

# Polarization and Pandering in Common-value Elections

Joseph McMurray\*

December 30, 2016

## Abstract

Standard election models robustly predict that candidates should adopt moderate policies, especially if they value winning office, and that this is good for voter welfare. Empirical polarization is therefore both inexplicable and troubling. This paper shows that if voter differences reflect beliefs about an underlying truth variable, rather than idiosyncratic preferences, then polarization is the robust equilibrium prediction, even when the value of winning is high, as both sides claim superiority. Moderation can then be undesirable, as a spatial form of pandering to popular opinion, though the greater danger is over-extremism, which can result from overconfidence.

JEL Classification Number D72, D82

Keywords: Voting, Elections, Ideology, Median Voter, Information Aggregation, Polarization, Jury Theorem, Public Opinion, Epistemic Democracy, Overconfidence

## 1 Introduction

The central economic model for analyzing electoral politics is the *private-value* spatial model pioneered by Hotelling (1929) and Downs (1957). In that model, each voter is characterized by a one-dimensional preference parameter, ranging from the liberal left to the conservative right, that designates his favorite policy.<sup>1</sup> To attract

---

\*Brigham Young University Economics Department. Email joseph.mcmurray@byu.edu. Thanks to Mark Fey, Ken Shotts, John Duggan, Navin Kartik, Gábor Virág, Stephane Wolton, Odilon Câmara, Jay Goodliffe, Adam Meirowitz, Rainer Schwabe, Roger Myerson, Tim Feddersen, Jean Guillaume Forand, Adam Dynes, Aniol Llorente-Saguer, Faruk Gul, and Chris Ellis, for their interest and suggestions.

<sup>1</sup>Throughout this paper, masculine pronouns refer to voters and feminine pronouns refer to candidates.

votes, candidates adopt moderate equilibrium policy positions. This is good for social welfare, because compromising between the left and right minimizes the average distance from voters' ideal policies, and thus the total disutility voters must endure from policies that they most despise. Empirically, however, it is counterfactual. For example, statistical studies of political behavior in the U.S. House, Senate, presidency, and state legislatures repeatedly find Democrats and Republicans to be clearly distinct from one another, and more extreme on average than voters.<sup>2</sup> Estimates by Poole and Rosenthal (1984) and Bafumi and Herron (2010) are especially extreme, with politicians resembling the most liberal and conservative voters in the electorate. This seems to be voters' perception, as well: in the eleven U.S. presidential elections between 1972 and 2012, for example, ninety percent of participants in the American National Election Studies (ANES) rated both major candidates as weakly farther from the center of a standard seven-point ideological scale than they rated themselves, while only eleven percent rated both candidates as weakly more moderate than themselves.<sup>3</sup> Numerous studies have tried to explain polarization theoretically, but as Section 2 explains, the most widely cited theories are problematic; especially when candidates value holding office, the convergence prediction has proven so robust that Roemer (2004) refers to the "tyranny of the median voter theorem". From the perspective of this literature, therefore, polarization remains both puzzling and troubling, reflecting some inexplicable political failure.

This paper takes a new look at candidate polarization, utilizing a *common-value* paradigm that has been largely overlooked by modern theory, but was introduced over two centuries ago by Condorcet (1785). In this model, voters view political decisions as if through the eyes of social planners, and thus share a common interest in doing whatever is truly best for society. Because information is imperfect, however, they disagree what the optimal policy is. The true optimum is modeled as an unknown random variable, and private opinions are modeled as signals, each correlated with the truth. The foundation for democracy in that setting is Condorcet's (1785) "jury" theorem, which Krishna and Morgan (2011) refer to as "the first welfare theorem of political economy": by the law of large numbers, collective opinion can reliably

---

<sup>2</sup>For example, see Poole and Rosenthal (1984), McCarty and Poole (1995), Ansolabehere, Snyder, and Stewart (2001), Jessee (2009, 2010, 2016), Bafumi and Herron (2010), Shor (2011), and Fowler and Hall (2016).

<sup>3</sup>This is also consistent with campaign rhetoric, where candidates emphasize their differences but rarely their similarities.

identify whichever of two policies is truly superior, even when individual opinions are highly unreliable. Condorcet’s original model features only a binary choice, but in McMurray (2016a) I extend this to a spatial environment where the optimal policy may lie anywhere in a continuum. As that paper explains, focusing on voter beliefs rather than preferences can explain various empirical behaviors that are puzzling from a private-value perspective.<sup>4</sup> The model below extends the analysis further by introducing two candidates who compete for office by choosing policy platforms that they commit to implement if elected, as in standard spatial literature.

Like the private-value literature, the analysis below considers several possible specifications of candidate motivations. In contrast with that literature, however, the robust result is that candidates polarize in equilibrium, adopting positions far left and far right of center. This can happen, for example, if candidates desire to do what is truly optimal for society, but hold opposite beliefs regarding what that is: relying on the jury theorem, each candidate then expects voters to discover the truth, and join her side. Even if candidates are ex ante identical, opposite beliefs can arise endogenously in equilibrium, because of a strategic calculus for candidates akin to the well-known “pivotal” calculus of voters: a candidate’s platform will only matter if she receives a majority of votes, which is most likely to happen when truth is on her side. Candidates whose policy preferences do not depend at all on the state of the world may polarize as well, because voter opinions are correlated in a way that limits the benefit of moderation.

If candidates are sufficiently office motivated then a *median opinion theorem* emerges, predicting that candidates converge to the center as in standard models. However, the welfare implication of this behavior is rather opposite that of the private-value case: rather than a welfare-enhancing compromise between competing extremes, convergence here can reflect a compromise of truth in pursuit of popularity—like the *pandering* results of Canes-Wrone, Herron, and Shotts (2001) and Maskin and Tirole (2004), but with a geometry that explains why, empirically, moderate candidates are often shunned. In some cases, compromise can even produce policies that are known

---

<sup>4</sup>For example, information considerations can explain why voter preferences update over time, why electoral outcomes are often lop-sided, why citizens on both sides of an issue expect to win, and why citizens who lack confidence in their opinions tend to remain ideologically moderate and to abstain from voting or to cast incomplete ballots. That paper also explains how a common interest can arise from low levels of voter altruism, amplified in large elections. Principal-agent models that treat the electorate monolithically as a single individual (see Besley, 2006, for a review) are most easily understood in a common-value framework, as well.

ex ante *not* to be optimal, such as a moderate-sized economic stimulus policy, when the optimal level of stimulus is either very large or very small.

In large elections, the logic of the jury theorem alleviates the need for candidates to pander to voters. As long as candidates place *some* utility weight on the policy outcome, therefore, they may be highly polarized, *in spite of* strong office motivation. In fact, depending on the specification, they may be just as polarized in large elections as they are when they care only about the policy outcome, and do not value winning at all. In a common-value environment, therefore, polarization emerges as the robust equilibrium prediction. That candidates resist the urge to pander can be good for welfare, but if candidates are overconfident in their policy beliefs, or have more extreme preferences than voters, then they will over-polarize. The welfare consequences of over-polarization are especially severe when there is *aggregate uncertainty* about the underlying state variable, meaning that voters possess incomplete information, even collectively.

The remainder of this paper is organized as follows. Section 2 begins by comparing this paper to the extensive literatures on electoral convergence and imperfect information. Section 3 formally introduces the model, and Sections 4 and 5 analyze equilibrium incentives for voters and candidates, respectively. Section 6 then discusses social welfare, and Section 7 concludes.

## 2 Literature

### 2.1 Electoral Convergence

Hotelling (1929) and Downs (1957) show that if voting is deterministic then candidates who seek only to win office should adopt identical platforms at the ideal point of the median voter. Calvert (1985) and Wittman (1977) show that policy motivated candidates should do the same, essentially because choosing policy requires winning first. Pundits often attribute polarization to the undue influence of extremists within either party, whether because moderate voters opt not to participate or because primary electorates are skewed relative to the general electorate, but either of these is problematic because party extremists themselves should favor candidate moderation, to avoid sacrificing the election to the opposing side.<sup>5</sup>

---

<sup>5</sup>Davis, Hinich, and Ordeshook (1970) show that the median voter theorem is robust whether voter abstention is allowed or not. Hirano, Snyder, Ansolabehere, and Hansen (2010) and McGhee

Wittman (1983, 1990), Hansson and Stuart (1984), Calvert (1985), Londregan and Romer (1993), and Roemer (1994) show that candidates do not converge when they are policy motivated and uncertain how citizens will vote, so subsequent literature has often attributed polarization to this combination of ingredients. However, Calvert (1985), Roemer (1994), and Banks and Duggan (2005) show that small levels of uncertainty should only produce a small degree of polarization: straying too far from the expected position of the median voter still surrenders policy control (with high probability) to one’s opponent.<sup>6</sup> Given the prevalence of public opinion polls, it seems unlikely that uncertainty should be so severe that candidates expect the median voter to be extremely liberal or extremely conservative. Moreover, if candidates are office- or vote-motivated then they should converge in spite of electoral uncertainty, as Hinich (1977, 1978), Coughlin and Nitzan (1981), Calvert (1985), Lindbeck and Weibull (1987), Enelow and Hinich (1989), Duggan (2000, 2006), and Banks and Duggan (2005) variously show.

Alesina (1988) attributes political extremism to the lack of credibility of campaign promises: once elected, a candidate can be as extreme as she wishes. As that paper shows, however, this logic is only valid to the extent that candidates or parties do not value reelection, as otherwise they should seek to establish moderate reputations. Even when reelection incentives are absent, candidates who are known to intrinsically prefer moderate policies should enjoy an electoral advantage over extremists, by the standard reasoning. A lack of credibility can also generate equilibria with polarized candidates in the entry models of Osborne and Slivinski (1996) and Besley and Coate (1997). However, such equilibria require near-ties (Eguia, 2007), which often do not occur empirically, and equilibria with little or no polarization also exist in these models, and are equally plausible.<sup>7</sup> In any case, if a lack of credibility explains why candidates who pretended to be moderate turn out not to be, it offers no explanation for candidates who openly advocate opposite extremes.

The above seem to be the most widely cited explanations for political polarization, but alternative theories abound, including minor party influence (Palfrey 1984,

---

et al. (2014) show empirically that polarization is not affected by the structure of state primary elections.

<sup>6</sup>A similar observation applies to the analysis of Coleman (1971), who shows that uncertainty regarding the general election makes voters in primary elections more willing to nominate extremists.

<sup>7</sup>The same is true of the multiple-candidate positioning game studied by Myerson and Weber (1993) where, in equilibrium, strategic voters ignore all but two candidates, who may have any policy positions, polarized or not.

Castanheira 2003, Callander and Wilson 2007, Brusco and Roy 2011); asymmetric ability, charisma, incumbency status, or media exposure (Bernhardt and Ingberman, 1985, Ansolabehere and Snyder, 2000, Groseclose, 2001, Aragonés and Palfrey, 2002, Gul and Pesendorfer, 2009, Soubeyran, 2009, Krasa and Polborn, 2010, 2012, and Matějka and Tabellini, 2015); informational asymmetries (Glaeser, Ponzetto, and Shapiro, 2005, Bernhardt, Duggan, and Squintani, 2007, and Asako, 2014); efforts to signal hidden types (Kartik and McAfee, 2007, Callander and Wilkie, 2007, Callander, 2008, and Kartik, Squintani, and Tinn, 2012); interactions across jurisdictions or between branches of government (Ortuño-Ortín, 1997, Alesina and Rosenthal, 2000, Eyster and Kittsteiner, 2007, Krasa and Polborn, 2015, and Polborn and Snyder, 2016); non-policy competition (Ashworth and Bueno de Mesquita, 2009, and Van Weelden, 2013); and convex voter utility (Kamada and Kojima, 2014). Each of these theories, of course, has strengths and weaknesses; by comparison, the explanation below is notable for its simplicity: if truth is on her side, a candidate need not move to the center because voters will instead come to her. In particular, this explanation does not require special electoral circumstances such as multiple elections, an incumbent, or third-party pressure.

## 2.2 Imperfect Information

Existing extensions of Condorcet’s (1785) model focus on electoral efficiency with correlated signals or other informational impediments (Ladha, 1992, 1993, Mandler, 2012, Dietrich and Spiekermann, 2013, and Pivato, 2016), alternative voting rules (Young, 1995, Feddersen and Pesendorfer, 1998, List and Goodin, 2001, and Ahn and Oliveros, 2016), and deviations from common value (Feddersen and Pesendorfer, 1997, 1999, Kim and Fey, 2007, Krishna and Morgan, 2011, and Bhattacharya, 2013), as well as on strategic incentives to vote insincerely (Austen-Smith and Banks, 1996, and Acharya and Meirowitz, 2016) and abstain from voting (Feddersen and Pesendorfer, 1996, Krishna and Morgan, 2012, and McMurray, 2013). For the most part, this work retains Condorcet’s binary structure or extends to a small number of alternatives and truth states. Many authors also explicitly restrict the scope of their analysis to committees or juries, agreeing with Black (1987, p. 163) that common values are “clearly inapplicable” to public elections. Most importantly, these models focus only on voting behavior, treating candidates’ characteristics and policy positions as exogenous. The same is true of McMurray (2016a), where I introduce

a truly spatial model of common-value elections, but specify candidates' policy positions exogenously. The model below is identical to that, except that candidates' policy positions are derived endogenously, for various assumptions about preferences and beliefs.

There are several existing information models in which candidates play active roles. Models of binary choice include those of Canes-Wrone, Herron, and Shotts (2001) and Maskin and Tirole (2004), which feature a single incumbent politician, and Heidhues and Lagerlof (2003), Laslier and Van der Straeten (2004), and Gratton (2014), which feature two candidates competing for office. More traditional spatial models include those of Schultz (1996), Martinelli (2001), Loertscher (2012), and Kartik, Squintani, and Tinn (2013), who all model voters preferences as idiosyncratic, but shifted in tandem according to a common shock. Pandering arises in the models of Canes-Wrone, Herron, and Shotts (2001), Maskin and Tirole (2004), and Loertscher (2012), as candidates implement policies that they believe to be inferior but popular. Candidates in Laslier and Van der Straeten (2004) and Gratton (2014) instead reveal their information truthfully, expecting voters to receive signals of their own. In Kartik, Squintani, and Tinn (2013) candidates *anti*-pander by deviating even further from voters' priors than their private information warrants, so as to appear confident and well-informed. In all of this literature, however, the focus is on whether or not candidates' policy choices reveal their private information to voters. In contrast, the model below focuses on aggregating *voter* information. This seems appropriate in that voters do not seem to rely heavily on inferences about candidates' private information.<sup>8</sup>

There are two existing models in which voter information influences candidate behavior.<sup>9</sup> In a binary setting, Harrington (1993) shows that an overconfident incumbent politician may choose the ex-ante unpopular policy, confident that voters will discover its merit and reelect her. Bernhardt, Duggan, and Squintani (2009a) consider an electorate with idiosyncratic policy preferences that shift in tandem with a truth variable that is realized after candidate positions are chosen but before voting takes place, leading policy-motivated candidates to deviate from the median voter's

---

<sup>8</sup>Presumably, a candidate also knows more than a typical voter knows, but much less than the electorate knows collectively.

<sup>9</sup>In addition to the election models listed here, Razin (2003) and McMurray (2016b) consider common-value models in which candidates adjust their policy positions *after* voting takes place, so that voting takes on a signaling role.

ideal point.<sup>10</sup> In both papers, deviations from the median voter’s preferences or prior beliefs are smaller if candidates are office motivated. The analysis below includes similar results for overconfident and policy- and office-motivated candidates, but includes two active candidates with a wider variety of motives, allows a continuum of policies and truth states, and treats large electorates. These generalizations are important as they highlight the extent of polarization, and how office motivation can be unimportant in large elections.

### 3 The Model

The information structure and voting behavior in this paper are identical to those modeled in McMurray (2016a). To begin, there are  $N$  citizens in an electorate, where, as in Myerson (1998),  $N$  is drawn at the beginning of the game from a Poisson distribution with mean  $n$ . There is an interval  $\mathcal{X} = [-1, 1]$  of policy alternatives, and the electorate must implement one of these, which will then provide a common utility to each citizen. The policy  $z$  provides the greatest utility possible, but its location is unknown; at the beginning of the game, nature draws  $z$  from the domain  $\mathcal{Z} \subseteq \mathcal{X}$ . If the state of the world is  $z$  but policy  $x$  is implemented then each citizen receives the following utility,

$$u(x, z) = -(x - z)^2 \tag{1}$$

which declines quadratically with the distance between  $x$  and  $z$ . This specification is convenient because expected utility is then similarly quadratic in the policy choice. In particular, this means that preferences are single-peaked, as in standard spatial voting models, and that the optimal policy choice is simply the expectation of  $z$ , conditional on any available information. The concavity of (1) also implies that voters are risk averse.

There are two important specifications of this model. The simpler of the two assumes *binary truth*, as in Condorcet’s (1785) original model. That is,  $\mathcal{Z} = \{-1, 1\}$ , meaning that the optimal policy lies at one of the two ends of the policy space. One application where this seems appropriate is macroeconomic policy: depending on whether Keynesian or more classical economic theory is closer to the

---

<sup>10</sup>With a similar source of divergence, Levy and Razin (2015) show that voter overconfidence may increase or decrease polarization.

truth, the ideal size of an economic stimulus policy is either quite large or quite small. A moderate-sized stimulus is also feasible—and could be desirable for avoiding catastrophic mistakes—but is known *ex ante* not to be optimal, *per se*. More broadly, Harrington (1993) proposes binary truth to describe voters’ deepest worldviews: if governments are either generally effective or generally ineffective at improving on market outcomes, for example, then the optimal policy may be either “extensive or minimal government intervention in the economy”. Moderate interventions may not ultimately be optimal, but hedge against error until one worldview emerges as clearly superior. For most applications, moderate policies cannot be ruled out as optimal, so it is more appropriate to assume *continuous truth*. In that case, let  $\mathcal{Z} = [-1, 1]$ , meaning that any feasible policy might also be optimal.<sup>11</sup>

Whether truth is binary or continuous, let  $z$  be distributed uniformly on  $\mathcal{Z}$ . The function

$$f(z) = \begin{cases} \frac{1}{2} & \text{if } z \in \mathcal{Z} \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

doubles conveniently as a density function for the case of continuous truth, and as a mass function for the case of binary truth.

An individual’s *hunch* regarding the location of the optimal policy can be modeled as a private signal  $s_i$  drawn from the same domain  $\mathcal{S} = \mathcal{Z}$  as the true optimum. The confidence that a voter places on his own hunch depends on his level of expertise on the policy question at hand.<sup>12</sup> Let  $q_i$  denote the quality of a citizen’s signal, drawn independently for each citizen (and independently from  $z$ ) from the domain  $\mathcal{Q} = [0, 1]$ , according to some common distribution  $G$ , which for simplicity is differentiable and has a strictly positive density  $g$ . Conditional on  $q_i = q$ , the distribution of  $s_i = s$  in state  $z$  is then given by the following,

$$h(s|q, z) = \frac{1}{2} (1 + qsz) \quad (3)$$

---

<sup>11</sup>In McMurray (2016a) I show that a continuous  $z$  is also appropriate when truth is binary but there is *aggregate uncertainty*. As Section 6 discusses below, this possibility has important consequences for the interpretation of welfare results.

<sup>12</sup>“Expertise”, “confidence”, and “information quality” are treated as synonymous here, and could derive from policy-relevant technical training, or simply from time spent thinking deeply about political issues. That voters might over- or under-estimate their own competency is an important possibility for future work to explore, but as Sunstein (2002) writes, “it is sensible to say that as a statistical matter, though not an invariable truth, people who are confident are more likely to be right”.

which, like (2), doubles conveniently as a density if truth is continuous and as a mass function if truth is binary. For binary truth,  $q_i$  corresponds to the correlation coefficient between  $s_i$  and  $z$ ; a citizen with  $q_i = 1$ , for example, observes  $z$  perfectly. With continuous truth, the correlation between  $s_i$  and  $z$  is only  $\frac{1}{3}q_i$ , so even the highest quality signals include substantial noise, but with either specification  $s_i$  is uniform on  $\mathcal{S}$  and its precision increases with  $q_i$ . The lowest quality signal reveals nothing: if  $q_i = 0$  then  $s_i$  and  $z$  are independent.

By Bayes' rule, a citizen's posterior belief about the optimal policy inherits the linearity of (2) and (3),

$$f(z|q, s) = \frac{h(s|q, z) g(q) f(z)}{\int_{\mathcal{Z}} h(s|q, z) g(q) f(z) dz} = \frac{1}{2}(1 + qsz) \equiv \frac{1}{2}(1 + \theta z) = f(z|\theta) \quad (4)$$

and depends on  $q_i$  and  $s_i$  only through the product  $\theta_i = q_i s_i$ .<sup>13</sup> Once again, (4) can be interpreted either as a density or a mass function (where integration in the latter case should be interpreted as a discrete summation over  $\mathcal{Z}$ ). Summing or integrating over  $\mathcal{Z}$ , a citizen's expectation of the optimal policy is then simply proportional to  $\theta_i$ , which can therefore be interpreted as a citizen's *ideology*.<sup>14</sup> The sign and magnitude of ideology depend on the sign and magnitude of  $s_i$ , and the magnitude also depends on a citizen's expertise  $q_i$ . Specifically, a citizen who lacks confidence in his opinions remains ideologically moderate, even if  $s_i$  is quite extreme. In McMurray (2016a) I show that this is true empirically, and also point out how this so naturally produces a spectrum of opinions: even if truth is binary, voter beliefs range continuously from fully embracing one side, to merely leaning in one direction or the other, to fully embracing the opposite side.

Citizens do not vote for policies directly. Instead, there are two candidates,  $A$  and  $B$ , who propose policy platforms  $x_A, x_B \in \mathcal{X}$  that they commit to implement if elected.<sup>15</sup> Observing these platforms, citizens then vote for either candidate.<sup>16</sup> A strategy  $v : \mathcal{Q} \times \mathcal{S} \rightarrow \{A, B\}$  in the voting subgame specifies a candidate choice  $j \in \{A, B\}$  for every realization  $(q, s) \in \mathcal{Q} \times \mathcal{S}$  of private information. Let  $\mathcal{V}$  denote the set of such strategies. Votes are cast simultaneously, and a winning candidate

<sup>13</sup>Formulated this way,  $\theta_i$  are *affiliated* with  $z$  in the sense of Milgrom and Weber (1982).

<sup>14</sup>For binary truth,  $E(z|q_i, s_i) = \theta_i$ ; for continuous truth,  $E(z|q_i, s_i) = \frac{1}{3}\theta_i$ .

<sup>15</sup>This extends McMurray (2016a), where candidate positions are specified exogenously. McMurray (2016b) extends further by relaxing the commitment assumption.

<sup>16</sup>For simplicity, abstention is not allowed here, but is treated in McMurray (2016a).

$w \in \{A, B\}$  is determined by majority rule, breaking a tie if necessary by a coin toss. The policy outcome is then the winning candidate's policy platform  $x_w$ . For a voter, therefore, expected utility for a given pair  $(x_A, x_B) \in \mathcal{X}^2$  of candidate platforms must average not only across the possible locations of the optimal policy, but also across the identity of the election winner, conditional on that optimum.

$$E[u(x, z)] = \int_{\mathcal{Z}} \left[ \sum_{w=A, B} u(x_w, z) \Pr(w|z) \right] f(z) dz \quad (5)$$

Implicitly, (5) depends (through  $\Pr(w|z)$ ) on the strategies used by every voter. If his peers all vote according to the strategy  $v \in \mathcal{V}$ , a citizen's *best response* is the strategy  $v^{br} \in \mathcal{V}$  that maximizes (5) for every realization  $(q, s) \in \mathcal{Q} \times \mathcal{S}$  of private information. A (symmetric) *Bayesian Nash equilibrium (BNE)* in the voting subgame is a strategy  $v^*$  that is its own best response.<sup>17</sup> Equilibrium voting behavior is analyzed in Section 4, taking a pair  $(x_A, x_B) \in \mathcal{X}^2$  of candidate platforms as given.

Section 5 analyzes candidate behavior under alternative assumptions regarding candidates' motivations and beliefs. One possibility is that, like ordinary citizens, candidates prefer policies as close as possible to whatever is truly optimal. This case is labeled below as *truth motivation*. Alternatively, it may be that candidates prefer specific policies, say  $\hat{x}_A$  and  $\hat{x}_B$ , regardless of the true state of the world. This case is labeled *policy motivation*. In the case of truth motivation, a candidate's optimal behavior will depend on her beliefs about  $z$ . The most natural assumption is that she receives a private signal of her own, but it also seems plausible that she might be overconfident in her information. The analysis below formalizes notions of *Bayesian* and *overconfident* candidates that approximate these cases, but also considerably simplify the analysis.<sup>18</sup> All types of candidates may also be *office motivated*, meaning that, regardless of the policy outcome, they perceive a benefit  $\beta \geq 0$  from winning office. If the magnitude of  $\beta$  is sufficiently large, a candidate is willing to make any policy concessions necessary in order to win. For any of these versions of the model, let  $\Sigma$  denote the set of complete voting strategies  $\sigma : \mathcal{X}^2 \rightarrow \mathcal{V}$ ,

---

<sup>17</sup>In games of Poisson population uncertainty, equilibrium symmetry is inevitable because the distribution of opponent behavior is the same for any two individuals within the game (unlike a game between a finite number of players), implying that a best response for one citizen is a best response for all.

<sup>18</sup>In particular, these specifications alleviate the need for voters to infer candidates' private information, as in the candidate revelation models listed in Section 2.

which specify subgame behavior for every possible pair  $(x_A, x_B) \in \mathcal{X}^2$  of candidate platforms. A *perfect Bayesian equilibrium (PBE)* is a triple  $(x_A^*, x_B^*, \sigma^*) \in \mathcal{X}^2 \times \Sigma$  such that  $\sigma^*(x_A, x_B)$  constitutes a BNE in the voting subgame associated with every platform pair  $(x_A, x_B) \in \mathcal{X}^2$ , and candidates' platform choices  $x_A^*$  and  $x_B^*$  maximize the appropriate objectives (given available information), taking the opposing platform and the voting strategy  $\sigma^*$  as given. Given the symmetry of the model, it is natural to focus further on equilibria that are *platform-symmetric*, meaning that  $x_A^* = -x_B^*$ .

## 4 Voting

### 4.1 Equilibrium Voting

In the subgame associated with a particular pair  $(x_A, x_B) \in \mathcal{X}^2$  of candidate platforms (where, without loss of generality,  $x_A \leq x_B$ ), the analysis of voting behavior in McMurray (2016a) directly applies. This section therefore merely reiterates the notation and results of that paper for subsequent use, and then extends the analysis to large elections. First, if citizens follow the voting strategy  $v \in \mathcal{V}$  then, in state  $z \in \mathcal{Z}$ , each votes for candidate  $j \in \{A, B\}$  with the following probability,

$$\phi(j|z) = \int_{\mathcal{Q}} \int_{\mathcal{S}} 1_{v(q,s)=j} h(s|q, z) g(q) dsdq \quad (6)$$

where the indicator function  $1_{v(q,s)=j}$  equals one if  $v(q, s) = j$  and zero otherwise. As Myerson (1998) explains,  $\phi(j|z)$  can also be interpreted as the expected vote share of candidate  $j$  in state  $z$ , and the numbers  $N_A$  and  $N_B$  of  $A$  and  $B$  votes are independent Poisson random variables with means  $n\phi(A|z)$  and  $n\phi(B|z)$ , respectively. The joint distribution of exactly  $a$  votes for candidate  $A$  and  $b$  votes for candidate  $B$  is therefore as follows.

$$\psi(a, b|z) = \frac{e^{-n\phi(A|z)-n\phi(B|z)}}{a!b!} [n\phi(A|z)]^a [n\phi(B|z)]^b \quad (7)$$

By the environmental equivalence property of Poisson games (Myerson 1998), a voter within the game reinterprets (7) as the distribution of votes cast by his peers; by voting himself, he can add one to either candidate's total. His own vote will be *pivotal* in the election (event  $\mathcal{P}$ ) if the candidates otherwise tie or one candidate trails by exactly one vote (and would win the tie-breaking coin toss); in terms of (7), this occurs with the following probability.

$$\begin{aligned}
\Pr(\mathcal{P}|z) &= \Pr(N_A = N_B|z) + \frac{1}{2}\Pr(N_A = N_B + 1|z) + \frac{1}{2}\Pr(N_B = N_A + 1|z) \\
&= \sum_{k=0}^{\infty} \left[ \psi(k, k|z) + \frac{1}{2}\psi(k, k+1|z) + \frac{1}{2}\psi(k+1, k|z) \right]
\end{aligned} \tag{8}$$

Since citizens seek to make the policy outcome as close as possible to  $z$ , and since those with more conservative ideologies believe  $z$  to be further to the right, whenever one citizen finds it optimal to vote for candidate  $B$ , more conservative citizens prefer to vote  $B$  as well. The implication of this is that, as Lemma 1 of McMurray (2016a) states, the best response to any subgame voting strategy  $v$  is *ideological*, meaning that there is an *ideology threshold*  $\tau \in \mathcal{X}$  such that citizens with ideology left of  $\tau$  vote  $A$  and those with ideology right of  $\tau$  vote  $B$ . Specifically, the ideology threshold that characterizes the unique best response to any voting strategy is given by the following<sup>19</sup>

$$\tau^{br} = \frac{\bar{x} - E(z|\mathcal{P})}{E(z^2|\mathcal{P}) - \bar{x}E(z|\mathcal{P})} = \frac{\bar{x} - E(z|\mathcal{P})}{V(z|\mathcal{P}) + E(z|\mathcal{P})^2 - \bar{x}E(z|\mathcal{P})} \tag{9}$$

This expression depends on the midpoint  $\bar{x} = \frac{x_A + x_B}{2}$  between the two candidates' platforms and on a citizen's expectation

$$E(z|\mathcal{P}) = \frac{\int_{\mathcal{Z}} z \Pr(\mathcal{P}|z) f(z) dz}{\int_{\mathcal{Z}} \Pr(\mathcal{P}|z) f(z) dz} \tag{10}$$

of the optimal policy, conditional on the event of a pivotal vote. When he casts his vote, of course, a citizen does not know whether his vote will be pivotal or not, but as Austen-Smith and Banks (1996) point out, a citizen should adopt a strategy that will be optimal in the event that his vote does turn out to be pivotal, as otherwise his behavior does not influence his utility.<sup>20</sup>

---

<sup>19</sup>The above definition of an ideological voting strategy does not specify how citizens behavior whose ideologies lie exactly at the ideology threshold. This is unimportant for the results below in that such ideology is realized with zero probability and since, in equilibrium, such a citizen is indifferent between voting  $A$  and voting  $B$ . However, claims of equilibrium uniqueness throughout this paper are only valid up to the specification of behavior for this indifferent type.

<sup>20</sup>In forming beliefs, voters should also infer whatever they can from candidates' platform choices. For simplicity, however, candidates' information is specified below in a stylized way such that their behavior conveys nothing useful to voters.

For any pair of distinct candidate platforms, Proposition 1 of McMurray (2016a) states the existence and uniqueness of an equilibrium strategy, characterized by the ideology threshold  $\tau^*$ . Proposition 1 of this paper extends that result to show that  $\tau^*$  increases with  $\bar{x}$ , and does not otherwise depend on candidates' platforms. Also,  $\tau^*$  is symmetric for symmetric values of  $\bar{x}$ . If candidates are equidistant from the center so that  $\bar{x} = 0$ , for example, then  $\tau^* = 0$  as well: that is, citizens simply vote  $A$  if  $\theta_i$  is negative and vote  $B$  if  $\theta_i$  is positive.

**Proposition 1** *There exists a unique function  $\tau^* : \mathcal{X} \rightarrow \mathcal{X}$  such that for any  $x_A, x_B \in \mathcal{X}$  with midpoint  $\bar{x}$  the ideological strategy  $v^*(\bar{x})$  characterized by the ideology threshold  $\tau^*(\bar{x})$  constitutes a BNE in the voting subgame. For  $x_A < x_B$ ,  $v^*(\bar{x})$  is the unique BNE. Moreover,  $\frac{d\tau^*(\bar{x})}{d\bar{x}} > 0$  and  $\tau^*(-\bar{x}) = -\tau^*(\bar{x})$ .*

Proposition 1 characterizes equilibrium voting for a fixed population parameter  $n$ . Since real-world electorates tend to be very large, the following section analyzes voting behavior in the limit as  $n$  grows large.

## 4.2 Large Elections

The numbers of votes that each candidate receives depends not only on the voting strategy, but on the many realizations of voters' private signals, which in turn depend on state of the world  $z$ . For an ideological strategy with ideology threshold  $\tau$ , define  $z_\tau$  to be the realization of  $z$  that minimizes  $|\phi(A|z) - \phi(B|z)|$ —that is, the state that equalizes candidates' vote shares as closely as possible. As the electorate grows large, the probability of a single vote being pivotal shrinks to zero, but when expected vote shares are equal, which occurs in state  $z_\tau$ , pivot probabilities shrink more slowly than in any other state. Accordingly, a voter who behaves as if his vote will be pivotal increasingly behaves as if  $z_\tau$  will be realized as the optimal policy.

If the number of citizens is large and a citizen's peers follow an ideological strategy with  $z_\tau < \bar{x}$ , then, by the above logic, he should vote  $A$  in response (since  $x_A$  is closer to  $z_\tau$  than  $x_B$  is) regardless of his private information; if  $z_\tau > \bar{x}$  then he should vote  $B$  in response. Either way, therefore, a citizen should be unwilling to adopt the ideology strategy of his peers. It must therefore be the case that, as  $n$  grows large, the equilibrium threshold  $\tau^*(\bar{x})$  adjusts so that the state of the world  $z_{\tau^*(\bar{x})}$  implied by the equilibrium threshold leaves voters indifferent between  $A$  and  $B$ , and therefore willing to follow their signals, as Lemma 1 now states.

**Lemma 1** For any  $\bar{x} \in \mathcal{X}$ , the limiting equilibrium threshold  $\lim_{n \rightarrow \infty} \tau^*(\bar{x})$  solves  $\phi(A|z = \bar{x}; \tau) = \phi(B|z = \bar{x}; \tau) = \frac{1}{2}$ .

Lemma 1 highlights how the pivotal voting calculus substantially evens out the vote shares of the two candidates, which is relevant for candidate incentives in Section 5. For example, suppose that truth is continuous and that  $q_i = 1$  for every voter, so that an individual's private expectation of the optimal policy is simply  $E(z|q_i, s_i) = \frac{1}{3}s_i$ , and suppose further that  $x_A = .9$  and  $x_B = 1$  (implying that  $\bar{x} = .95$ ). That is, both candidates are so conservative that even the most conservative voter (i.e.,  $s_i = 1$  and  $E(z|q_i, s_i) = \frac{1}{3}$ ) prefers candidate  $A$ , who is slightly less extreme. Lemma 1 implies that, in spite of this lopsided support, the equilibrium threshold adjusts in large elections to solve  $\phi(B|z = \bar{x}; \tau) = \int_{\tau}^1 \frac{1}{2}(1 + s\bar{x}) ds = \frac{1}{2}$ , or  $\tau^* \approx .4$ . Thus, in equilibrium, candidate  $B$ 's vote share may range anywhere from  $\phi(B|z = -1; \tau = .4) = \int_{.4}^1 \frac{1}{2}(1 - s) ds = .09$  to  $\phi(B|z = 1; \tau = .4) = \int_{.4}^1 \frac{1}{2}(1 + s) ds = .51$ ; on average, candidate  $B$  expects about 30% of the votes. Intuitively, this balancing occurs because a vote is most likely to be pivotal when the quality difference between candidates—and therefore the difference in vote shares—is smaller than expected. If the only citizens who voted for candidate  $B$  were those with extremely far-right signals, for example, then a pivotal vote would be unlikely except when  $z$  happens to be extremely far right—which is precisely the scenario that makes candidate  $B$  more attractive than candidate  $A$ .

In McMurray (2016a) I show that Condorcet's (1785) binary jury theorem extends in a natural way to this spatial environment: specifically, the candidate whose platform is closest to the policy that is truly optimal almost surely wins a large election. That normative result is relevant in its own right, but also suggests an alternative intuition for the equilibrium balancing predicted above. If  $x_A = .9$  and  $x_B = 1$ , for example, then it is highly likely that  $x_A$  is superior to  $x_B$ . If citizens all merely voted naively for the candidate who seems superior, however, then none would vote for candidate  $B$ , and  $A$  would win *even* in the few states of the world where a  $B$  victory is optimal (namely, any state  $z > .95$ ). When  $z = .95$  exactly, policies at 9 and at 1 generate equivalent utility. In that case, the median signal realization is approximately .4. For maximal efficiency, therefore, a social planner instructs citizens with signals lower than .4 to vote  $A$  and instructs voters with signals above .4 to vote  $B$ . In this way,  $A$ 's vote share exceeds 50% precisely when  $z < .95$ , and  $B$ 's vote share exceeds 50% precisely when  $z > .95$ . The planner's recommendation cannot

be improved upon by any individual voter, this behavior constitutes an equilibrium voting strategy, so in McLennan (1998).

The jury theorem is a normative result, but in McMurray (2016a) I argue that it also sheds light on empirical facets of voter behavior, such as the broad support for using majority rule, and the tendency for political rhetoric to appeal to public opinion for support. Most importantly for this paper, it explains a *consensus effect* whereby individuals on both sides of an issue expect to belong to the majority: “in essence, a citizen who concludes that one policy is better than another predicts that other reasonable citizens, after weighing the evidence, will come to the same conclusion.” As I report in that paper, for example, 96% of ANES survey respondents who ultimately voted Democrat in the 2012 U.S. presidential election had earlier predicted a Democratic victory, while 83% of those who voted Republican had predicted a Republican victory. In what follows, similar reasoning leads a candidate who believes she is on the side of truth expects to be rewarded with votes.

## 5 Candidates

Having characterized voters’ equilibrium response to any pair of candidate platforms, this section proceeds to analyze the incentives this creates for candidates choosing policy platforms. Recall from Section 3 that a voting strategy  $\sigma \in \Sigma$  in the complete game specifies voting behavior  $v \in \mathcal{V}$  for the subgames associated with every pair  $(x_A, x_B) \in \mathcal{X}^2$  of candidate platforms. In particular, let  $\sigma_{\tau^*}$  denote the voting strategy that induces an ideological voting strategy in every subgame, with ideology thresholds given by  $\tau^*(\bar{x})$ , where  $\bar{x}$  is the midpoint between  $x_A$  and  $x_B$  and  $\tau^*$  is the function identified in Proposition 1. An equilibrium  $(x_A^*, x_B^*, \sigma^*) \in \mathcal{X}^2 \times \Sigma$  in the complete game consists of platform positions for both candidates and a voting strategy, and Proposition 1 implies that a necessary condition for equilibrium is that  $\sigma^*$  corresponds to  $\sigma_{\tau^*}$  in every subgame for which  $x_A \neq x_B$ .

There are various incentives that might plausibly motivate candidates’ policy decisions. This section considers several of these. Section 5.1 begins by assuming that candidates are *truth motivated*, meaning that they favor policies as close as possible to whatever is truly optimal. This is the same motivation as other citizens, consistent with the fact that candidates are citizens first, and could reflect an intrinsic interest in the public good, or could stem from a more selfish desire to establish a favorable

legacy. Section 5.2 then considers the possibility of *policy motivated* candidates, who prefer specific policies regardless of what is socially optimal. This, too, could be intrinsic, or might somehow reflect capture by special interests. The key result of either specification is equilibrium polarization, the extent of which is illustrated with numerical examples in Section 5.3. All of these sections presume that the intrinsic benefit of winning office is  $\beta = 0$ ; Section 5.4 then considers  $\beta > 0$ , meaning that candidates are also *office motivated*.

## 5.1 Truth Motivation

If candidates seek to do what is truly optimal then their behavior depends on where they believe  $z$  to be located. As noted in Section 3, the most straightforward assumption would be that candidates receive private signals regarding the optimal policy. Given the high levels of self-confidence that candidates exude during campaigns, it also seems appropriate to consider the possibility that candidates are overconfident, ascribing greater accuracy to their own policy opinions than is warranted. Section 5.1.1 and Section 5.1.2 approximate these cases, beginning with the simpler case of *overconfident* candidates and then proceeding to candidates who are *Bayesian*. For tractability and for emphasis, beliefs are modeled as being first stronger and then weaker than is plausible. Reality is likely somewhere between these two extremes, but the extreme cases prove surprisingly similar, suggesting that similar behavior would arise in an intermediate model, as well.

### 5.1.1 Overconfident Candidates

This section treats candidates who are *overconfident* in their private opinions regarding the location of the optimal policy. For simplicity, overconfidence is modeled in an extreme way: candidate  $A$  believes with probability one that the optimal policy is  $\theta_A$ , while candidate  $B$  believes with probability one that the optimal policy is  $\theta_B$  (where  $\theta_A < \theta_B$ ). Such beliefs are implausibly strong, but circumvent the need for a candidate to update her beliefs in response to the behavior of her opponent, or of voters. This is not only makes the analysis tractable, but also makes the consequences of overconfidence the most transparent.<sup>21</sup> With truth motivation and overconfidence modeled this way, the expected utility  $EU_j^{TO}$  of candidate  $j \in \{A, B\}$  can be written

---

<sup>21</sup>This is also the case considered by Harrington (1993).

as follows,

$$EU_j^{TO} = \sum_{w=j,-j} u(x_w, \theta_j) \Pr(w|z = \theta_j) + \beta \Pr(w = j|z = \theta_j) \quad (11)$$

where the utility  $u(x_w, z)$  associated with the winning candidate's platform and the probability  $\Pr(w|z)$  of that candidate winning are both evaluated at  $z = \theta_j$ . The second term in (11) reflects the possibility of office motivation, which Section 5.4 considers, but for now let  $\beta = 0$ .

From (11) it is clear that the trade-off faced by a candidate is fundamentally the same as in standard probabilistic voting models such as Wittman (1983) and Calvert (1985): a candidate's ideal policy is  $\theta_j$ , and moving toward this policy improves her utility, conditional on her winning the election, but moving toward her opponent attracts additional voters (given the monotonicity of  $\tau^*(\bar{x})$  with respect to either platform, observed in Proposition 1), thus increasing the probability of winning. Even if a candidate does not care about winning per se, this is valuable because she believes that her own policy platform is superior to her opponent's. In equilibrium, it cannot be the case that candidates adopt their ideal policies  $\theta_A$  and  $\theta_B$ , because the first-order utility loss from deviating slightly from these is zero, while the payoff gain from improving the chance of victory is strictly positive. It also cannot be the case that platforms coincide, however, because a candidate could then deviate toward her preferred policy position, making herself better off if she wins and no worse off if she loses. In other words, by standard reasoning, a candidate's equilibrium policy position lies strictly between her opponent's position and the policy  $\theta_j$  that she believes to be optimal. Theorem 1 states this formally, and points out that if candidates are *symmetrically overconfident*, meaning that  $\theta_A = -\theta_B$ , then given the other symmetry of the model, equilibrium can also be platform-symmetric; in fact, there is exactly one such equilibrium.

**Theorem 1** *If candidates are overconfident with  $\beta = 0$  then  $(x_A^*, x_B^*, \sigma^*) \in \mathcal{X}^2 \times \Sigma$  is a PBE only if  $\theta_A < x_A^* < x_B^* < \theta_B$ . If candidates are symmetrically overconfident then, for any  $n$ , there is exactly one PBE that is platform-symmetric. For any sequence of equilibria,  $\lim_{n \rightarrow \infty} x_j^* = \theta_j$  for  $j = A, B$ .*

While the basic logic of Theorem 1 is quite standard, the extent of polarization is not, as Section 2 notes: in standard probabilistic voting models, uncertainty regarding

the location of the median voter gives candidates leeway to move a little bit in their desired directions, but unless this uncertainty is quite severe, candidates must cater to the approximate location of the median voter, and therefore remain close to one another. In contrast, the last part of Theorem 1 makes clear that, for overconfident candidates, polarization is quite pronounced: in large elections, candidates do not moderate at all from the policies  $\theta_A$  and  $\theta_B$  that they most prefer. This more dramatic polarization simply stems from the jury theorem: when the electorate is large, majority opinion will almost surely favor the candidate whose policy platform is truly superior. When she is confident that her platform is superior, therefore, a candidate can also be confident that she will win. Thus, policy concessions are unnecessary. This is especially stark when truth is binary, so that  $\theta_A = -1$  and  $\theta_B = 1$ , and candidates adopt positions at opposite extremes of the policy space.

### 5.1.2 Bayesian Candidates

With overconfidence specified as above, one or both of the candidates must hold incorrect beliefs about the true state of the world. This section instead considers candidates who are *Bayesian*, meaning that they start from the correct prior and update rationally from any available information, using Bayes' rule. As noted above, the most straightforward source of information for a candidate would be a private signal of her own. In equilibrium, as will be shown below, a candidate can also infer information from voters. In fact, voting behavior turns out to be surprisingly informative. To emphasize this, candidates' private signals are actually not modeled here at all; thus, voters are candidates' *only* source of information.<sup>22</sup> Since she receives no private signal, the expected utility of candidate  $j \in \{A, B\}$  can be written as follows,

$$EU_j^{TB} = \int_{\mathcal{Z}} \left[ \sum_{w=j,-j} u(x_w, z) \Pr(w|z) \right] f(z) dz + \beta \Pr(w = B) \quad (12)$$

which differs from (11) only in that it now integrates over all possible realizations of the state variable  $z$ . As before, the second term in (12) reflects the possibility of office motivation, which Section 5.4 considers, but for now let  $\beta = 0$ .

---

<sup>22</sup>This also simplifies the analysis considerably, avoiding the need for voters to infer candidates' private signals from their platform positions, or for candidates to infer information from one another, or anticipate how much of their information will be inferred by voters.

With no private signals and no exogenous differences such as incumbency status, management experience, or charisma, Bayesian candidates are ex ante identical. Starting from the same prior, therefore, their basic inclination would be to adopt the same policy platform—namely, the center of the policy interval. With ideological voting, however, candidate  $A$  wins the election only when voters receive more liberal signals than conservative signals, which tends to be the case when the optimal policy is liberal; similarly, candidate  $B$  tends to win only when the optimal policy is conservative. Upon winning the election, therefore, the two candidates form different posterior beliefs regarding the location of the optimal policy.

This ex post inference may seem irrelevant for candidates' platform decisions, which must be made before voting takes place. However, a candidate's policy choice will only affect her utility if she wins the election. Thus, as Theorem 2 states, she chooses a policy platform  $x_j^* = E(z|w = j)$  in equilibrium that will be optimal in the event that she wins, given the updated beliefs that she expects to form in that event.<sup>23</sup> As above, the symmetry of the model is such that equilibrium platforms may be symmetric, which by Proposition 1 prompts symmetric voting; there is exactly one such equilibrium for any  $n$ , and when  $n$  is large,  $A$  simply infers that  $z$  is negative while  $B$  infers that  $z$  is positive, and the candidates polarize accordingly.<sup>24</sup>

**Theorem 2** *If candidates are Bayesian with  $\beta = 0$  then  $(x_A^*, x_B^*, \sigma^*) \in \mathcal{X}^2 \times \Sigma$  is a PBE only if  $x_A^* = E(z|w = A) < 0 < E(z|w = B) = x_B^*$ . For every  $n$ , there is exactly one platform-symmetric PBE, and for the sequence of these equilibria,  $\lim_{n \rightarrow \infty} (x_{A,n}^*, x_{B,n}^*) = (E(z|z < 0), E(z|z > 0))$ .*

The extent of polarization exhibited in Theorem 2 is substantial, and this is remarkable given that candidates are ex ante identical. With continuous truth, for example, Footnote 14 points out that citizens with the most extreme signal realizations favor policies  $-\frac{1}{3}$  and  $\frac{1}{3}$ . In large elections, however, Theorem 2 predicts that

---

<sup>23</sup>If candidates adopted identical platforms then voters would be indifferent between  $x_A$  and  $x_B$ , so voting need not be ideological and  $E(z|w = A) = E(z|w = B) = 0$  would be possible. This cannot occur in equilibrium, however, because deviations by either candidate would send voters into a different subgame in which voting *is* ideological, thus justifying the deviation. Considerations such as these rather complicate the equilibrium characterization: in evaluating positions  $x_B$  and  $x_B'$ , for example, candidate  $B$  must compare what she will learn if voters express a preference for  $x_B$  over  $x_A$  with what she will learn if they express a preference for  $x_B'$  over  $x_A$ .

<sup>24</sup>Given the symmetry of the model, it seems reasonable to conjecture that equilibria with asymmetric platforms do not exist. With ex-ante identical candidates, there could also be an equilibrium with  $B$  on the left and  $A$  on the right, but in that case Theorem 2 can be viewed simply as a relabeling of the candidates.

candidates adopt platforms close to  $E(z|z < 0) = -\frac{1}{2}$  and  $E(z|z > 0) = \frac{1}{2}$ . Such polarization, even relative to the most polarized voters, is consistent with the estimates of Bafumi and Herron (2010) and with the ANES perceptions documented in Section 1. With binary truth,  $E(z|z < 0) = -1$  and  $E(z|z > 0) = 1$ , implying that candidates diverge to the far extremes of the policy space.

The inference that leads to such polarization in Theorem 2 can be viewed as a “pivotal” calculus for candidates. That is, a candidate’s platform is only pivotal when she wins, so strategic candidates restrict attention to this event, and update their beliefs accordingly. That strategic voters should restrict attention to pivotal events has been recognized at least since Downs (1957), and that this focus can dramatically alter voting behavior in games of private information has been recognized at least since Austen-Smith and Banks (1996) and Feddersen and Pesendorfer (1996, 1998, 1999). The analogous calculus for candidates seems not to have been documented in existing literature, but Theorem 2 shows that its consequences can be similarly substantial. In fact, the  $N$  pieces of voter information are so informative when  $N$  is large that increasing this to  $N + 1$  would have little additional impact. In that sense, the assumption that candidates have no private information of their own is unrealistic, but innocuous.<sup>25</sup>

Empirically, it is not clear whether candidates actually perform the strategic calculus described above, any more than it is clear whether voters really take into account the informational implications of a pivotal vote. For example, politicians rarely admit any inkling that their policy proposals might be sub-optimal. One possibility is that candidates are sufficiently overconfident that rationally inferring voters’ information has little impact on their posteriors. Another (not mutually exclusive) possibility is that the structure of voter information is more complex than in the model above, so that majority opinion in reality is more fallible than the jury theorem suggests, and therefore pushes candidates’ beliefs less dramatically. It is not unreasonable for candidates to be at least a little more confident when they feel bolstered by popular opinion, even if this is subconscious.<sup>26</sup> In any case, to the extent that candidates are

---

<sup>25</sup>This basic logic would hold even if the quality of a candidate’s own signal were higher than the average quality of voters’ signals, and would be strengthened further if candidates could infer each other’s contradictory signals from their policy platforms.

<sup>26</sup>Sobel (2006) uses essentially this logic, for example, to explain why group decisions in experimental settings are often more extreme than individual opinions. Glaeser and Sunstein (2009) make similar arguments, while also emphasizing the possible importance of non-Bayesian cognitive mistakes that are not modeled here.

not already certain of the optimal policy, it would be irrational to ignore the informational content of winning the election, just as it is irrational for voters to ignore the informational content of a pivotal vote. The analysis of this section makes clear how remarkably strong such theoretical considerations can potentially be.

## 5.2 Policy Motivation

Whatever their beliefs, Section 5.1 assumes that candidates share voters' objective. This section considers an alternative possibility, which is that voters' and candidates' incentives are misaligned, whether intrinsically or because candidates are somehow beholden to special interests. Specifically, candidates  $A$  and  $B$  are *policy motivated* if they prefer policies  $\hat{x}_A \in \mathcal{X}$  and  $\hat{x}_B \in \mathcal{X}$ , respectively, regardless of the state of the world. Utility is still quadratic as in (1), but with  $z$  replaced by  $\hat{x}_j$ . Expected utility can then be written as follows.

$$EU_j^P = \sum_{w=j,-j} u(x_w, \hat{x}_j) \Pr(w) + \beta \Pr(w = j) \quad (13)$$

This is similar to the expression (11) for overconfident candidates, merely evaluating disutility in terms of the distance to  $\hat{x}_j$  rather than  $\theta_j$ , but differs in that a candidate's perceived probability  $\Pr(w = j)$  of winning the election is not related to her preferred policy position. As before, the second term in (13) reflects the additional possibility of office motivation, which Section 5.4 considers, but for now let  $\beta = 0$ .

Unlike truth motivated candidates who are overconfident or Bayesian, a policy motivated candidate has no reason to expect voters to be on her side. To have any influence over policy, therefore, she must cater to voters' beliefs. As in the probabilistic voting literature, however, candidates cannot converge entirely in equilibrium: if platforms were identical, a candidate could deviate toward her preferred policy, making herself no worse off if she loses but better off if she wins. Thus, Theorem 3 states that  $\hat{x}_A < x_A^* < x_B^* < \hat{x}_B$  in equilibrium. If candidates are *symmetrically policy motivated*, meaning that  $\hat{x}_A = -\hat{x}_B$ , then there again exists exactly one equilibrium that is platform-symmetric.

**Theorem 3** *If candidates are policy motivated with  $\hat{x}_A < \hat{x}_B$  and  $\beta = 0$  then  $(x_A^*, x_B^*, \sigma^*) \in \mathcal{X}^2 \times \Sigma$  is a PBE only if  $\hat{x}_A < x_A^* < x_B^* < \hat{x}_B$ . If candidates are symmetrically policy motivated then, for any  $n$ , there is exactly one platform-symmetric*

PBE, and for this sequence of equilibria,  $\lim_{n \rightarrow \infty} x_j^* = \begin{cases} \frac{\hat{x}_j}{1+|\hat{x}_j|} & \text{if } \mathcal{Z} = [-1, 1] \\ \hat{x}_j & \text{if } \mathcal{Z} = \{-1, 1\} \end{cases}$  for  $j = A, B$ .

The basic result of non-convergence is familiar from the probabilistic voting literature, but Theorem 3 predicts a much higher degree of polarization than in private-value models, as the numerical examples of Section 5.3 will make clear. In large elections, equilibrium platforms are no closer to the center than they are to candidates' preferred policies; with binary truth, in fact, candidates do not moderate from their preferred positions at all. Evidently, then, polarization is a robust prediction of the common-value spatial model: the precise degree of polarization depends on candidates' motivations and beliefs and on the nature of uncertainty, but substantial polarization occurs in every case.

### 5.3 Numerical Examples

In proving that candidates' equilibrium platforms do not coincide, Theorems 1 through 3 all draw on the logic of the private-value probabilistic voting literature: if platforms were the same, one candidate could deviate toward her ideal point and be no worse off if she loses, but better off if she wins. The purpose of this section is to illustrate by way of numerical examples that, while the basic logic is the same, the degree of equilibrium polarization differs dramatically between private-value and common-value models. To facilitate computation, all of the examples of this section assume that the number  $2n + 1$  of voters is known and odd, rather than following a Poisson distribution.<sup>27</sup>

Like the common-value model with policy motivated candidates, the private-value examples computed in this section assume that voter utility  $u_i(x) = -(x - \hat{x}_i)^2$  is quadratic in the distance between the policy implemented and the voter's privately preferred policy  $\hat{x}_i$ . Candidate utility is given by the same function  $u_j(x) = -(x - \hat{x}_j)^2$ , with candidate ideal points  $\hat{x}_A = -1$  and  $\hat{x}_B = 1$  at the far extremes of the policy space. As Duggan (2013) notes, there are two canonical specifications of probabilistic voting: in the *stochastic preference* models of Wittman (1983) and Calvert (1985), candidates cannot observe voters' precise ideal points, but know the

---

<sup>27</sup>Thus, (7) and (8) must be replaced by  $\psi(a, b|z) = \frac{2n!}{a!b!} \phi(A|z)^a \phi(B|z)^b$  and  $\Pr(\mathcal{P}|z) = \frac{2n!}{n!n!} \phi(A|z)^n \phi(B|z)^n$ , respectively.

distribution from which  $\hat{x}_i$  are drawn. In the examples below,  $\hat{x}_i$  is uniformly distributed on  $[-1, 1]$ . The *stochastic partisanship* models of Hinich (1978) and Lindbeck and Weibull (1987) assume that, in addition to policy utility, voters are biased in favor of one candidate or the other, each receiving a benefit  $\alpha_i$  if candidate  $A$  is elected, and voting for candidate  $B$  if and only if  $u(x_A, \theta_i) + \alpha_i > u(x_B, \theta_i)$ . For simplicity, the computations below assume that  $\alpha_i$  is an i.i.d. draw from a uniform distribution. For the sake of emphasis, the domain of  $\alpha_i$  is the interval  $[-4, 4]$  (where a negative  $\alpha_i$  reflects a bias for  $B$ ), calibrated such that biases can be large enough to make even the most extreme citizen vote for a candidate at the opposite extreme.

Column 1 of Table 1 displays the equilibrium positions of policy-motivated candidates in a private-value election with stochastic voter preferences. In small elections, uncertainty regarding the location of the median voter leads candidates to take up positions that are rather polarized. As the electorate grows large, however, the distribution of the median voter converges to its expected location, so equilibrium polarization declines. Column 2 of Table 1 displays equilibrium candidate positions for a private-value model with both stochastic preferences and stochastic partisanship. This generates more substantial polarization than the previous case, but once again, polarization declines as the electorate grows large. With at least 100,000 voters, for example, candidates are more moderate in equilibrium than 94% of the electorate, in spite of the huge partisan biases.

Equilibrium Candidate Positions								
Model:	Private Value		Continuous Truth			Binary Truth		
	Stochastic Preference	Stochastic Partisanship	Over-confident	Bayesian	Policy-motivated	Over-confident	Bayesian	Policy-motivated
Voters	1	2	3	4	5	6	7	8
3	0.400	0.549	0.142	0.123	0.104	0.778	0.688	0.675
5	0.348	0.511	0.227	0.152	0.157	0.862	0.793	0.777
11	0.270	0.451	0.395	0.212	0.245	0.958	0.931	0.923
15	0.241	0.427	0.474	0.241	0.282	0.980	0.965	0.961
25	0.199	0.389	-	-	-	0.996	0.993	0.993
51	0.149	0.340	-	-	-	1.000	1.000	1.000
101	0.111	0.296	-	-	-	1.000	1.000	1.000
1,001	0.038	0.180	-	-	-	1.000	1.000	1.000
10,001	0.012	0.106	-	-	-	1.000	1.000	1.000
100,001	0.004	0.061	-	-	-	1.000	1.000	1.000

Note: Table entries list the absolute value of candidates' equilibrium platform positions, for various specifications of the model.

Table 1

By way of contrast, columns 3 through 8 of Table 1 display equilibrium policy positions for overconfident candidates (with  $\theta_A = -1$  and  $\theta_B = 1$ ), Bayesian candidates,

and policy-motivated candidates (with  $\hat{x}_A = -1$  and  $\hat{x}_B = 1$ ), for both continuous and binary truth. In contrast with the private-value cases above, candidates become *more* polarized as the electorate grows large, not less. With binary truth this is especially rapid: even with only 15 voters, candidates are more extreme than 96% of the electorate, regardless of the precise specification of candidate motivations and beliefs. With continuous truth, polarization is less pronounced but still grows with the number of voters, quickly exceeding that of the private-value model.<sup>28</sup>

It would be possible to increase the degree of polarization in the private-value examples above to any level, of course, simply by adding uncertainty to the model. For candidates to take highly polarized positions, however, they must place reasonable probability on the median voter being highly liberal or highly conservative.<sup>29</sup> In a private-value context, it is unclear why uncertainty should be so severe: a simple preference parameter should be straightforward to ascertain in pre-election polls.<sup>30</sup> In a common-value setting, by contrast, voting behavior is tied to voter beliefs, which are much more volatile: polls could measure voter opinions on a certain date, but uncertainty would persist as to how these opinions will continue to evolve, right up until election day. In that sense, an information model provides a plausible and intuitive rationale for why elections are so unpredictable, as highlighted in the recent U.S. presidential election and U.K. referendum to leave the European Union.<sup>31</sup>

Columns 5 and 8 of Table 1 are of special interest in that polarization is high even though candidate preferences are specified just as in columns 1 and 2. In particular, policy motivated candidates do not care about the state variable  $z$ , except that it introduces uncertainty regarding the location of the median voter. In all of these models, voter ideal points are drawn from a known distribution; the key distinction is

---

<sup>28</sup>Computational difficulties limit the size of electorates for which examples can be computed, but limiting policy positions for overconfident, Bayesian, and policy-motivated candidates, respectively, are given in Theorems 1 through 3 as  $\theta_B = 1$ ,  $E(z|z > 0) = 0.5$ , and  $\frac{\hat{x}_B}{1+|\hat{x}_B|} = 0.5$ , respectively.

<sup>29</sup>If the median voter were known to lie between  $-\varepsilon$  and  $+\varepsilon$ , for example, candidates' equilibrium positions would have to be less extreme than  $-2\varepsilon$  and  $+2\varepsilon$ , lest one candidate deviate to the center and win with certainty.

<sup>30</sup>Drawing voter ideal points randomly from a known distribution is the canonical formulation of probabilistic voting, but in a truer mapping to reality, candidates would estimate the unknown population median by sampling voters via pre-election polls, as in Bernhardt, Duggan, and Squintani (2009b). Small polls would leave much uncertainty, so that candidates might indulge in extremism, but both sides would have incentive to sample more voters, thereby reducing uncertainty—and with it polarization—just as in columns 1 and 2 of Table 1.

<sup>31</sup>See [www.nytimes.com/2016/11/09/us/politics/hillary-clinton-donald-trump-president.html](http://www.nytimes.com/2016/11/09/us/politics/hillary-clinton-donald-trump-president.html) and [www.nytimes.com/2016/06/25/world/europe/britain-brexit-european-union-referendum.html](http://www.nytimes.com/2016/06/25/world/europe/britain-brexit-european-union-referendum.html).

that in columns 1 and 2 these draws are *independent* of one another, while in columns 5 and 8 they are *correlated* (since each is correlated with  $z$ ). From candidates' perspective, the source of correlation in voter preferences is unimportant. It may seem, therefore, that the assumption of common values is unnecessary: even in a purely private-value setting, assuming a correlation between voters' ideal points would generate polarization similar to that of columns 5 and 8. Without the common-value assumption, however, there is no reason why voter preferences should be correlated. That is, why should knowing the preferences of one voter shed any light on the preferences of his neighbor, if not because their interests are fundamentally linked? In fact, if voter preferences are correlated then the correlation structure actually *defines* a common value. Formally, de Finetti's (1990) theorem from probability theory states that symmetrically correlated (or *exchangeable*) random variables always can be viewed as mutually independent, conditional on a latent variable; in this application, that latent variable could be interpreted as the object of common value,  $z$ .

## 5.4 Office Motivation

The analysis above assumes candidates do not care about winning the election *per se*; winning is merely a means to the end of implementing favorable policy outcomes. Much of the spatial voting literature has started from the opposite assumption of *office motivation*: candidates' policy promises are merely a means to the end of winning the election. Most likely, candidates have mixed motivations, desiring good policies but at the same time hoping to be the one to implement them. To allow this possibility, this section leaves candidates' policy preferences as specified above, but assumes that a candidate also receives a positive benefit  $\beta > 0$  if she wins office. This accommodates both the case in which  $\beta$  is so large that it can compensate for any undesirable policy outcome, as well as moderate values of  $\beta$ , for which neither policy nor office motivations strictly dominate. Candidates' entry decisions are not modeled here but it seems reasonable to conjecture that, for candidates who are drawn into the election,  $\beta$  might be substantial.

In private-value models such as Hotelling (1929), Downs (1957), and Calvert (1985), candidates move toward each other to attract voters whose ideal points lie between the two platforms. The model here is very different, but the incentives are similar: moving toward each other attracts voters whose estimates of  $z$  lie between the two platforms. If office motivation is strong enough, therefore, then full convergence

occurs, as Theorem 4 now states.<sup>32</sup> For smaller values of  $\beta$ , full convergence does not occur but there is a unique platform-symmetric equilibrium, and for that equilibrium polarization decreases with  $\beta$ , a prediction that is common in private-value literature.<sup>33</sup> Theorem 4 is labeled as the median *opinion* theorem to emphasize that voter ideologies, which are all-important in determining political behavior, are in this model actually only approximations of a more fundamental preference, so that, as Section 6 emphasizes below, this familiar behavior has new implications for social welfare.

**Theorem 4 (Median Opinion Theorem)** *If candidates are overconfident, Bayesian, or policy motivated then there exists  $\bar{\beta}$  such that if  $\beta \geq \bar{\beta}$  then  $(x_A^*, x_B^*, \sigma^*) \in \mathcal{X}^2 \times \Sigma$  is a PBE only if  $x_A^* = x_B^*$ . Moreover, such an equilibrium exists. If  $\beta < \bar{\beta}$  and candidates are symmetrically overconfident, Bayesian, or symmetrically policy motivated, then there is a unique platform-symmetric PBE  $(x_A^*, x_B^*, \sigma^*) \in \mathcal{X}^2 \times \Sigma$  and, in this equilibrium, polarization  $|x_B^* - x_A^*|$  strictly decreases in  $\beta$ .*

On its surface, Theorem 4 might seem to suggest that the polarization predicted in Sections 5.1 and 5.2 is not robust: candidates who are not office motivated polarize to varying degrees depending on their specific motivations, but all types of candidates converge when the benefit of winning is sufficiently high. This is especially important if the desire to win office is candidates' primary objective, as observers of elections commonly assume. However, this reasoning is incomplete: whereas Theorem 4 fixes the population size  $n$  and considers arbitrarily large values of  $\beta$ , Theorem 5 now fixes  $\beta$  and, for the various policy preferences treated above, analyzes behavior in the limit as  $n$  grows large. The precise consequences of this depend on the precise specification of candidate incentives and whether truth is binary or continuous, but the common theme is that, to varying degrees, candidates still polarize even when the benefit of winning office is quite large—even arbitrarily large, in some cases.

**Theorem 5** *If truth is binary then  $\lim_{n \rightarrow \infty} (x_{A,n}^*, x_{B,n}^*)$  is the same for all  $\beta \geq 0$ , whether candidates are overconfident, Bayesian, or policy-motivated. If truth is continuous then  $(x_{A,n}^*, x_{B,n}^*)$  approaches  $(\theta_A, \theta_B)$  if candidates are overconfident, and for*

---

<sup>32</sup>If candidates are truth motivated and Bayesian or policy motivated, convergence must occur at the center of the policy interval. This is possible with overconfident candidates as well, but there is a range of possible points of convergence in that case, where both candidates expect to be winning.

<sup>33</sup>For example, see Alesina (1988) and Bernhardt, Duggan, and Squintani (2009a).

the unique sequence of platform-symmetric equilibria, approaches  $(E(z|z < 0) + \frac{\beta}{4}, E(z|z > 0) - \frac{\beta}{4})$  if candidates are Bayesian, and  $(\frac{\hat{x}_A + \frac{\beta}{4}}{1 + |\hat{x}_A|}, \frac{\hat{x}_B - \frac{1}{4}\beta}{1 + |\hat{x}_B|})$  if candidates are symmetrically policy-motivated.

According to Theorem 5,  $\beta$  may not matter at all in large elections. Specifically, if truth is binary or if candidates are overconfident then platforms are just as polarized when  $\beta$  is large as they are in Theorems 2 and 3, where  $\beta = 0$ . In other words, candidates in these situations are willing to promise any policy in order to get elected, but in equilibrium adopt the same polarized positions that they would choose if they didn't value winning at all. For overconfident candidates, the logic for this result is fundamentally the same as the logic underlying polarization in Sections 5.1 and 5.2, where  $\beta = 0$ : if a candidate knows for certain that truth is on her side then policy concessions are unnecessary, because proposing what she believes to be truly optimal is a strategy that, in large elections, is all but guaranteed to win. When truth is binary, candidates of all types come to a similar conclusion, because there are no intermediate states of the world where policy concessions will make a difference: if truth turns out to be on one side, a candidate will win no matter how polarized she is; if truth is against her, she will lose no matter how much she moderates her policy stance.

For continuous truth and candidates who are either policy motivated or Bayesian, the implication of Theorem 5 is that polarization in large elections strictly decreases in  $\beta$ , just as polarization in finite elections does. Even in these cases, however, candidates remain at least somewhat polarized unless  $\beta$  is quite large. For Bayesian candidates, full convergence requires that  $\beta$  exceed 2, meaning that winning the election compensates for a policy outcome that is a distance of  $\sqrt{2} \approx 1.4$ —or 70% of the length of the policy interval—from the optimal policy. For policy motivated candidates, full convergence requires that  $\beta$  exceed  $4|\tilde{x}_j|$ , which can be as high as 4, meaning that it compensates for a policy distance of 2, which is the entire length of the policy interval. In other words, office motivation must be strong enough to compensate for *any* undesirable policy (even though the amount by which they actually compromise is only half this length).

For reasonable values of  $\beta$ , Theorem 5 implies that polarization remains quite substantial. Suppose, for example, that  $\beta = \frac{1}{4}$ , which is large enough to compensate a candidate for a policy that is a distance of  $\frac{1}{2}$ —or 25% of the length of the policy interval—from her most preferred policy. In that case, truth motivated candidates

adopt policy positions at  $\pm\frac{7}{16} \approx .44$  in large elections, while policy motivated candidates adopt positions at  $\pm\frac{15}{32} \approx .47$ . Neither is much less polarized than  $\pm.5$ , which, according to Theorems 2 and 3, are the platforms that would prevail if  $\beta$  were zero.

Taken together, the various results of Theorem 5 make clear that, in contrast with the intuition that may seem suggested by Theorem 4, and in contrast with the large private-value literature, polarization in this setting is quite robust to the addition of office motivation. Fixing a population size, it is true that  $\beta$  can be increased to the point that policy platforms completely converge. Fixing any  $\beta$ , however,  $n$  can also be increased to the point that candidates revert completely (or almost completely) to their original polarized positions. In large elections, therefore, substantial polarization can persist for very large—in some cases, arbitrarily large—values of  $\beta$ .

## 6 Welfare

The analysis above has focused on characterizing candidates' equilibrium behavior. This section analyzes the implications of such behavior for social welfare. Since citizens and truth-motivated candidates share a common objective and since elections are zero-sum for other types of candidates (at least if their preference or belief biases are symmetric), it is uncontroversial to measure welfare simply by the expected utility  $E[u(x_n^*, z)]$  of an individual citizen. This averages over the various realizations both of the state variable  $z$  and of the policy outcome  $x_n^*$ , where candidates' policy positions depend on the expected number of citizens  $n$  in combination with the equilibrium voting strategy, and the identity of the election winner depends on the realized number of citizens  $N$  and on the various realizations of each citizen's private information.

Standard private-value models highlight the benefit of convergence: centrist policies compromise between the competing interests of the left and right, thus minimizing the total disutility that voters suffer from a policy that is far from their bliss points.<sup>34</sup> In that light, the theoretical prediction that competition for office should drive candidates—who might otherwise prefer extreme policies—toward each other and toward the political center is sometimes viewed as a sort of “invisible hand” of politics. On the other hand, as Ansolabehere, Snyder, and Stewart (2001) note,

---

<sup>34</sup>This is formalized by Davis and Hinich (1968). If utility functions are tent-shaped or quadratic, for example, then total utility is maximized at the median voter's or mean voter's ideal point, respectively; generically, the utilitarian optimum lies in the interior of the policy space.

empirical evidence of polarization must then be interpreted as evidence of political failure.

In this common-value environment, Proposition 2 states that welfare is maximized by the equilibrium behavior of candidates who are Bayesian, with  $\beta = 0$ , even though these positions are highly polarized. In other words, equilibrium polarization is good for social welfare in that case, in contrast with the private-value paradigm. In the case of binary truth, in fact, policy platforms approach opposite extremes of the policy space in large elections, but the policy outcome converges exactly to  $z$ .

**Proposition 2** *There exists a strategy vector  $(x_A^*, x_B^*, \sigma^*) \in \mathcal{X}^2 \times \Sigma$  that maximizes  $E_{x,z}[u(x, z); x_A, x_B, \sigma]$ . Moreover,  $\sigma^*(x_A^*, x_B^*) = \sigma_{\tau^*}(x_A^*, x_B^*)$ , and  $x_A^*$  and  $x_B^*$  correspond to the PBE platforms for candidates who are Bayesian, with  $\beta = 0$ . If truth is binary then, for the optimal strategy vector,  $|x_w^* - z| \rightarrow_{a.s.} 0$ .*

The proof of Proposition 2 relies on an observation by McLennan (1998), that in a common-value environment, whatever is socially optimal is also individually optimal. Intuitively, the benefit of polarization is that it tailors the policy choice to the situation: if the optimal policy is on the left then citizens can elect the liberal candidate; if it is on the right they can elect the conservative. The same logic drives a similar result in the model of Bernhardt, Duggan, and Squintani (2009a), where idiosyncratic preferences are shifted by a common “shock”, so that the median voter prefers a menu of two similar alternatives, rather than one. While the intuition is similar, the implications here are much stronger, because the optimal policy may turn out to be far from the center, so the optimal level of polarization can be quite high.

A particularly stark example of this is the case of binary truth, such as the economic stimulus example introduced in Section 3: if one candidate favors a large economic stimulus while her opponent favors a small level of stimulus, competition for votes can produce moderate stimulus as a compromise outcome. However, such an outcome is known by all *not* to be optimal. In that sense, the convergence result of Theorem 4 can be viewed as a spatial version of the “pandering” behavior described in the binary settings of Canes-Wrone, Herron, and Shotts (2001) and Maskin and Tirole (2004), where candidates knowingly adopt inferior policies that appeal to voters’ mistaken opinions.

The spatial version of pandering exhibits a geometry that has empirical implications that are not apparent in binary models. Specifically, pandering pandering

takes the form of clinging to the safety of the center when bold steps in the appropriate direction would have greater social benefit. This can explain why moderate politicians are often viewed disparagingly. In the U.S., for example, moderate candidates are often derided as DINO or RINO (i.e. Democrats- or Republicans-in-name-only) for compromising on their parties' ideals in pursuit of popularity. In 2000, third-party U.S. presidential candidate Ralph Nader publicly criticized Republicans and Democrats for being "look alike parties", "Tweedledum and Tweedledee".<sup>35</sup> These concerns are not new: almost two centuries ago, Tocqueville (1835, p. 175) wrote in praise of political parties that "cling to principles rather than to their consequences". Half a century ago, the American Political Science Association (1950) issued a manifesto calling for "responsible parties" who believe that "putting a particular candidate into office is not an end in itself", and advocating to "keep parties apart" in order to "provide the electorate with a proper range of choice". Such recommendations are odd if convergence to the center is optimal for voters, as in standard preference models, but can be easily understood in the context of the information model above, as a call for truth- over office-motivation.

Understandably, pundits tend to blame polarization on the rogue candidates who adopt such polarized positions, and on voters' inability to reign them in. That voters hold moderate candidates in contempt, however, together with the fact that polarization occurs for candidates with all types of motivations, suggests that political polarization should be attributed more squarely to voters. This offers perspective for why, across states, extremists tend to outperform moderates in congressional primaries (Brady, Han, and Pope, 2007, and Hall and Snyder, 2013), extremism imposes only a minor handicap in general elections (Ansolabehere, Snyder, and Stewart, 2001, Canes-Wrone, Brady, and Cogan, 2002, and Cohen, McGrath, Aronow, and Zaller, 2016), especially in like-minded jurisdictions (Hall, 2015). Presidential elections seem no different: in the most recent example, Donald Trump and Hillary Clinton were each extreme in their own ways, but even so, many primary voters favored Ted Cruz or Bernie Sanders, who were universally viewed as even more conservative and liberal, respectively.<sup>36</sup> Evidently, voters seek not a moderate who can attract centrist

---

<sup>35</sup>See <http://www.cbsnews.com/news/nader-assails-major-parties/> (accessed 12/22/2016).

<sup>36</sup>In 2012, many Republicans explicitly reported viewing Mitt Romney as the most likely candidate to beat President Barack Obama in the general election, but "not conservative enough", and so voted for Rick Santorum, Newt Gingrich, or Ron Paul instead (see <http://elections.nytimes.com/2012/primaries/states/ohio/exit-polls>).

voters, but a bold and confident champion who can convince the electorate of her superior policy position. Such is the popular perception of Franklin Roosevelt, Barry Goldwater, George McGovern, Ronald Reagan, Barack Obama, and other prominent presidential candidates.

It is remarkable that identical behavior could have such opposite welfare implications in private- and common-value settings. This underscores the importance of identifying the right model of behavior, since behavior can be observed, but welfare implications can only be inferred from a model. With a standard model in mind, for example, policy makers might take actions to curb polarization, expecting welfare to increase accordingly. Such actions might even be effective, in that polarization decreases empirically, but if the standard paradigm is actually mistaken, then welfare might decrease instead of increase. Of course, being too moderate is not the only sin a candidate can commit: candidates are often criticized for being too extreme and uncompromising, as well. This, too, has a clear rationale within the context of the model above: with continuous truth, overconfident candidates who believe the optimal policy to be extreme adopt platform policies that are more extreme than those adopted by Bayesian candidates.<sup>37</sup> According to Proposition 2, Bayesian candidates' platforms maximize social welfare, implying that overconfident candidates' more polarized positions are overly extreme.

To a certain extent, it may be possible to increase or decrease the benefit  $\beta$  that candidates perceive from winning an election, for example by adjusting office holder salaries. However, whether this is desirable or not depends on whether excessive pandering or overconfidence is a bigger problem: raising  $\beta$  might improve welfare by mitigating over-extremism, for example, or could reduce welfare by feeding candidates' desires to pander. In any case, if candidates are overconfident or if truth is binary then the implication of Theorem 5 is that changing  $\beta$  may have a negligible effect on the equilibrium level of polarization, one way or the other.

The result that candidates polarize in spite of office motivation suggests that pandering is not of great concern in large elections. As noted above, Harrington (1993) proposes binary truth to model the broadest ideological conflict between “extensive or minimal government intervention in the economy”; in that case, a comparison of columns 6 and 7 of Table 1 suggests that over-extremism is of little concern as well,

---

<sup>37</sup>By similar logic, overconfident candidates with overly moderate ideology could be *under*-polarized. Policy motivated candidates could be over- or under-polarized, as well (i.e., if  $\hat{x}_j$  is extreme and  $\beta$  is low or if  $\hat{x}_j$  is moderate and  $\beta$  is high, respectively).

as Bayesian candidates are almost just as extreme as overconfident candidates are. However, the analysis above interprets  $z$  as the exactly optimal policy; as I explain in McMurray (2016a), another possibility is that there is *aggregate uncertainty*, meaning that  $z = E(z^*)$  is only an *approximation* of an optimal policy  $z^*$  that remains uncertain even after all private information is pooled. This is important because, as that paper explains,  $z$  may be continuous even if  $z^*$  is binary.<sup>38</sup> If so, the relevant column from Table 1 is 4 rather than 7, revealing over-polarization (relative to column 6) to be much more severe. The essential problem is that aggregate uncertainty warrants additional caution: overconfident candidates treat the world as black-and-white when, in reality, the evidence is inconclusive. In the U.S. and many other places, elections tend to be quite close in most years; this suggests that expectations  $E(z|w = A)$  and  $E(z|w = B)$  should not differ greatly from each other, and therefore that the optimal level of polarization should be rather low, implying that the substantial polarization observed empirically is indeed excessive.

## 7 Conclusion

That candidates should converge to the political center, and that this behavior is good for voters, is one of the most robust predictions in all of political economic theory, so empirical polarization is one of the deepest and most troubling paradoxes. This paper has synthesized the Downsian model with the long-overlooked paradigm of Condorcet (1785), showing that if voter disagreements reflect differences of opinion regarding an underlying truth variable, rather than fundamentally intractable conflicts of interest, then the paradox goes away: for a variety of assumptions about candidate motivations and beliefs, polarization is now the robust equilibrium prediction, as each candidate trusts voters to reward her for being on the side of truth.<sup>39</sup> In particular, polarization may be high even when candidates have strong desires to win. Convergence loses its utilitarian appeal, instead reflecting a spatial version of pandering,

---

<sup>38</sup>A moderate  $z$  then indicates that  $z^* = -1$  and  $z^* = 1$  are equally likely, while  $z$  close to  $-1$  or  $1$  indicates that  $z^* = -1$  or  $z^* = 1$  with high probability. Thus, for example, a *skeptic* who observes a high-quality private signal that  $z$  is moderate advocates moderate policies, not because he believes them to ultimately be optimal or because he lacks information generally, but because he finds the evidence on either side unconvincing.

<sup>39</sup>Beyond the immediate context of elections, a similar mechanism might explain dysfunction within polarized legislatures, as each side refuses to bargain, expecting to gain seats in a subsequent election.

consistent with the disdain that is often observed empirically. Pandering incentives all but disappear in large elections, however, leaving overextremism—rooted either in overconfidence, or in deviant preferences—as the greater threat. Overextremism is especially problematic when uncertainty is most severe, and unfortunately, need not respond to obvious policy interventions such as adjusting office holder salaries.

The information structure above constitutes almost a best-case scenario for information aggregation: voter signals are informative and conditionally independent, and voters are fully aware of their cognitive limitations, and Bayesian in their use of information. As in Condorcet’s (1785) original model, this makes large electorates nearly infallible. This is a natural starting point, but in the real world it is easy to worry that voters might suffer from biases or irrationalities that impair their collective judgment.<sup>40</sup> Ortoleva and Snowberg (2015) present evidence, for example, that voters themselves are typically overconfident. Thus, extending the information structure is an important direction for future exploration.

In binary settings, information impediments reduce the accuracy of public opinion, but not completely: quite generally, public opinion remains at least positively correlated with the truth.<sup>41</sup> If the same proves true for spatial environments then the logic above should generalize to still produce an equilibrium tendency toward polarization. In fact, voter irrationality may give a candidate even less control over her electoral success, so voters’ cognitive limitations may well produce an even greater degree of equilibrium polarization. In any case, forces similar to those highlighted above are likely to operate: an incentive to moderate in pursuit of votes, offset by an incentive to follow the perceived course of truth, in hopes that voters come around—with mixed implications for voter welfare.

Even more important than the polarization paradox, this paper sheds light on deeper questions regarding the fundamental nature of public disagreement (i.e., competing interests or competing opinions) and the basic value of democracy (i.e., preference aggregation or information aggregation). That such questions matter is underscored by the fact that the median voter theorem and median opinion theorem predict identical behavior, but with opposite welfare implications. With that in mind, the

---

<sup>40</sup>Otherwise, for example, voters should reach a consensus *ex post*, once electoral results are announced, as I explain in McMurray (2016a).

<sup>41</sup>For example, Ladha (1992, 1993), Dietrich and Spiekermann (2013), and Pivato (2016) suppose that voters make correlated errors; Triossi (2013) analyzes what happens when large elections eliminate voters’ incentive to acquire information.

model above also has value as a building block for future analysis. Building on the present model, for example, I show in McMurray (2016b) how large margins of victory can convey additional information about the state of the world, leading truth-motivated candidates to adjust their policy positions ex post, consistent with the popular notion of electoral *mandates*. In that case, voting takes on a signaling role that provides a rationale for otherwise puzzling behavior such as abstaining out of protest, or supporting minor parties who are unlikely to win the election. In McMurray (2016c) I show how the common-value model of this paper can extend to multiple dimensions, which is a well-known limitation of private-value models, and show how logical correlations across issues shape the endogenous bundling of policy positions, so that a single ideological dimension emerges.

## A Appendix

**Proof of Proposition 1.** The proof of Proposition 1 of McMurray (2016a) shows that the best response ideology threshold  $\tau^{br}(\tau)$  to an ideological strategy with ideology threshold  $\tau$  decreases with  $\tau$ , and using that fact shows if  $x_A < x_B$  then there exists a unique fixed point  $\tau^* = \tau^{br}(\tau^*)$  that characterizes an ideological strategy that is its own best response. From (9) it can be seen that, for any  $\tau \in \mathcal{X}$ ,  $\tau^{br}(\tau)$  depends on  $x_A$  and  $x_B$  only through the midpoint  $\bar{x}$ ; accordingly, the same fixed point  $\tau^*(\bar{x})$  characterizes the unique equilibrium response to any pairs of candidate platforms with the midpoint  $\bar{x}$ . (If  $x_A = x_B$  then any voting strategy—including the ideological strategy characterized by  $\tau^*(\bar{x})$ —constitutes a BNE.) From (9) it is clear that  $\tau^{br}(\tau)$  also increases in  $\bar{x}$ , for any  $\tau$ ; since  $\tau^{br}(\tau)$  decreases in  $\tau$  but increases in  $\bar{x}$  for any  $\tau$ , the fixed point  $\tau^* = \tau^{br}(\tau^*)$  increases in  $\bar{x}$ , as claimed.

For an ideological strategy, (6) can be rewritten as follows,

$$\phi(A|z; \tau) = \int_{-1}^{\tau} \int_{\mathcal{Q}} \int_{\mathcal{S}} 1_{qs=\theta} g(q) \frac{1}{2} (1 + qs) dsdq d\theta \quad (14)$$

$$\phi(B|z; \tau) = \int_{\tau}^1 \int_{\mathcal{Q}} \int_{\mathcal{S}} 1_{qs=\theta} g(q) \frac{1}{2} (1 + qs) dsdq d\theta \quad (15)$$

and symmetry can be seen by the following,

$$\begin{aligned}
\phi(A|-z; -\tau) &= \int_{-1}^{-\tau} \int_{\mathcal{Q}} \int_{\mathcal{S}} 1_{qs=\theta} g(q) \frac{1}{2} (1 - qsz) dsdq d\theta \\
&= \int_{\tau}^1 \int_{\mathcal{Q}} \int_{\mathcal{S}} 1_{qs=-\theta} g(q) \frac{1}{2} (1 + qsz) dsdq d\theta \\
&= \int_{\tau}^1 \int_{\mathcal{Q}} \int_{\mathcal{S}} 1_{qs=\theta} g(q) \frac{1}{2} (1 + qsz) dsdq d\theta = \phi(B|z; \tau)
\end{aligned}$$

where the second and third equalities follow from replacing  $\theta$  and then  $s$  with variables of opposite sign. By (7) through (10),  $\phi(A|-z; -\tau) = \phi(B|z; \tau)$  translates into symmetric pivot probabilities (i.e.  $\Pr(\mathcal{P}|-z; -\tau) = \Pr(\mathcal{P}|-z; \tau)$ ) and therefore produces symmetric expectations  $E(z|\mathcal{P}; -\tau) = -E(z|\mathcal{P}; \tau)$  and  $E(z^2|\mathcal{P}; -\tau) = E(z^2|\mathcal{P}; \tau)$ . If  $\tau^{br}(\tau^*; \bar{x}) = \tau^*$ , therefore, then from (9) it is clear that

$$\tau^{br}(-\tau^*; -\bar{x}) = \frac{-\bar{x} + E(z|\mathcal{P})}{E(z^2|\mathcal{P}) - \bar{x}E(z|\mathcal{P})} = -\tau^{br}(\tau^*; \bar{x}) = -\tau^*.$$

In other words,  $\tau^*(-\bar{x}) = -\tau^*(\bar{x})$ . ■

**Proof of Lemma 1.** For any  $\tau$ , differentiating (14) and (15) with respect to  $z$  yields the following.

$$\begin{aligned}
\frac{\partial \phi(A|z; \tau)}{\partial z} &= \int_{-1}^{\tau} \int_{\mathcal{Q}} \int_{\mathcal{S}} 1_{qs=\theta} g(q) \frac{1}{2} (qs) dsdq d\theta = E(\theta|\theta < \tau) \Pr(\theta < \tau) \\
\frac{\partial \phi(B|z; \tau)}{\partial z} &= \int_{\tau}^1 \int_{\mathcal{Q}} \int_{\mathcal{S}} 1_{qs=\theta} g(q) \frac{1}{2} (qs) dsdq d\theta = E(\theta|\theta > \tau) \Pr(\theta > \tau)
\end{aligned}$$

These must sum to  $E(\theta) = 0$ , implying that  $\phi(A|z; \tau)$  decrease in  $z$  and  $\phi(B|z; \tau)$  increases in  $z$ . The difference  $\phi(A|z; \tau) - \phi(B|z; \tau)$  therefore decreases in  $z$ , implying that  $z_{\tau} = \arg \min_z |\phi(A|z; \tau) - \phi(B|z; \tau)|$  is well-defined for any  $\tau$ . For any  $z$ , (14) and (15) also increase and decrease in  $\tau$ , respectively, implying that  $z_{\tau}$  is  $-1$  for  $\tau$  sufficiently low and is  $1$  for  $\tau$  sufficiently high, and otherwise strictly increases in  $\tau$ .

As  $n$  grows large,  $\Pr(\mathcal{P}|z)$  decreases to zero for any  $z$ , but as Myerson (2000) shows, the magnitude of  $\Pr(\mathcal{P}|z)$  is largest for  $z = z_{\tau}$ , implying that it shrinks at rate  $\frac{1}{\sqrt{n}}$  in this state and at rate  $e^{-n}$  in all others. Thus,  $f(z|\mathcal{P})$  converges to a degenerate distribution with unit mass on  $z_{\tau}$ , implying that  $E(z|\mathcal{P}) \rightarrow z_{\tau}$  and  $V(z|\mathcal{P}) \rightarrow 0$ . Note that  $z_{\tau}$  has the same sign as  $\tau$ , because  $\phi_A(0; 0) = \frac{1}{2}$  and  $\phi_A(z; \tau)$  is increasing in  $\tau$  but decreasing in  $z$ . Therefore, the right-hand side of

$$(9) \text{ converges to } \frac{-1}{z_{\tau}}, \text{ implying that } \tau^{br}(\tau, \bar{x}) \rightarrow \begin{cases} 1 & \text{if } z_{\tau} < \bar{x} \\ 0 & \text{if } z_{\tau} = \bar{x} \\ -1 & \text{if } z_{\tau} > \bar{x} \end{cases}. \text{ Let } \tau_{\bar{x}} \text{ denote the}$$

solution to  $z_\tau = \bar{x}$ , which is unique since  $z_\tau$  increases in  $\tau$ . For  $n$  large enough,  $\tau^{br}(\tau_{\bar{x}} - \varepsilon) > \tau_{\bar{x}} + \varepsilon$  and  $\tau^{br}(\tau_{\bar{x}} + \varepsilon) < \tau_{\bar{x}} - \varepsilon$ . Since  $\tau^{br}(\tau)$  decreases in  $\tau$ , this implies that  $\tau_{\bar{x}} - \varepsilon < \tau_n^* < \tau_{\bar{x}} + \varepsilon$ . In other words,  $\tau_n^*$  converges to  $\tau_{\bar{x}}$ , thereby solving  $\phi_A(z; \tau) = \phi_B(z; \tau) = \frac{1}{2}$  for  $z = \bar{x}$ . ■

**Proof of Theorem 1.** Proposition 1 implies that  $(x_A^*, x_B^*, \sigma^*) \in \mathcal{X}^2 \times \Sigma$  is a PBE only if  $\sigma^*$  is equivalent to the ideological strategy  $\sigma_{\tau^*}$  in every subgame with  $x_A \neq x_B$ . It cannot be the case in equilibrium that  $x_A^*$  is closer to  $\theta_B$  than  $x_B^*$  is, because in that case, candidate  $B$  could improve her welfare by mimicking  $A$ 's platform. It also cannot be the case in equilibrium that  $x_B^*$  is more extreme than  $\theta_B$ , because if that were so then, by moderating her position to  $\theta_B$ , candidate  $B$  could improve her odds of winning, and also her utility conditional on winning. Symmetrically,  $x_A^*$  cannot be more extreme than  $\theta_A$ . Together, these observations imply that  $\theta_A \leq x_A^* \leq x_B^* \leq \theta_B$ .

Imposing  $\beta = 0$  and differentiating (11) for candidate  $B$  with respect to her own platform yields the following.

$$\begin{aligned} \frac{\partial EU_B^{TO}}{\partial x_B} &= -2(x_B - \theta_B) \Pr(w = B|z = \theta_B) + \sum_{j=A,B} u(x_j, \theta_B) \frac{\partial}{\partial x_B} \Pr(w = j|z = \theta_B) \\ &= 2(\theta_B - x_B) \Pr(w = B|z = \theta_B) \\ &\quad + [u(x_B, \theta_B) - u(x_A, \theta_B)] \frac{\partial}{\partial x_B} \Pr(w = B|z = \theta_B) \end{aligned} \tag{16}$$

The result that  $\theta_A \leq x_A^* \leq x_B^* \leq \theta_B$  in equilibrium implies that the first term in this sum is weakly positive while the second term is weakly negative. For both terms to be zero, it must be the case that  $x_A^* = x_B^* = \theta_B$ , but this cannot occur in equilibrium because the symmetric condition for candidate  $A$  requires that  $x_A^* = x_B^* = \theta_A$ , and by assumption  $\theta_A < \theta_B$ . For the sum to be zero, therefore, the first term must be strictly positive and the second term must be strictly negative, implying (together with the symmetric conditions for candidate  $A$ ) that  $\theta_A < x_A^* < x_B^* < \theta_B$  in equilibrium.

If  $\theta_A = -\theta_B$  and  $x_A = -x_B$  then the two candidates' incentives are symmetric, implying that their best response strategies satisfy  $x_A^{br} = -x_B^{br}$ . Thus, for any  $x \in [0, 1]$ , a best response to the symmetric platform pair  $(x_A, x_B) = (-x, x)$  is another symmetric platform pair  $(x_A^{br}, x_B^{br}) = (-x^{br}, x^{br})$ . Restricting attention to symmetric platform pairs, candidate  $B$ 's expected utility is continuous in  $x$  over the compact set  $[0, 1]$ , but increases in  $x$  when  $x = 0$  and decreases in  $x$  when  $x = \theta_B$ . By the intermediate value theorem, then, there exists an intermediate  $0 < x^* < \theta_B$  such that  $(x_A^*, x_B^*) = (-x^*, x^*)$  constitutes its own best response and therefore (together with the voting strategy  $\sigma_{\tau^*}$ ) characterizes a PBE. Uniqueness follows because  $(x_A, x_B) = (-x, x)$  implies that  $\bar{x} = 0$  for any  $x$ , so that neither  $\Pr(w = A|z)$  nor  $\frac{\partial}{\partial x_A} \Pr(w = A|z = \theta_A)$  changes with  $x$ . Substituting into (17) and differentiating

with respect to  $x$  therefore yields

$$\begin{aligned} & -x \Pr(w = B|z = \theta_B) + [-2(x - \theta_B) + 2(x + \theta_B)] \frac{\partial}{\partial x_B} \Pr(w = B|z = \theta_B) \\ = & -x \Pr(w = B|z = \theta_B) + 4\theta_B \frac{\partial}{\partial x_B} \Pr(w = B|z = \theta_B), \end{aligned}$$

which is strictly negative. Thus, there exists only one pair  $(-x^*, x^*)$  satisfying  $\frac{\partial EU_B^{TO}}{\partial x_B} = 0$ .

For any  $n$ , candidate  $B$  could deviate to  $x_B = \theta_B$  and receive the following utility.

$$\begin{aligned} E[u(x, z) | z = \theta_B] &= u(x_A^*, \theta_B) \Pr(w = A|z = \theta_B; x_A = x_A^*, x_B = \theta_B) \\ &\quad + u(\theta_B, \theta_B) \Pr(w = B|z = \theta_B; x_A = x_A^*, x_B = \theta_B) \quad (18) \end{aligned}$$

Since  $\lim_{n \rightarrow \infty} \Pr(w = B|z = \theta_B; x_A = x_A^*, x_B = \theta_B) = 1$  (by Proposition 3 of McMurray, 2016a), a sequence of such deviations would yield utility  $u(\theta_B, \theta_B)$  in the limit. But for every  $n$ ,  $x_B = x_B^*$  is a best response to  $x_A^*$ , and so provides weakly greater utility than  $x_B = \theta_B$ . This implies that equilibrium utility approaches  $u(\theta_B, \theta_B)$  as well. This is possible only if  $\lim_{n \rightarrow \infty} x_B^* = \theta_B$ . By symmetric arguments,  $\lim_{n \rightarrow \infty} x_A^* = \theta_A$  as well. ■

Lemma A1 is a technical result that is used in Theorem 2.

**Lemma A1** *A candidate's win probability  $\Pr(w = j)$  increases with her expected vote share  $\phi(j)$  (and decreases with her opponent's expected vote share  $\phi(-j)$ ). If voting is ideological then, for any  $z$ ,  $\phi(A|z)$  and  $\phi(B|z)$  increase and decrease, respectively, in the ideology threshold  $\tau$ .  $E(z|w = j)$  increases in  $\tau$ , as well, and for any  $\tau$ ,  $E(z|w = A) < 0 < E(z|w = B)$ .*

**Proof.** Write the difference in win probabilities for the two candidates as

$$\begin{aligned} \Pr(w = B) - \Pr(w = A) &= \sum_{a=0}^{\infty} \Pr(N_A = a) [\Pr(N_B > a) - \Pr(N_B < a)] \\ &= \sum_{b=0}^{\infty} \Pr(N_B = b) [\Pr(N_A < b) - \Pr(N_A > b)]. \end{aligned}$$

Since the distributions of  $N_A$  and  $N_B$  are increasing in  $\phi(A)$  and  $\phi(B)$ , respectively, in the sense of first-order stochastic dominance, the first of these expressions is increasing in  $\phi(B)$  and the second is decreasing in  $\phi(A)$ . Since  $\Pr(w = A) + \Pr(w = B) = 1$ , this establishes the first claim.

That  $\phi(A|z; \tau)$  and  $\phi(B|z; \tau)$  increase and decrease in  $\tau$ , respectively, for any  $z$ ,

is clear from (14) and (15). These expressions can also be rewritten as follows,

$$\begin{aligned}\phi(A|z) &= \Pr(\theta < \tau) [1 + zE(\theta|\theta < \tau)] \\ \phi(B|z) &= \Pr(\theta > \tau) [1 + zE(\theta|\theta > \tau)]\end{aligned}$$

and from these it can be seen that, for positive  $z$ , the ratio

$$\frac{\phi(B|z)}{\phi(B|-z)} = \frac{1 + E(\theta|\theta > \tau)}{1 - E(\theta|\theta > \tau)}$$

exceeds 1 and increases in  $\tau$ , so by the first part of the lemma,  $\frac{\Pr(w=B|z)}{\Pr(w=B|-z)}$  exceeds 1 and increases in  $\tau$  as well. The latter implies that  $E(z|w=B, |z|=\bar{z}) = \frac{\bar{z}\Pr(w=B|\bar{z}) - \bar{z}\Pr(w=B|-\bar{z})}{\Pr(w=B|\bar{z}) + \Pr(w=B|-\bar{z})}$  is positive and increases in  $\tau$ . Integrating over  $\bar{z}$ ,  $E(z|w=B)$  is positive for any  $\tau$  and also increases in  $\tau$ . Symmetric arguments establish that  $E(z|w=A)$  also increases in  $\tau$  as well, but is negative for any  $\tau$ . ■

**Proof of Theorem 2.** Setting  $\beta = 0$  and differentiating (12) for candidate  $B$  with respect to her own platform yields the following (as long as  $x_A \neq x_B$ , so that, by Proposition 1, voting behavior is uniquely characterized by the ideological strategy  $\sigma_{\tau^*}$ ),

$$\begin{aligned}\frac{\partial EU_B^{TB}}{\partial x_B} &= E_z \left[ \frac{\partial u(x_B, z)}{\partial x_B} \Pr(w=B|z) \right] + E_z \left[ \sum_{j=A,B} u(x_j, z) \frac{\partial \Pr(w=j|z)}{\partial \tau^*(\bar{x})} \frac{\partial \tau^*(\bar{x})}{\partial \bar{x}} \frac{\partial \bar{x}}{\partial x_B} \right] \\ &= E_z [2(z - x_B) \Pr(w=B|z)] + \frac{\partial E[u(x, z)]}{\partial \tau^*(\bar{x})} \frac{\partial \tau^*(\bar{x})}{\partial \bar{x}} \frac{\partial \bar{x}}{\partial x_B} \\ &= 2 \Pr(w=B) [E(z|w=B) - x_B]\end{aligned}\tag{19}$$

where the final equality follows because, by Proposition 3 of McMurray (2016a), the equilibrium ideology threshold  $\tau^*$  maximizes  $E[u(x, z)]$ , implying that  $\frac{\partial E[u(x, z)]}{\partial \tau^*(\bar{x})} = 0$ .

For any voting strategy  $-1 < E(z|w=B) < 1$ , which implies that if  $x_B = -1$  then (19) is positive and  $B$  prefers to move to the right, while if  $x_B = 1$  then (19) is negative and  $B$  prefers to move to the left. Thus, a best response  $x_B^{br}$  to  $x_A$  (and to the equilibrium voting response  $\sigma^*$ ) that satisfies  $x_B^{br} \neq x_A$  requires that (19) equal zero, which is the case if and only if  $x_B^{br} = E(z|w=B)$ . Similarly, a best response  $x_A^{br} \neq x_B$  requires that  $x_A^{br} = E(z|w=A)$ . If  $x_A = x_B$  then voting need not be ideological, but non-ideological voting cannot produce higher utility. Thus,  $x_j^{br} = E(z|w=j)$  is the best response for either candidate, and  $x_j^* = E(z|w=j)$  for  $j = A, B$  is a necessary condition for a PBE. With ideological voting, Lemma A1 implies that  $E(z|w=A) < 0 < E(z|w=B)$ .<sup>42</sup>

<sup>42</sup>These inequalities could also be reversed in equilibrium, but then candidates could be relabeled so that  $A$  is to the left of  $B$ , as the statement of the theorem presumes.

For any pair of symmetric platforms  $x_A = -x_B$ , the midpoint  $\bar{x} = 0$  lies exactly at the center of the policy interval, so by Proposition 1, voters' equilibrium response is characterized by an ideological strategy with ideology threshold  $\tau^*(0) = 0$ . By the symmetry of the model, this implies that candidates form symmetric expectations  $E(z|w = A) = -E(z|w = B)$ , and therefore symmetric platforms  $x_A^* = -x_B^*$ . Together with the equilibrium voting strategy  $\sigma^*$ , these constitute a PBE, and by Theorem 2, this is the only pair of symmetric platforms that can be sustained when  $\tau = 0$ .

The limit result follows because, with symmetric platforms for all  $n$ ,  $\tau^* = 0$  for all  $n$ . The expression (6) therefore reduces to  $\phi(A|z) = \Pr(s < 0|z)$  and  $\phi(B|z) = \Pr(s > 0|z)$ . If  $z < 0$ , therefore, then  $\phi(A|z) > \frac{1}{2} > \phi(B|z)$ , so Proposition 3 of McMurray (2016a) implies that  $\lim_{n \rightarrow \infty} \Pr(w = A|z) = 1$ , while if  $z > 0$  then these inequalities are reversed, so  $\lim_{n \rightarrow \infty} \Pr(w = B|z) = 1$ . Thus,  $f(z|w = A)$  and  $f(z|w = B)$  converge to  $f(z|z < 0)$  and  $f(z|z > 0)$ , respectively, and  $E(z|w = A)$  and  $E(z|w = B)$  therefore converge to  $E(z|z < 0)$  and  $E(z|z > 0)$ . ■

**Proof of Theorem 3.** According to Proposition 1, the ideological strategy  $\sigma_{\tau^*}$  characterizes equilibrium voting behavior for all platform pairs, and does so uniquely when  $x_A \neq x_B$ . Candidates' probabilities of winning are monotonic in the ideology threshold  $\tau^*$  which, according to Proposition 1, is monotonic in the midpoint  $\bar{x} = \frac{x_A + x_B}{2}$  between the candidates, and therefore monotonic in both  $x_A$  and  $x_B$ . If candidates are policy motivated then clearly  $\hat{x}_A \leq x_A^* \leq x_B^* \leq \hat{x}_B$  in equilibrium, as otherwise one candidate could improve her expected utility by deviating either to her opponent's policy position or to her own ideal point.

Candidate  $B$ 's expected utility can be rewritten from (13) as follows,

$$EU_B^P = u(x_A, \hat{x}_B) \Pr(w = A) + u(x_B, \hat{x}_B) \Pr(w = B)$$

and differentiating with respect to her own platform  $x_B$  yields the following.

$$\begin{aligned} \frac{\partial EU_B^P}{\partial x_B} &= \frac{\partial u(x_B, \hat{x}_B)}{\partial x_B} \Pr(w = B) + [u(x_B, \hat{x}_B) - u(x_A, \hat{x}_B)] \frac{\partial \Pr(w = B)}{\partial \bar{x}} \frac{\partial \bar{x}}{\partial x_B} \\ &= -2(x_B - \hat{x}_B) \Pr(w = B) + 2(x_B - x_A) (\hat{x}_B - \bar{x}) \frac{\partial \Pr(w = B)}{\partial \bar{x}} \frac{\partial \bar{x}}{\partial x_B} \end{aligned} \quad (20)$$

If  $x_A = x_B$  then the difference in brackets is zero, so this first-order condition is satisfied only if  $x_B = \hat{x}_B$ . An analogous first-order condition for candidate  $A$  is satisfied only if  $x_A = \hat{x}_A$ , however, and  $\hat{x}_A = x_A = x_B = \hat{x}_B$  cannot be satisfied since, by assumption,  $\hat{x}_A < \hat{x}_B$ . Thus, equilibrium requires  $x_A^* < x_B^*$ . Since moving to the right of  $x_A$  cedes votes to candidate  $A$ , thus lowering  $\Pr(w = B)$ , the second term in (20) is negative. Since  $\Pr(w = B)$  is positive, the sum equals zero only if  $\frac{\partial u(x_B, \hat{x}_B)}{\partial x_B}$  is also positive, implying that  $x_B^* < \hat{x}_B$  in equilibrium, or that  $B$  is less conservative

than she would like to be. With symmetric considerations for candidate  $A$ , this implies that  $\hat{x}_A < x_A^* < x_B^* < \hat{x}_B$ , as claimed.

To see the uniqueness of symmetric equilibrium platforms, suppose that  $(x_A, x_B) = (-x, x)$  for some  $x \geq 0$ . As  $x$  changes,  $\Pr(w = B)$  and  $\frac{\partial \Pr(w=B)}{\partial \bar{x}}$  do not change, but  $\frac{\partial u(x_B=x, \hat{x}_B)}{\partial x_B}$  and  $u(x_B, \hat{x}_B) - u(x_A, \hat{x}_B)$  are both linear in  $x$ . This implies that the right-hand side of (20) is linear in  $x$  as well, and therefore equals zero for a unique value  $x^*$ , implying that  $(x_A^*, x_B^*) = (-x^*, x^*)$  constitutes the unique pair of symmetric equilibrium platforms.<sup>43</sup> According to Lemma 1,  $\tau_n^*(\bar{x})$  approaches the solution  $\tau_\infty^*(\bar{x})$  to  $z_\tau = \bar{x}$ . That is, in the limit as  $n$  grows large, if  $z = \bar{x}$  then the expected vote shares  $\phi[A|z = \bar{x}; \tau_\infty^*(\bar{x})] = \phi[B|z = \bar{x}; \tau_\infty^*(\bar{x})]$  for the two candidates will be exactly the same; if  $z$  turns out to be less than  $\bar{x}$  then  $\phi[A|z; \tau_\infty^*(\bar{x})] < \phi[B|z; \tau_\infty^*(\bar{x})]$  and if  $z$  turns out to exceed  $\bar{x}$  then  $\phi[A|z; \tau_\infty^*(\bar{x})] > \phi[B|z; \tau_\infty^*(\bar{x})]$ . In a large election, the candidate with the larger expected vote share almost surely wins (see Myerson, 2002).

With continuous truth, the result that in large elections  $A$  wins almost surely if  $z < \bar{x}$  and  $B$  wins almost surely if  $z > \bar{x}$  implies that  $\lim_{n \rightarrow \infty} \Pr(w = B) = 1 - F(\bar{x}) = \frac{1-\bar{x}}{2}$  and therefore that  $\lim_{n \rightarrow \infty} \frac{\partial \Pr(w=B)}{\partial \bar{x}} = \frac{\partial}{\partial \bar{x}} \left( \frac{1-\bar{x}}{2} \right) = -\frac{1}{2}$ .<sup>44</sup> In a platform-symmetric equilibrium  $x_A = -x_B$  and  $\bar{x} = 0$ , so (20) therefore converges to the following,

$$\lim_{n \rightarrow \infty} \frac{\partial E[u(x, \hat{x}_B)]}{\partial x_B} = -x_B + \hat{x}_B - x_B \hat{x}_B \quad (21)$$

and the limit  $x_B^*$  of a sequence of solutions  $x_{B,n}^*$  to  $\frac{\partial E[u(x, \hat{x}_B)]}{\partial x_B} = 0$  must satisfy  $\lim_{n \rightarrow \infty} \frac{\partial E[u(x, \hat{x}_B)]}{\partial x_B} = 0$ , implying that  $x_B^* = \frac{\hat{x}_B}{1+\hat{x}_B}$ . An analogous derivation for  $A$  yields  $x_A^* = \frac{\hat{x}_A}{1-\hat{x}_A}$ .

With binary truth, the result that in large elections  $A$  wins almost surely if  $z < \bar{x}$  and  $B$  wins almost surely if  $z > \bar{x}$  implies that  $\lim_{n \rightarrow \infty} \Pr(w = B) = \Pr(z > \bar{x}) = \frac{1}{2}$  for any  $\bar{x}$ , and therefore that  $\lim_{n \rightarrow \infty} \frac{\partial \Pr(w=B)}{\partial \bar{x}} = 0$ . In that case,  $\lim_{n \rightarrow \infty} \frac{\partial E[u(x, \hat{x}_B)]}{\partial x_B} = 0$  if and only if  $\frac{\partial u(x_B, \hat{x}_B)}{\partial x_B} = 0$ , or  $x_B^* = \hat{x}_B$ . By an analogous derivation,  $x_A^* = \hat{x}_A$ . ■

**Proof of Theorem 4.** According to Proposition 1, the ideological strategy  $\sigma_{\tau^*}$  characterizes equilibrium voting behavior for all platform pairs, and characterizes the unique equilibrium voting behavior for distinct pairs  $x_A \neq x_B$ . Therefore, the

<sup>43</sup>Alternatively, if the right-hand side of (20) is positive for all  $x$  then  $(x_A^*, x_B^*) = (-1, 1)$  constitutes the unique pair of symmetric equilibrium platforms.

<sup>44</sup>The latter result also relies on the implicit function theorem, as  $\tau_n^{br}(\tau; \bar{x})$  is continuously differentiable in  $\bar{x}$  and  $n$ , so the solution  $\tau_n^*(\bar{x})$  to the fixed point problem  $\tau = \tau_n^{br}(\tau; \bar{x})$  is continuously differentiable in  $\bar{x}$  and  $n$  as well, and therefore so is  $\Pr[w = B; \tau_n^{b*}(\bar{x})]$ .

derivative of (12) generalizes from (17) to

$$\begin{aligned} \frac{\partial EU_B^{TO}}{\partial x_B} &= 2(\theta_B - x_B) \Pr(w = B|z = \theta_B) \\ &\quad + [u(x_B, \theta_B) - u(x_A, \theta_B) + \beta] \frac{\partial \Pr(w = B|z = \theta_B)}{\partial \tau^*} \frac{\partial \tau^*(\bar{x})}{\partial \bar{x}} \frac{\partial \bar{x}}{\partial x_B} \end{aligned} \quad (22)$$

if candidates are overconfident, generalizes from (19) to

$$\frac{\partial EU_B^{TB}}{\partial x_B} = 2 \Pr(w = B) [E(z|w = B) - x_B] + \beta \frac{\partial \Pr(w = B)}{\partial \tau^*} \frac{\partial \tau^*(\bar{x})}{\partial \bar{x}} \frac{\partial \bar{x}}{\partial x_B} \quad (23)$$

if candidates are Bayesian, and generalizes from (20) to

$$\begin{aligned} \frac{\partial EU_B^P}{\partial x_B} &= 2 \Pr(w = B) (\hat{x}_B - x_B) \\ &\quad + [\beta + u(x_B, \hat{x}_B) - u(x_A, \hat{x}_B)] \frac{\partial \Pr(w = B)}{\partial \tau^*} \frac{\partial \tau^*(\bar{x})}{\partial \bar{x}} \frac{\partial \bar{x}}{\partial x_B} \end{aligned} \quad (24)$$

if candidates are policy motivated. In all three cases, this derivative decreases in  $\beta$ , because  $\frac{\partial \Pr(w=B|z=\theta_B)}{\partial \tau^*}$  and  $\frac{\partial \Pr(w=B)}{\partial \tau^*}$  are negative while  $\frac{\partial \tau^*(\bar{x})}{\partial \bar{x}}$  and  $\frac{\partial \bar{x}}{\partial x_B}$  are positive. For any pair  $(x_A, x_B) \in \mathcal{X}^2$  of platforms, the other terms in the expression for the derivative are finite, so there exists a threshold  $\bar{\beta}_{x_A, x_B}$  sufficiently large that for all  $\beta > \bar{\beta}_{x_A, x_B}$  the derivative is negative. The set of platform pairs is compact and  $\bar{\beta}_{x_A, x_B}$  is continuous in the platform pair, so there exists a maximum  $\bar{\beta} = \max_{(x_A, x_B)} \bar{\beta}_{x_A, x_B}$ , and for any  $\beta > \bar{\beta}$  the derivative is negative, meaning that candidate  $B$  does not want to move  $x_B$  away from  $x_A$ . Symmetrically,  $A$  does not want to move  $x_A$  away from  $x_B$ . Thus, if  $\beta > \bar{\beta}$  then there is no PBE with distinct platforms  $x_A < x_B$ .

While neither candidate wishes to move away from her opponent, a candidate might have incentive to “leap frog” her opponent, to attract more votes: if the two candidates converge to a position in which  $B$  wins with probability lower than one-half then  $B$  can move  $x_B$  just to the left of  $x_A$  and win with greater than  $\frac{1}{2}$  probability instead. For the cases of overconfident or policy motivated candidates, this implies that the unique PBE is  $x_A^* = x_B^* = 0$  (together with the voting strategy  $\sigma^* = \sigma_{\tau^*}$ ). For the case of overconfident candidates, there is a range of platforms  $x$  in which  $x_A^* = x_B^* = x$  can be sustained in equilibrium (including  $x_A^* = x_B^* = 0$ ), because each candidate believes that she is already on the side that will win with probability exceeding  $\frac{1}{2}$ .

For any symmetric platform pair  $(x_A, x_B) = (-x, x)$ , the midpoint is  $\bar{x} = 0$  and the voter response threshold is  $\tau^*(\bar{x}) = 0$ , regardless of the magnitude of  $x$ , implying that  $\Pr(w = B)$  and  $\frac{\partial \Pr(w=B)}{\partial \tau^*} \frac{\partial \tau^*}{\partial \bar{x}} \frac{\partial \bar{x}}{\partial x_B}$  do not depend on the magnitude of  $x$ . The utility differences  $u(x, \theta_B) - u(-x, \theta_B) = 4\theta_B x$  and  $u(x, \hat{x}_B) - u(-x, \hat{x}_B) = 4\hat{x}_B x$

are linear in  $x$ , implying that (22) through (24) are linear in  $x$ . For each of these motivations, therefore, there is a unique  $x^* \in [0, 1]$  such that the best responses for  $A$  and  $B$ , respectively, to any pair  $(-x, x)$  of symmetric platforms are  $x_A^* = -x^*$  and  $x_B^* = x^*$ . Thus,  $(x_A^*, x_B^*) = (-x^*, x^*)$  (together with  $\sigma^* = \sigma_{\tau^*}$ ) constitute the unique PBE with symmetric platforms. Since  $\frac{\partial \Pr(w=B|z)}{\partial \tau}$  is negative and  $\frac{\partial \tau(\bar{x})}{\partial \bar{x}}$  and  $\frac{\partial \bar{x}}{\partial x_B}$  are positive, (22) through (24) are all decreasing in  $\beta$ . If  $\beta < \bar{\beta}$ , therefore, then, as  $\beta$  increases, the platform  $(-x^*, x^*)$  that previously constituted an equilibrium now produces a negative  $\frac{\partial E[u(x,z)]}{\partial x_B}$  (and, symmetrically, a positive  $\frac{\partial E[u(x,z)]}{\partial x_A}$ ), implying that the new equilibrium platform pair has a lower value of  $x^*$ . ■

**Proof of Theorem 5.** If candidates are overconfident then, for any  $n$ , the utility (18) derived by deviating to  $x_B = \theta_B$  in response to candidate  $A$ 's equilibrium platform  $x_{A,n}^*$  generalizes to include an additional term.

$$\begin{aligned} \frac{\partial EU_B^{TO}}{\partial x_B} &= u(x_{A,n}^*, \theta_B) \Pr(w = A | z = \theta_B; x_A = x_{A,n}^*, x_B = \theta_B) \\ &\quad + [u(\theta_B, \theta_B) + \beta] \Pr(w = B | z = \theta_B; x_A = x_{A,n}^*, x_B = \theta_B) \end{aligned}$$

Since  $\lim_{n \rightarrow \infty} \Pr(w = B | z = \theta_B; x_A = x_{A,n}^*, x_B = \theta_B) = 1$  by Proposition 3 of McMurray (2016a), a sequence of such deviations yields expected utility  $u(\theta_B, \theta_B) + \beta$  in the limit. This is the maximum utility possible, but  $B$ 's equilibrium policy position is a best response to  $x_A^*$ , and so must generate utility that is at least as high, thus requiring  $\lim_{n \rightarrow \infty} x_{B,n}^* = \theta_B$ . This result is independent of  $\beta$ , and holds whether truth is continuous or binary.

If candidates are Bayesian then the equilibrium condition (19) generalizes to include an additional term.

$$\frac{\partial EU_B^{TB}}{\partial x_B} = 2 \Pr(w = B) [E(z | w = B) - x_B] + \beta \frac{\partial \Pr(w = B)}{\partial \bar{x}} \frac{\partial \bar{x}}{\partial x_B}$$

Symmetric platforms imply that  $\tau^*(\bar{x}) = 0$  and therefore that  $\Pr(w = B) = \frac{1}{2}$ , and, as the proof of Theorem 2 shows, that  $E(z | w = B)$  approaches  $E(z | z > 0)$  as  $n$  grows large. If truth is continuous then the proof of Theorem 3 shows that  $\frac{\partial \Pr(w=B)}{\partial \bar{x}}$  approaches  $-\frac{1}{2}$  as  $n$  grows large, so  $\frac{\partial E[u(x,z)]}{\partial x_B}$  approaches  $E(z | z > 0) - x_B^* - \frac{1}{4}\beta$ , which is zero if and only if  $x_{B,n}^*$  approaches  $x_B^* = E(z | z > 0) - \frac{1}{4}\beta$ . If truth is binary then  $\frac{\partial \Pr(w=B)}{\partial \bar{x}}$  approaches zero instead, so  $\frac{\partial EU_B^{TB}}{\partial x_B}$  approaches  $E(z | z > 0) - x_B^*$ , which is zero if and only if  $x_{B,n}^*$  approaches  $x_B^* = E(z | z > 0)$ , regardless of  $\beta$ .

If candidates are policy motivated then the derivative of expected utility generalizes from (20) to include an extra term.

$$\frac{\partial EU_B^P}{\partial x_B} = -2(x_B - \hat{x}_B) \Pr(w = B) + [2(x_B - x_A)(\hat{x}_B - \bar{x}) + \beta] \frac{\partial \Pr(w = B)}{\partial \bar{x}} \frac{\partial \bar{x}}{\partial x_B}$$

If truth is binary then  $\frac{\partial \Pr(w=B)}{\partial \bar{x}}$  converges to 0, so maintaining that  $\frac{\partial EU_B^P}{\partial x_B} = 0$  requires that  $x_{B,n}^*$  approach  $\hat{x}_B$ . A similar derivation for candidate  $A$  implies that  $x_{A,n}^*$  approaches  $\hat{x}_A$ . For any sequence of platform-symmetric equilibria,  $\bar{x} = 0$  and  $\tau^*(\bar{x}) = 0$ , so  $\Pr(w = B) = \frac{1}{2}$ . If truth is continuous then  $\frac{\partial \Pr(w=B)}{\partial \bar{x}}$  then converges in large elections to  $-\frac{1}{2}$ , as above, so  $\frac{\partial EU_B^P}{\partial x_B}$  approaches  $-x_B + \hat{x}_B - x_B \hat{x}_B - \frac{1}{4}\beta$ , which is zero if and only if  $x_{B,n}^*$  approaches  $x_B^* = \frac{\hat{x}_B - \frac{1}{4}\beta}{1 + \hat{x}_B}$ . ■

**Proof of Proposition 2.** Drawing on the common-value logic of McLennan (1998), Proposition 3 of McMurray (2016a) states that, for any  $n$ , an optimal response  $v_n^*$  by voters to any pair  $(x_A, x_B) \in \mathcal{X}^2$  of policy platforms exists and constitutes a BNE in the voting subgame. By Proposition 1 of this paper, therefore,  $v_n^*$  is given by the ideological strategy  $\sigma_{\tau_n^*}(x_A, x_B)$ , evaluated at the platform pair. The optimal combination of voter and candidate behavior can then be obtained by maximizing over the set  $\mathcal{X}^2$  of platform pairs. Since this set is compact and expected utility is continuous in both platforms, an optimal platform pair  $(x_{A,n}^*, x_{B,n}^*) \in \mathcal{X}^2$  exists by the extreme value theorem. Together with any voting strategy  $\sigma_n^*$  that implements  $\sigma_{\tau_n^*}(x_A^*, x_B^*)$  in the appropriate subgame, this constitutes an optimal strategy vector. For the policy platform pair  $(x_{A,n}^*, x_{B,n}^*)$  to maximize expected utility, given the voting strategy  $\sigma_n^*$ , however,  $x_{A,n}^*$  must maximize expected utility given  $x_{B,n}^*$  and  $\sigma_n^*$ , and  $x_{B,n}^*$  must maximize expected utility given  $x_{A,n}^*$  and  $\sigma_n^*$ . In other words,  $x_{A,n}^*$  and  $x_{B,n}^*$  must be equilibrium platforms in a game with candidates who are Bayesian, for  $\beta = 0$ , as claimed.

For any  $n$ , let  $(x_A, x_B) = (-1, 1)$  and let citizens follow the ideological strategy with ideology threshold  $\tau = 0$ . In that case, if truth is binary, then, by Proposition 3 of McMurray (2016a),  $\Pr(w = A|z = -1)$  and  $\Pr(w = B|z = 1)$  both tend to one as  $n$  grows large, so expected utility approaches  $\frac{1}{2}u(-1, -1) + \frac{1}{2}u(1, 1) = 0$ . The optimal strategy vector provides weakly greater utility than this, implying that  $x_A$  and  $x_B$  converge to  $-1$  and  $1$  in that case, as well. Since the superior of these wins with probability approaching one, the winning policy  $x_{w,n}$  converges almost surely to  $z$ . ■

## References

- [1] Acharya, Avidit and Adam Meierowitz. 2016. “Sincere Voting in Large Elections.” *Games and Economic Behavior*, forthcoming.
- [2] Ahn, David S., and Santiago Oliveros. 2016. “Approval Voting and Scoring Rules with Common Values.” Manuscript, UC Berkeley and University of Essex.
- [3] Alesina, Alberto. 1988. “Credibility and Policy Convergence in a Two-Party System with Rational Voters.” *American Economic Review*, 78(4): 796-805.
- [4] Alesina, Alberto and George-Marios Angeletos. 2005. “Fairness and Redistribution”. *American Economic Review*, 95(4): 960-980.

- [5] Alesina, Alberto and Howard Rosenthal. 2000. "Polarized Platforms and Moderate Policies with Checks and Balances." *Journal of Public Economics*, 75: 1-20.
- [6] American Political Science Association. 1950. "Toward a More Responsible Two-Party System: A Report of the Committee on Political Parties." *American Political Science Review*, 44(3, Part 2 supplement).
- [7] Ansolabehere, Stephen, and James M. Snyder, Jr. 2000. "Valence Politics and Equilibrium in Spatial Election Models." *Public Choice*, 103: 327-336.
- [8] Ansolabehere, Stephen, James M. Snyder, Jr. and Charles Stewart, III. 2001. "Candidate Positioning in U.S. House Elections." *American Journal of Political Science*, 45(1): 136-159.
- [9] Aragonés, Enriqueta and Thomas R. Palfrey. 2004. "Mixed Equilibrium in a Downsian Model with a Favored Candidate." *Journal of Economic Theory*, 103: 131-161.
- [10] Asako, Yasushi. 2014. "Campaign Promises as an Imperfect Signal: How Does an Extreme Candidate Win Against a Moderate Candidate?" *Journal of Theoretical Politics*, 27(4): 613-649.
- [11] Ashworth, Scott and Ethan Bueno de Mesquita. 2009. "Elections with Platform and Valence Competition." *Games and Economic Behavior*, 67: 191-216.
- [12] Austen-Smith, David and Jeffrey S. Banks. 1996. "Information Aggregation, Rationality, and the Condorcet Jury Theorem." *The American Political Science Review*, 90(1): 34-45.
- [13] Bafumi, Joseph, and Michael C. Herron. 2010. "Leapfrog Representation and Extremism: A Study of American Voters and Their Members in Congress," *American Political Science Review*, 104(3): 519-542.
- [14] Banks, Jeffrey S. and John Duggan. 2005. "Probabilistic Voting in the Spatial Model of Elections: The Theory of Office-motivated Candidates," in *Social Choice and Strategic Decisions: Essays in Honor of Jeffrey S. Banks*, D. Austen-Smith and J. Duggan, eds., New York: Springer-Verlag.
- [15] Bernhardt, Dan, John Duggan, and Francesco Squintani. 2009a. "The Case for Responsible Parties." *American Political Science Review*, 103(4): 570-587.
- [16] Bernhardt, Dan, John Duggan, and Francesco Squintani. 2009b. "Private Polling in Elections and Voter Welfare." *Journal of Economic Theory* 144: 2021-2056.
- [17] Besley, Timothy, and Stephen Coate. 1997. "An Economic Model of Representative Democracy", *Quarterly Journal of Economics*, 112(1): 85-114.
- [18] Bhattacharya, Sourav. 2013. "Preference Monotonicity and Information Aggregation in Elections." *Econometrica*, 81(3): 1229-1247.
- [19] Brady, David W., Hahrie Han, and Jeremy C. Pope. 2007. "Primary Elections and Candidate Ideology: Out of Step with the Primary Electorate?" *Legislative Studies Quarterly*, 32(1): 79-105.
- [20] Brusco, Sandro, and Jaideep Roy. 2011. "Aggregate Uncertainty in the Citizen Candidate Model Yields Extremist Parties." *Social Choice and Welfare*, 36: 83-104.
- [21] Buera, Francisco J., Alexander Monge-Naranjo, and Giorgio E. Primiceri. 2011. "Learning the Wealth of Nations". *Econometrica*, 79(1): 1-45.

- [22] Callander, Steven. 2008. "Political Motivations." *Review of Economic Studies*, 64: 671-697.
- [23] Callander, Steven, and Simon Wilkie. "Lies, Damned Lies, and Political Campaigns." *Games and Economic Behavior*, 60: 262-286.
- [24] Callander, Steven, and Catherine H. Wilson. 2007. "Turnout, Polarization, and Duverger's Law." *The Journal of Politics*, 69(4): 1045-1056.
- [25] Calvert, R. 1985. "Robustness of the Multidimensional Voting Model: Candidate Motivations, Uncertainty, and Convergence." *American Journal of Political Science*, 29: 69-95.
- [26] Canes-Wrone, Brandice, David W. Brady, and John F. Cogan. 2002. "Out of Step, Out of Office: Electoral Accountability and House Members' Voting." *American Political Science Review*, 96(1): 2002.
- [27] Canes-Wrone, Brandice, Michael C. Herron, and Kenneth W. Shotts. 2001. "Leadership and Pandering: A Theory of Executive Policymaking." *American Journal of Political Science*, 45(3): 532-550.
- [28] Castanheira, Micael. 2003. "Why Vote for Losers?" *Journal of the European Economic Association*, 1(5): 1207-1238.
- [29] Chambers, Christopher P., and Paul J. Healy. 2012. "Updating Toward the Signal." *Economic Theory*, 50:765-786.
- [30] Coleman, James S. 1971. "Internal Processes Governing Party Positions in Elections." *Public Choice* 11: 35-60.
- [31] Condorcet, Marquis de. 1785. *Essay on the Application of Analysis to the Probability of Majority Decisions*. Paris: De l'imprimerie royale. Trans. Iain McLean and Fiona Hewitt. 1994.
- [32] Coughlin, Peter J. and Shmuel Nitzan. 1981. "Electoral Outcomes with Probabilistic Voting and Nash Social Welfare Maxima." *Journal of Public Economics*, 15: 113-122.
- [33] Crawford, Vincent P. and Joel Sobel. 1982. "Strategic Information Transmission." *Econometrica*, 50(6): 1431-1451.
- [34] Davis, Otto A. and Melvin J. Hinich. 1968. "On the Power and Importance of the Mean Preference in a Mathematical Model of Democratic Choice." *Public Choice*, 5: 59-72.
- [35] Davis, Otto A., Melvin J. Hinich, and Peter C. Ordeshook. 1970. "An Expository Development of a Mathematical Model of the Electoral Process." *American Political Science Review*, 64(2): 426-448.
- [36] de Fenetti, Bruno. 1990. *Theory of Probability*, vol. 2, New York: Wiley.
- [37] Dietrich, Franz and Kai Spiekermann. "Epistemic Democracy with Defensible Premises." *Economics and Philosophy*, 29: 87-120.
- [38] Downs, Anthony. 1957. *An Economic Theory of Democracy*. New York: Harper and Row.
- [39] Duggan, John. 2000. "Equilibrium Equivalence Under Expected Plurality and Probability of Winning Maximization." Manuscript, University of Rochester.

- [40] Duggan, John. 2006. "Candidate Objectives and Electoral Equilibrium," in *The Oxford Handbook of Political Economy*, eds., Barry R. Weingast and Donald A. Wittman, Oxford: Oxford University Press.
- [41] Duggan, John. 2013. "A Survey of Equilibrium Analysis in Spatial Models of Elections." Manuscript, University of Rochester.
- [42] Eguia, Jon X. 2007. "Citizen Candidates Under Uncertainty." *Social Choice and Welfare*, 29: 317-331.
- [43] Enelow, James M. and Melvin J. Hinich. 1989. "A General Probabilistic Spatial Theory of Elections." *Public Choice*, 61(2): 101-113.
- [44] Eyster, Erik and Thomas Kittsteiner. 2007. "Party Platforms in Electoral Competition with Heterogeneous Constituencies." *Theoretical Economics*, 2: 41-70.
- [45] Feddersen, Timothy J. and Wolfgang Pesendorfer. 1996. "The Swing Voter's Curse." *American Economic Review*, 86(3): 408-424.
- [46] Feddersen, Timothy J. and Wolfgang Pesendorfer. 1997. "Voting Behavior and Information Aggregation in Elections with Private Information." *Econometrica*, 65(5): 1029-1058.
- [47] Feddersen, Timothy J. and Wolfgang Pesendorfer. 1998. "Convicting the Innocent: The Inferiority of Unanimous Jury Verdicts under Strategic Voting." *The American Political Science Review*, 92(1): 23-35.
- [48] Feddersen, Timothy J. and Wolfgang Pesendorfer. 1999. "Abstention in Elections with Asymmetric Information and Diverse Preferences." *The American Political Science Review*, 93(2): 381-398.
- [49] Fowler, Anthony and Andrew B. Hall. 2016. "The Elusive Quest for Convergence." *Quarterly Journal of Political Science*, 11:131-149.
- [50] Glaeser, Edward L., Giacomo A. M. Ponzetto, and Jesse M. Shapiro. 2005. "Strategic Extremism: Why Republicans and Democrats Divide on Religious Values." *The Quarterly Journal of Economics*, 120(4): 1283-1330.
- [51] Glaeser, Edward L. and Cass R. Sunstein. 2009. "Extremism and Social Learning." *Journal of Legal Analysis*, 1(1): 263-324.
- [52] Goren, Paul. 1997. "Political Expertise and Issue Voting in Presidential Elections." *Political Research Quarterly*, 50: 387-412.
- [53] Gratton, Gabriele. "Pandering and Electoral Competition." *Games and Economic Behavior* 84: 163-179.
- [54] Groseclose, Tim. 2001. "A Model of Candidate Location When One Candidate Has a Valence Advantage." *American Journal of Political Science*, 45(4): 862-886.
- [55] Gul, Faruk, and Wolfgang Pesendorfer. 2009. "Partisan Politics and Election Failure with Ignorant Voters." *Journal of Economic Theory*, 144: 146-174.
- [56] Hall, Andrew B. 2015. "What Happens When Extremists Win Primaries?" *American Political Science Review*, 109(1): 18-42.
- [57] Hall, Andrew B. and James M. Snyder, Jr. 2015. "Candidate Ideology and Electoral Success." Manuscript, Stanford University.
- [58] Hansson, Ingemar and Charles Stuart. 1984. "Voting Competitions with Interested Politicians: Platforms Do Not Converge to the Preferences of the Median Voter." *Public Choice*, 44: 431-441.

- [59] Harrington, Joseph E., Jr. 1993. "Economic Policy, Economic Performance, and Elections." *The American Economic Review*, 83(1): 27-42.
- [60] Heidhues, Paul and Johan Lagerlöf. 2003. "Hiding Information in Electoral Competition." *Games and Economic Behavior*, 42: 48-74.
- [61] Hinich, Melvin J. 1977. "Equilibrium in Spatial Voting: The Median Voter Result is an Artifact", *Journal of Economic Theory*, 16(2): 208-219.
- [62] Hinich, Melvin J. 1978. "The Mean Versus the Median in Spatial Voting Games," in P. Ordeshook, ed., *Game Theory and Political Science*, New York: NYU Press.
- [63] Hirano, Shigeo, James M. Snyder, Jr., Stephen Ansolabehere, and John Mark Hansen. 2010. "Primary Elections and Partisan Polarization in the U.S. Congress." *Quarterly Journal of Political Science*, 5: 169-191.
- [64] Hotelling, Harold. 1929. "Stability in Competition." *Economic Journal*, 39(153): 41-57.
- [65] Jessee, Stephen A. 2009. "Spatial Voting in the 2004 Presidential Election." *American Political Science Review*, 103(1): 59-81.
- [66] Jessee, Stephen A. 2010. "Partisan Bias, Political Information and Spatial Voting in the 2008 Presidential Election." *Journal of Politics*, 72(2): 327-340.
- [67] Jessee, Stephen A. 2016. "(How) Can We Estimate the Ideology of Citizens and Political Elites on the Same Scale?" *American Journal of Political Science*, forthcoming.
- [68] Kamada, Yuichiro and Fuhito Kojima. 2014. "Voter Preferences, Polarization, and Electoral Policies." *American Economic Journal: Microeconomics*, 6(4): 203-236.
- [69] Kartik, Navin, and R. Preston McAfee. 2007. "Signaling Character in Electoral Competition." *American Economic Review*, 97(3): 852-870.
- [70] Kartik, Navin, Francesco Squintani, and Katrin Tinn. 2013. "Information Revelation and (Anti-) Pandering in Elections". Manuscript, Columbia University.
- [71] Kim, Jaehoon and Mark Fey. 2007. "The Swing Voter's Curse with Adversarial Preferences." *Journal of Economic Theory*, 135: 236-252.
- [72] Krasa, Stefan, and Mattias Polborn. 2010. "Competition Between Specialized Candidates." *American Political Science Review*. 104(4): 745-765.
- [73] Krasa, Stefan, and Mattias Polborn. 2012. "Political Competition Between Differentiated Candidates." *Games and Economic Behavior*, 76: 249-271.
- [74] Krasa, Stefan, and Mattias Polborn. 2015. "Political Competition in Legislative Elections." Manuscript, University of Illinois.
- [75] Krishna, Vijay, and John Morgan. 2011. "Overcoming Ideological Bias in Elections". *Journal of Political Economy*, 119(2): 183-211.
- [76] Krishna, Vijay, and John Morgan. 2012. "On the Benefits of Costly Voting." *Journal of Economic Theory*, forthcoming.
- [77] Ladha, Krishna K. 1992. "The Condorcet Jury Theorem, Free Speech, and Correlated Votes," *American Journal of Political Science*, 36(3): 617-634.
- [78] Ladha, Krishna K. 1993. "Condorcet's Jury Theorem in Light of de Finetti's Theorem: Majority-Rule Voting with Correlated Votes." *Social Choice and Welfare*, 10(1): 69-85.

- [79] Ladha, Krishna K. 1995. "Information Pooling through Majority-Rule Voting: Condorcet's Jury Theorem with Correlated Votes." *Journal of Economic Behavior and Organization*, 26: 353-372.
- [80] Laslier, Jean-François and Karine Van der Straeten. 2004. "Electoral Competition Under Imperfect Information." *Economic Theory*, 24: 419-446.
- [81] Levy, Gilat and Ronny Razin. 2015. "Does Polarisation of Opinions Lead to Polarisation of Platforms? The Case of Correlation Neglect." *Quarterly Journal of Political Science*, 10: 321-355.
- [82] Lindbeck, Assar and Jörgen W. Weibull. 1987. "Balanced-Budget Redistribution as the Outcome of Political Competition." *Public Choice*, 52(3): 273-297.
- [83] List, Christian and Robert E. Goodin. 2001. "Epistemic Democracy: Generalizing the Condorcet Jury Theorem." *The Journal of Political Philosophy*, 9(3): 277-306.
- [84] Loertscher, Simon. "Location Choice and Information Transmission." Working paper, University of Melbourne.
- [85] Martinelli, Cesar. 2001. "Elections with Privately Informed Parties and Voters." *Public Choice*, 108: 147-167.
- [86] Maskin, Eric and Jean Tirole. 2004. "The Politician and the Judge: Accountability in Government". *American Economic Review*, 94(4): 1034-1054.
- [87] Matějka, Filip and Guido Tabellini. 2015. "Electoral Competition with Rationally Inattentive Voters." CERGE-EI Working Paper Series # 552.
- [88] McCarty, Nolan M. and Keith T. Poole. 1995. "Veto Power and Legislation: An Empirical Analysis of Executive and Legislative Bargaining from 1961 to 1986." *Journal of Law, Economics, and Organization*, 11(2): 282-312.
- [89] McGhee, Eric, Seth Masket, Boris Shor, Steven Rogers, and Nolan McCarty. 2014. "A Primary Cause of Partisanship? Nomination Systems and Legislator Ideology." *American Journal of Political Science*, 58(2): 337-351.
- [90] McLennan, Andrew. 1998. "Consequences of the Condorcet Jury theorem for Beneficial Information Aggregation by Rational Agents." *American Political Science Review*, 92(2): 413-418.
- [91] McMurray, Joseph C. 2013. "Aggregating Information by Voting: The Wisdom of the Experts versus the Wisdom of the Masses." *The Review of Economic Studies*, 80(1): 277-312.
- [92] McMurray, Joseph C. 2016a. "Ideology as Opinion: A Spatial Model of Common-value Elections." *American Economic Journal: Microeconomics*, forthcoming.
- [93] McMurray, Joseph C. 2016b. "Signaling in Elections: Mandates, Minor Parties, and the Signaling Voter's Curse." *Games and Economic Behavior*, forthcoming.
- [94] McMurray, Joseph C. 2016c. "Why the Political World is Flat: An Endogenous Left-Right Spectrum in Multidimensional Political Conflict." Manuscript, Brigham Young University.
- [95] Milgrom, Paul R. and Robert J. Weber. 1982. "A Theory of Auctions and Competitive Bidding." *Econometrica*, 50(5): 1089-1122.
- [96] Myerson, Roger. 1998. "Population Uncertainty and Poisson Games." *International Journal of Game Theory*, 27: 375-392.

- [97] Myerson, Roger. 2000. "Large Poisson Games." *Journal of Economic Theory*, 94: 7-45.
- [98] Myerson, Roger. 2002. "Comparison of Scoring Rules in Poisson Voting Games." *Journal of Economic Theory*, 103: 219-251.
- [99] Myerson, Roger B. and Robert J. Weber. 1993. "A Theory of Voting Equilibria." *American Political Science Review*, 87(1): 102-114.
- [100] Nitzan, Shmuel and Jacob Paroush. 1982. "Optimal Decision Rules in Uncertain Dichotomous Choice Situations." *International Economic Review*, 23(2): 289-297.
- [101] Ortoleva, Pietro and Erik Snowberg. 2015. "Overconfidence in Political Behavior." *American Economic Review*, 105(2): 504-535.
- [102] Ortuño-Ortín, Ignacio. "A Spatial Model of Political Competition and Proportional Representation." *Social Choice and Welfare*, 14: 427-438.
- [103] Osborne, Martin J., and Al Slivinski. 1996. "A Model of Political Competition with Citizen-Candidates." *Quarterly Journal of Economics*, 111(1): 65-96.
- [104] Palfrey, Thomas R. 1984. "Spatial Equilibrium with Entry." *The Review of Economic Studies*, 51(1): 139-156.
- [105] Palfrey, Thomas R. and Keith T. Poole. 1987. "The Relationship between Information, Ideology, and Voting Behavior." *American Journal of Political Science*, 31(3): 511-530.
- [106] Pivato, Marcus. 2016. "Epistemic Democracy with Correlated Voters." Manuscript, Université de Cergy-Pontoise.
- [107] Polborn, Mattias and James M. Snyder. 2016. "Intra-Party Disagreement and Inter-Party Polarization." Manuscript, University of Illinois and Harvard University.
- [108] Poole, Keith T. and Howard Rosenthal. 1984. "U.S. Presidential Elections 1968-80: A Spatial Analysis." *American Journal of Political Science*, 28(2): 282-312.
- [109] Razin, Ronny. 2003. "Signaling and Election Motivations in a Voting Model with Common Values and Responsive Candidates." *Econometrica*, 71(4): 1083-1119.
- [110] Roemer, John E. 1994. "A Theory of Policy Differentiation in Single Issue Electoral Politics," *Social Choice and Welfare*, 11: 355-380.
- [111] Roemer, John E. 2004. "Modeling Party Competition in General Elections." Manuscript, Yale University.
- [112] Romer, Thomas. 1975. "Individual Welfare, Majority Voting, and the Properties of a Linear Income Tax." *Journal of Public Economics*, 4: 163-185.
- [113] Schultz, Christian. "Polarization and Inefficient Policies." *Review of Economic Studies* 63: 331-344.
- [114] Shor, Boris. 2011. "All Together Now: Putting Congress, State Legislatures, and Individuals in a Common Ideological Space to Assess Representation at the Macro and Micro Levels." Manuscript, University of Chicago.
- [115] Sobel, Joel. 2006. "Information Aggregation and Group Decisions." Manuscript, University of California, San Diego.
- [116] Soubeyran, Raphaël. 2009. "Does a Disadvantaged Candidate Choose an Extremist Position?" *Annals of Economics and Statistics*, 93/94: 327-348.
- [117] Sunstein, Cass R. 2002. "The Law of Group Polarization." *The Journal of Political Philosophy*, 10(2): 175-195.

- [118] Tocqueville, Alexis de. 1835. *Democracy in America*. Trans. Henry Reeve. Ed. Phillips Bradley. New York: Alfred A. Knopf. 1945.
- [119] Triossi, Matteo. 2013. "Costly Information Acquisition. Is it Better to Toss a Coin?" *Games and Economic Behavior*, 82: 169-191.
- [120] Van Weelden, Richard. "Candidates, Credibility, and Re-election Incentives." *Review of Economic Studies*, 80: 1622-1651.
- [121] Wittman, Donald. 1977. "Candidates with Policy Preferences: A Dynamic Model." *Journal of Economic Theory*, 14: 180-89.
- [122] Wittman, Donald. 1983. "Candidate Motivation: A Synthesis of Alternative Theories." *American Political Science Review*, 77(1): 142-157.
- [123] Wittman, Donald. 1990. "Spatial Strategies when Candidates Have Policy Preferences," in *Advances in the Spatial Theory of Voting*, ed. James M. Enelow and Melvin J. Hinich, New York: Cambridge University Press.
- [124] Young, H. Peyton. 1995. "Optimal Voting Rules." *The Journal of Economic Perspectives*, 9(1): 51-64.