(\widetilde{w}\right)+u\left(\delta_v+\mu+a\right)+\sigma_\gamma\int_{\widetilde{w}}^{\infty}\gamma h\left(\gamma\right)d\gamma. \qquad (6)
\end{aligned}$$

Define the following auxiliary functions:

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

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

Then the expected utility is (6) is the sum of $\Phi\left(a;\mu,\delta_v\right)$ and $\Psi\left(a;\mu,\delta_v\right)$. Differentiating these functions with respect to a gives

$$\begin{aligned}
\frac{d}{da}\Phi\left(a;\mu,\delta_v\right) \;=\;& \left[-u'\left(\delta_v+\mu-a\right)-u'\left(\delta_v+\mu+a\right)\right]H\left(\widetilde{w}\right) \\
& +\left[u\left(\delta_v+\mu-a\right)-u\left(\delta_v+\mu+a\right)\right]h\left(\widetilde{w}\right)\frac{d\widetilde{w}}{da}, \qquad (7)
\end{aligned}$$

$$\frac{d}{da}\Psi\left(a;\mu,\delta_v\right)=u'\left(\delta_v+\mu+a\right)-\sigma_\gamma\widetilde{w}h\left(\widetilde{w}\right)\frac{d\widetilde{w}}{da}, \tag{8}$$

where

$$\frac{d\widetilde{w}}{da}=\frac{1}{\sigma_\gamma}\left[-u'\left(\mu-a\right)-u'\left(\mu+a\right)\right].$$

The second-order derivatives with respect to a are given by

$$\begin{aligned}\frac{d^2}{da^2}\Phi\left(a;\mu,\delta_v\right) &= \left[u''\left(\delta_v+\mu-a\right)-u''\left(\delta_v+\mu+a\right)\right]H\left(\widetilde{w}\right)\\ &\quad+2\left[-u'\left(\delta_v+\mu-a\right)-u'\left(\delta_v+\mu+a\right)\right]h\left(\widetilde{w}\right)\frac{d\widetilde{w}}{da}\\ &\quad+\left[u\left(\delta_v+\mu-a\right)-u\left(\delta_v+\mu+a\right)\right]\left\{h'\left(\widetilde{w}\right)\frac{d\widetilde{w}}{da}+h\left(\widetilde{w}\right)\frac{d^2\widetilde{w}}{da^2}\right\},\end{aligned} \tag{9}$$

$$\frac{d^2}{da^2}\Psi\left(a;\mu,\delta_v\right)=u''\left(\delta_v+\mu+a\right)-\sigma_\gamma\left\{h\left(\widetilde{w}\right)\left[\frac{d\widetilde{w}}{da}\right]^2+\widetilde{w}h'\left(\widetilde{w}\right)\left[\frac{d\widetilde{w}}{da}\right]^2+\widetilde{w}h\left(\widetilde{w}\right)\frac{d^2\widetilde{w}}{da^2}\right\}, \tag{10}$$

where

$$\frac{d^2\widetilde{w}}{da^2}=\frac{1}{\sigma_\gamma}\left[u''\left(\mu-a\right)-u''\left(\mu+a\right)\right].$$

When evaluated at $a=0$,

$$\widetilde{w}=\frac{1}{\sigma_\gamma}\left[u\left(\mu\right)-u\left(\mu\right)\right]=0,$$

$$\left.\frac{d\widetilde{w}}{da}\right|_{a=0}=\frac{1}{\sigma_\gamma}\left[-u'\left(\mu\right)-u'\left(\mu\right)\right]=-\frac{2u'\left(\mu\right)}{\sigma_\gamma}.$$

The first-order derivatives in (7) and (8) then become

$$\left.\frac{d}{da}\Phi\left(a;\mu,\delta_v\right)\right|_{a=0}=-2u'\left(\delta_v+\mu\right)H\left(0\right).$$

$$\left.\frac{d}{da}\Psi\left(a;\mu,\delta_v\right)\right|_{a=0}=u'\left(\delta_v+\mu\right).$$

Hence,

$$\begin{aligned}\left.\frac{dE\left[U_v\left(a\right)\right]}{da}\right|_{a=0} &= \int_{-\infty}^{\infty}\left\{\left.\frac{d}{da}\Phi\left(a;\mu,\delta_v\right)\right|_{a=0}+\left.\frac{d}{da}\Psi\left(a;\mu,\delta_v\right)\right|_{a=0}\right\}f\left(\mu\right)d\mu\\ &= \left[1-2H\left(0\right)\right]\int_{-\infty}^{\infty}u'\left(\delta_v+\mu\right)f\left(\mu\right)d\mu.\end{aligned}$$

It follows that if $H\left(0\right)=1/2$, then

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

Next, we turn to the second-order condition at $a = 0$. First note that,

$$\left.\frac{d^2\widetilde{w}}{da^2}\right|_{a=0} = \frac{1}{\sigma_\gamma}\left[u''(\mu) - u''(\mu)\right] = 0.$$

The expressions in (9) and (10) can now be simplified to become

$$\left.\frac{d^2}{da^2}\Phi(a;\mu,\delta_v)\right|_{a=0} = 8u'(\delta_v + \mu)\, h(0)\, \frac{u'(\mu)}{\sigma_\gamma}.$$

$$\left.\frac{d^2}{da^2}\Psi(a;\mu,\delta_v)\right|_{a=0} = u''(\delta_v + \mu) - 4h(0)\, \frac{\left[u'(\mu)\right]^2}{\sigma_\gamma}.$$

Hence,

$$\left.\frac{d^2E\left[U_v(a)\right]}{da^2}\right|_{a=0} = \frac{4h(0)}{\sigma_\gamma}\int_{-\infty}^{\infty} u'(\mu)\left[2u'(\delta_v + \mu) - u'(\mu)\right] f(\mu)\, d\mu + \int_{-\infty}^{\infty} u''(\delta_v + \mu)\, f(\mu)\, d\mu.$$

The second integral is strictly negative due to the strict concavity of $u(\cdot)$. The sign of the first integral is ambiguous in general. However, this term can be made arbitrarily small by raising the value of σ_γ. Thus, as $\sigma_\gamma \to \infty$,

$$\left.\frac{d^2E\left[U_v(a)\right]}{da^2}\right|_{a=0} \to \int_{-\infty}^{\infty} u''(\delta_v + \mu)\, f(\mu)\, d\mu < 0.$$

This proves that $a = 0$ is a solution of (P1).

We now establish the uniqueness of $a = 0$. Recall that as $\sigma_\gamma \to \infty$, $\widetilde{w} \to 0$ and $d\widetilde{w}/da \to 0$ for any (a, μ). Using these, we can simplify the first-order derivatives in (7) and (8) to become

$$\frac{d}{da}\Phi(a;\mu,\delta_v) = \left[u'(\delta_v + \mu - a) - u'(\delta_v + \mu + a)\right] H(0),$$

$$\frac{d}{da}\Psi(a;\mu,\delta_v) = u'(\delta_v + \mu + a).$$

Hence, any interior solution $\widetilde{a} > 0$, if exists, must satisfy the first-order condition

$$\frac{dE\left[U_v(a)\right]}{da} = \Lambda(\widetilde{a};\delta_v) \equiv \int_{-\infty}^{\infty} \left\{\left[u'(\delta_v + \mu - \widetilde{a}) + u'(\delta_v + \mu + \widetilde{a})\right]\right\} f(\mu)\, d\mu = 0.$$

Note that (i) $\Lambda(0;\delta_v) = 0$, for all δ_v (the first result) and (ii)

$$\frac{d}{da}\Lambda(a;\delta_v) = \int_{-\infty}^{\infty} \left\{\left[u''(\delta_v + \mu - a) + u''(\delta_v + \mu + a)\right]\right\} f(\mu)\, d\mu < 0,$$

for all $a \geq 0$ and for all δ_v. These two observations together imply

$$\Lambda(0; \delta_v) = 0 > \Lambda(a; \delta_v),$$

for all $a > 0$, which rules out any interior solution when σ_γ is sufficiently large. This completes the proof of Proposition 1.

Proof of Proposition 2

Suppose $u(x) = -x^2$. Then we have

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

and the auxiliary functions $\Phi(a; \mu, \delta_v)$ and $\Psi(a; \mu, \delta_v)$ defined in the previous proof can be simplified to become

$$\Phi(a; \mu, \delta_v) = 4(\delta_v + \mu) aH(\widetilde{w}),$$
$$\Psi(a; \mu, \delta_v) = -(\delta_v + \mu + a)^2 + \sigma_\gamma \int_{\widetilde{w}}^{\infty} \gamma h(\gamma) d\gamma.$$

Their first-order derivatives are now given by

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

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

It follows that

$$\begin{aligned}
\frac{dE[U_v(a)]}{da} &= \int_{-\infty}^{\infty}\left[\frac{d}{da}\Phi(a; \mu, \delta_v) + \frac{d}{da}\Psi(a; \mu, \delta_v)\right] f(\mu) d\mu \\
&= 4\int_{-\infty}^{\infty}\left[(\delta_v + \mu) H(\widetilde{w}) + \delta_v \widetilde{w} h(\widetilde{w})\right] f(\mu) d\mu - 2(\delta_v + \mu) \\
&= 4\left\{\int_{-\infty}^{\infty} \mu H(\widetilde{w}) f(\mu) d\mu + \delta_v \int_{-\infty}^{\infty} H(\widetilde{w}) f(\mu) d\mu + \delta_v \int_{-\infty}^{\infty} \widetilde{w} h(\widetilde{w}) f(\mu) d\mu\right\} \\
&\quad -2(a + \delta_v) \qquad (11)
\end{aligned}$$

We will evaluate each of the three integrals inside the curly brackets. The first one can be expressed as[7]

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

[7]Here we use the notation $\varpi(\mu)$ to showcase the dependence of ϖ on μ.

Note that in general $\widetilde{w}(\mu)$ is an odd function, i.e.,

$$\widetilde{w}(\mu) = \frac{1}{\sigma_\gamma}\left[u(\mu - a) - u(\mu + a)\right] = -\widetilde{w}(-\mu).$$

This is true even without using the quadratic function. Define $z = -\mu$, then

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

The second line uses the assumption that $f(\cdot)$ is symmetric at zero. Hence,

$$\begin{aligned}
\int_{-\infty}^{\infty} \mu H[\widetilde{w}(\mu)] f(\mu) d\mu &= \int_{0}^{\infty} \mu H[\widetilde{w}(\mu)] f(\mu) d\mu - \int_{0}^{\infty} \mu H[-\widetilde{w}(\mu)] f(\mu) d\mu \\
&= \int_{0}^{\infty} \mu \left\{H[\widetilde{w}(\mu)] - H[-\widetilde{w}(\mu)]\right\} f(\mu) d\mu \\
&= \int_{0}^{\infty} \mu \left\{2H[\widetilde{w}(\mu)] - 1\right\} f(\mu) d\mu \\
&= 2\int_{0}^{\infty} \mu H[\widetilde{w}(\mu)] f(\mu) d\mu - \int_{0}^{\infty} \mu f(\mu) d\mu, \qquad (12)
\end{aligned}$$

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

Next consider the second integral inside the brackets in (11), which is

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

Define $z = -\mu$, then

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

Hence,

$$\int_{-\infty}^{\infty} H[\widetilde{w}(\mu)] f(\mu) d\mu = \int_{0}^{\infty} \left\{H[\widetilde{w}(\mu)] + H[-\widetilde{w}(\mu)]\right\} f(\mu) d\mu = \frac{1}{2}. \qquad (13)$$

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

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

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

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

Hence,

$$\int_{-\infty}^{\infty} \widetilde{w}(\mu) h[\widetilde{w}(\mu)] f(\mu) d\mu = 0. \tag{14}$$

Substituting (12)-(14) into (11) gives

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

This shows that the first-order condition for any turning points (either maximum or minimum) is independent of δ_v. In other words, all voters will make the same choices. In addition,

$$\left.\frac{dE[U_v(a)]}{da}\right|_{a=0} = 8\int_{0}^{\infty} \mu H(0) f(\mu) d\mu - 4\int_{0}^{\infty} \mu f(\mu) d\mu = 0. \tag{15}$$

This confirms 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

$$\begin{aligned}\frac{d^2E[U_v(a)]}{da^2} &= 8\int_{0}^{\infty} \mu h[\widetilde{w}(\mu)]\left(\frac{4\mu}{\sigma_\gamma}\right) f(\mu) d\mu - 2 \\ &= \frac{32}{\sigma_\gamma}\int_{0}^{\infty} \mu^2 h\left(\frac{4\mu}{\sigma_\gamma}a\right) f(\mu) d\mu - 2.\end{aligned}$$

$$\frac{d^3E[U_v(a)]}{da^3} = \frac{128}{\sigma_\gamma^2}\int_{0}^{\infty} \mu^3 h'\left(\frac{4\mu}{\sigma_\gamma}a\right) f(\mu) d\mu < 0. \tag{16}$$

The last inequality shows 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 (16) implies that

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

This, together with (15), implies

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

These show that $E\left[U_v\left(a\right)\right]$ and $dE\left[U_v\left(a\right)\right]/da$ are both decreasing concave functions in a for all $a > 0$. A graphical illustration of $dE\left[U_v\left(a\right)\right]/da$ is shown in Figure 3(a). Hence, $a = 0$ is the unique global maximiser for all types of voters.

Next, consider the case in which

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

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

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

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

In sum, we have learned three things about $dE\left[U_v\left(a\right)\right]/da$ when $\sigma_\gamma < \sigma_{\min}$. First, 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 $a > 2\int_0^\infty \mu f\left(\mu\right)d\mu$. By the intermediate value theorem, there exists a unique value $a^* > 0$, which is strictly less than $2\int_0^\infty \mu f\left(\mu\right)d\mu$, such that

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

Hence, $a^* > 0$ is the unique global maximiser for all types of voters. A graphically illustration of this result is shown in Figure 3(b).

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

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

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

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

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

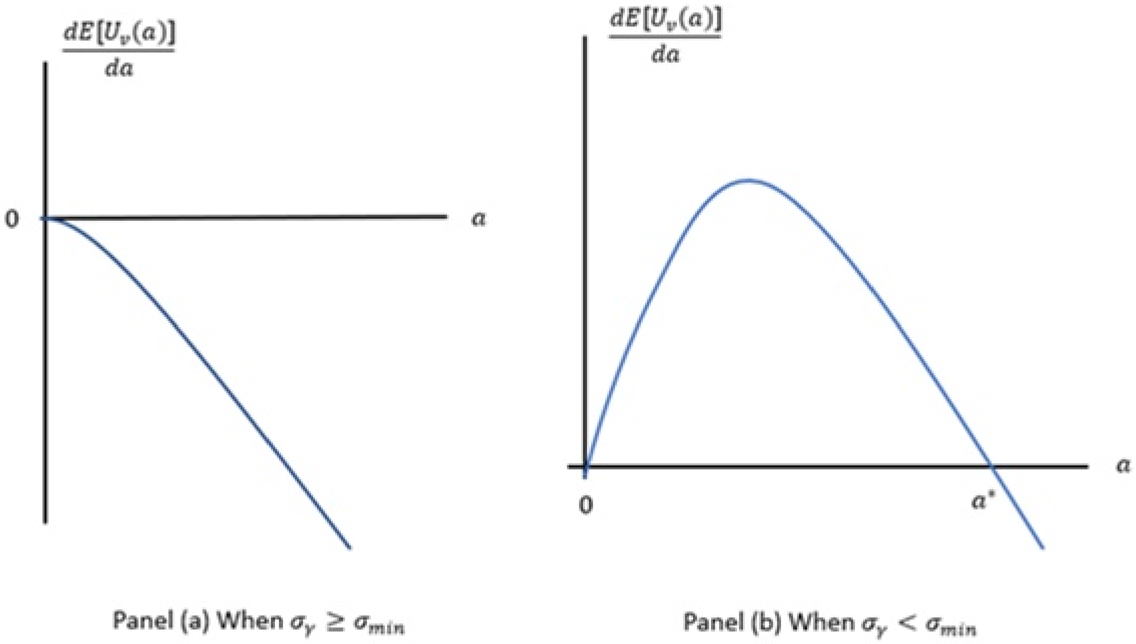

Figure 3: Illustration of Proposition 2

Proposition 2.

Proof of Proposition 3

Since U_R is continuous, in any symmetric equilibrium, the optimal policy of R must satisfy the following first-order condition given $a_L = -a$:

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

for $a_R > 0$, and $\frac{dU_R}{da_R} \leq 0$ for $a_R = 0$.

First, we rule out a symmetric equilibrium where $a_L = a_R = 0$. In this case,

$$\frac{dU_R}{da_R} = b \int_{-\infty}^{\infty} h(0) \frac{1}{\sigma_\gamma} u'(-\mu)\ f(\mu) d\mu + \frac{1}{2} u'(-\psi) > 0$$

since the first term is 0 and the second term is positive. Therefore, R would be better off choosing a policy to the right of 0.

Assume the best response of R, i.e. $a_R(a_L)$ is single-valued, i.e. is a function. Consider the best response function $a_R(-x_L) : [0, \psi] \to [0, \psi]$. This function is continuous due to continuity of U_R on the relevant range. Therefore, by Brouwer's fixed theorem, there exists a fixed point a^*. The pair $(a_L, a_R) = (-a^*, a^*)$ constitutes a symmetric equilibrium since $a_L(a^*) = -a^*$ by definition and $a_R(-a^*) = a^*$ by symmetry.

Assume now that parties are purely office-motivated. Their payoff is b if they win the election, and 0 otherwise. Parties are therefore maximizing their probability of winning. Given $a_L = -a$, party R maximizes

$$U_R(-a, a_R) = b \int_{-\infty}^{\infty} H\left(\frac{1}{\sigma_\gamma}(u(a_R - \mu) - u(-a - \mu))\right) f(\mu) d\mu.$$

The first-order condition is given by

$$\frac{dU_R}{da_R} = b \int_{-\infty}^{\infty} h\left(\frac{1}{\sigma_\gamma}(u(a_R - \mu) - u(-a - \mu))\right) \frac{1}{\sigma_\gamma} u'(a_R - \mu) f(\mu) d\mu = 0$$

for $a_R > 0$, and $\frac{dU_R}{da_R} \leq 0$ for $a_R = 0$.

For any $a_L = -a < 0$, $\frac{dU_R}{da_R} < 0$ for all $a_R \geq 0$ since the values of μ for which $u'(a_R - \mu)$ is negative are more heavily weighted. Therefore, the only potential equilibrium is $a_R = a_L = 0$. This is indeed an equilibrium since the first-order condition is satisfied:

$$b \int_{-\infty}^{\infty} h(0) \frac{1}{\sigma_\gamma} u'(-\mu) f(\mu) d\mu = 0,$$

and since $\frac{dU_R(0,a_R)}{da_R} < 0$ for all $a_R > 0$. This completes the proof.

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*, 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*, 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*, 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.