8-1-2013

The Impact of Political Uncertainty: A Robust Control Approach

Robert Baumann
College of the Holy Cross, rbaumann@holycross.edu

Justin Svec
College of the Holy Cross, jsvec@holycross.edu

The Impact of Political Uncertainty: A Robust Control Approach

By

Robert Baumann and Justin Svec

July 2013

COLLEGE OF THE HOLY CROSS, DEPARTMENT OF ECONOMICS
FACULTY RESEARCH SERIES, PAPER NO. 13-06*

Department of Economics
College of the Holy Cross
Box 45A
Worcester, Massachusetts 01610
(508) 793-3362 (phone)
(508) 793-3708 (fax)

http://www.holycross.edu/departments/economics/website

*All papers in the Holy Cross Working Paper Series should be considered draft versions subject to future revision. Comments and suggestions are welcome.

The Impact of Political Uncertainty: A Robust Control Approach

By
Robert Baumann[†]
College of the Holy Cross

and

Justin Svec[††]
College of the Holy Cross

July 2013

Abstract

In this paper, we examine how candidate uncertainty affects the policy platforms chosen in a uni-dimensional, two candidate Downsian spatial model. The candidates, we assume, do not know the true distribution of voters. Following the robust control literature, candidates respond to this uncertainty by applying a max-min operator to their optimization problem. This approach, consistent with findings within the behavioral economics literature, protects the candidate by ensuring that her expected utility never falls too far, regardless of the true voter distribution. We show that this framework produces policy convergence between the two candidates but there is a multiplicity of possible policy platforms upon which the candidates could settle, some of which could be quite distant from the median voter. These results are robust to the timing of the game and the level of uncertainty faced by the candidates. We argue that our model explains drift, which is our term for changing political beliefs over time. While drift may be caused by shifting attitudes or demographics, we show that drift could also be the result of candidate uncertainty.

.

JEL Classification Codes: H00, D78, D84

Keywords: Robust control, candidate uncertainty, voting, spatial model

---

[†]Department of Economics, Box 192A, College of the Holy Cross, Worcester, MA 01610-2395, 508-793-3879 (phone), 508-793-3708 (fax), rbaumann@holycross.edu

[††] Department of Economics, Box 87A, College of the Holy Cross, Worcester, MA 01610-2395, 508-793-3875 (phone), 508-793-3708 (fax), jsvec@holycross.edu

1 Introduction

A common approach to modeling the behavior of political candidates comes from the spatial model developed by Downs (1957) and Black (1958). In that model, voters are assumed to have single-peaked preferences over a uni-dimensional policy space, candidates can commit to their policy platforms, and both candidates and voters have perfect information. Given these assumptions, then in an election with two candidates, both candidates announce policy platforms that reflect the median voter's preferred policy. However, anecdotal evidence from elections seems inconsistent with this intuitive result. Consequently, researchers soon began modifying the theory in an attempt to better reflect reality.

An early criticism of the model was that the assumption of perfect information was unrealistic. Uncertainty, it was argued, could arise for two reasons. First, candidates could be unsure of the distribution of voters along the policy space or second, voters could be unsure about the policy positions of the candidates. The first type of uncertainty is easy to imagine. While polling attempts to gauge the electorate's preferences, it only uses a subset of the population and random sampling is difficult. Even when polling is reliable, voter turnout across elections can be highly variable as we detail later in this paper. These factors make it difficult for the political candidate to know the median voter's preferred policy and therefore she might choose a platform that does not maximize the likelihood of winning the election. The second type of uncertainty often reflects a distrust among voters that a candidate will be willing or able to keep her campaign promises once in office.

In studies that incorporate uncertainty, candidates or voters are typically assumed to know the probability model that characterizes this randomness.[1] For example, in Enelow and Hinich (1989), candidates face uncertainty about the preferences of the voters. That is, the candidates only know the preferences of the voters with error. Importantly, the authors assume that the candidates know the true distribution of this error term, which then allows a straight-forward calculation of the candidates' expected vote shares. While the intention of this approach is to model uncertainty, we believe these models require more information than a candidate is likely to have.

In our paper, we take a different approach. As we detail below, one merit of this alternative approach is that there is a long line of behavioral research that is consistent with the assumed behavior. Specifically, we

[1] The behavioral literature distinguishes between two types of uncertainty. The first type, called "risk," refers to a situation in which the agent does not know how the environment's randomness will be resolved, but does know the distribution over this randomness. Thus, risk refers to a situation corresponding to the 'known unknown.' The second type, called "uncertainty", refers to a situation in which the agent knows neither how the randomness will be resolved nor the distribution over the randomness. Thus, uncertainty refers to a situation corresponding to the 'unknown unknown.'
To our knowledge, all papers in the spatial voting literature that deal with "uncertainty" actually assume an environment of "risk". Our paper, we believe, is the first to assume an environment in which candidates face "uncertainty".

assume that the political candidates do not know the true voter distribution, similar to the assumption made in Enelow and Hinich (1989). Rather, the candidates are endowed with a set of alternative voter distributions that they believe could characterize the true distribution. Following the robust control literature and in stark contrast to other papers on candidate uncertainty, we assume that the candidates do not place a prior distribution over these alternative voter distributions. That is, the candidates do not assign particular weights reflecting their subjective belief that each alternative voter distribution is true. Assuming the candidates are uncertainty averse and have a preference for robustness, the candidates apply a max-min operator to their optimization problem. This operator induces each candidate to choose the policy platform that maximizes her subjective expected utility, where the expectation is taken with respect to the worst-case voter distribution within the set of alternative voter distributions. This behavior protects the candidate by ensuring that her subjective expected utility never falls too far, regardless of the true voter distribution.

With this setup, we show that the Nash equilibrium in the candidate game involves both candidates choosing the same policy platform, i.e. convergence is optimal. However, the novel aspect of this convergence is that there is a multiplicity of possible policy platforms upon which the candidates could settle. Intuitively, the candidates are no longer tethered to the median voter's preferred policy because the candidates are unsure as to the location of the median voter. Further, the set of possible equilibria weakly grows with a candidate's uncertainty level. Thus, model uncertainty combined with uncertainty aversion leads to a multiplicity of equilibria, all of which involve the candidates announcing the same policy platform, and some of which could be quite distant from the median voter's ideal point. We can show that these results are robust to both the timing of the game and the level of the candidates' uncertainty.

Because of the presence of multiple equilibria, our model can be used to explain changes in political party beliefs over time, a phenomenon that we term "drift". In our approach, drifting political positions need not come from the electorate changing its views or demographic composition. Rather, drifting political positions could be the result of candidates responding to model uncertainty.

Our paper proceeds as follows. In Section 2, we review the pertinent literature. Section 3 provides some justification for our modeling choices. The theoretical model and our key results are contained in Section 4, while Section 5 further explores our results using three simple numerical examples. We conclude in Section 6.

2 Literature review

As stated in the introduction, Downs (1957) and Black (1958) are the focal point for studies of candidate policy choice. Their application of Hotelling's (1929) spatial model assumes single-peaked preferences over a one-dimensional policy space, candidate commitment to policy choices, and full information of the distribution of voter-preferred policies. In an election with two candidates, both choose policy platforms that reflect the median voter's preferred policy.

Several theoretical extensions have been proposed to this baseline spatial voting model. As it is the subject of this paper, we begin with the papers that incorporate uncertainty. The literature focuses on two types of uncertainty: voter uncertainty of candidate policies and candidate uncertainty of voter preferences. Voter uncertainty can occur in several ways. The easiest to imagine is voter skepticism that a candidate will keep her campaign promises. Alternatively, the candidate may be unable to achieve her campaign promises because of resistance from another branch of government (for example, Banks (1990) is flexible to either manifestation). Another common way to model this type of uncertainty is to assume candidates have private ideological policy preferences (Wittman (1983)). In these models, candidates are maximizing a utility function that includes ideological policy preferences and winning the election. Candidate uncertainty is typically modeled by giving candidates either a probability distribution over an error term (Enelow and Hinich (1989)), a probability distribution of the median voter's location (Aragones and Palfrey (2002, 2005)), or a subjective probability of winning (Calvert (1985)).

A second common extension is to incorporate intangible characteristics of the candidates (termed "valence"; see Stokes (1963), Groseclose (2001), Aragones and Palfrey (2002), Kartik and McAfee (2007), Schofield (2007), among others). In these studies, it is possible for candidates to choose positions away from each other and/or the median voter in order to exploit a valence advantage or avoid an opponent who has one.

Another modification is to assume that the policy space is multi-dimensional. These studies (Calvert (1985), Enelow and Hinich (1989), Aragones and Palfrey (2002), Schofield (2007), among others) are notable because multi-dimensional policy space usually does not produce pure strategy Nash equilbria, although convergence is still possible. We also note that these extensions are not mutually exclusive in the literature. For example, Enelow and Hinich (1989) model both types of uncertainty and note their findings are robust to multi-dimensional preferences.

Although policy choices are difficult to measure numerically, several empirical tests of convergence exist

(Fiorina (1974), Poole and Rosenthal (1984), Adams and Merrill (1999), and a series of studies by Erikson and Wright (1980, 1989, 1993, and 1997)). Most of these studies find that candidate policy choices often diverge from their opponent and/or the likely position of the median voter.

In addition to the spatial voting literature, our paper is related to the burgeoning robust control literature. In this literature, mostly concentrated within macroeconomics and finance, researchers relax the assumption of rational expectations and examine whether the same theoretical conclusions hold under uncertainty and uncertainty aversion. One of the most vibrant applications of robust control has been to re-analyze optimal monetary policy; some examples are Dennis (2010), Dennis, Leitemo, and Soderstrom (2009), Hansen and Sargent (2008), Leitemo and Soderstrom (2008), and Levin and Williams (2003). Subsequent research, including Karantounias (2013) and Svec (2012, 2013), has applied the same framework to fiscal policy models. Robust control has also been used to explain the equity premium; see Hanson, Barillas, and Sargent (2009) for an example. Finally, there has been a recent push to explore the impact of model uncertainty and uncertainty aversion in other fields. For one example in the field of education, see Congdon-Hohman, Nathan, and Svec (2013).

3 Voter turnout and uncertainty aversion

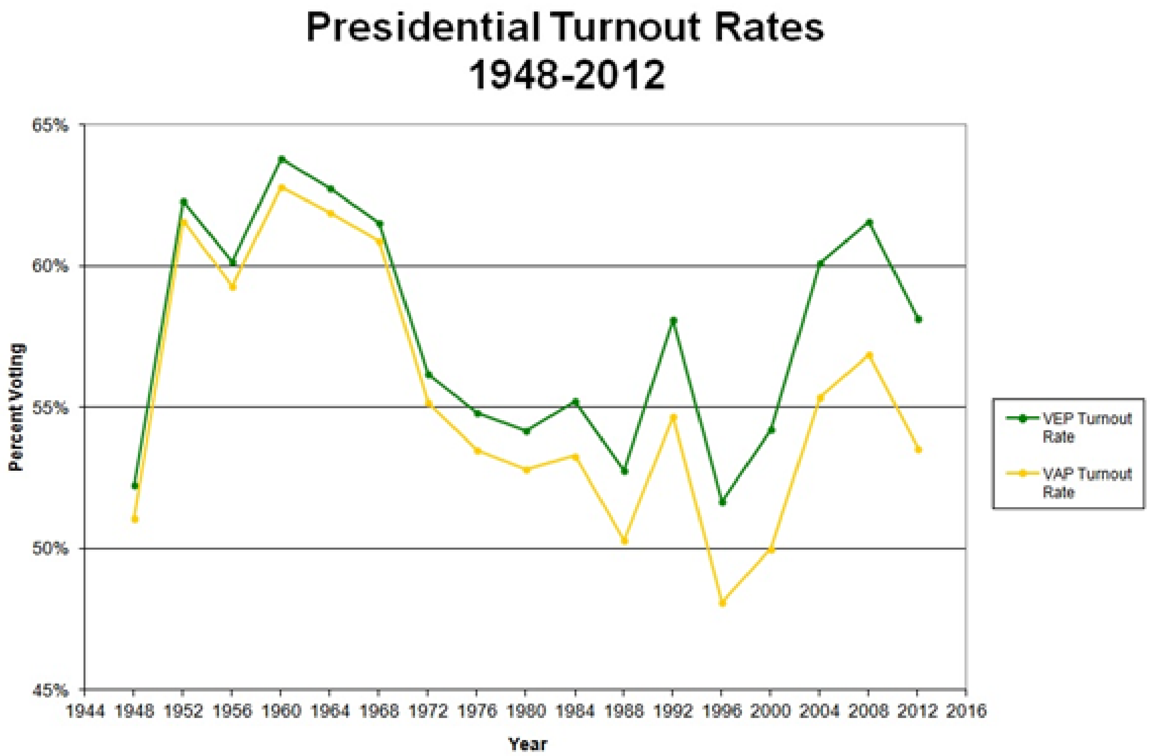
Before we describe the structure of our model, it would be helpful to pause in order to further discuss and support our underlying assumptions. The premise of our model is that political candidates face uncertainty about the distribution of voters' preferred policies and that they respond to this uncertainty as if they are uncertainty averse. As we argue below, we believe that both of these assumptions are reasonable.

The first assumption, that candidates are uncertain about the true distribution of voters' preferred policy points, seems plausible because candidates face two distinct forms of uncertainty: the candidates are likely to be unsure about both the electorate's preferences and the likelihood of specific types of voters to vote. Polling and campaign donations can be used to assuage the first form of uncertainty. Both of these actions are signals from potential voters about their preferences. However, variation in voter turnout increases the noise from these signals, and unfortunately for candidates voter turnout can vary considerably over time.

The United States Election Project[2] has tracked voter turnout in presidential elections since 1948. Figure 1 presents a graph of two measures of voter turnout: voting age population (VAP in the graph) and voting eligible population (VEP, which excludes those from the voting age population who are ineligible to vote).

[2] McDonald (2013)

Not only is the range of national turnout rates large, but there seems to be little predictive power of turnout rates from election to election.



However, presidential elections are determined by the electoral college, which shifts the focus of candidates to a handful of swing states. Thus, one could argue that the swings in voter turnout at the national level are less important than the volatility in voter turnout in the swing states. Nevertheless, even focusing on swing states, a significant amount of variation in voter turnout rates remains. Take Ohio, for example. Ohio's near-average mix of demographics along with the sixth-highest number of electoral votes has made this a hotly contested state in presidential elections. Over the five presidential elections from 1996 to 2012, the turnout rate in Ohio (relative to registered voters) varied from 63.65% (2000) to 71.77% (2008).[3] This translates to a large swing in the number of votes given an average of 7.7 million registered voters over this time period.

The second assumption, that agents respond to uncertainty as if they are uncertainty averse, is consistent

[3] The source of these data is the Ohio Secretary of State. These voter turnout rates are taken relative to registered voters, which some (such as the United States Election Project) warn against since they usually include registered voters who have moved. However, most of these warnings are for comparisons across states and we have no reason to believe that any fix would significantly diminish the variation of voter turnout.

with both a long line of behavioral research and the robust control literature. In one of the first papers documenting uncertainty aversion, Ellsberg (1961) runs a series of experiments in which subjects must choose between a gamble with unknown odds and a gamble with known odds. The author shows that most subjects reject gambles with unknown odds in favor of gambles with known odds. This result is particularly surprising because, if we assume that the subjects had any fixed prior over the unknown gamble as would be assumed in a Bayesian approach, the subjects would have been weakly better off by choosing the unknown-odds gamble. Ellsberg rationalizes this behavior by modifying standard expected utility theory in two ways: first, he assumes the agent believes that there is a set of probability models that could characterize the unknown-odds gamble and second, for any bet made, the agent worries that the worst case probability model will result, ensuring that she would lose the bet. These modifications have helped inform the theoretical formalization of behavior under model uncertainty.

Subsequent research further explored this finding of uncertainty aversion. In an early review of the literature, Camerer and Weber (1992) discuss papers that show that uncertainty aversion characterizes people's behavior in a number of different environments, including betting on natural events and trading in markets. Finally, there is considerable recent experimental evidence that supports the finding of uncertainty aversion; examples include Halevy (2007), Bossaerts, Ghirardato, Guarnaschelli, and Zame (2010), Ahn, Choi, Gale, and Kariv (2011), and Abdellaoui, Baillon, Placidoa, and Wakker (2011). Thus, given the extensive literature documenting people's uncertainty aversion, we believe that our second assumption is particularly relevant.

4 Theoretical model

In this section, we formulate our model of candidate competition under model uncertainty. This model is a modification of the spatial model of Downs (1957) and Black (1958).

Assume that voters have single-peaked preferences over policies and that the policy space is the unidimensional line from $[0, 1]$. Each voter cares only about the policies that will be enacted by the winning candidates.[4] Also, assume that there are two candidates for political office. The sole goal of the candidates is to get elected, and so they choose their policy platforms in order to maximize the number of votes they receive. The candidates are assumed to have access to some commitment technology, implying that the winner of the election enacts the policy that she announces as candidate. Finally, the politician that wins

[4] We abstract away from any difference in candidate valence, a characteristic that is at the heart of Groseclose (2001), Kartik and McAfee (2007), and Schofield (2007).

the majority of votes gets elected; if there's a tie, then a coin is flipped to determine which politician is elected.

If we assume that both candidates have full information as to the true distribution of voters' preferred policies on the spectrum $[0, 1]$, then the unique Nash equilibrium in the candidate game is for both candidates to announce the policy that conforms to the median voter's preferred policy. This is the result described in the Median Voter Theorem.

In this paper, though, our goal is to analyze how model uncertainty influences the equilibrium in the candidate game. To this end, we assume that both politicians are endowed with an approximating model that specifies the distribution of voters' preferred points along the uni-dimensional spectrum from $[0, 1]$. For simplicity, both politicians are endowed with the same approximating model, labeled $f(x)$, where x represents a particular policy and $f(x)$ represents the mass of voters with that policy as their ideal policy. The politicians, however, are not confident that this approximating model correctly characterizes the true probability distribution of the voters. They worry that other probability models could potentially characterize the distribution of voters' preferred policies. In order to ensure that these alternative models conform to some degree with the approximating model, we place restrictions on what types of alternative models are allowed. To do so, we follow the robust control literature, and in particular, Hansen and Sargent (2005, 2007).

We assume that each member of the set of alternative models is absolutely continuous with respect to the politician's approximating model. This implies that the politicians only fear models that correctly put no weight on zero probability outcomes. That is, the alternative models could indicate different proportions of voters at each $x \in [0, 1]$, as long as the mass of the voters under the approximating probability model is between zero and one. Further, the assumption of absolute continuity implies that the Radon-Nikodym theorem holds, which indicates that there exists a measurable function m such that the expectation of a random variable X under the alternative models can be rewritten in terms of the approximating probability model:

$$\widetilde{E}[X] = E[mX]$$

where $\widetilde{E}$ is the subjective expectations operator. To guarantee that each alternative model is a legitimate probability model, we assume $E[m] = 1, \forall i$.

Using the function m, we can now define the distance between the alternative and approximating probability models to be the entropy:

$$\epsilon(m) \equiv E[m \log m],$$

a measure that is convex and grounded. Following the robust control literature, we will use this distance measure to define the politicians' multiplier preferences. The multiplier preferences characterize how the politicians value each possible position along the policy spectrum.

Without loss of generality, let us focus on candidate r, taking as given the location of candidate l for the moment. For a given platform of the opponent l, candidate r has three possible positions: she can choose $r > l$, $r < l$, or $r = l$. Each of these possibilities leads to different objective functions for candidate r. We will go through each of these in turn, but initially we focus on the first possibility. Assuming that r chooses a policy platform to the right of l and that the voters have symmetric preferences, her goal is to choose the policy platform that maximizes the following objective function:

$$\min_{m(x)} \left\{ \int_{\frac{r+l}{2}}^{1} m(x) f(x) \, dx + \theta \int_{0}^{1} m(x) \log m(x) f(x) \, dx - \theta\lambda \left[\int_{0}^{1} m(x) f(x) \, dx - 1 \right] \right\}$$

The first term in this objective function describes the mass of voters that the candidate expects to vote for her. The candidate knows that all voters whose preferred policy point is at least $\frac{r+l}{2}$ will vote for her. But, due to the uncertainty, candidate r is not sure about how many voters are included in that set. To this end, she tilts her approximating model with the multiplicative term, $m(x)$, discussed earlier. In effect, candidate r believes that the mass of voters whose preferred policy is in between two points a and b is $\int_{a}^{b} m(x) f(x) \, dx$. If $m(x) > 1$ ($m(x) < 1$) $\forall x \in [a, b]$, then the candidate believes that there is a larger (smaller) mass of voters in that policy space than is true under the approximating model. The second term in this objective function is a penalty term that effectively constrains the size of the endogenous probability tilting.[5] Finally, the third term is the legitimacy constraint, which guarantees that the probabilities associated with each alternative voter distribution sum to unity.

A critical parameter in this objective function is $\theta > 0$. θ is a penalty parameter that indexes the degree to which candidate r is uncertain about the distribution of the voters. A small θ indicates that the candidate is not penalized too harshly for tilting her probability model away from the approximating model. The *min* operator then yields a set of $m(x)$ that diverge greatly from one. The resulting probabilities $\{m(x) f(x)\}$ are distant from the approximating model. Thus, a small θ indicates that the candidate is very unsure

[5] We will discuss this term in greater detail in the following paragraph.

about the approximating model and so fears a large set of alternative models. A large θ means that the candidate faces a sizable penalty for tilting her probability model away from the approximating model. As a result, the *min* operator yields a set of $m(x)$ close to unity, implying that the worst-case alternative model is close to the approximating model. Thus, a large θ signifies that the candidate has more confidence about the underlying measure and so fears only a small set of alternative models. As $\theta \to \infty$, this model collapses to the rational expectations framework discussed above.

To solve for the candidate's optimal platform, we must first solve the minimization problem within this objective function. In this step, candidate r fears that, for a given policy platform to the right of l, the worst-case distribution of voters within the set of possible alternative distributions happens to be correct. Note that for each policy platform considered, the candidate could consider a different worst-case distribution. The solution that results from this minimization is the politician's worst-case voter distribution for a given platform to the right of l.

The first order conditions from this objective function are

$$\begin{aligned} m(x), \text{ when } x &\in \left[\frac{r+l}{2}, 1\right] : 1 + \theta\left[1 + \log m(x)\right] - \theta\lambda = 0 \\ m(x), \text{ when } x &\in \left[0, \frac{r+l}{2}\right) : \theta\left[1 + \log m(x)\right] - \theta\lambda = 0 \\ \lambda &: \int_0^1 m(x) f(x)\, dx - 1 = 0 \end{aligned}$$

Combining the first two FOCs with the legitimacy constraint, we can derive the following values for $m(x)$:

$$\begin{aligned} m(x) &= \frac{\exp\left[-\frac{1}{\theta}\right]}{\int_0^{\frac{r+l}{2}} f(x)\, dx + \exp\left[-\frac{1}{\theta}\right] \int_{\frac{r+l}{2}}^1 f(x)\, dx} < 1, \text{ for all x} \in \left[\frac{r+l}{2}, 1\right] \\ m(x) &= \frac{1}{\int_0^{\frac{r+l}{2}} f(x)\, dx + \exp\left[-\frac{1}{\theta}\right] \int_{\frac{r+l}{2}}^1 f(x)\, dx} > 1, \text{ for all } x \in \left[0, \frac{r+l}{2}\right) \end{aligned}$$

These values represent the candidate's subjective probability tilitings of the voter distribution for each value of x. That is, candidate r fears that the true percentage of voters who have a preferred policy within $x \in [a, b]$ is $\int_a^b m(x) f(x)\, dx$.

We can use these probability tiltings to determine the candidate's subjective expected utility for each choice of $r > l$. To do this, plug the optimal values of $m(x)$ into the candidate's multiplier preferences. With some manipulation, it can be shown that the candidate's objective function simplifies to the following:

$$-\theta \log \left[\int_{0}^{\frac{r+l}{2}} f(x)\, dx + \exp\left[-\frac{1}{\theta}\right] \int_{\frac{r+l}{2}}^{1} f(x)\, dx \right]$$

This robust objective function allows us to determine the candidate's optimal policy platform for given values of l and θ while still assuming that $r > l$.

Candidate r, though, might not choose a policy platform to the right of l. Instead, if candidate r chooses a policy platform to the left of l, then the objective of the candidate is to maximize the following function:

$$\min_{m(x)} \left\{ \int_{0}^{\frac{r+l}{2}} m(x)\, f(x)\, dx + \theta \int_{0}^{1} m(x) \log m(x)\, f(x)\, dx - \theta\lambda \left[\int_{0}^{1} m(x)\, f(x)\, dx - 1 \right] \right\}$$

Following the same steps as outlined above, we would find that the robust objective function is

$$-\theta \log \left[\exp\left[-\frac{1}{\theta}\right] \int_{0}^{\frac{r+l}{2}} f(x)\, dx + \int_{\frac{r+l}{2}}^{1} f(x)\, dx \right]$$

for given values of l and θ and assuming that $r < l$.

Finally, the third category of possible policy platforms for candidate r is to set $r = l$. With this choice, regardless of the value of θ and the position of l, candidate r can guarantee a subjective expected utility equal to $\frac{1}{2}$. This is because the voters cannot distinguish between the two candidates and therefore each candidate expects to win half of the votes.

Collecting these results, we can write down the subjective expected utility for candidate r for any possible platform chosen, for a particular value of θ, and for a fixed platform of candidate l:

$$\text{Subjective expected utility} = \begin{cases} -\theta \log \left[\int_0^{\frac{r+l}{2}} f(x)\, dx + \exp\left[-\frac{1}{\theta}\right] \int_{\frac{r+l}{2}}^1 f(x)\, dx \right], \text{ if } r > l \\ -\theta \log \left[\exp\left[-\frac{1}{\theta}\right] \int_0^{\frac{r+l}{2}} f(x)\, dx + \int_{\frac{r+l}{2}}^1 f(x)\, dx \right], \text{ if } r < l \\ \frac{1}{2}, \text{ if } r = l \end{cases} \quad (1)$$

Notice the difference between this objective function under model uncertainty and the analogous one under rational expectations. (1) must be divided into three components because the candidate's worst-case voter distribution depends on the policy platform under consideration. That is, if the candidate considers choosing a platform to the right of her opponent, she fears that the true voter distribution will be weighted towards the left end of the spectrum; if the candidate considers choosing a platform to the left of her opponent, then she fears that the true voter distribution will be weighted towards the right end of the spectrum. These fears are consistent with the type of behavior exhibited in Ellsberg (1961) and the subsequent literature on uncertainty aversion. Only when $r = l$ does uncertainty aversion not imply that the candidate fears a harmful voter distribution. This is because, when $r = l$, the candidate knows that she and her political opponent are identical, meaning that she has guaranteed herself half the vote.

We can use (1) to derive the optimal policy platform of candidate r. These equations suggest that candidate r's optimal platform depends upon three factors: the approximating distribution of voters, the position of candidate l, and candidate r's level of uncertainty, θ. In the theorem below, we establish r's optimal choice of policy for given values of l and θ.

Theorem 1 *Define* $\underline{l}$ *through the equation* $\frac{1}{2} = -\theta \log \left[\int_0^{\underline{l}} f(x)\, dx + \exp\left[-\frac{1}{\theta}\right] \int_{\underline{l}}^1 f(x)\, dx \right]$ *and* $\bar{l}$ *through the equation* $\frac{1}{2} = -\theta \log \left[\exp\left[-\frac{1}{\theta}\right] \int_0^{\bar{l}} f(x)\, dx + \int_{\bar{l}}^1 f(x)\, dx \right]$. *Also, assume that* f *is continuous. Then, if* $l \in \left[\underline{l}, \bar{l}\right]$, $r = l$; *if* $l \notin \left[\underline{l}, \bar{l}\right]$, r *will choose a policy platform that is* $\epsilon > 0$ *closer to the median voter (under the approximating model) than* l.

To help with the flow of the paper, we prove this theorem in the appendix.

This theorem has two parts. First, the theorem introduces the bounds $\underline{l}$ and $\bar{l}$. All values of l within these two bounds are within what we will call candidate r's "range of uncertainty." This range is important because if candidate l chooses a policy platform within candidate r's range of uncertainty, then candidate r believes that the median voter's preferred policy could be on either side of l. However, if candidate l chooses a policy platform that is outside of the range of uncertainty, candidate r believes that the median voter is strictly on one side of candidate l. Thus, candidate r believes that all policies in the interior of her range of uncertainty might represent the preferred policy point of the median voter, while all policies outside of the range cannot be the median voter's preferred policy point.

The second part of the theorem indicates what policy platform candidate r should choose, conditional on the location of l. If l is within r's range of uncertainty, then the theorem indicates that candidate r should optimally choose $r = l$. This behavior is intuitive because candidate r believes that the median voter could be on either side of l. As a result, if candidate r would choose $r \neq l$, her feared, worst-case voter distribution would imply that the median voter was on the other side of candidate l, leading to an electoral loss for candidate r. To prevent this outcome, then, candidate r chooses the same policy platform as candidate l.

If, though, candidate l chooses a policy platform outside of candidate r's range of uncertainty, then policy divergence is optimal. This is because candidate r can reject the idea that the median voter could be on either side of candidate l. Consequently, candidate r finds it optimal to choose a platform that is slightly closer to the median voter under the approximating model. This choice leads candidate r to believe that she will receive the majority of votes in the election.

One additional point is worth noting about candidate r's range of uncertainty: the size of candidate r's range of uncertainty depends positively on the size of that candidate's uncertainty. That is, as θ falls and candidate r's uncertainty grows, then candidate r's range of uncertainty grows as well. This means that there is a wider range of policy platforms that candidate l could choose such that candidate r would choose $r = l$.

Up to this point, we have analyzed candidate r's optimal policy platform for a given choice of l. Now, we turn to candidate l's optimal platform for a given choice of candidate r. As both candidates are assumed to have the same approximating model, then after switching the terms r and l, (1) and Theorem 1 also apply to candidate l. This implies that candidate l also has a range of uncertainty, which depends on her value of θ. Further, if candidate r chooses a policy platform within that range, then candidate l should choose $l = r$. If, though, candidate r chooses a policy outside of that range, then candidate l will choose a policy that is slightly closer to the median voter under the approximating model.

Combining each candidate's best response for a given position of the opponent, we can now characterize the Nash equilibria in this candidate game.

Proposition 1 *Any combination $r = l$ where both $r \in [\underline{r}, \overline{r}]$ and $l \in \left[\underline{l}, \overline{l}\right]$ is a Nash equilibrium in the candidate game. Further, there is no other pure strategy Nash equilibrium in the candidate game.*

Proof. Suppose $r \in [\underline{r}, \overline{r}]$ and $l \in \left[\underline{l}, \overline{l}\right]$. Then, using Theorem 1, we see that candidate r's best response is to set $r = l$ and candidate l's best response is to set $l = r$. As neither candidate has the incentive to deviate from their announced policy platform, then we have a Nash equilibrium. Depending on the size of both candidates' uncertainty, θ, there could be an infinite number of Nash equilibria or the equilibrium could be unique.

To show that there exists no other pure strategy Nash equilibria in the game, consider without loss of generality an equilibrium in which $l \notin \left[\underline{l}, \overline{l}\right]$. If this is the case, then candidate r's best response is to move slightly closer to the median voter than l. Depending on whether r is within candidate l's range of uncertainty, candidate l might either set $l = r$ or move closer to the median voter. Either way, as long as $l \notin \left[\underline{l}, \overline{l}\right]$, then candidate r would continue to want to move even closer to the median voter. Thus, an equilibrium in which $l \notin \left[\underline{l}, \overline{l}\right]$ cannot exist. As this was without loss of generality, the same implication holds for candidate r: there exists no equilibrium in which $r \notin [\underline{r}, \overline{r}]$. Taken together, this implies that the only pure strategy Nash equilibrium involves both $r \in [\underline{r}, \overline{r}]$ and $l \in \left[\underline{l}, \overline{l}\right]$ and therefore $r = l$. ■

This proposition suggests that, by introducing model uncertainty and uncertainty aversion into the Downsian spatial model, it is optimal for both candidates to announce the same policy platform. This policy convergence is a common finding in the literature across a number of different assumptions; see Aragones and Palfrey (2002) and Enelow and Hinich (1989) for additional examples. One novel implication of our model is that, if both candidates have a finite θ, there is a range of possible Nash equilibria upon which the candidates could converge. This multiplicity stems from the fact that neither candidate is confident about the location of the median voter's preferred policy point. As a consequence, both candidates are willing to coordinate on any policy platform within a larger set so that they can both guarantee that they garner no worse than 50% of the vote, regardless of the true voter distribution.

Another implication of our model is that model uncertainty offers a potential justification as to why candidates' platforms can change over time, even if the median voter's preferred policy has remained fixed. To see this, consider the situation in which the median voter's location does not move from election to election. In this situation, the baseline model of Downs (1957) would predict that all candidates will announce the same policy platform across time. However, faced with the same situation, uncertain candidates might coordinate on a different equilibrium, as described in the proposition. Intuitively, the candidates are not tethered to the median voter's preferred policy because the candidates are unsure about that location.

It is important to note that the results described in this paper hold for any combination of the candidates' uncertainty. That is, suppose one believes that one candidate faces a large amount of uncertainty while the other candidate is relatively confident in the approximating model. The set of Nash equilibria described in the proposition still hold: the only pure strategy Nash equilibria in this particular game would involve both candidates converging on the same policy platforms within the intersection of both candidates' ranges of uncertainty. The fact that one candidate is less uncertain than the other merely reduces the size of the set of possible Nash equilibria. In fact, Downs (1957) should be viewed as a special case of this more general model, where both candidates are characterized by $\theta \to \infty$. Moreover, the same Nash equilibrium as in Downs (1957) would hold and be unique if only one of the candidates had full information, while the other faced model uncertainty.

One might wonder whether our model assumes that each candidate knows the other's uncertainty level, θ. The answer is no: each candidate need only know her own level of uncertainty and the approximating model. Given these pieces of information, each candidate can construct her range of uncertainty, which then determines her best response function.

Finally, we would like to highlight that the set of Nash equilibria described in the proposition holds regardless of whether the game is played simultaneously or sequentially. To see this, understand that the model we described above assumed that the game was played simultaneously and that each candidate, in constructing her best response function, merely considered the possibility that the other chose a particular policy platform. Well, instead of assuming that the candidates consider the other's possible action, we could have assumed that the candidate views the actual action of the opponent. This change would then lead to the same theorem and the same proposition as we have described. Thus, our results are independent of the timing of the game.

5 Numerical examples

In this section, we present three numerical examples to better illustrate the results described above. In all three examples, we assume that the approximating probability distribution of candidate r is uniform on

$[0, 1]$. This assumption of uniformity allows us to re-write (1) as

$$\text{Subjective expected utility} = \begin{cases} -\theta \log \left[\left(\frac{r+l}{2} \right) + \exp\left(-\frac{1}{\theta} \right) \left(1 - \frac{r+l}{2} \right) \right], \text{ if } r > l \\ -\theta \log \left[\exp\left(-\frac{1}{\theta} \right) \left(\frac{r+l}{2} \right) + \left(1 - \frac{r+l}{2} \right) \right], \text{ if } r < l \\ \frac{1}{2}, \text{ if } r = l \end{cases}$$

In these examples, we will assume particular values of θ and l. Given these values, we will then determine candidate r's optimal policy platform.

In the first example, we will assume that candidate r has a large degree of uncertainty and l is relatively close to the median voter. We will show that it is optimal for candidate r to set $r = l$. In example 2, we will keep the same value for θ, but assume that l is further away from the median voter. This will lead candidate r to choose a policy such that $r \neq l$. Finally, in example 3, we will return to our original assumption about the location of l, but now assume that candidate r faces a smaller degree of uncertainty. In this case, just like in example 2, we will show that candidate r will choose a policy in which $r \neq l$.

Example 1: High uncertainty, moderate l

Suppose that, in addition to the uniformity assumption, $\theta = 1$ and $l = \frac{2}{5}$. Given that $\theta = 1$, we can show that candidate r's range of uncertainty has the following bounds: $\underline{l} \approx 0.3775$ and $\bar{l} \approx 0.6225$. These bounds imply that candidate l has chosen a policy platform that is within candidate r's range of uncertainty. Consequently, we should expect to see that it is optimal for candidate r to set $r = l$.

To see whether this is indeed true, let's examine candidate r's subjective expected utility for different choices of r. If $r = l$, then candidate r's subjective expected utility is $\frac{1}{2}$. If $r = l + \epsilon$ where $\epsilon > 0$ is small, then her subjective expected utility approaches 0.4769. This is the highest expected utility that candidate r can hope to achieve by choosing a policy platform to the right of l. If $r = l - \epsilon$, then her subjective expected utility approaches 0.2915. Again, this is the highest expected utility that candidate r can hope to achieve by choosing a policy platform to the left of l. As $\frac{1}{2}$ is the highest subjective expected utility that candidate r can achieve across all of her possible policy choices, she would choose to set $r = l$.

Example 2: High uncertainty, extreme l

In this example, we still assume that $\theta = 1$, but now we assume that $l = \frac{1}{3}$. This policy chosen by candidate l is outside of candidate r's range of uncertainty, and so we should expect to see that candidate r's optimal policy announcement is to set $r = l + \epsilon$.

As before, candidate r has 3 options. If she chooses $r = l$, then her subjective expected utility is $\frac{1}{2}$. If she sets $r = l + \epsilon$, then her utility approaches 0.5471. This utility falls if candidate r chooses any value of r such that $r > l + \epsilon$. Finally, if she chooses $r = l - \epsilon$, her utility approaches 0.2366. Again, her utility falls if she chooses an $r < l - \epsilon$. Thus, to maximize her subjective expected utility, candidate r chooses $r = l + \epsilon$, for some small value of $\epsilon > 0$.

Example 3: Low uncertainty, moderate l

In our final example, we will assume that $\theta = 10$ and $l = \frac{2}{5}$. Given this level of uncertainty, we can show that candidate r's range of uncertainty has the following bounds: $\underline{l} \approx 0.4875$ and $\bar{l} \approx 0.5125$. These bounds imply that candidate l has chosen a policy platform that is outside of candidate r's range of uncertainty. As such, we should expect to see that candidate r will choose $r = l + \epsilon$.

If candidate r chooses $r = l$, then her subjective expected utility is $\frac{1}{2}$. If she chooses $r = l + \epsilon$, then her utility is 0.5879. This utility falls as candidate r moves further to the right. Finally, if candidate r chooses $r = l - \epsilon$, her utility is 0.3881. This utility falls as candidate r moves further to the left. Thus, candidate r will choose to set $r = l + \epsilon$ for some small $\epsilon > 0$.

6 Conclusion

In this paper, we relax the assumption that political candidates in the Downsian voting model understand the true distribution of voters. Instead, we assume that they believe that there is a set of alternative voter distributions, each element of which could potentially characterize the true voter distribution. This type of uncertainty implies that the politicians understand the relationship between a citizen's type and her vote, but that they do not know how many citizens of that type will show up to the polls to vote. To us, this type of uncertainty seems plausible, as there are hundreds of polls taken in the years before a Presidential election, but the actual distribution of voters varies dramatically from election to election, both at the national level and at the state level.

With this assumption of candidate uncertainty, we derive two main results. First, all pure strategy Nash equilibria involve policy convergence. Second, there is a multiplicity of possible equilibria that surround the median voter under the approximating model. Put another way, the candidates can coordinate on any policy that is close enough to the median voter, where "close enough" depends upon the size of the candidates' uncertainty. If their uncertainty is small, then the range of possible equilibria is small; if their uncertainty is large, then there is a large range of possible equilibria.

To review the logic behind this result, imagine a sequential game between candidates l and r. Suppose l

initially chooses a particular policy platform that is sufficiently close but not necessarily equal to the median voter. Then, r has three possibilities. First, r could choose a platform to the right of l. Given deterministic voting behavior, r knows with certainty that she will have support from all voters to the right of l, will have no support from voters to the left of l, and split the votes in between l and r. But, r doesn't know the true distribution of voters, and therefore cannot know whether such a policy is a winning strategy. Further, r is uncertainty averse, meaning that the candidate is worried that the true distribution of voters is heavily weighted to the left of l. The analogous logic applies if r chooses a platform to the left of l. These forces push r's policy choice to l's policy choice, i.e. convergence.

This logic, though, assumes that l has chosen an initial point sufficiently close to the median voter. If instead l chooses a point that is sufficiently far from the median voter, then candidate r would not choose to exactly match l's platform. The reason behind this result is that r's uncertainty is not great enough for her to believe that the distribution is one in which many voters lie close to the extreme point of l. Consequently, candidate r chooses a platform that is slightly closer to the median voter under the approximating model, and so there is policy divergence. Since l is guaranteed a loss with this announcement, this policy divergence is not a Nash equilibrium.

References Cited

[1] Abdellaoui, M., A. Baillon, L. Placido, and P. Wakker 2011. The rich domain of uncertainty: Source functions and their experimental implementation. *American Economic Review* 101, 695-723.

[2] Adams, J. and S. Merrill 1999. Modeling party strategies and policy representation in multiparty elections: Why are strategies so extreme? *American Journal of Political Science* 43, 765-791.

[3] Ahn, D., S. Choi, D. Gale, and S. Kariv 2008. Estimating ambiguity aversion in a portfolio choice experiment. Working paper.

[4] Aragones, E. and T. Palfrey. 2002. Mixed equilibrium in a Downsian model with a favored candidate. *Journal of Economic Theory* 103, 131-161.

[5] Aragones, E. and T. Palfrey. 2005. Spatial competition between two candidates of different quality: The effects of candidate ideology and private information. Social Choice and Strategic Behavior: Essays in Honor of Jeffrey Scot Banks. Austen-Smith, D. and J. Duggan, eds. Springer: Berlin.

[6] Banks, J. S., 1990. A model of electoral competition with incomplete information. *Journal of Economic Theory* 50(2), 309–325.

[7] Black, D., 1958. *The Theory of Committees and Elections*, New York, Cambridge University Press.

[8] Bossaerts, P., P. Ghirardato, S. Guarnaschelli, and W. Zame 2010. Ambiguity in asset markets: Theory and experiment. *Review of Financial Studies* 23, 1325-1359.

[9] Calvert, R., 1985. Robustness of the Multidimensional Voting Model, Candidate Motivations, Uncertainty and Convergence. *Journal of Political Science* 39, 69-95.

[10] Camerer, C., Weber, R., 1992. Recent Developments in Modeling Preferences: Uncertainty and Ambiguity. *Journal of Risk and Uncertainty* V, 325-370.

[11] Congdon-Hohman, J., A. Nathan, and J. Svec. 2013. Student Uncertainty and Major Choice: A Robust Control Approach, working paper.

[12] Dennis, R., K. Leitemo, and U. Soderstrom 2009. Methods for Robust Control. *Journal of Economic Dynamics and Control* 33, 1604-1616.

[13] Dennis, R., 2010. How robustness can lower the cost of discretion. *Journal of Monetary Economics*, 57, 653-667.

[14] Downs, A., 1957. *An Economic Theory of Democracy*, New York, Harper & Row.

[15] Ellsberg, D., 1961. Risk, Ambiguity and the Savage Axioms. *Quarterly Journal of Economics* LXXV, 643-669.

[16] Enelow, J. and M. Hinich 1989. A general probabilistic spatial theory of elections. *Public Choice* 61, 101-113.

[17] Erikson, R. and G. Wright 1980. Policy representation of constituency interests. *Political Behavior* 2, 91-106.

[18] Erikson, Robert S., and Gerald C. Wright, Jr. 1989. "Voters, Candidates, and Issues in Congressional Elections." In Congress Reconsidered, 4th ed., ed. Lawrence C. Dodd and Bruce I. Oppenheimer. Washington: Congressional Quarterly.

[19] Erikson, R. and G. Wright 1993. Voters, candidates, and issues in congressional elections. *Congress Reconsidered*, 5th edition, Washington DC, Congressional Quarterly Press.

[20] Erikson, R. and G. Wright 1997. Voters, candidates, and issues in congressional elections. *Congress Reconsidered*, 6th edition, Washington DC, Congressional Quarterly Press.

[21] Fiorina, M., 1974. *Representatives, Roll Calls, and Constituencies*, Lexington, MA, Lexington.

[22] Groseclose, T., 2001. A model of candidate location when one candidate has a valence advantage. *American Journal of Political Science* 45, 862-886.

[23] Halevy, Y., 2007. Ellsberg revisited: An experimental study. *Econometrica* 75, 503-536.

[24] Hanson, L., F. Barillas, and T. Sargent 2009. Doubts or Variability? *Journal of Economic Theory* 144, 2388-2418.

[25] Hansen, L. and T. Sargent 2005. Robust estimation and control under commitment. *Journal of Economic Theory* 124, 258–301.

[26] Hansen, L. and T. Sargent 2007. Recursive robust estimation and control without commitment. *Journal of Economic Theory* 136, 1–27.

[27] Hansen, L. and T. Sargent 2008. *Robustness*. Princeton, Princeton University Press.

[28] Hotelling, H., 1929. Stability in competition. *The Economic Journal* 39, 41-57.

[29] Karantounias, A. 2013. Managing pessimistic expectations and fiscal policy. *Theoretical Economics* 8, 193-231.

[30] Kartik N. and R. McAfee 2007. Signaling character in electoral competition. *American Economic Review* 97, 852-870.

[31] Leitemo, K. and U. Söderström 2008. Robust monetary policy In the new keynesian framework. *Macroeconomic Dynamics* 12, 126-135.

[32] Levin, A. and J. Williams 2003. Robust monetary policy with competing reference models. *Journal of Monetary Economics* 50, 945-975.

[33] McDonald, M., 2013. Presidential voter turnout rates, 1948-2012. *United States Elections Project*, accessed 4-1-2013.

[34] Poole, K. and H. Rosenthal 1984. U.S. presidential elections 1968-1980: A spatial analysis. *American Journal of Political Science* 28, 283-312.

[35] Schofield, N., 2007. The mean voter theorem: necessary and sufficient conditions to a convergent equilibrium. *Review of Economic Studies* 74, 1083-1119.

[36] Stokes, D., 1963. Spatial models of party competition. *American Political Science Review* 57, 368-377.

[37] Svec, J. 2012. Optimal fiscal policy with robust control. *Journal of Economic Dynamics and Control* 36, 349-368.

[38] Svec, J. 2013. Optimal Capital Taxation and Consumer Uncertainty. College of the Holy Cross working paper.

[39] Wittman, D., 1983. Candidate Motivation: A Synthesis of Alternatives. *American Political Science Review* 77, 142-157.

7 Appendix

In this appendix, we prove Theorem 1. This proof involves three main steps. In the first step, we sign the derivative of candidate r's subjective expected utility. This derivative will allow us to say that, conditional on choosing a platform to the right of l, candidate r is best off choosing a platform as close as possible to l and, conditional on choosing a platform to the left of l, candidate r is best off choosing a platform as close as possible to l. This will allow us to narrow our field of possible best responses to platforms close to candidate l's platform. The second step then shows that if either $l = \underline{l}$ or $l = \bar{l}$, then $r = l$ is optimal. Step 3 compares the subjective expected utility of candidate r at particular values of r to the expected utility obtained at the bounds of candidate r's range of uncertainty. In doing so, we will be able to say whether candidate r's utility is greater or less than $\frac{1}{2}$, which then allows us to know whether that candidate is better off by setting $r = l$. We will repeat this second step multiple times to complete the proof.

For ease, we re-write candidate r's subjective expected utility below:

$$\text{Subjective expected utility} = \begin{cases} -\theta \log \left[\int\limits_0^{\frac{r+l}{2}} f(x)\,dx + \exp\left[-\frac{1}{\theta}\right] \int\limits_{\frac{r+l}{2}}^{1} f(x)\,dx \right], \text{ if } r > l \\ -\theta \log \left[\exp\left[-\frac{1}{\theta}\right] \int\limits_0^{\frac{r+l}{2}} f(x)\,dx + \int\limits_{\frac{r+l}{2}}^{1} f(x)\,dx \right], \text{ if } r < l \\ \frac{1}{2}, \text{ if } r = l \end{cases}; \tag{2}$$

Step 1: Determine the sign of $\frac{\partial U}{\partial r}$ conditional on the location of r relative to l.

If we assume that $r > l$, then candidate r's subjective expected utility is

$$-\theta \log \left[\int\limits_0^{\frac{r+l}{2}} f(x)\,dx + \exp\left[-\frac{1}{\theta}\right] \int\limits_{\frac{r+l}{2}}^{1} f(x)\,dx \right]$$

Given this, we can show that

$$\frac{\partial U}{\partial r} = \frac{-\frac{\theta}{2} f\left(\frac{r+l}{2}\right)\left[1 - \exp\left(-\frac{1}{\theta}\right)\right]}{\int\limits_0^{\frac{r+l}{2}} f(x)\,dx + \exp\left[-\frac{1}{\theta}\right] \int\limits_{\frac{r+l}{2}}^{1} f(x)\,dx} < 0$$

This derivative implies that, conditional on $r > l$, candidate r is always better off by remaining close to l

than announcing a platform far away from l. Doing the same thing assuming that $r < l$, we see that

$$\frac{\partial U}{\partial r} = \frac{-\frac{\theta}{2} f\left(\frac{r+l}{2}\right)\left[\exp\left(-\frac{1}{\theta}\right) - 1\right]}{\int_0^{\frac{r+l}{2}} f(x)\,dx + \exp\left[-\frac{1}{\theta}\right] \int_{\frac{r+l}{2}}^{1} f(x)\,dx} > 0$$

Again, conditional on $r < l$, candidate r is better off by remaining close to l.

Step 2: If either $l = \underline{l}$ or $l = \bar{l}$, then $r = l$ is optimal

This will be a proof by contradiction. Suppose that, without loss of generality, $l = \underline{l}$, and it was optimal for candidate r to choose $r = \underline{l} - \epsilon$ for some small $\epsilon > 0$. This must mean that two conditions hold:

$$\begin{aligned} -\theta \log\left[\exp\left[-\frac{1}{\theta}\right] \int_0^{\underline{l}} f(x)\,dx + \int_{\underline{l}}^{1} f(x)\,dx\right] &\geq \frac{1}{2} \\ -\theta \log\left[\int_0^{\underline{l}} f(x)\,dx + \exp\left[-\frac{1}{\theta}\right] \int_{\underline{l}}^{1} f(x)\,dx\right] &= \frac{1}{2} \end{aligned}$$

Re-writing these conditions, we get

$$\begin{aligned} \exp\left[-\frac{1}{\theta}\right] \int_0^{\underline{l}} f(x)\,dx + \int_{\underline{l}}^{1} f(x)\,dx &\leq \exp\left(-\frac{1}{2\theta}\right) \\ \int_0^{\underline{l}} f(x)\,dx + \exp\left[-\frac{1}{\theta}\right] \int_{\underline{l}}^{1} f(x)\,dx &= \exp\left(-\frac{1}{2\theta}\right) \end{aligned}$$

Adding these two conditions together, we see

$$2\exp\left(-\frac{1}{2\theta}\right) \geq 1 + \exp\left(-\frac{1}{\theta}\right)$$

which then implies that

$$\left(\exp\left(-\frac{1}{2\theta}\right) - 1\right)^2 \leq 0$$

which is a contradiction. Thus, if $l = \underline{l}$, then it cannot be optimal for candidate r to choose $r = \underline{l} - \epsilon$. Also, given our definition of $\underline{l}$, we know that if $r = \underline{l} + \epsilon$, then candidate r's subjective expected utility equals $\frac{1}{2}$. Finally, using our derivatives from Step 1, we know that candidate r's subjective expected utility falls as she moves either to the left of $r = \underline{l} - \epsilon$ or to the right of $r = \underline{l} + \epsilon$. Taken together, we see that if $l = \underline{l}$, then candidate r's optimal response is $r = \underline{l}$. The same pattern of the proof applies to $\bar{l}$, so we suppress that proof for brevity.

Step 3: Utility comparison

1. Show that if $l \in [\underline{l}, \bar{l}]$, $r = l$. This is again a proof by contradiction. Suppose that $l \in [\underline{l}, \bar{l}]$, but that candidate r's optimal response is $r \neq l$. Assume first that candidate r's best response is $r = l + \epsilon$. Then, two conditions hold:

$$\begin{aligned}
-\theta \log \left[\int_0^l f(x)\, dx + \exp\left[-\frac{1}{\theta}\right] \int_l^1 f(x)\, dx \right] &\geq \frac{1}{2} \\
-\theta \log \left[\int_0^{\underline{l}} f(x)\, dx + \exp\left[-\frac{1}{\theta}\right] \int_{\underline{l}}^1 f(x)\, dx \right] &= \frac{1}{2}
\end{aligned}$$

which implies that

$$\int_0^{\underline{l}} f(x)\, dx + \exp\left[-\frac{1}{\theta}\right] \int_{\underline{l}}^1 f(x)\, dx \geq \int_0^l f(x)\, dx + \exp\left[-\frac{1}{\theta}\right] \int_l^1 f(x)\, dx$$

But, with finite θ, then $\exp\left[-\frac{1}{\theta}\right] \in (0,1)$. This means that the left hand side is smaller than the right hand side because the left hand side places a smaller weight on a larger mass of the voter distribution. Thus, we have a contradiction. Combining this result with the derivative from Step 1, we know that if $l \in [\underline{l}, \bar{l}]$, $r \leq l$.

Assume that candidate r's best response is $r = l - \epsilon$. Then, two conditions hold:

$$\begin{aligned}
-\theta \log \left[\exp\left[-\frac{1}{\theta}\right] \int_0^l f(x)\, dx + \int_l^1 f(x)\, dx \right] &\geq \frac{1}{2} \\
-\theta \log \left[\exp\left[-\frac{1}{\theta}\right] \int_0^{\bar{l}} f(x)\, dx + \int_{\bar{l}}^1 f(x)\, dx \right] &= \frac{1}{2}
\end{aligned}$$

which implies that

$$\exp\left[-\frac{1}{\theta}\right] \int_0^{\bar{l}} f(x)\, dx + \int_{\bar{l}}^1 f(x)\, dx \geq \exp\left[-\frac{1}{\theta}\right] \int_0^l f(x)\, dx + \int_l^1 f(x)\, dx$$

But, this cannot hold because the left hand side places a smaller weight on more mass than does the right hand side. Thus, we have a contradiction. Combining this result with the derivative from Step 1 and with the earlier work in Step 3, we know that if $l \in \left[\underline{l}, \overline{l}\right]$, $r = l$.

2. Show that if $l \notin \left[\underline{l}, \overline{l}\right]$, candidate r will choose a policy platform that is $\epsilon > 0$ closer to the median voter (under the approximating model) than l. This is again a proof by contradiction. Suppose that $l < \underline{l}$, and candidate r chooses $r = l - \epsilon$. Then, we know that the following two conditions hold:

$$\begin{aligned} -\theta \log \left[\exp\left[-\frac{1}{\theta}\right] \int_0^l f(x)\,dx + \int_l^1 f(x)\,dx \right] &\geq \frac{1}{2} \\ -\theta \log \left[\exp\left[-\frac{1}{\theta}\right] \int_0^{\overline{l}} f(x)\,dx + \int_{\overline{l}}^1 f(x)\,dx \right] &= \frac{1}{2} \end{aligned}$$

These conditions imply

$$\exp\left[-\frac{1}{\theta}\right] \int_0^{\overline{l}} f(x)\,dx + \int_{\overline{l}}^1 f(x)\,dx \geq \exp\left[-\frac{1}{\theta}\right] \int_0^l f(x)\,dx + \int_l^1 f(x)\,dx$$

But, this cannot hold because the left hand side places a smaller weight on more of the mass than does the right hand side. Thus, we have a contradiction. Combining this result with the derivative from Step 1, then we know it is not optimal for candidate r to choose $r < l$ if $l < \underline{l}$.

Now, we know that if $l < \underline{l}$, then candidate r will set $r \geq l$. But, we still must show that it is optimal for candidate r to choose $r = l + \epsilon$. To do this, suppose the opposite; namely, $r = l$ is the optimal policy announcement for candidate r. Then, the following conditions must hold:

$$\begin{aligned} -\theta \log \left[\int_0^l f(x)\,dx + \exp\left[-\frac{1}{\theta}\right] \int_l^1 f(x)\,dx \right] &\leq \frac{1}{2} \\ -\theta \log \left[\int_0^{\underline{l}} f(x)\,dx + \exp\left[-\frac{1}{\theta}\right] \int_{\underline{l}}^1 f(x)\,dx \right] &= \frac{1}{2} \end{aligned}$$

These conditions imply

$$\int_0^{\underline{l}} f(x)\,dx + \exp\left[-\frac{1}{\theta}\right] \int_{\underline{l}}^1 f(x)\,dx \leq \int_0^l f(x)\,dx + \exp\left[-\frac{1}{\theta}\right] \int_l^1 f(x)\,dx$$

But, this cannot hold because the left hand side places a smaller weight on less mass than does the right hand side, meaning that the left hand side is greater than the right hand side. Thus, we have a contradiction. Combining these results with the derivative from Step 1, we can see that it is optimal for candidate r to choose $r = l + \epsilon$ if $l < \underline{l}$.

The proof is similar for $l > \bar{l}$, so we suppress it. Taken together, we have now proven each component of Theorem 1.