Polarization and Pandering in Common Interest Elections

Joseph McMurray*

January 18, 2019

Abstract

This paper analyzes candidate positioning in common interest elections, meaning that voter differences reflect private estimates of what is best for society, not idiosyncratic tastes. Centrist candidates have a competitive advantage, but may be bad for welfare. An extreme candidate can still win if truth is on her side, though, so for a variety of model specifications, candidates polarize in equilibrium, even when each wants very badly to win.

JEL Classification Number D72, D82

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

1 Introduction

The reigning paradigm for analyzing elections is the spatial model pioneered by Hotelling (1929) and Downs (1957), where the *median voter theorem* predicts that, to attract votes, candidates will adopt moderate equilibrium policy positions. The original theorem assumes office-motivated candidates, but Calvert (1985) shows that policy-motivated candidates who do not care at all about winning behave the same way, no matter how extreme their policy preferences, because controlling policy requires winning first. From a welfare perspective, this is good: centrist policies have

*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 Dynes, Martin Osborne, Adam Meirowitz, Rainer Schwabe, Roger Myerson, Tim Feddersen, Jean Guillaume Forand, Aniol Llorente-Saguer, Faruk Gul, Chris Ellis, Sourav Bhattacharya, Jean-Pierre Benoît, and David Myatt for their interest and suggestions.

utilitarian appeal, minimizing the total disutility voters experience from policies that are far from their ideal points.[1] In that light, the competitive force that drives extremists to the center can be viewed as the "invisible hand" of politics. Empirically, however, this convergence seems not to occur. In statistical studies of the U.S. House, Senate, presidency, and state legislatures, for example, elected officials remain as polarized as the most extremely liberal and conservative voters in the electorate.[2] This matches voter perception: across eleven U.S. presidential elections (1972-2012), ninety percent of participants in the American National Election Studies (ANES) rated both major candidates as weakly more extreme than they rated themselves on a seven-point ideological scale; only eleven percent rated both candidates as weakly less extreme than themselves.[3] From the standard perspective, polarization is both puzzling and troubling, suggesting some inexplicable failure of democracy.

Repeatedly, efforts to account for polarization have shown the median voter logic to be robust. When the precise median is unknown, for example, office-motivated candidates still converge.[4] Policy motivated candidates no longer do,[5] but in that case polarization is only minimal, unless uncertainty is severe.[6] Polarization is often attributed to the need to win support from primary election voters, donors, and activists who belong to extreme factions within a candidate's party, and to motivate this group to voter turnout. However, withholding support from moderate candidates merely sacrifices victory to the opposing side, so factions motivated by winning or by policy outcomes should not behave this way. Even if they do, the same extremism that mobilizes a candidate's own core supporters should also mobilize the opposition, who will now suffer greater disutility if they lose.[7]

[1] See Davis and Hinich (1968).

[2] For example, see Poole and Rosenthal (1984), Alvarez and Nagler (1995), McCarty and Poole (1995), Ansolabehere, Snyder, and Stewart (2001), Jessee (2009, 2010, 2016), Bafumi and Herron (2010), Shor (2011), and Fowler and Hall (2016).

[3] It also matches campaign rhetoric, where candidates trumpet their differences but rarely their similarities.

[4] See Hinich (1977, 1978), Coughlin and Nitzan (1981), Calvert (1985), Lindbeck and Weibull (1987), Enelow and Hinich (1989), Duggan (2000, 2006), Banks and Duggan (2005), and Bernhardt, Duggan, and Squintani (2009a).

[5] See the "probabilistic" voting models of Hinich (1978), Wittman (1983), Hansson and Stuart (1984), and Calvert (1985).

[6] Calvert (1985), Roemer (1994), Banks and Duggan (2005), and Bernhardt, Duggan, and Squintani (2007) show this formally using continuity arguments. See also the discussion in Section 7.3.

[7] For various versions of these insights, see Davis, Hinich, and Ordeshook (1970), Aranson and Ordeshook (1972), Coleman (1972), Aldrich (1983), Baron (1994), and Glaeser, Ponzetto, and Shapiro (2005). Moderation should also mobilize factions within a candidate's party whose policy prefer-

Because enforcement is difficult, candidates might promise centrist policies but implement extreme policies, once elected. In repeated elections, however, candidates should voluntarily moderate themselves, to establish centrist reputations.[8] Extreme candidates have a stronger incentive to run for office,[9] but if voters know candidates' preferences then intrinsically moderate candidates should win elections, by the standard reasoning. Polarized equilibria can therefore exist in entry models, but equilibria with little or no polarization are just as likely to prevail.[10] In the end, convergence remains so obstinately robust that Roemer (2004) refers to the "tyranny of the median voter theorem."[11]

By treating voter differences as fundamental taste parameters, the spatial models above all embrace a *private interest* paradigm. To shed new light on polarization, this paper instead turns to the *common interest* paradigm of Condorcet (1785), where voters unanimously prefer to do whatever is truly best for society, but are unsure what that is.[12] Formally, the true optimum is modeled as a random variable, and voters' individual opinions as to what is best are modeled as private signals, each imperfectly correlated with the truth. With only two policies, Condorcet's "jury"

ences are less extreme. Empirical evidence summaried by Barber and McCarty (2013) cast primary elections as an unlikely source of polarization, along with partisan redistricting, which is typically thought of as similarly making candidates beholden to extremists.

[8] See Alesina (1988). Even 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 policies.

[9] See Grosser and Palfrey (2014).

[10] See Osborne and Slivinski (1996) and Besley and Coate (1997).

[11] Recent literature offers dozens of additional explanations for polarization, including third party entry (Palfrey 1984, Castanheira 2003, Callander and Wilson 2007, and Brusco and Roy 2011), candidate asymmetries (Bernhardt and Ingberman 1985, Ansolabehere and Snyder 2000, Maravall-Rodriguez 2006, Krasa and Polborn 2010, Krasa and Polborn 2012, and Matějka and Tabellini 2018), uncertain voting (Banks 1990, Bernhardt, Duggan, and Squintani 2007, Ashworth and Bueno de Mesquita 2009, Gul and Pesendorfer 2012, and Yuksel 2018), interactions across legislative jurisdictions or branches of government (Ortuño-Ortín 1997, Alesina and Rosenthal 2000, Eyster and Kittsteiner 2007, Krasa and Polborn 2015, and Polborn and Snyder 2017), signaling (Kartik and McAfee 2007, Callander and Wilkie 2007, Callander 2008, Asako 2014, and Kartik, Squintani, and Tinn 2015), unaware voters (Gul and Pesendorfer 2009, Galeotti and Mattozzi 2011, Aragonès and Xefteris IER 2017), electoral discipline (Van Weelden REStud 2013), and convex voter utility (Kamada and Kojima 2014). Detailed discussions of each of these theories is beyond the scope of this paper, but none has been as influential as the theories listed above. Many of these theories only plausibly generate low levels of polarization, and many rely on special assumptions or circumstances that can apply in some settings but not universally.

[12] As I explain in McMurray (2017a), this is not to say that conflicts of interest do not exist or are unimportant, but that large elections amplify voter altruism, so that voters act as social planners, even when they are almost entirely selfish.

theorem points out that majority opinion can reliably identify the true optimum, even when individual signals are extremely noisy.[13] In McMurray (2017a) I extend that model to a spectrum of policy alternatives, using this to explain several empirical features of elections that are puzzling from a private interest perspective.[14]

Private interest literature considers various motivations that candidates might hold, and the analysis below does the same. The main specification assumes that, like voters, candidates seek to maximize social welfare. An extension considers candidates with deviant preferences. In either case, candidates may also value winning the election and holding office, and the benefit of this may be large or small relative to policy utility. Whether a candidate seeks to do what is socially optimal or merely seeks to please voters who do so, her beliefs about which policy is optimal also become relevant.[15] The most straightforward assumption would be that candidates form Bayesian beliefs, but this would require each voter and candidate to reason about the private signals of the others, thus introducing higher-order complexities that make the model intractable. Instead, therefore, the analysis below considers *underconfident* candidates, whose incentives are similar to Bayesians' but more clear-cut, along with *overconfident* candidates, who deviate from the Bayesian case in the opposite direction, thus giving perspective on the range of behavior possible in the Bayesian case.

Naturally, some features of a common interest spatial model resemble private interest spatial models or binary common interest models. In spite of their common interest, for example, idiosyncratic private signals lead every voter to favor a different policy. This produces a familiar trade-off for a candidate: moving toward the policy that she likes best improves her utility conditional on winning, but moving toward her opponent's platform makes her more likely to win, by attracting voters who favor policies between the two platforms. At the same time, however, the rare event of

[13] More specifically, the jury theorem states that if private signals are each positively correlated with the truth but, conditional on the truth, statistically independent, then as the number of voters grows large, majority opinion identifies the better of two policies with probability appoaching one. Krishna and Morgan (2011) call this "the first welfare theorem of political economy."

[14] Specifically, this framework can explain why voters often favor policies contrary to their own interests, change their policy opinions, and expend effort trying to persuade others; why voters on both sides of an issue tend to expect to win; why margins of victory are often large; and why voters who lack confidence in their information tend to remain politically moderate, and often abstain from voting. It also closely matches public discourse, where voters and candidates emphasize the universal appeal of their favored policies.

[15] Throughout this paper, feminine pronouns refer to candidates and masculine pronouns refer to voters.

a pivotal vote conveys information as in binary models of information aggregation, and taking this into account generalizes the jury theorem, so that no matter where two candidates locate, the one whose platform is truly superior wins the election with high probability.

The central result of the analysis below is that candidates polarize in equilibrium much more than in a private interest setting, adopting policy positions far left and far right of center. The logic for this is clearest when candidates are truth motivated and overconfident, with opposite beliefs about what is truly best for voters. The candidate who thinks the best policy is far left of center then expects voters to realize this as well. If they do, they will favor her policy over her opponent's (perhaps by a wide margin) even if her opponent is more centrist. Underconfident candidates have no exogenous reason to believe that the optimal policy is far left or far right, but form extreme beliefs endogenously, because of a pivotal calculus analogous to that performed by voters: specifically, her platform choice will only matter if she wins, which will only occur when truth is on her side. Even candidates who are ex ante identical polarize, therefore—potentially as much as those who are most overconfident. Such polarization from both over- and underconfident candidates suggests that Bayesian candidates would polarize substantially, too.

Since moving toward a candidate's opponent always wins votes, a median *opinion* theorem states that candidates who want badly enough to win adopt identical equilibrium platforms at the political center. In that sense, the centrifugal force highlighted in private interest literature operates here as well. However, this no longer has the same utilitarian appeal. In fact, it can produce policies that are known ex ante *not* to be optimal, such as unfunded programs. Such *pandering* to voters' mistaken opinions is reminiscent of the binary models of Canes-Wrone, Herron, and Shotts (2001) and Maskin and Tirole (2004), but the richer geometry here can explain why political compromise is sometimes viewed with disdain, as sacrificing truth for popularity.

In large elections, the logic of the jury theorem reduces the advantage of centrist platforms, by inducing aggregate uncertainty about how voters will behave. Because of this, candidates whose preferences differ from voters' polarize, as well. To a lesser degree, this remains true if voters overlook the informational content of pivotal voting. In contrast with private interest models, polarization also remains highly robust to office motivation. For some specifications, in fact, candidates are willing to make *any* policy concession necessary to win but remain just as polarized as if they cared nothing about winning. Moreover, this polarization can be substantial. As discussed below, these observations have important consequences for welfare.

2 Related Literature

There are three groups of papers that view elections as truth-seeking mechanisms. Explicit extensions of Condorcet's (1785) model focus exclusively on voting: specifically, informational efficiency given informational impediments,[16] alternative voting rules,[17] deviations from common value,[18] and strategic incentives to vote insincerely[19] or abstain.[20] 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 the common interest assumption is "clearly inapplicable" to public elections. In McMurray (2017a) I explore a truly spatial model of policy choice, but all of this literature focuses on voter behavior alone, modeling candidate platforms as exogenous.

A second group of papers focuses on whether or not candidates reveal their private information to voters. Binary models 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. Schultz (1996), Martinelli (2001), Loertscher (2012), and Kartik, Squintani, and Tinn (2013) consider private interest spatial models with a standard continuum of idiosyncratic preferences, but shifted together in the direction of a common shock. Pandering arises in Canes-Wrone, Herron, and Shotts (2001), Maskin and Tirole (2004), and Loertscher (2012), as candidates implement policies that are popular but inferior, whereas candidates in Laslier and Van der Straeten (2004) and Gratton (2014) reveal their private information completely, aware that voters may discover the truth. In

[16] See Ladha (1992, 1993), Mandler (2012), Dietrich and Spiekermann (2013), Pivato (2016), and Barelli, Bhattacharya, and Siga (2017).

[17] See Young (1995), Feddersen and Pesendorfer (1998), List and Goodin (2001), and Ahn and Oliveros (2016).

[18] See Feddersen and Pesendorfer (1997, 1999), Kim and Fey (2007), Krishna and Morgan (2011), and Bhattacharya (2013, 2018).

[19] See Austen-Smith and Banks (1996) and Acharya and Meirowitz (2016).

[20] See Feddersen and Pesendorfer (1996), Krishna and Morgan (2012), and McMurray (2013).

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 contrast with all of this literature, the analysis below considers how candidate positioning is influenced by *voter* information. This is appropriate in that a candidate presumably knows more than a typical voter, but much less than the electorate collectively.

There are three papers that study candidate positioning in light of voter information. In a private interest setting with binary valence, Bernhardt, Duggan, and Squintani (2009a) show that convergence by office motivated candidates can prevent the median voter from utilizing updated information about which policy he prefers. In binary common interest settings, Harrington (1993) shows that an overconfident incumbent politician trusts voters to learn the truth, and thus resists the temptation to pander by implementing popular but inferior policies, and Prato and Wolton (2017) show that information aggregation can fail, as office motivated candidates converge on whatever is favored ex ante, thereby delivering voters a degenerate policy menu. In contrast with these papers, the model below considers a continuum of policy alternatives and a continuum of truth states, thus becoming the first to analyze candidate positioning in a common interest spatial model.[21] Such richness has empirical merit, relates more directly to private interest literature, and, most importantly, is essential for exploring the extent of polarization.

3 The Model

3.1 Voters

There are N voters in an electorate, where, as in Myerson (1998), N is drawn from a Poisson distribution with mean n, who must implement one policy from an interval $X = [-1, 1]$ of alternatives. The policy z maximizes social welfare, and voters unanimously prefer policies as close as possible to this, but its location is unknown; at the beginning of the game, nature draws z from the domain $Z \subseteq X$. Specifically,

[21] Razin (2003) and McMurray (2017b) analyze common interest spatial models as well, but candidates adjust their policy positions *after* voting takes place, so that voting takes on a signaling role instead.

if policy x is implemented then each voter receives the following utility,

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

which declines quadratically with the distance between x and z.[22] This specific functional form does not seem important, intuitively, for any of the results below, but makes analysis tractable. In particular, expected utility is then quadratic in x, and maximized at the expectation of z, conditional on any available information. The concavity of (1) also implies that voters are risk averse.

There are two important cases of this model. The simpler of the two assumes *binary truth*, as in Condorcet's (1785) original model. That is, $Z = \{-1, 1\}$, 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 truth, the ideal size of an economic stimulus policy is either very large or very small. A moderate-sized stimulus is also feasible—and could be desirable for hedging against 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."

For many applications, moderate policies might be optimal, so it is more appropriate to assume *continuous truth*. In that case, let $Z = [-1, 1]$, meaning that any feasible policy might also be optimal.[23] Whether truth is binary or continuous, let z be distributed uniformly on Z. The function

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

doubles conveniently as a density or a mass function, thus accommodating either specification.

[22] Identical preferences are not essential for the resuts below, but if preferences have both private interest and common interest components then, as discussed below, the common interest component must be substantial.

[23] In McMurray (2017a) 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.

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 $S = Z$ as the true optimum. How confident a voter feels about his hunch depends on how much he knows generally about the policy question at hand. Let q_i denote the quality[24] of a voter's signal, drawn independently for each voter (and independently from z) from the domain $Q = [0, 1]$, according to some common distribution G which, for simplicity, is differentiable and has a strictly positive density g and mean $\bar{q}$. Conditional on $q_i = q$, the distribution of $s_i = s$ in state z is then given by the following,

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

which, like (2), doubles conveniently as a density if truth is continuous and as a mass function if truth is binary. The linearity of (3) does not seem intuitively important for any of the results below, but does seem essential for keeping things tractable once a voter updates his private beliefs to condition on the event of a pivotal vote.[25] This also provides a useful parameterization of the impact of expertise. For binary truth, q_i gives the correlation coefficient between s_i and z; a voter with $q_i = 1$, for example, observes z perfectly. With continuous truth, this correlation is only $\frac{1}{3}q_i$, so even the highest quality signals include substantial noise. Either way, the precision of s_i increases with q_i, and the lowest quality signal reveals nothing: if $q_i = 0$ then s_i and z are independent. Also, s_i is uniform on S.

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

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

[24]The terms *expertise*, *confidence*, and *information quality* are used interchangeably to refer to q_i, and could derive from policy-relevant technical training, or simply from time spent thinking deeply about an issue. That voters might miscalculate their own competency is an important possibility for future work to explore, but as Sunstein (2002) writes, "as a statistical matter, though not an invariable truth, people who are confident are more likely to be right".

[25]Specifically, the threshold structure of equilibrium relies on posterior beliefs that are monotonic in a voter's signal, *after* he updates to account for his vote being pivotal. It seems reasonable that this should hold generally, but pivotal events depend so intricately on model parameters that it is difficult to verify, without an especially simple signal structure like that of (3). The extensive symmetry of the model serves the same purpose.

and depends on q_i and s_i only through the product $\theta_i = q_i s_i$.[26] Once again, (4) can be interpreted as either a density or a mass function. Summing or integrating over Z, a voter's expectation of the optimal policy is then simply proportional to θ_i, which can therefore be interpreted as a voter's *ideology*.[27] The sign and magnitude of ideology depend on the sign and magnitude of s_i, and the magnitude also depends on a voter's expertise q_i. Specifically, a voter who lacks confidence in his opinions remains ideologically moderate, even if s_i is quite extreme. In McMurray (2017a) I show that this is true empirically,[28] 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.

The voter information and incentives described above are exactly as in McMurray (2017a). In that paper, however, voters choose between two exogenously specified alternatives; here, the menu of policies is endogenous: two candidates, A and B, with beliefs described below, propose policy platforms $x_A, x_B \in X$ that they commit to implement if elected. Observing these platforms, voters then vote for either candidate.[29] A strategy $v : Q \times S \to \{A, B\}$ in the voting subgame specifies a candidate choice $j \in \{A, B\}$ for every realization $(q, s) \in Q \times S$ of private information. Let V denote the set of such strategies. Votes are cast simultaneously, and a winning candidate $w \in \{A, B\}$ 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. When his peers all vote according to the strategy $v \in V$, a voter's *best response* is the strategy $v^{br} \in V$ that maximizes $E_{w,z}\left[u\left(x_w, z\right)\right]$ for every realization $(q, s) \in Q \times S$ of private information. A (symmetric) *Bayesian Nash equilibrium (BNE)* in the voting subgame is a strategy v^* that is its own best response.[30]

3.2 Candidates

Candidates are assumed to be *truth motivated*, meaning that they prefer policies as close as possible to z, maximizing the same objective (1) as voters. Such public

[26] Formulated this way, θ_i are *affiliated* with z in the sense of Milgrom and Weber (1982).

[27] For binary truth, $E\left(z|q_i, s_i\right) = \theta_i$; for continuous truth, $E\left(z|q_i, s_i\right) = \frac{1}{3}\theta_i$.

[28] For more recent evidence of a causal effect of information on polarization, see Garz (2018).

[29] Abstention is treated in McMurray (2017a), and would likely produce similar results here, but substantially complicate the model.

[30] With Poisson population uncertainty, BNE are necessarily symmetric (Myerson, 1998).

spirit could be intrinsic or could reflect a more selfish desire to develop a favorable legacy or reputation. Either way, this makes the model parsimonious in the sense that candidates' motivations are fundamentally the same as ordinary voters' (like the "citizen candidates" of Osborne and Slivinski, 1996, and Besley and Coate, 1997). On the other hand, many observers worry that candidates do not share voters' interests; as an extension of the baseline model, therefore, Section 7.1 considers candidates who are *selfishly policy motivated*, meaning that they favor specific policies that privilege themselves (or favored interest groups), regardless of what is best for voters. That specification is also useful for comparing the results below with existing literature. With either type of policy motivation, candidates may also be *office motivated*, meaning that they receive a benefit $\beta \geq 0$ (in prestige or other perks) from winning office, regardless of the policy outcome. Relative to policy utility, β can be large or small.

In addition to preferences, policy choices depend on candidates' beliefs about what is truly optimal. The most straightforward assumption would be that candidates are *Bayesian*, meaning that they start from the same prior belief as voters, receive private signals of their own, and update according to Bayes' rule. Alternatively, many observers worry that candidates are overconfident. To the extent that platforms reflect candidates' private information, however, the best response to one candidate's platform by the opposing candidate or by voters should take into account the private information that likely prompted that candidate's platform choice. Similarly, the private signals that determine voter behavior should be taken into account both by candidates and by other voters. In equilibrium, then, a candidate must anticipate what her opponent believes that one voter believes that other voters believe that she herself believes, and so on. Such higher-order beliefs seem hopelessly intractable, so the analysis below instead specifies beliefs in two extreme ways that avoid (some of) these complexities. These deviate from the Bayesian case in opposite directions, with hopes of revealing the range of behavior that might plausibly arise in the intractable Bayesian case.

Section 5.1 considers candidates who are *overconfident*. Such candidates are characterized by exogenous ideology parameters $\theta_A, \theta_B \in X$ (where, for ease of exposition, $\theta_A < \theta_B$), each representing a policy that one candidate feels certain is socially optimal. That is, candidate $j \in \{A, B\}$ believes (perhaps wrongly) that $\Pr(z = \theta_j) = 1$. Such beliefs are mutually irreconcilable, of course, and implausibly extreme, but this specification has two important virtues. First, it makes the analysis tractable by eliminating the need for voters to infer candidates' private information from their platform choices, and the need for candidates to infer information from one another. Second, precisely because it is so extreme, it provides a useful benchmark. The expected utility of an overconfident candidate $j \in \{A, B\}$ can be written as follows,

$$EU_j^O = \sum_{w=j,-j} u(x_w, \theta_j) \Pr(w|z = \theta_j) + \beta \Pr(w = j|z = \theta_j) \tag{5}$$

where the utility $u(x_w, z)$ associated with the winning candidate's platform and the conditional probability $\Pr(w|z)$ of either candidate winning are both evaluated at $z = \theta_j$. The first term in (5) captures the policy benefit a particular platform choice; the second captures the office benefit.

Section 5.2 assumes that candidates are *underconfident*, possessing no private information about the location of z beyond the commonly held prior, and whatever they can infer from voting behavior about voters' private signals. Equivalently, underconfident candidates could be thought of as receiving private signals, but undervaluing their own opinions so extremely as to discard their signals completely and act on the basis of the prior alone.[31] Like the specification of overconfidence, this is implausibly extreme, but makes the analysis tractable by avoiding the need for voters to infer candidates' signals, or for a candidate to infer her opponent's signal. Moreover, I show below that, in equilibrium, underconfident candidates infer a surprisingly large amount of information from voters—to the extent that explicitly adding candidate signals to the model would likely have little impact on equilibrium behavior. The expected utility of candidate $j \in \{A, B\}$ can be written as follows,

$$EU_j^U = \int_Z \left[\sum_{w=j,-j} u(x_w, z) \Pr(w|z) \right] f(z)\, dz + \beta \Pr(w = j) \tag{6}$$

which differs from (5) in that it now integrates over all possible realizations of z.

Whatever candidates' beliefs, let Σ denote the set of complete voting strategies $\sigma : X^2 \to V$, which specify subgame behavior for every possible pair $(x_A, x_B) \in X^2$ of candidate platforms. A (symmetric) *perfect Bayesian equilibrium (PBE)* is a triple $(x_A^*, x_B^*, \sigma^*) \in X^2 \times \Sigma$ such that $\sigma^*(x_A, x_B)$ constitutes a (symmetric) BNE in the

[31] The present model also applies if candidates receive private signals but announce them publically (and truthfully) before the election. In that case, the common prior f can merely be reintpreted as the common posterior, updated to account for this public information.

voting subgame associated with every platform pair $(x_A, x_B) \in X^2$, and candidates' platform choices x_A^* and x_B^* maximize either (5) or (6), taking the opposing platform x_{-j} and the voting strategy σ 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 Equilibrium

4.1 Voting Subgame

This section analyzes equilibrium voting in the subgame associated with an arbitrary pair $x_A \leq x_B$ of candidate platforms. If voters follow the voting strategy $v \in V$ then, in state $z \in Z$, each votes for candidate $j \in \{A, B\}$ with the following probability,

$$\phi(j|z) = \int_Q \int_S 1_{v(q,s)=j} h(s|q,z) g(q) \, ds dq \tag{7}$$

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. By the environmental equivalence property, a voter within the game reinterprets N_A and N_B as the numbers of votes cast by his peers; by voting himself, he can add one to either total.

Section 3 points out that, given the functional forms above, a voter's expectation $E(z|q_i, s_i)$ of the optimal policy depends only on his ideology $\theta_i = q_i, s_i$, and is monotonic in this value. It is therefore natural to focus on ideological voting strategies, as defined in Definition 1, meaning that a voter with ideology left of some ideology threshold τ vote A, while those with ideology right of τ vote B.

Definition 1 *A strategy v_τ is ideological, with ideology threshold $\tau \in X$, if* $v(\theta) = \begin{cases} A & \text{if } \theta < \tau \\ B & \text{if } \theta > \tau \end{cases}$.

A vote is unlikely to be *pivotal* (event P), meaning that it reverses the election outcome, but unless this occurs, a voter's behavior has no impact on his utility, so as Austen-Smith and Banks (1996) point out, a voter should optimally condition their behavior on this event. Instead of merely supporting the candidate closest to $E(z|q_i, s_i)$, therefore, a voter optimally supports the candidate who is closest to $E(z|P, q_i, s_i)$. It is this pivotal updating that makes a general model intractable, which is why Section 3 employs such specific functional forms.[32] Given these simplifications, however, the pivotal voting calculus does not alter the basic observation that voters with more conservative signals believe the optimal policy to be further to the right, and are thus more willing to support candidate B over candidate A. Accordingly, I show in McMurray (2017a) that the best response to any subgame voting strategy is ideological. That paper proves the existence and uniqueness of an equilibrium ideology threshold τ^*.[33] Lemma 1 now extends that result slightly, stating that τ^* is an increasing function of the midpoint $\bar{x} = \frac{x_A + x_B}{2}$ between the two candidates, and does not otherwise depend on candidates' platforms. Proofs of this and other formal results are presented in the appendix.

Lemma 1 *There exists a unique function $\tau^* : X \to X$ such that, for any $x_A, x_B \in X$ with midpoint $\bar{x}$, the ideological strategy $v_{\tau^*(\bar{x})}$ 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 $\tau^*(-\bar{x}) = -\tau^*(\bar{x})$.*

The last part of Lemma 1 states that τ^* is symmetric for symmetric values of $\bar{x}$. If platforms are symmetric so that $\bar{x} = 0$, for example, then $\tau^*(\bar{x}) = 0$ as well, meaning that voters simply vote A if θ_i is negative and vote B if θ_i is positive. This is useful because, empirically, many voters seem incapable of pivotal updating (Esponda and Vespa, 2014). This could undermine the theoretical prediction that voters condition on the event of a pivotal vote, except that, in this model, voters who condition on θ_i alone and those who condition on both P and θ_i behave identically.

4.2 Large Elections

Lemma 1 characterizes equilibrium voting for a fixed population parameter n. Since real-world electorates tend to be very large, this section analyzes voting behavior in the limit. To this end, first note that the number of votes that each candidate

[32]In addition to inferring information from other voters, a citizen should infer whatever he can from candidates' platform choices. This is not relevant below, however, because candidate beliefs are specified in a stylized way that conveys nothing useful.

[33]Uniqueness here, and in subsequent results, is up to the specification of behavior for the zero-measure set of voters with ideology $\theta = \tau$ right at the ideology threshold. In equilibrium, such voters are indifferent between voting A and voting B.

receives depends not only on the voting strategy, but on the realizations of voters' many private signals, which in turn depend on the state of the world z. For an ideological strategy with ideology threshold τ, define z_τ to be the realization of z that minimizes $|\phi(A|z) - \phi(B|z)|$—that is, the state that equalizes candidates' expected vote shares as closely as possible. The probability of a single vote being pivotal shrinks to zero in state z_τ, but at a slow rate; in all other states, it shrinks exponentially. Accordingly, a voter who behaves as if his vote will be pivotal increasingly behaves as if z_τ will be realized as the optimal policy.

If the number of voters is large and a voter's peers follow an ideological strategy with ideology threshold τ such that $z_\tau < \bar{x}$, then, by the above logic, he should vote A in response (since x_A is closer to z_τ than x_B is) regardless of his private information; if $z_\tau > \bar{x}$ then he should vote B in response. Either way, a voter should be unwilling to adopt the ideological strategy of his peers. It must therefore be the case that, as n grows large, the equilibrium threshold $\tau_n^*(\bar{x})$ adjusts so that the implied state of the world leaves voters indifferent between A and B, and therefore willing to follow their signals, as Lemma 2 now states.

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

Lemma 2 highlights how the pivotal voting calculus substantially evens out candidates' vote shares, an issue that is important for candidate incentives in Section 5. As an example, let truth be continuous with $q_i = 1$ for every voter, and let $x_A = .8$ and $x_B = 1$ (with midpoint $\bar{x} = .9$). In that case, even the most conservative voter only favors the policy $E(z|q_i, s_i) = \frac{1}{3}$, and thus views both candidates as too conservative. Since A is less extreme, she would win unanimously under sincere voting. This would likely be appropriate, because $z < .9$, but if $z > .9$ then candidate B would lose even though her platform is superior.

In this example, a strategic voter is more willing to vote for candidate B than a sincere voter is, reasoning that his own vote will only be pivotal when candidate B has received more votes than he had expected, meaning that more of his peers than expected had received conservative signals, suggesting that z is farther right than he expected, thus making candidate B's policy platform more attractive. In large elections, the equilibrium threshold adjusts to solve $\phi(B|z = \bar{x}; \tau_\infty^*) = \int_{\tau_\infty^*}^{1} \frac{1}{2}(1 + s\bar{x})\, ds = \frac{1}{2}$, or $\tau_\infty^* \approx .4$. Candidate B's vote share then ranges from about 9% when $z = -1$ to about 52% when $z = 1$. Precisely when z exceeds .9, candidate B wins the election. In McMurray (2017a) I derive an important consequence of this, restated here as Lemma 3, which is that Condorcet's (1785) jury theorem extends to this spatial environment. That is, no matter where the two candidates locate, the one whose policy platform is actually superior wins almost surely when n is large. This follows from McLennan's (1998) observation that, in common interest games, individuals cannot improve on a socially optimal strategy profile.

Lemma 3 (McMurray (2017a)) *For any n and any $(x_A, x_B) \in X^2$, $v_n^* \in V$ is a BNE in the voting subgame if and only if it maximizes $E\left[u\left(x_w, z\right); n, v\right]$. Moreover, under v_n^*, $w \rightarrow_{a.s.} \arg\max_j u\left(x_j, z\right)$.*

The jury theorem is a normative result, but in McMurray (2017a) I argue that it also sheds light on empirical facets of voter behavior, such as the broad support for using majority rule, the tendency to view popular support as evidence of superiority, and a *consensus effect* whereby individuals on both sides of an issue expect to belong to the majority.[34] As that paper reports, 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 essence, a voter who decides that one candidate's policy is better than the other's expects other voters, after weighing the evidence, to reach the same conclusion. This reasoning is important for the analysis below, as well, as a candidate who believes she is on the side of truth expects voters to recognize her superiority, and reward her with votes.

5 Candidates

Having characterized voters' equilibrium response to any platform pair, this section proceeds to analyze what incentives this creates for candidates in choosing platforms. Let $\sigma_{\tau^*} \in \Sigma$ denote the strategy that induces ideological voting in every subgame, with ideology thresholds given by the function $\tau^*(\bar{x})$ identified in Lemma 1. An equilibrium $(x_A^*, x_B^*, \sigma^*) \in X^2 \times \Sigma$ in the complete game consists of platform

[34] Hence also the dismay that many feel when a candidate who loses the popular vote takes office, as after the 2016 U.S. presidential election.

positions for both candidates and a voting strategy, where, by Lemma 1, a necessary condition is that σ^* corresponds to σ_{τ^*} in every subgame for which $x_A \neq x_B$. Sections 5.1 and 5.2 assume that $\beta = 0$, meaning that candidates are entirely policy motivated. Section 5.3 then generalizes these results.

5.1 Overconfident Candidates

If candidates are overconfident then, as Section 3 describes, each believes that $z = \theta_j$, and seeks to maximize (5). If $\beta = 0$, this objective function makes clear that the trade-off such candidates face is fundamentally the same as in standard private interest models with policy motivation and probabilistic voting, such as Wittman (1983) and Calvert (1985): moving toward her opponent improves a candidate's chance of winning office—which is desirable even if she doesn't value winning *per se*, as long as she prefers her own platform policy to her opponent's—but, conditional on winning, moving toward her ideal policy θ_j increases utility. In equilibrium, it cannot be the case that candidates adopt their ideal policies θ_A and θ_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. By standard reasoning, then, a candidate's equilibrium policy position lies strictly between her opponent's position and the policy 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, which by Lemma 1 induces symmetric voting. There is exactly one such equilibrium.

Theorem 1 *If candidates are overconfident and $\beta = 0$ then $(x_A^*, x_B^*, \sigma^*) \in X^2 \times \Sigma$ 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 $\left(x_{A,n}^*, x_{B,n}^*\right)$ of PBE platform pairs, $\lim_{n\to\infty} x_j^* = \theta_j$ for $j \in \{A, B\}$.*

While the basic logic of Theorem 1 is quite standard, the extent of polarization is not, as Section 7.3 makes clear below: in standard probabilistic voting models, uncertainty about 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 still cater to the estimated median voter, thus remaining close to one another. With standard formulations, candidates converge asymptotically in large elections. In contrast, Theorem 1 makes clear that overconfident candidates polarize substantially, especially in large elections, where they propose the policies θ_A and θ_B that they most prefer, and do not moderate at all. This more dramatic polarization 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 each candidate believes her own platform is superior, therefore, each is confident that she will win, and that policy concessions are unnecessary. This is especially stark when candidate ideologies lie at opposite ends of the policy space, and produce policies that are equally extreme. For example, this is bound to be the case when truth is binary.

5.2 Underconfident Candidates

As Section 3 describes, candidates who are truth motivated but underconfident start from the same prior belief as voters and have no private information of their own. They therefore seek to maximize (6). With no exogenous differences such as incumbency status, ability, or charisma, candidates' basic inclination is to adopt the same policy platform. Since candidates are risk averse and have no information beyond the prior, their natural policy choice would be at the center of the policy interval. As Theorem 2 states, however, this does not occur in equilibrium: if $\beta = 0$, candidates adopt policy positions with opposite signs and, at least in large elections, are highly polarized. With continuous truth, for example, platforms approach $-\frac{1}{2}$ and $\frac{1}{2}$, even though the most extreme voters only favor policies $-\frac{1}{3}$ and $\frac{1}{3}$ (see Footnote 27); with binary truth, candidates polarize to the far extremes -1 and 1 of the policy space. As noted in Section 1, this is consistent with empirical evidence that candidates are as polarized as the most extreme voters. Given the symmetry of the model, platforms may also be symmetric; for any n, there is exactly one such equilibrium.[35]

[35] 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 **??** can be viewed simply as a relabeling of the candidates. Given the symmetry of the model, equilibria with asymmetric platforms seem unlikely to exist, though this is mere conjecture.

Theorem 2 *If candidates are underconfident and $\beta = 0$ then $(x_A^*, x_B^*, \sigma^*) \in X^2 \times \Sigma$ is a PBE only if $x_A^* = E(z|w = A) < 0 < E(z|w = B) = x_B^*$. For every n, exactly one platform-symmetric PBE $(x_{A,n}^*, x_{B,n}^*, \sigma_n^*)$ exists. These equilibria satisfy $\lim_{n\to\infty} (x_{A,n}^*, x_{B,n}^*) = (E(z|z < 0), E(z|z > 0))$.*

It is remarkable that candidates who are identical and risk averse should polarize so substantially. The key observation underlying Theorem 2 is that candidates' optimal behavior depends on a pivotal calculus analogous to that performed by voters. Through this pivotal calculus, candidates infer additional information from voters' behavior. The two candidates infer different information from one another, however, and therefore differ endogenously. To see how this works, suppose that candidates adopt distinct platforms from one another, so that voting is ideological. In that case, the smaller z is, the more likely voters are to receive signals below the ideology threshold, and to vote for candidate A, making candidate A more likely to win the election. The larger z is, the more likely it is that candidate B will win. Ex ante, candidates' common expectation of the optimal policy is $E(z) = 0$, but if candidate A wins the election then she revises her beliefs about z downward; if B wins, she revises her beliefs upward. As a consequence, $E(z|w = A) < 0 < E(z|w = B)$. When choosing a policy platform, a candidate does not yet know if she will win the election. If she does not win, however, then her platform choice will not influence her utility. Thus, just like a voter who takes into account information that is not universally true but will be true in the event that his vote is pivotal, a candidate takes the event of winning the election into account in advance, so that if she does win, her policy position will be optimal.[36]

Another way to understand Theorem 2 is in light of the fact that truth motivated candidates have the same motivation and same beliefs as one another, and as a social planner who seeks to maximize voter utility. Anticipating that the better platform will surely win a large election, a planner would divide the state space into two equal regions, and assign the two candidates to take positions $E(z|z < 0)$ and $E(z|z > 0)$ that are optimal within the two regions. If candidates colluded in choosing their

[36]Intuitively, it may seem that there should be an additional equilibrium, with $x_A = x_B = 0$. After all, convergent platforms would leave voters indifferent between election outcomes, and therefore indifferent between voting strategies. In particular, voters could then follow a strategy that is unrelated to their private signals, in which case candidates would learn nothing useful about z from the event of winning, and would have no reason to polarize. That intuition is invalid, however, because a candidate who deviates to a different platform would trigger ideological voting, and thus infer information about z that justifies the deviation.

platforms, these are the platforms they would adopt. Theorem 2 assumes that candidates optimize independently, but since they share the same objectives and information, they adopt the same behavior as if they had optimized jointly. Section 6 relies on similar reasoning to show that the platforms highlighted in Theorem 2 are optimal from voters' perspective, as well.[37]

Specified this way, underconfidence implies that each voter in the electorate has more information about which policy is optimal than either candidates has. This is unrealistic in that candidates are voters themselves, and have invested in a career of policy making. Explicitly adding candidate signals would make the model intractable, as described above, by introducing higher order beliefs,[38] but the behavior of Bayesian candidates presumably lies somewhere between that of overconfident and underconfident candidates, so the result that both of these candidate types polarize suggests that Bayesian candidates would likely polarize, as well. In fact, Bayesian candidates may actually behave quite similarly to the underconfident candidates of this section. Suppose, for example, that candidates' policy positions are monotonic in their signals, so that by observing candidates' platforms, each voter would infer both candidates' signals. The event of a pivotal vote already takes into account the many signals of a voter's N peers, so increasing this number to $N+2$ should have little impact, and the updated conditional expectation $E(z|P, s_i, s_A, s_B)$ should be similar to $E(z|P, s_i)$.[39] Similarly, a candidate's updated expectation $E(z|w=j, s_j)$ (or even $E(z|w=j, s_j, s_{-j})$, if she can somehow infer her opponent's signal) should be similar to $E(z|w=j)$, since the event $w=j$ of winning already takes voters' N signals into account.[40]

[37] In McMurray (2017b) I show that, ex post, welfare motivated candidates have an incentive to incorporate additional information from realized vote totals, and adjust their policy positions accordingly. Since voters share these candidates' objective, they have no reason to object. In that sense, welfare motivation undermines the assumption that platform commitments are binding. If not all candidates are welfare motivated, however, a culture of requiring candidates to uphold their campaign promises could be an important safeguard against the occasional candidate with selfish policy motives (considered in Section 7.1), who may be inclined to deviate ex post to policy positions that are horrible from voters' perspective.

[38] Since candidates move before citizens, adding candidate signals might also introduce incentives to misrepresent their information and manipulate voter beliefs.

[39] Moreover, if his N peers are already accounting for s_A and s_B, a voter's own adjustment can be minimal.

[40] In fact, if the N voters are already taking s_A and s_B into account, a candidate's own adjustment can be minimal.

It is not clear whether the pivotal considerations above matter empirically for candidates, any more than it is clear whether pivotal considerations matter for voters. It is rare, for example, for candidates to describe learning from voters, or to admit any possibility that their current policy opinions might later prove to be wrong. One possibility is that the forces highlighted here operate in the real world, but a more complex information structure makes majority opinion more fallible, and therefore pushes candidates' beliefs less dramatically. After all, it does seem reasonable for candidates to be at least a little more confident when they feel bolstered by popular opinion, even if this is subconscious.[41] Candidates might also be overconfident, as assumed in Section 5.1, and less willing to update their beliefs than they should be. In any case, existing literature makes clear that pivotal considerations are relevant for optimal voting behavior, and can have dramatic consequences. The analysis here makes clear that pivotal considerations are relevant to candidates who join voters in the search for truth, and can have equally dramatic consequences, to the point that ex-ante identical candidates polarize more than any voter, and virtually ignore any initial opinions of their own.

5.3 Office Motivation

Sections 5.1 and 5.2 assume that $\beta = 0$, meaning that candidates do not care at all about winning, per se. Presumably, however, candidates in the real world value the prestige and other perks of office, in addition to caring about policy outcomes.[42] To allow that possibility, this section allows $\beta > 0$. A recurring theme from private interest literature is that office motivation decreases polarization.[43] In general, equilibrium platforms balance a candidate's desire to move toward her ideal policy and the need to move toward her opponent, to attract the votes of voters whose bliss points lie between the two platforms. As office motivation increases, the latter incentive dominates and a candidate moves toward the center.

In a common interest setting, voters ultimately share the same ideal point. Nev-

[41] Sobel (2006) uses essentially this logic 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 not modeled here.

[42] Entry decisions are not modeled here, but intuitively, positive β may be an important reason why individuals run for office.

[43] For example, see Alesina (1988) and Bernhardt, Duggan, and Squintani (2009a).

ertheless, Theorem 3 states that the familiar logic holds: polarization decreases with β until eventually candidate platforms exactly coincide. As before, equilibrium platforms balance the desire to move toward what a candidate believes is the best policy (θ_j if a candidate is overconfident, and $E(z|w=j)$ if she is underconfident) with the desire to attract the votes of voters whose private opinions tentatively favor policies that lie between the two platforms. As β increases, the latter incentive dominates and a candidate moves toward the center. Theorem 3 is labeled the median *opinion* theorem to emphasize that, while the behavior is familiar, the underlying source of voter heterogeneity differs from standard models. As Section 6 emphasizes below, this has important consequences for social welfare.

Theorem 3 (Median Opinion Theorem) *There exists $\bar{\beta}$ such that if $\beta \geq \bar{\beta}$ then $(x_A^*, x_B^*, \sigma^*) \in X^2 \times \Sigma$ is a PBE only if $x_A^* = x_B^*$. Moreover, such an equilibrium exists. If $\beta < \bar{\beta}$ and candidates are underconfident or symmetrically overconfident then a unique platform-symmetric PBE $(x_A^*, x_B^*, \sigma^*) \in X^2 \times \Sigma$ exists and, in this equilibrium, polarization $|x_B^* - x_A^*|$ strictly decreases in β.*

Given the importance of public policy decisions, arbitrarily large values of β may not be relevant empirically. With quadratic policy utility, for example, $\beta \geq 4$ makes a candidate willing to implement the *worst* policy in X, in order to win. Even smaller values of β such as 1 or $\frac{1}{4}$ may be implausibly large, implying that, to win, a candidate is willing to implement policies that deviate from the optimum up to a distance of 1 or $\frac{1}{2}$—that is, 50% or 25% of the length of X. In that light, Theorem 4 offers a different perspective into office motivation. Instead of analyzing large β for fixed n, this theorem analyzes large n for (arbitrarily high but) fixed β, thus generalizing Theorems 1 and 2. In doing so, Theorem 4 shows polarization to be quite robust.

Theorem 4 *If truth is binary or candidates are overconfident then $\lim_{n\to\infty} x_{j,n}^*$ is the same for all $\beta \geq 0$, for $j \in \{A, B\}$. If truth is continuous and candidates are underconfident then (for the unique sequence of platform-symmetric equilibria) $\lim_{n\to\infty} x_{B,n}^* = \max\left\{\frac{1}{2} - \frac{\beta}{4}, 0\right\}$ and $\lim_{n\to\infty} x_{A,n}^*$ is symmetric.*

If truth is continuous and candidates are underconfident then, according to Theorem 4, candidates only adopt identical platforms if $\beta \geq 2$, meaning that winning compensates for being a distance of $\sqrt{2}$ from what is optimal—or 70% of the length of the policy interval. If $\beta = \frac{1}{4}$ then candidates adopt positions at $\pm\frac{7}{16} \approx .44$, which is only barely less polarized than the positions at $\pm.5$ that purely policy motivated candidates would adopt (by Theorem 2). If candidates are overconfident or truth is binary then candidates are just as polarized when β is large as they are when $\beta = 0$.

For overconfident candidates, the logic behind 4 is essentially the same as the logic behind Theorem 1: if truth is on her side, the jury theorem guarantees (in large elections) that truth will prevail, making policy concessions unnecessary. Even if she is so motivated by the perks of office that she is willing to promise the worst policy in X in order to win, then, a candidate refuses to compromise at all in equilibrium, confident that doing what she believes is right will be a winning strategy. Similar logic applies to underconfident candidates, who do not know the precise location of z, but become increasingly certain, conditional on winning, that z is left of center or right of center. These candidates do moderate somewhat from their desired policy positions, because if one remained inflexible then the other could moderate somewhat in response, thereby winning office in additional states of the world—namely, for interior realizations of z. With binary truth, such states never occur, so, in large elections, moderation provides no benefit.

Together, Theorems 3 and 4 help make sense of three empirical observations that are otherwise difficult to reconcile. The first is that, in general elections, moderate candidates tend to have a competitive advantage. Hall (2015) presents causal evidence of this, for example, using regression discontinuities from primaries. By itself, this observation seems to validate the familiar logic of the median voter theorem. However, the second observation is that, for some reason, candidates choose to remain polarized, as Section 1 notes, thus foregoing this advantage. Polarization is often attributed to candidates' efforts to please extremist supporters who serve as donors, activists, and primary voters. In 2012 in the U.S., for example, many Republican primary voters acknowledged Mitt Romney as the most likely candidate to defeat incumbent president Barack Obama in the general election, but voted for Rick Santorum, Newt Gingrich, or Ron Paul instead, because Romney was "not conservative enough."[44] According to Brady, Han, and Pope (2007) and Hall and Snyder (2013), this pattern is typical: primary elections tend to favor extremists. If extremist supporters favor extreme candidates can indeed explain why candidates do not moderate, but begs the related question of why these extremist voters penalize

[44] For example, see http://elections.nytimes.com/2012/primaries/states/ohio/exit-polls. Similarly, in 2016, many Democratic and Republican primary voters viewed Hillary Clinton or Donald Trump as too moderate, and so voted for Bernie Sanders or Ted Cruz, instead.

moderation, given its competitive advantage.

The analysis above provides a new but simple reason why candidates and their supporters should resist moderation: if a candidate is on the side of truth, she can win the election even when her platform is more extreme than her opponent's. Put differently, Theorem 4 implies that the competitive benefit of moderation is more limited than other models suggest. The third empirical observation is precisely that: extreme candidates often beat moderates, as Ansolabehere, Snyder, and Stewart (2001), Canes-Wrone, Brady, and Cogan (2002), and Cohen, McGrath, Aronow, and Zaller (2016) variously point out, so the empirical benefit of moderation is positive, but not large.

Related to these three observations is a fourth observation, which is that, empirically, candidates who prefer one candidate over the other also tend to expect that candidate to win. In McMurray (2017a), for example, I document that "96 percent of ANES survey respondents who would later vote Democrat also predicted a Democratic victory, while 83 percent of those who would ultimately vote Republican predicted a Republican victory." The model above does not formally include primary elections, but in McMurray (2017a) I show that primary election voters tend to be very confident in their policy opinions. In light of the evidence summarized here, these voters seem the most prone—perhaps like the overconfident candidates modeled above—to view an extreme candidate as invincible.

6 Welfare

In private interest settings, centrist policies reflect a compromise between the competing interests on the left and on the right. This maximizes social welfare, by minimizing the total disutility that voters suffer from a policy that is far from their bliss points.[45] In that light, the prediction that competition for office should drive candidates toward the center, even when their intrinsic policy preferences are extreme, can be viewed as an "invisible hand" of politics. On the other hand, empirical evidence that candidates remain polarized must then be interpreted as some kind of political failure, as Ansolabehere, Snyder, and Stewart (2001) point out. A standard perspective might therefore motivate efforts to curb polarization,

[45] Davis and Hinich (1968) show this formally. If voters' loss functions are linear or quadratic, for example, then the utilitarian optimum lies respectively at the median or mean of the distribution of voter bliss points. Generically, it lies between the lowest and highest bliss points.

for example by raising office holder salaries, or otherwise increasing the benefit β that candidates perceive from winning.

Bernhardt, Duggan, and Squintani (2009a) add a caveat to the standard perspective. In their otherwise standard private-interest probabilistic voting model, a common shock shifts the distribution of bliss points, so that the median voter is either left of center or right of center. To maximize welfare in that case, candidates would need to cater to the two possible median locations, thereby allowing the median voter to secure his realized bliss point. Candidates who are purely policy motivated are overly extreme in that case, so a small amount of office motivation is beneficial, but candidates who are too office motivated do not polarize enough. As long as the shock is small in magnitude, relative to voters' private policy preferences, neither over- nor under-polarization is of great concern, because ideal policy positions remain close to the center, and therefore to the positions of overly office motivated candidates, and overly policy motivated candidates over-polarize only slightly, as in standard probabilistic voting models. Caveat notwithstanding, therefore, political compromise retains its basic appeal.

Theorem 5 analyzes social welfare under pure common interest. Since the benefit of winning office is zero-sum between the candidates, it is uncontroversial to measure welfare simply by the expected utility $E\left[u\left(x^*_{w,n}, z\right)\right]$ of an individual voter (which coincides with candidates' policy preferences). If voters are underconfident with $\beta = 0$, they share voters' interests exactly. As McLennan (1998) points out, behavior that is socially optimal is also individually optimal in that case, and therefore prevails in the unique equilibrium. The last part of the theorem simply points out further that, if truth is binary, optimal policy positions converge precisely to z as the electorate grows large.

Theorem 5 *For any n, there exists a strategy vector $\left(x^*_{A,n}, x^*_{B,n}, \sigma^*_n\right) \in X^2 \times \Sigma$ that maximizes $E_{N,q,s,w,z}\left[u\left(x^*_{w,n}, z\right)\right]$, and if candidates are underconfident with $\beta = 0$ then this vector also constitutes a PBE. If truth is binary then, for the optimal strategy vector, $|x^*_w - z| \rightarrow_{a.s.} 0$.*

Theorem 5 implies that, as long as candidates are underconfident, no amount of office motivation is desirable. As in Bernhardt, Duggan, and Squintani (2009a), then, office motivation reduces welfare by making candidates overly centrist. The fundamental value of polarization is also the same as in that paper, namely that it allows voters to tailor their policy choice more closely to the perceived state of the world. In this model, however, centrist policies no longer hold any utilitarian appeal at all: truth might lie in the center, but if it doesn't, extreme policies better serve the entire electorate. The optimal policy positions are therefore not just off-centered, but highly polarized—for the signal structure above, in fact, more polarized than any voter in the electorate. In that sense, the median opinion theorem above and the median voter theorems of existing literature predict identical behavior, but with opposite welfare implications. The contrast between Theorem 5 and results from existing literature is most stark when truth is binary. To end an economic recession, for example, candidates might disagree whether an economic "stimulus" policy should be large or small. In that case, competition for votes might produce moderate stimulus as a compromise, but this is known universally *not* to be optimal. Similar forces could generate a clearly sub-optimal, under-funded implementation of any policy or program.

That candidates can win votes by adopting inferior policies is reminiscent of the *pandering* models of Canes-Wrone, Herron, and Shotts (2001) and Maskin and Tirole (2004). However, those models are binary. The richer structure here makes clear that, in a spatial setting, pandering amounts to moderation—foregoing extreme policies that might dramatically improve welfare, and instead clinging to the safety of the political center. Recognizing that a centrist platform benefits a candidate at the expense of society, Tocqueville (1835, p. 175) wrote in praise of political parties that "cling to principles rather than to their consequences". More recently, the American Political Science Association (1950) issued a manifesto advocating to "keep parties apart," calling for "responsible parties" who believe that "putting a particular candidate into office is not an end in itself". Such recommendations are odd in the context of standard private interest models, where convergence is socially optimal, but fit neatly into the common interest paradigm above, as a call for truth motivation over office motivation.

This welfare perspective also relates to candidates' reluctance to compromise, as discussed in Section 5.3. In defense of his notoriously extreme policy positions, for example, 1964 Republican presidential candidate Barry Goldwater proclaimed that "extremism in the defense of liberty is no vice," while "moderation in the pursuit of justice is no virtue."[46] Similarly, 2000 Green party presidential candidate Ralph

[46] See https://www.washingtonpost.com/wp-srv/politics/daily/may98/goldwaterspeech.htm.

Nader criticized Republicans and Democrats for being "look alike parties", "Tweedledum and Tweedledee".[47] Extremists in either of the major U.S. political parties often disparage moderates in the same party with pejorative labels DINO or RINO (i.e. Democrats- or Republicans-in-name-only).

If candidates are underconfident then the policy implication of Theorem 5 is that β should be *lowered*, not raised. On the other hand, Theorem 4 suggests that excessive office motivation may be of limited concern, as polarization remains largely robust, implying that underconfident candidates already largely resist the urge to pander. Furthermore, Theorem 5 also implies that candidates who hold extreme and overconfident beliefs can over-polarize.[48] In fact, close elections suggest that available evidence only weakly favors the winning side, so the distance between $E(z|w=A)$ and $E(z|w=B)$ should be small. As in private interest models, therefore, combating polarization may be desirable. According to Theorem 4, however, increasing the perks of office cannot accomplish his goal, as overconfident candidates still polarize even when they want very badly to win.

If truth is binary, as Harrington (1993) suggests it might be for the most fundamental ideological questions, then much of this discussion may seem moot, because Theorems 1 and 2 then predict that overconfident and underconfident candidates adopt identical policies at ± 1. So far, however, the analysis above has interpreted z as the exactly optimal policy. Alternatively, I explain in McMurray (2017a) that there may *aggregate uncertainty* about the optimal policy, meaning that $z = E(z^*)$ is a mere approximation of an optimal policy z^* that remains uncertain even after all private information is pooled. In particular, z is continuous even when z^* is binary.[49] In that case, underconfident candidates should adopt platforms at $E(z|z<0) = -\frac{1}{2}$ and $E(z|z>0) = \frac{1}{2}$ in large elections, but candidates who are overconfident that $z^* = -1$ or $z^* = 1$ are liable to take positions at the far extremes of the policy interval.

[47] See http://www.cbsnews.com/news/nader-assails-major-parties/ (accessed 12/22/2016).

[48] If candidates are overconfident with moderate policy beliefs, they may under-polarize, as well.

[49] In that case, z close to zero indicates that $z^* = -1$ and $z^* = 1$ are equally likely, while z close to -1 means that $z^* = -1$ with high probability. A voter with high q_i and s_i close to the center can be interpreted as a *skeptic*, meaning that he recognizes that the true optimum is ultimately either $z^* = -1$ or $z^* = 1$, but perceives that available information is too ambiguous to warrant decisive policy moves in either direction.

7 Extensions

7.1 Selfish Policy Motivation

As an alternative to the assumption of truth motivation in Section 3, this section considers *selfishly policy motivated* (or *selfish*) candidates, who prefer policies $\hat{x}_A, \hat{x}_B \in X$ that privilege themselves (or favored interest groups), regardless of what is best for voters. Utility

$$u_j(x) = -(x - \hat{x}_j)^2 \tag{8}$$

is quadratic as in (1), but is now maximized at $\hat{x}_j$ instead of at z. Selfishly policy motivated candidates do not care about what is optimal for voters, but z is still relevant because it determines voter behavior. Beliefs therefore matter, as well. This section assumes that candidates are underconfident; as explained in Section 2, the case of Bayesian candidates would likely be similar.[50] Expected utility

$$EU_j^P = \int_Z \left[\sum_{w=j,-j} u(x_w, \hat{x}_j) \Pr(w|z) \right] f(z)\, dz + \beta \Pr(w = j) \tag{9}$$

is therefore the same as in (6), except with disutility measuring the distance to $\hat{x}_j$ rather than to z.

As in (5) and (6), the first and second term in (9) reflect policy motivation and office motivation, respectively. Once again, these motivations are in conflict: moving toward her ideal point improves a candidates' utility if she wins, but moving toward her opponent makes her more likely to win. If β is sufficiently high, it can be shown using the logic of Theorem 3 that candidates adopt identical platforms $x_A^* = x_B^* = 0$ at the center of the policy space. If β is sufficiently low then, for finite n, the logic of Theorems 1 and 2 instead guarantees that $\hat{x}_A < x_A^* < x_B^* < \hat{x}_B$: a candidate does not mimic her opponent because deviating toward her ideal policy would then make her better off if she wins, and no worse off if she loses. She also does not propose her favorite policy, because deviating from this would then improve her chance of winning, with zero marginal disutility. For symmetrically motivated candidates, Theorem 6

[50] This presumes that, even if a candidate's signal is more accurate than the typical voter's, it is not more accurate than N voters' pooled information, which seems reasonable. Martinelli (2001) shows that selfish candidates who are better informed than the electorate polarize more substantially, taking advantage of voters' trust that they know what is in voters' interest better than voters themselves do. See also Kartik, Squintani, and Tinn (2015).

now characterizes equilibrium behavior in the limit, as n grows large.

Theorem 6 *If candidates are selfishly policy motivated with* $\hat{x}_A = -\hat{x}_B$ *then, for any* n*, there is exactly one platform-symmetric PBE* $\left(x^*_{A,n}, x^*_{B,n}, \sigma^*\right)$*. Moreover,* $\lim_{n\to\infty} x^*_{B,n} = \begin{cases} \min\left\{\frac{\hat{x}_B - \frac{1}{4}\beta}{1+\hat{x}_B}, 0\right\} & \text{if } Z = [-1,1] \\ \hat{x}_B & \text{if } Z = \{-1,1\} \end{cases}$ *and* $\lim_{n\to\infty} x^*_{A,n}$ *is symmetric.*

Theorem 6 makes clear that if truth is continuous then office and policy motivations both remain relevant no matter how large the electorate grows. On one hand, a candidate is never as extreme as she would be if she could guarantee victory. This is true even if $\beta = 0$, so that candidates are purely policy motivated, because controlling policy requires winning first. On the other hand, candidates remain quite polarized, even when β is rather substantial. If β is as high as $\frac{1}{4}$, for example, so that winning compensates for policy distances amounting to 25% of the policy interval, then candidates with the most extreme preferences adopt policies at $\pm\frac{15}{32} \approx 0.47$ in large elections. These are only slightly less polarized than the positions ± 0.5 they would adopt if β were zero. Complete convergence would require $\beta \geq 4$, meaning that a candidate is willing to commit to her *least* favorite policy in order to win. Polarization is even greater if truth is binary: in that case, equilibrium policy positions converge to $\hat{x}_A$ and $\hat{x}_B$ in large elections.

The reason for polarization in Theorems 1 and 2 was that each candidate believed (exogenously or endogenously) that truth is on her side. A selfish candidate ultimately polarizes for a similar reason. The location of z does not change a selfish candidate's policy preferences, but still matters because it matters to voters. If truth is solidly on her side, voters will likely discern as much (especially in large elections), so she will likely win the election even if her policy position is more extreme than her opponent's. If truth strongly favors her opponent, she will likely lose, even if her own position is more centrist than her opponent's. Moderating her policy position thus only improves her chance of winning when z lies almost exactly between x_A and x_B. With continuous truth, such states exist but occur with low probability, so the benefit of moderation is low. With binary truth, intermediate values of z cannot occur, so moderation has no benefit at all. In equilibrium, then, a candidate merely adopts the policy that she most prefers, and hopes that she is lucky, so that z favors her over her opponent.[51]

[51] The logic of Theorems 1 and 6 can be combined to show that candidates who are selfish and

7.2 Sincere Voting

The model of Section 3 assumes that voters vote strategically. In Section 4, these voters make strategic inference from the event of a pivotal vote, and as that section explains, this contributes to the polarization derived in Section 5. In laboratory settings, however, voters persistently prove unable to make such pivotal inferences (Esponda and Vespa, 2014). Accordingly, this section explores the alternative possibility that voters vote *sincerely*. That is, each simply votes A if $s_i < \bar{x}$ and votes B otherwise. Beyond any empirical merit, this specification is useful both for clarifying the importance of strategic voting for polarization, and for making the analysis above more comparable with existing literature.

If candidate platforms are symmetric then sincere voters and strategic voters behave identically, of course, voting A if $s_i < 0$ and B if $s_i > 0$. Intuitively, then, it might seem irrelevant whether voters are strategic or sincere. However, equilibrium can only be sustained if candidates lack the incentive to deviate to other policy positions, and deviations from symmetry elicit different responses from sincere voters than from strategic voters. For some model specifications, therefore, sincere and strategic voting do generate different incentives for candidates.[52] For sincere voting and generic midpoint $\bar{x} = \frac{x_A + x_B}{2}$, candidate j's expected vote share is given by the following, instead of (7).

$$\phi(j|z) = \int_Q \int_{\bar{x}}^1 h(s|q,z)\, g(q)\, ds dq$$

Whether candidates are truth motivated and overconfident, as assumed in Section 5.1, truth motivated and underconfident, as assumed in Section 5.2, or selfishly policy motivated and underconfident, as assumed in Section 7.1, sincere and strategic

overconfident polarize in large elections, as well. Confident that $z = \theta_j$, candidate j adopts a policy as close to $\hat{x}_j$ as possible, subject to the constraint of being closer to θ_j than her opponent is. It is also straightforward to show that candidates with a mix of truth motivation and selfish policy motivation simply adopt policies in between the positions of candidates with truth motivation alone and those with selfish policy motivation alone.

[52] An alternative information structure that is commonly used assumes that private signals are normally distributed around the truth. In that case, Martinelli (2001) points out that sincere voting and strategic voting are the same in large elections, because conditioning on the event of a pivotal vote reveals to a voter that the median signal realization within the electorate (which perfectly reveals the truth) is his own. Generically, the efficiency of sincere voting relies on this being the case. For strategic voting, however, the logic of McLennan (1998) guarantees efficiency as long as *any* voting strategy is efficient.

voting produce the same basic trade off: moving toward her preferred policy position improves a candidate's utility if she wins, but makes winning less likely. For any of these cases, the logic of Theorems 1 through 3 therefore implies for any finite electorate that $x_A^* = x_B^*$ if β is sufficiently high but $x_A^* < x_B^*$ if β is sufficiently low. What happens when n grows large depends somewhat on the specification, as Theorem 7 now outlines.

Theorem 7 *If voting is sincere then, for any sequence* $\left(x_{A,n}^*, x_{B,n}^*, \sigma_n^*\right)$ *of platform-symmetric PBE:*

1. *If candidates are truth motivated and overconfident then* $\lim_{n\to\infty} x_{B,n}^* = \theta_B$*.*
2. *If candidates are truth motivated and underconfident then* $\lim_{n\to\infty} x_{B,n}^* = \begin{cases} \min\left\{\frac{1}{2} - \frac{1}{2\bar{q}}\beta, 0\right\} & \text{if } Z = [-1,1] \\ 1 & \text{if } Z = \{-1,1\} \end{cases}$.
3. *If candidates are symmetrically selfish and underconfident then* $\lim_{n\to\infty} x_{B,n}^* = \begin{cases} \min\left\{\frac{\hat{x}_B - \frac{1}{2\bar{q}}\beta}{1+\frac{2}{\bar{q}}\hat{x}_B}, 0\right\} & \text{if } Z = [-1,1] \\ \hat{x}_B & \text{if } Z = \{-1,1\} \end{cases}$.

In all of these cases, $\lim_{n\to\infty} x_{A,n}^*$ *is symmetric.*

Comparing Theorem 7 with Theorems 4 and 6 makes clear that, if truth is binary or candidates are overconfident, sincere and strategic voting produce identical responses. These are precisely the cases where polarization does not depend on office motivation. With continuous truth and underconfident candidates, however, sincere voting mutes polarization, whatever candidates' motivations. The logic for this result is that, with strategic voting, candidates polarize because moderation only has benefit in the rather unlikely event that z is almost equidistant from the two platforms; for very high or very low realizations of z, a candidate will either win with high probability or lose with high probability, whether she moderates or not. Sincere voting weakens the dependence of voting behavior on z, so even when z turns out to be extreme, there are citizens who happened to draw moderate signals. Adopting a moderate policy stance attracts these voters. For this same reason, polarization is lowest when the average quality $\bar{q}$ of voter information is close to zero. In that case, signals follow roughly the same uniform distribution regardless of z, so there are always moderate voters that can be attracted by a moderate policy platform.

Though polarization changes across specifications, the conclusion of Theorem 5 does not: as long as platforms are symmetric around the origin, sincere voting is identical to strategic voting (at least on the equilibrium path), and therefore every bit as informative. If platforms are asymmetric then the jury theorem need not hold: the candidate closer to the center has an electoral advantage, even when her opponent's platform is closer to the true realization of z. This, too, is especially important if the average quality of voter information is low.

7.3 Private Interest

The model above is one of pure common interest, in that voters ultimately prefer the same policy. In contrast, the assumption that pervades most existing literature is pure private interest, meaning that one voter's preferred policy is unrelated to another's. No doubt the real world includes a mix of common and private interests, but comparing the two extremes sheds light on the forces that might dominate in a hybrid model. This also helps clarify the novelty of the results above. Repeatedly, for example, that analysis balanced the same forces that operate in private interest, probabilistic voting models such as Wittman (1983) and Calvert (1985): moving toward a candidate's desired policy improves her utility conditional on winning, but moving toward her opponent makes winning more likely, which is important even when $\beta = 0$, because controlling policy requires winning first. Conveniently, the model above can be reinterpreted as one of private interest, simply by treating s_i as private bliss points $\hat{x}_i$, each drawn from the same uniform distribution on $[-1, 1]$, instead of as signals of a common interest, and by letting G be degenerate on $q_i = 0$ so that these preferences are independent of one another.

With stochastic voter preferences, as long as office motivation is not too strong, the familiar logic of Wittman (1983) and Calvert (1985) guarantees that candidates adopt distinct platforms in equilibrium: deviating from her opponent results in the same policy outcome if she loses, but a superior outcome if she wins. Moreover, if a candidate expects her opponent to deviate from the center then she herself can deviate further from the center without fear of losing, in turn making her opponent willing to deviate further, and so on. Intuitively, then, it may seem that polarization should be high in equilibrium, even if uncertainty is mild, with candidates taking extreme positions, but still expecting to win with close to 50% probability, just as if they had both been more centrist. With such compelling logic, and since imperfect information seems inevitable, it is no wonder that polarization is so frequently attributed to electoral uncertainty. Ultimately, however, this intuition is incorrect, as Theorem 8 now states: with private interests, polarization is negligible in large elections, even for $\beta = 0$. Thus, while the fundamental forces at work in common and private interest models are the same, the magnitudes of these forces are not.

Theorem 8 *Let G be degenerate on $q = 0$. For any β, if candidates are selfishly policy motivated and voting is sincere then, for any sequence $\left(x^*_{A,n}, x^*_{B,n}, \sigma^*_n\right)$ of PBE, $\lim_{n\to\infty} x^*_{A,n} = \lim_{n\to\infty} x^*_{B,n} = 0$.*

The rationale for Theorem 8 is simple: as the electorate grows large, the realized location of the median voter converges probabilistically to the median of the distribution from which ideal points are drawn, which is 0 in this case. As this happens, it remains true that polarized candidates each win with equal probability, and that marginal deviations from these polarized positions only slightly raise or lower the chance of winning. In addition to marginal policy adjustments, however, an extreme candidate has the option of jumping clear to the political center. If her opponent polarizes, doing so almost surely wins the election.[53] In equilibrium, then, each candidate remains sufficiently moderate that centrist policies will not guarantee her opponent victory.[54] As the range of possible locations of the median voter shrinks, so does the range of platforms that can win against a centrist.

Specifying probabilistic voting this way, candidates lack information about individual voters but have virtually perfect information about the electorate as a whole. In reality, of course, candidates have no way of knowing the macro distribution of preferences, other than surveying individual voters to find out, as in Bernhardt, Duggan, and Squintani (2009b). Such a model does not nest so neatly into the framework above, but standard probability theory makes clear that the distribution of the realized sample median converges in large samples to the population median, even if the latter is unknown. As candidates conduct larger and larger polls, therefore, their uncertainty regarding the location of the true median voter should dissipate, and polarization should decline just as in Theorem 8, vanishing in the limit.[55] Preventing

[53] A candidate prefers a centrist policy outcome with certainty over a lottery yielding one of the extremes because of the concavity of utility.

[54] If the median voter were known with certainty to lie between $-\varepsilon$ and ε, for instance, symmetrical platforms could not be more polarized than -2ε and 2ε in equilibrium, lest one candidate deviate to the center and win with certainty.

[55] Candidates have incentive to conduct large polls, since a more accurate estimate of the median voter's location gives a candidate a competitive advantage over her opponent. Voters also have incentive to actively make their interests known, thus lowering candidates' cost of acquiring this information.

convergence in the limit would require *aggregate uncertainty*, meaning that the precise location of the median voter remains unknowable, even after polls grow large. For polarization to be substantial, this aggregate uncertainty must be severe. In a common interest setting, aggregate uncertainty stems naturally from the inherent unobservability of the truth variable.[56]

To generate substantial polarization, existing literature sometimes assumes that the *median voter* is drawn from a uniform distribution on $[-1, 1]$. In the context of the model of Theorem 8, this amounts to an electorate with $n = 1$. For $\beta = 0$, it can be shown in that case platforms approach $-.5$ and $.5$ in the limit. To believe that the median voter might lie at either extreme of the policy space, however, a candidate must place non-trivial probability on *half* of the electorate being bunched at that extreme. In a private interest setting with even minimal polling, this level of uncertainty seems implausibly severe. An uncertain median voter is sometimes interpreted as the result of a common-interest "valence" shock that shifts the entire distribution of ideal points to the left or right, but if common-interest considerations are small relative to voters' idiosyncratic preferences (as the term *valence* suggests), this is insufficient for sustaining polarization. On the other hand, if common interest considerations are large enough to sustain true polarization, the model amounts to one of essentially common interests, just as above.

Another way to generate polarization in a private interest model would be to assume that $\hat{x}_i$ are *correlated* draws from the population distribution, rather than i.i.d. As in the common interest specification, this would increase polarization by reducing the benefit of moderation: if one voter turns out to be on her side, others likely will be as well, so moderation will be unnecessary; if one voter has preferences opposite her own, others likely will too, and moderation will not be enough to save her campaign. On the surface, this might seem to make common interest unnecessary for polarization. However, assuming correlated preferences amounts to *defining* a common interest. To see this, note that candidates are unlikely to know more about one voter than another, meaning that the distribution of ideal points should

[56] In principle, polling could eliminate uncertainty in a common interest setting as well, by revealining voters' realized opinions. After a poll is conducted, however, voters are liable to continue receiving new information (or re-thinking the information they had received previously) up until the day of the election. A candidate may therefore remain polarized even after a poll indicates that her platform is far from the realized median opinion, because she expects that as voters continue to learn, they will come around to her side.

be *exchangeable*.[57] If so, de Finetti's (1980) theorem states that the $\hat{x}_i$ can be reinterpreted as mutually independent, conditional on a latent variable. In this application, that latent variable can be interpreted as the object z of common interest. In that sense, it is without loss of generality to assume a common interest component to voters' ideal points, but once again, polarization will only be substantial if this common interest component is sufficiently important, relative to the idiosyncratic portion of voters' preferences.

Together with earlier results, Theorem 8 helps quantify the importance of various model assumptions for polarization. For example, suppose that candidates have extreme policy preferences (i.e., $\hat{x}_A = -1$ and $\hat{x}_B = 1$) and do not care about winning office (i.e., $\beta = 0$). With private interest voting (i.e., degenerate G), candidates do not polarize at all in large elections. With common interest voting (say, with G uniform on $[0, 1]$, so that $\bar{q} = .5$), equilibrium positions instead approach $(-.2, .2)$, by Theorem 7. If voters are strategic instead of sincere then, by Theorem 6, platforms instead approach $(-.5, .5)$, making candidates more extreme than any voter in the electorate. If candidates are truth motivated instead of selfishly policy motivated, they do the same, by Theorem 2. If candidates are overconfident with $\theta_A = -1$ and $\theta_B = 1$, they instead take positions at -1 and 1, by Theorem 1. When truth is binary instead of continuous, any specification with common interest voting produces the same. Robustness to office motivation follows the same pattern: with private interest voting, candidates do not polarize for any β; with common interest but sincere voting, they converge when $\beta \geq 1$; with strategic voting, convergence requires $\beta \geq 4$; if truth motivated candidates are underconfident they converge only if $\beta \geq 2$; if they are overconfident or truth is binary then they polarize for any β.

From all of this, it is clear that individual assumptions about beliefs and preferences each contribute to equilibrium polarization, but the key ingredient for substantial polarization is a substantial common component to voter interests. Voter interests need not correlated perfectly; in fact, private interest could interact in interesting ways with the logic above. Consider, for example, a hybrid of the models above, where voter ideal points $\hat{x}_i + z$ consist of both idiosyncratic and common components, as in Bernhardt, Duggan, and Squintani (2009a)—say, both uniform on $[-1, 1]$—with analogous candidate ideal points $-1 + z$ and $1 + z$. In large elections,

[57] Formally, exchangeability means that the joint distribution of $\hat{x}_i$ is constant with respect to permutations of the set of voters.

the logic of Feddersen and Pesendorfer (1999) assures that the candidate closer to the median voter's realized ideal policy $0 + z$ will win with high probability. With symmetric platforms, therefore, candidates would still form posterior expectations $E(z|w = A) = E(z|z < 0) = -.5$ and $E(z|w = B) = E(z|z > 0) = .5$, as in Section 5.2, but, given their more extreme preferences, would take up platforms at -1.5 and 1.5. In that sense, idiosyncratic differences can greatly exacerbate the polarization that arises under pure common interest, even though, as the discussion above makes clear, a model with idiosyncratic preferences alone yields no polarization at all.

8 Conclusion

Synthesizing the insights of private- and common-interest literature, this paper has proposed a solution to the puzzle of why candidates remain so polarized and forego the competitive advantage that compromising would give. Ironically, what makes candidates so polarized is the unity of voters, which translates into aggregate uncertainty about the location of the median. As long as the common component of voter interests is substantial, polarization is robust to a variety of specific modeling assumptions. Equilibrium policy positions balance the desire for better policies against the need to attract votes, just as in standard probabilistic voting models, but in large elections, a candidate who sides with truth is sure to win in spite of her extremism. Recognizing this, candidates remain highly polarized, even when they want very badly to win.[58]

One of the key virtues of a theoretical model is its ability to link behavior, which is observable, to welfare, which is not. The prevailing view is that centrist policies maximize welfare, but the analysis above offers an opposite perspective: political compromise can also reflect the sacrifice of truth for popularity, producing policies that are known *not* to be optimal. That identical behavior can have opposite welfare implications underscores the importance of identifying which model is more accurate. Ultimately, under-polarization seems to be of limited concern, both because candidates largely resist the urge to pander and because, in close elections, the evidence favoring the minority side is nearly as strong as the evidence favoring the majority,

[58] The analysis above focuses on elections, but a similar mechanism might explain dysfunction and gridlock in polarized legislatures, where both sides resist compromise, each confident that, by siding with truth, they will gain additional support in future elections.

so the platforms $E(z|z<0)$ and $E(z|z>0)$ adopted by truth-motivated, underconfident candidates should be similar, implying that the optimal level of polarization is low. In that sense, polarization is less mysterious than before but no less troubling, revealing candidates as severely overconfident,[59] or putting special interests ahead of voters'. Unfortunately, such over-polarization may be difficult to remedy.

The more important message from the analysis above may simply be that, if empirical behavior matches voters who share a substantially common interest, electoral institutions should be designed and evaluated for their informational efficiency, rather than for other factors. However, one significant challenge for a common interest perspective is persistent voter disagreement: if elections are efficient, as in the model above, then all should agree ex post, and voters who learn that they belong to the minority should infer that they were wrong, update their opinions, and join the majority. On a smaller scale, any two individuals who exchange information should update their beliefs and develop a common posterior, based on their mutual information. Empirically, of course, achieving consensus is not nearly so easy. This suggests more elaborate information structure as an important direction for future exploration.[60] In reality, for example, policy decisions depend on a myriad of factors, so resolving disagreements requires determining which factors produced the discrepancy.[61] If voters communicate with one another or acquire information from similar sources then they may also be prone to correlated errors in judgment, in contrast with the conditionally independent signals assumed above. It is also possible that voters start from different beliefs about the signal generating process, which Acemoglu,

[59] Candidate overconfidence seems natural, given that, as Caplan (2007) and Ortoleva and Snowberg (2015) document, overconfidence among voters is rampant. Entry is not modeled here, but it also seems intuitive that individuals who are the most confident in their policy opinions might be the ones most inclined to run for office.

[60] Alternatively, it is tempting to discard the entire common interest paradigm and attribute political differences to fundamental, immutable tastes. As I argue in McMurray (2017a), however, this response is premature: empirically, individuals maintain unpopular opinions on non-political questions as well, where no interests are at stake, suggesting that conflicts of interest are not the (only) source of disagreement.

[61] Literature on *judgment aggregation* makes this more concrete (see List, 2012). Suppose, for example, that a policy is beneficial if and only if claims 1 and 2 are both true. The "doctrinal paradox" of Kornhauser and Sager (1986) is that a majority of voters might oppose the policy even though (different) majorities believe each of the claims. In that case, learning only that the majority opposes the policy may lead one individual to oppose the policy as well, but another with access to more information about the individual claims may continue to support the policy. In that sense, those with different information may draw opposite conclusions from observing the majority policy decision.

Chernozhukov, and Yildiz (2009) show can sustain disagreements indefinitely.

Extensions such as these may have important consequences for polarization. On one hand, informational limitations that impede voter agreement may also weaken the jury theorem.[62] If candidates can no longer be certain that truth will prevail, polarization may be less pronounced. On the other hand, erroneous swings in public opinion may be another important source of aggregate uncertainty.[63] Sustained voter disagreements also open the possibility of sustained candidate disagreements, which seem to be of central importance in real-world elections. Indeed, public debates and other campaign activities can best be understood as opportunities for a candidate to attempt to persuade voters (and for voters to attempt to persuade one another) that her own policy proposals are socially optimal. On the other hand, no matter how complex these extensions become, the simple conditionally i.i.d. signal structure above seems likely to remain useful as a benchmark.[64]

The focus in existing literature on private interest models leaves much to be explored in a common interest domain, and the analysis above points in several useful directions. For example, the result that candidates with different beliefs and motivations behave differently in equilibrium begs the question of how these various candidate types fare against one another, and who has the incentive to run for office in the first place. This may depend importantly on electoral institutions such as primary elections. Intuitively, the forces identified above for general elections seem likely to be relevant in primaries, as well. It seems natural, for example, that voters who are the most confident—and even overconfident—in their policy beliefs might be the most inclined to participate in primary elections, and also the most inclined to send an extremist nominee to the general election, confident that general election voters will recognize her extreme platform to be optimal. For voters and underconfident or Bayesian candidates, pivotal events should still be relevant.[65] In primary

[62] Ladha (1992, 1993), Dietrich and Spiekermann (2013), and Pivato (2016) show that public opinion becomes less accurate when individual voters make correlated errors. Mailath and Samuelson (2018) show a related result in a model of information exchange, given heterogeneous beliefs about the signal generating process.

[63] If voter errors were perfectly correlated, for example, then public opinion would swing every time a new piece of information was discovered.

[64] In particular, this specification preserves the connection between voter opinions and underlying truth, which is essential for welfare analysis.

[65] Though pivotal events are more compliated in primary elections: a vote that is pivotal for candidate A_1 over A_2 within the liberal party's primary election might still have the effect of switching the eventual policy outcome from x_{A_2} to x_{A_1}, for example, but alternatively might change the eventual

and general elections alike, another open question is how candidate behavior changes when voters have to learn not only about whose policy position is truly superior, but also which candidate is superior in non-policy "valence" aspects such as honesty or management skill.

Beyond the specific question of polarization, I show in McMurray (2017b) that the size of a winning candidate's margin of victory conveys useful information about the precise location of the socially optimal policy. If she responds to this by adjusting her policy position after taking office, voting takes on a signaling role, giving voters a novel reason to vote for minor parties who are unlikely to win, or to abstain from voting to protest candidates with undesirable policy positions. In McMurray (2018) I show that the common interest spatial model also extends readily to multiple dimensions, which is a well-known limitation of private interest models. In that paper, logical correlations across issues shape the endogenous bundling of policy positions, which can explain how a single, consistent ideological dimension emerges even when political decisions are inherently multidimensional.

A Appendix

Proof of Lemma 1. In terms of (7), the joint distribution of exactly a votes for candidate A and b votes for candidate B is simply the product

$$\psi\left(a,b|z\right)=\frac{e^{-n\phi(A|z)-n\phi(B|z)}}{a!b!}\left[n\phi\left(A|z\right)\right]^{a}\left[n\phi\left(B|z\right)\right]^{b} \tag{10}$$

of Poisson probabilities. A vote is pivotal if the candidates otherwise tie or if one candidate trails by exactly one vote (and would win the tie-breaking coin toss); in terms of (10), this occurs with the following probability.

$$\begin{aligned}\Pr\left(P|z\right) &= \Pr\left(N_A=N_B|z\right)+\frac{1}{2}\Pr\left(N_A=N_B+1|z\right)+\frac{1}{2}\Pr\left(N_B=N_A+1|z\right)\\ &= \sum_{k=0}^{\infty}\left[\psi\left(k,k|z\right)+\frac{1}{2}\psi\left(k,k+1|z\right)+\frac{1}{2}\psi\left(k+1,k|z\right)\right] \qquad (11)\end{aligned}$$

policy outcome from x_{A_2} to x_B (because A_2 was competitive in the general election while A_1 was not) or from x_B to x_{A_1} (because A_1 was competitive and A_2 was not), or may have no effect on the eventual policy outcome (either because A_1 and A_2 are both destined to prevail over B, or both destined to lose).

In terms of these variables, Lemma 1 of McMurray (2017a) states that the best response is ideological, with the following ideology threshold,

$$\tau^{br} = \frac{\bar{x} - E(z|P)}{E(z^2|P) - \bar{x}E(z|P)} = \frac{\bar{x} - E(z|P)}{V(z|P) + E(z|P)^2 - \bar{x}E(z|P)} \tag{12}$$

This expression depends on the midpoint $\bar{x}$ between the two candidates' platforms and on a voter's expectation

$$E(z|P) = \frac{\int_Z z \Pr(P|z) f(z) \, dz}{\int_Z \Pr(P|z) f(z) \, dz} \tag{13}$$

of the optimal policy, conditional on the event of a pivotal vote.

The proof of Proposition 1 of McMurray (2017a) shows that the best response ideology threshold $\tau^{br}(\tau)$ to an ideological strategy with ideology threshold τ decreases with τ, and using that fact shows that 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 (12) it can be seen that, for any $\tau \in X$, $\tau^{br}(\tau)$ depends on x_A and x_B only through the midpoint $\bar{x}$; accordingly, an ideological strategy v_{τ^*} with ideology threshold τ^* 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 τ^*—constitutes a BNE.) From (12) it is clear that $\tau^{br}(\tau)$ also increases in $\bar{x}$, for any τ; since $\tau^{br}(\tau)$ decreases in τ but increases in $\bar{x}$ for any τ, the fixed point $\tau^* = \tau^{br}(\tau^*)$ increases in $\bar{x}$, as claimed.

For an ideological strategy, (7) can be rewritten as follows.

$$\phi(A|z;\tau) = \int_{-1}^{\tau} \int_Q \int_S 1_{qs=\theta} g(q) \frac{1}{2}(1 + qsz) \, ds dq d\theta \tag{14}$$

$$\phi(B|z;\tau) = \int_{\tau}^{1} \int_Q \int_S 1_{qs=\theta} g(q) \frac{1}{2}(1 + qsz) \, ds dq d\theta \tag{15}$$

From these it is straightforward to show that $\phi(A|-z;-\tau) = \phi(B|z;\tau)$, which by (10) through (13) translates into symmetric pivot probabilities (i.e. $\Pr(P|-z;-\tau) = \Pr(P|-z;\tau)$) and therefore symmetric expectations $E(z|P;-\tau) = -E(z|P;\tau)$ and $E(z^2|P;-\tau) = E(z^2|P;\tau)$. If $\tau^{br}(\tau^*;\bar{x}) = \tau^*$, therefore, then from (12) it is clear that

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

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

Proof of Lemma 2. For any τ, differentiating (14) and (15) with respect to z

yields the following.

$$\begin{aligned}
\frac{\partial \phi (A|z;\tau)}{\partial z} &= \int_{-1}^{\tau} \int_Q \int_S 1_{qs=\theta} g(q) \frac{1}{2}(qs)\, dsdqd\theta = E(\theta|\theta<\tau)\Pr(\theta<\tau) \\
\frac{\partial \phi (B|z;\tau)}{\partial z} &= \int_{\tau}^{1} \int_Q \int_S 1_{qs=\theta} g(q) \frac{1}{2}(qs)\, dsdqd\theta = E(\theta|\theta>\tau)\Pr(\theta>\tau)
\end{aligned}$$

These must sum to $E(\theta)=0$, implying that $\phi(A|z;\tau)$ decreases 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 τ. For any z, (14) and (15) also increase and decrease in τ, respectively, implying that z_τ is -1 for τ sufficiently low and is 1 for τ sufficiently high, and otherwise strictly increases in τ.

As n grows large, $\Pr(P|z)$ decreases to zero for any z, but as Myerson (2000) shows, the magnitude of $\Pr(P|z)$ 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, $f(z|P)$ converges to a degenerate distribution with unit mass on z_τ, implying that $E(z|P)\to z_\tau$ and $V(z|P)\to 0$. Note that z_τ has the same sign as τ, because $\phi_A(z;\tau)$ is increasing in τ but decreasing in z, and $\phi_A(0;0)=\frac{1}{2}$. Therefore, the right-hand side of (12) converges to $\frac{-1}{z_\tau}$, implying that $\lim_{n\to\infty}\tau_n^{br}(\tau,\bar{x}) = \begin{cases} 1 \text{ if } z_\tau<\bar{x} \\ 0 \text{ if } z_\tau=\bar{x} \\ -1 \text{ if } z_\tau>\bar{x} \end{cases}$. Let $\tau_{\bar{x}}$ denote the solution to $z_\tau=\bar{x}$, which is unique since z_τ increases in τ. For any ε there is an n large enough such that $\tau_n^{br}(\tau_{\bar{x}}-\varepsilon)>\tau_{\bar{x}}+\varepsilon$ and $\tau_n^{br}(\tau_{\bar{x}}+\varepsilon)<\tau_{\bar{x}}-\varepsilon$. Since $\tau_n^{br}(\tau)$ decreases in τ, this implies that $\tau_{\bar{x}}-\varepsilon<\tau_n^*<\tau_{\bar{x}}+\varepsilon$. In other words, τ_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. Lemma 1 implies that $(x_A^*, x_B^*, \sigma^*)\in X^2\times\Sigma$ is a PBE only if σ^* is equivalent to the ideological strategy σ_{τ^*} in every subgame with $x_A\neq x_B$. It cannot be the case in equilibrium that x_A^* is closer to θ_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 θ_B, because if that were so then, by moderating her position to θ_B, candidate B could improve her odds of winning, and also her utility conditional on winning. Symmetrically, x_A^* cannot be more extreme than θ_A. Together, these observations imply that $\theta_A\le x_A^*\le x_B^*\le\theta_B$.

Imposing $\beta=0$ and differentiating (5) for candidate B with respect to her own platform yields the following.

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

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 $\left(x_A^{br}, x_B^{br}\right) = \left(-x^{br}, x^{br}\right)$. Restricting attention to symmetric platform pairs $(-x, x)$, 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 σ_{τ^*}) characterizes a PBE. Uniqueness follows because $(x_A, x_B) = (-x, x)$ implies that $\bar{x} = 0$ for any x, so neither $\Pr(w=j|z)$ nor $\frac{\partial}{\partial x_B}\Pr(w=j|z)$ changes with x. Substituting into (16) and differentiating with respect to x therefore yields

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

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

For any $x > 0$, $\Pr(w=B|z=\theta_B)$ approaches 1 as n grows large for policy pairs in any neighborhood of $(-x, x)$. This implies that $\lim_{n\to\infty}\frac{\partial}{\partial x_B}\Pr(w=B|z=\theta_B) = 0$, and therefore that $\lim_{n\to\infty}\frac{\partial EU_B^O}{\partial x_B}$ is positive for any $x < \theta_B$. Thus, $\lim_{n\to\infty} x_{B,n}^* = \theta_B$. By symmetric arguments, $\lim_{n\to\infty} x_{A,n}^* = \theta_A$. ■

Proof of Theorem ??. Setting $\beta = 0$ and differentiating (6) for candidate B with respect to her own platform yields the following (as long as $x_A \neq x_B$, so that, by Lemma 1, voting behavior is uniquely characterized by the ideological strategy σ_{τ^*}),

$$\begin{aligned}
\frac{\partial EU_B^U}{\partial x_B} &= E_z\left[\frac{\partial u(x_B,z)}{\partial x_B}\Pr(w=B|z)\right] + E_z\left[\sum_{j=A,B} u(x_j,z)\frac{\partial \Pr(w=j|z)}{\partial \tau^*(\bar{x})}\frac{\partial \tau^*(\bar{x})}{\partial \bar{x}}\frac{\partial \bar{x}}{\partial x_B}\right] \\
&= E_z\left[2(z-x_B)\Pr(w=B|z)\right] + \frac{\partial E[u(x,z)]}{\partial \tau^*(\bar{x})}\frac{\partial \tau^*(\bar{x})}{\partial \bar{x}}\frac{\partial \bar{x}}{\partial x_B} \\
&= 2\Pr(w=B)\left[E(z|w=B)-x_B\right] \qquad (17)
\end{aligned}$$

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

For any voting strategy, it must be true that $-1 < E(z|w=B) < 1$, which implies that if $x_B = -1$ then (17) is positive and B prefers to move to the right, while if $x_B = 1$ then (17) 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 σ^*) that satisfies $x_B^{br} \neq x_A$ requires that (17) 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, it is straightforward to show that $\Pr(w=B|z)$ increases in z, and therefore that $E(z|w=B) > E(z) = 0$. By symmetric arguments, $E(z|w=A) < 0$.

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 Lemma 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 σ^*, 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 (7) 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 Lemma 3 implies that $\lim_{n\to\infty}\Pr(w=A|z) = 1$. If $z > 0$ then these inequalities are reversed, so $\lim_{n\to\infty}\Pr(w=B|z) = 1$. 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 Lemma 1, the ideological strategy σ_{τ^*} 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 derivative of

(6) generalizes from (16) to the following if candidates are overconfident

$$\begin{aligned}\frac{\partial EU_B^O}{\partial x_B} &= 2\left(\theta_B - x_B\right)\Pr\left(w = B|z = \theta_B\right) \\ &\quad + \left[u\left(x_B, \theta_B\right) - u\left(x_A, \theta_B\right) + \beta\right]\frac{\partial \Pr\left(w = B|z = \theta_B\right)}{\partial \tau^*}\frac{\partial \tau^*\left(\bar{x}\right)}{\partial \bar{x}}\frac{\partial \bar{x}}{\partial x_B} \quad (18)\end{aligned}$$

and generalizes from (17) to the following if candidates are underconfident.

$$\frac{\partial EU_B^U}{\partial x_B} = 2\Pr\left(w = B\right)\left[E\left(z|w = B\right) - x_B\right] + \beta\frac{\partial \Pr\left(w = B\right)}{\partial \tau^*}\frac{\partial \tau^*\left(\bar{x}\right)}{\partial \bar{x}}\frac{\partial \bar{x}}{\partial x_B} \quad (19)$$

In both cases, this derivative decreases in β, 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 platform pair $(x_A, x_B) \in X^2$, 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 any platform $x_B > x_A$ is dominated by $x_B = x_A$. Thus, if $\beta > \bar{\beta}$ then there is no PBE with distinct platforms $x_A < x_B$. The logic of the proof of Theorems 1 and 2 still guarantee that $x_A^* \leq x_B^*$ in equilibrium, though, so $x_A^* = x_B^*$.

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 right of the origin, so that ideological voting makes B less likely to win than A, candidate B can move x_B infinitesimally to the left of x_A instead, and win with greater than $\frac{1}{2}$ probability (with the same policy utility). For the case of underconfident 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 x for which platforms $x_A^* = x_B^* = x$ 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 (by Lemma 1) 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 (18) and (19) are linear in x. In both cases, 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 $x_A^{br} = -x^*$ and $x_B^{br} = x^*$. 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, (18) through (**??**) are all decreasing in β. If $\beta < \bar{\beta}$, therefore, then, as β 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 4. candidates are overconfident with $\beta \neq 0$ then $\frac{\partial EU_B^O}{\partial x_B}$ includes a term that was not present in (16).

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

As before, however, $\lim_{n\to\infty}\frac{\partial}{\partial x_B}\Pr(w = B|z = \theta_B) = 0$ for any pair (x_A, x_B) of distinct platforms. In that case, then, $\lim_{n\to\infty}\frac{\partial EU_B^O}{\partial x_B}$ is positive for any $x_B < \theta_B$, implying that $\lim_{n\to\infty} x^*_{B,n} = \theta_B$ for any sequence $x^*_{B,n}$ of PBE platforms. By symmetric arguments, $\lim_{n\to\infty} x^*_{A,n} = \theta_A$.

If candidates are underconfident then the equilibrium condition (17) generalizes to include an additional term.

$$\frac{\partial EU_B^U}{\partial x_B} = 2\Pr(w = B)[E(z|w = B) - x_B] + \beta\frac{\partial \Pr(w = B)}{\partial \bar{x}}\frac{\partial \bar{x}}{\partial x_B} \tag{20}$$

Symmetric platforms imply that $\tau^*(\bar{x}) = 0$ and therefore that $\Pr(w = B) = \frac{1}{2}$. As the proof of Theorem 2 shows, $E(z|w = B)$ also approaches $E(z|z > 0)$ as n grows large. 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 $\lim_{n\to\infty}\Pr(w = B) = 1 - F(\bar{x}) = \frac{1-\bar{x}}{2}$ and therefore that $\lim_{n\to\infty}\frac{\partial \Pr(w=B)}{\partial \bar{x}} = \frac{\partial}{\partial \bar{x}}\left(\frac{1-\bar{x}}{2}\right) = -\frac{1}{2}$.[66] The right-hand side of (20) therefore approaches $E(z|z > 0) - x^*_{B,\infty} - \frac{1}{4}\beta$, which is zero if and only if $x^*_{B,n}$ approaches $x^*_{B,\infty} = E(z|z > 0) - \frac{1}{4}\beta$. If β exceeds 2 then this expression is negative, implying an equilibrium platform pair at $(0, 0)$ no matter how large the number of voters grows. If truth is binary then, regardless of β, $\frac{\partial \Pr(w=B)}{\partial \bar{x}}$ instead approaches zero, so $\frac{\partial EU_B^U}{\partial x_B}$ approaches $E(z|z > 0) - x^*_{B,\infty}$, which is zero if and only if $x^*_{B,n}$ approaches $x^*_{B,\infty} = E(z|z > 0)$. ■

Proof of Theorem 5. Lemma 3 states that, for any n, an optimal response v_n^* by voters to any pair $(x_A, x_B) \in X^2$ of policy platforms exists and constitutes a BNE in the voting subgame. By Lemma 1, 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 X^2 of

[66] $\Pr\left[w = B; \tau_n^{b*}(\bar{x})\right]$ is continuously differentiable in $\bar{x}$ and n because $\tau_n^{br}(\tau; \bar{x})$ 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 by the implicit function theorem.

platform pairs. Since this set is compact and expected utility is continuous in both platforms, an optimal platform pair $\left(x_{A,n}^*, x_{B,n}^*\right) \in X^2$ exists by the extreme value theorem. Together with any voting strategy σ_n^* that implements $\sigma_{\tau_n^*}\left(x_A^*, x_B^*\right)$ in the appropriate subgame, this constitutes an optimal strategy vector. For the policy platform pair $\left(x_{A,n}^*, x_{B,n}^*\right)$ to maximize expected utility, given the voting strategy σ_n^*, however, $x_{A,n}^*$ must maximize expected utility given $x_{B,n}^*$ and σ_n^*, and $x_{B,n}^*$ must maximize expected utility given $x_{A,n}^*$ and σ_n^*. In other words, $x_{A,n}^*$ and $x_{B,n}^*$ must be equilibrium platforms in a game with candidates who are underconfident, for $\beta = 0$, as claimed.

For any n, let $(x_A, x_B) = (-1, 1)$ and let voters follow the ideological strategy with ideology threshold $\tau = 0$. If truth is binary then, by Lemma 3, $\Pr_n(w = A|z = -1)$ and $\Pr_n(w = B|z = 1)$ 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. ■

Proof of Theorem 6. If candidates are selfishly policy motivated then the differentiating (9) yields the following.

$$\frac{\partial EU_B^P}{\partial x_B} = -2\left(x_B - \hat{x}_B\right)\Pr(w = B) + \left[2\left(x_B - x_A\right)\left(\hat{x}_B - \bar{x}\right) + \beta\right]\frac{\partial \Pr(w = B)}{\partial \bar{x}}\frac{\partial \bar{x}}{\partial x_B} \tag{21}$$

If truth is binary then $\frac{\partial \Pr(w=B)}{\partial \bar{x}}$ approaches 0, so maintaining $\frac{\partial EU_B^P}{\partial x_B} = 0$ requires that $x_{B,n}^*$ approach $\hat{x}_B$ as n grows large. A similar derivation for candidate A implies that $x_{A,n}^*$ approaches $\hat{x}_A$. For any sequence of platform-symmetric equilibria, $\bar{x}_n = 0$ and $\tau^*(\bar{x}_n) = 0$, so $\Pr_n(w = B) = \frac{1}{2}$. If truth is continuous then $\frac{\partial \Pr(w=B)}{\partial \bar{x}}$ approaches $-\frac{1}{2}$ in large elections, as in the proof of Theorem 4, so $\frac{\partial EU_B^P}{\partial x_B}$ approaches $-x_{B,\infty}^* + \hat{x}_B - x_{B,\infty}^*\hat{x}_B - \frac{1}{4}\beta$. For β sufficiently large, this expression is negative for $x_{B,\infty}^* = 0$, implying an equilibrium platform pair at $(0,0)$ no matter how large the electorate grows. Otherwise, $\frac{\partial EU_B^P}{\partial x_B} = 0$ if and only if $x_{B,n}^*$ approaches $x_{B,\infty}^* = \frac{\hat{x}_B - \frac{1}{4}\beta}{1+\hat{x}_B}$. Symmetrically, $\frac{\partial EU_A^P}{\partial x_A} = 0$ if and only if $x_{A,n}^*$ approaches $x_{A,\infty}^* = \frac{\hat{x}_A + \frac{1}{4}\beta}{1-\hat{x}_A}$. For symmetrically selfishly motivated candidates, $\hat{x}_A = -\hat{x}_B$, so these policy positions have equal magnitude. ■

Proof of Theorem 7. If voting is sincere then the probability of voting B in state z is given by the following,

$$\phi(B|z) = \int_0^1 \int_{\bar{x}}^1 \frac{1}{2}(1 + qsz)\, ds dq = \frac{1}{2}(1 - \bar{x}) + \frac{1}{4}\bar{q}z\left(1 - \bar{x}^2\right)$$

where $\bar{q} = E(q)$. As the number of voters grows large, candidate B's realized vote share approaches $\phi(B|z)$. This constitutes a majority if z exceeds $\frac{2\bar{x}}{\bar{q}(1-\bar{x}^2)}$. If a candidate is truth motivated and overconfident with $\theta_B > \frac{2\bar{x}}{\bar{q}(1-\bar{x}^2)}$ then she is certain she will win. In that case, $\lim_{n\to\infty} \frac{\partial}{\partial x_B} \Pr(w = B|z = \theta_B) = 0$ as long as $x_A < x_B < \theta_B$, so $x^*_{B,n}$ approaches θ_B, as in the proof of Theorem 4.

Since candidate B wins in large elections whenever z exceeds $\frac{2\bar{x}}{\bar{q}(1-\bar{x}^2)}$, an underconfident candidate perceives this probability to approach the following when truth is continuous.

$$\lim_{n\to\infty} \Pr(w = B) = \frac{1}{2}\left[1 - \frac{2\bar{x}}{\bar{q}(1-\bar{x}^2)}\right]$$

Differentiating yields the following,

$$\lim_{n\to\infty} \frac{\partial}{\partial \bar{x}} \Pr(w = B) = -\frac{1+\bar{x}^2}{\bar{q}(1-\bar{x}^2)^2}$$

which reduces to $\lim_{n\to\infty} \frac{\partial}{\partial \bar{x}} \Pr(w = B) = -\frac{1}{\bar{q}}$ when candidates' platforms are symmetric. In that case, the limit of (20) reduces to $\lim_{n\to\infty} \frac{\partial EU^O_B}{\partial x_B} = E(z|z > 0) - x_B - \frac{1}{2\bar{q}}\beta = \frac{1}{2} - x_B - \frac{1}{2\bar{q}}\beta$, which has a solution at $\lim_{n\to\infty} x^*_{B,n} = \frac{1}{2} - \frac{1}{2\bar{q}}\beta$. If this expression is negative then platforms converge to $\lim_{n\to\infty} x^*_{A,n} = -\lim_{n\to\infty} x^*_{B,n} = 0$. When truth is binary, z exceeds $\frac{2\bar{x}}{\bar{q}(1-\bar{x}^2)}$ with probability $\frac{1}{2}$. The derivative of this with respect to her own platform is zero, so (20) instead reduces to $\lim_{n\to\infty} \frac{\partial EU^O_B}{\partial x_B} = E(z|z > 0) - x_B$, which has a solution at $\lim_{n\to\infty} x^*_{B,n} = E(z|z > 0) = 1$.

Since $\lim_{n\to\infty} \frac{\partial}{\partial \bar{x}} \Pr(w = B) = -\frac{1}{\bar{q}}$ with continuous truth, (21) approaches $\lim_{n\to\infty} \frac{\partial EU^P_B}{\partial x_B} = -x_B + \hat{x}_B - (4x_B\hat{x}_B + \beta)\frac{1}{2\bar{q}}$ when $x_A = -x_B$. If there is a solution to $\frac{\partial EU^P_B}{\partial x_B} = 0$, therefore, then $\lim_{n\to\infty} x^*_{B,n} = \frac{\hat{x} - \frac{1}{2\bar{q}}\beta}{1 + \frac{2}{\bar{q}}\hat{x}_B}$. If this expression is negative then platforms converge to $\lim_{n\to\infty} x^*_{A,n} = -\lim_{n\to\infty} x^*_{B,n} = 0$. When truth is binary, $\lim_{n\to\infty} \frac{\partial}{\partial \bar{x}} \Pr(w = B) = 0$, as above, so (21) instead approaches $\lim_{n\to\infty} \frac{\partial EU^P_B}{\partial x_B} = -x_B + \hat{x}_B$, which has a unique solution at $\lim_{n\to\infty} x^*_{B,n} = \hat{x}_B$.

If $\hat{x}_A = -\hat{x}_B$ then, in all of the above, $\lim_{n\to\infty} x^*_{A,n} = -\lim_{n\to\infty} x^*_{B,n}$ by symmetric arguments. ■

Proof of Theorem 8. For any n, let $\hat{x}_{m,n}$ denote the realized median of $(\hat{x}_i)^n_{i=1}$. Since voters vote sincerely, candidate B wins the election if and only if $\hat{x}_{m,n} > \bar{x}$. As n grows large, $\hat{x}_m$ converges in distribution to 0, which is the median of the distribution of ideal points, implying that $\lim_{n\to\infty} \Pr(w = B) = \left\{ \begin{array}{l} 0 \text{ if } \bar{x} < 0 \\ 1 \text{ if } \bar{x} > 0 \end{array} \right.$. The election outcome no longer depends on z, so for candidate B, expected utility reduces from (9) to the following,

$$\lim_{n\to\infty} EU^P_B = 1_{\bar{x}>0}\, u(x_A, \hat{x}_B) + 1_{\bar{x}<0}\left[u(x_B, \hat{x}_B) + \beta\right]$$

and expected utility for candidate A is analogous. This is the same utility function faced by a candidate who knows that the median voter is located at 0, so the remainder of the proof follows the logic of Calvert (1985). ■

References

[1] Acemoglu, Daron, Victor Chernozhukov, and Muhamet Yildiz. 2009. "Fragility of asymptotic agreement under Bayesian learning." Working paper, Massachusetts Institute of Technology.

[2] Acharya, Avidit and Adam Meirowitz. 2016. "Sincere Voting in Large Elections." *Games and Economic Behavior*, forthcoming.

[3] Ahn, David S., and Santiago Oliveros. 2016. "Approval Voting and Scoring Rules with Common Values." Working paper, UC Berkeley and University of Essex.

[4] Alesina, Alberto. 1988. "Credibility and Policy Convergence in a Two-Party System with Rational Voters." *American Economic Review*, 78(4): 796-805.

[5] Alesina, Alberto and Howard Rosenthal. 2000. "Polarized Platforms and Moderate Policies with Checks and Balances." *Journal of Public Economics*, 75: 1-20.

[6] Alvarez, R. Michael and Jonathan Nagler. 1995. "Economics, Issues and the Perot Candidacy: Voter Choice in the 1992 Presidential Election." *American Journal of Political Science*, 39(3): 714-744.

[7] American Political Science Association. 1950. "Toward a More Responsible Two-Party System: A Report of the Comittee on Political Parties." *American Political Science Review*, 44(3, Part 2 supplement).

[8] Ansolabehere, Stephen, and James M. Snyder, Jr. 2000. "Valence Politics and Equilibrium in Spatial Election Models." *Public Choice*, 103: 327-336.

[9] 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.

[10] Aragonès, Enriqueta and Dimitrios Xefteris. 2017. "Imperfectly Informed Voters and Strategic Extremism." *International Economic Review*, 58(2): 439-471.

[11] Aranson, Peter H., and Peter C. Ordeshook. 1972. "Spatial Strategies for Sequential Elections," in R. G. Niemi and H. F. Weisberg, eds., *Probability Models of Collective Decision-Making*, Columbus: Merrill.

[12] 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.

[13] Ashworth, Scott and Ethan Bueno de Mesquita. 2009. "Elections with Platform and Valence Competition." *Games and Economic Behavior*, 67: 191-216.

[14] 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.

[15] 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.

[16] Banks, Jeffrey S. 1990. "A Model of Electoral Competition with Incomplete Information." *Journal of Economic Theory*, 50: 309-325.

[17] 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.

[18] Barber, Michael J. and Nolan McCarty. 2013. "Causes and Consequences of Polarization," in *Negotiating Agreement in Politics*, ed. Jane Mansbridge and Cathie Jo Martin. Washington D.C.: American Political Science Association.

[19] Baron, David P. 1994. "Electoral Competition with Informed and Uninformed Voters." *American Political Science Review*, 88(1): 33-47.

[20] Barelli, Paulo, Sourav Bhattacharya, and Lucas Siga. 2017. "On the Possibility of Information Aggregation in Large Elections." Working paper, University of Rochester, Royal Holloway University of London, and New York University.

[21] Bernhardt, Daniel M., John Duggan, and Francesco Squintani. 2009a. "The Case for Responsible Parties." *American Political Science Review*, 103(4): 570-587.

[22] Bernhardt, Daniel M., John Duggan, and Francesco Squintani. 2009b. "Private Polling in Elections and Voter Welfare." *Journal of Economic Theory* 144: 2021-2056.

[23] Bernhardt, Daniel M. and Daniel E. Ingberman. 1985. "Candidate Reputations and the 'Incumbency Effect'". *Journal of Public Economics*, 27: 47-67.

[24] Besley, Timothy, and Stephen Coate. 1997. "An Economic Model of Representative Democracy", *Quarterly Journal of Economics*, 112(1): 85-114.

[25] Bhattacharya, Sourav. 2013. "Preference Monotonicity and Information Aggregation in Elections." *Econometrica*, 81(3): 1229-1247.

[26] Bhattacharya, Sourav. 2018. "Condorcet Jury Theorem in a Spatial Model of Elections." Working paper, Royal Holloway University of London.

[27] 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.

[28] Brusco, Sandro, and Jaideep Roy. 2011. "Aggregate Uncertainty in the Citizen Candidate Model Yields Extremist Parties." *Social Choice and Welfare*, 36: 83-104.

[29] Buera, Francisco J., Alexander Monge-Naranjo, and Giorgio E. Primiceri. 2011. "Learning the Wealth of Nations". *Econometrica*, 79(1): 1-45.

[30] Callander, Steven. 2008. "Political Motivations." *Review of Economic Studies*, 64: 671-697.

[31] Callander, Steven, and Simon Wilkie. "Lies, Damned Lies, and Political Campaigns." *Games and Economic Behavior*, 60: 262-286.

[32] Callander, Steven, and Catherine H. Wilson. 2007. "Turnout, Polarization, and Duverger's Law." *The Journal of Politics*, 69(4): 1045-1056.

[33] Calvert, R. 1985. "Robustness of the Multidimensional Voting Model: Candidate Motivations, Uncertainty, and Convergence." *American Journal of Political Science*, 29: 69-95.

[34] 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.

[35] 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.

[36] Caplan, Bryan. 2007. *The Myth of the Rational Voter: Why Democracies Choose Bad Policies*. Princeton, NJ: Princeston University Press.

[37] Castanheira, Micael. 2003. "Why Vote for Losers?" *Journal of the European Economic Association*, 1(5): 1207-1238.

[38] Coleman, James S. 1972. "The Positions of Political Parties in Elections," in R. G. Niemi and H. F. Weisberg, eds., *Probability Models of Collective Decision-Making*, Columbus: Merrill.

[39] 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.

[40] Coughlin, Peter J. and Shmuel Nitzan. 1981. "Electoral Outcomes with Probabilistic Voting and Nash Social Welfare Maxima." *Journal of Public Economics*, 15: 113-122.

[41] 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.

[42] 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.

[43] de Finetti, Bruno. 1980. "Foresight; its Logical Laws, its Subjective Sources," in *Studies in Subjective Probability*, Eds. H. E. Kyberg and H. E. Smoker, pp. 93-158, New York: Dover.

[44] Dietrich, Franz and Kai Spiekermann. "Epistemic Democracy with Defensible Premises." *Economics and Philosophy*, 29: 87-120.

[45] Downs, Anthony. 1957. *An Economic Theory of Democracy*. New York: Harper and Row.

[46] Duggan, John. 2000. "Equilibrium Equivalence Under Expected Plurality and Probability of Winning Maximization." Working paper, University of Rochester.

[47] 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.

[48] Duggan, John. 2013. "A Survey of Equilibrium Analysis in Spatial Models of Elections." Working paper, University of Rochester.

[49] Eguia, Jon X. 2007. "Citizen Candidates Under Uncertainty." *Social Choice and Welfare*, 29: 317-331.

[50] Enelow, James M. and Melvin J. Hinich. 1989. "A General Probabilistic Spatial Theory of Elections." *Public Choice*, 61(2): 101-113.

[51] Esponda, Ignacio and Emanuel Vespa. 2014. "Hypothetical Thinking and Information Extraction in the Laboratory." *American Economic Journal: Microeconomics*, 6(4): 180-202.

[52] Eyster, Erik and Thomas Kittsteiner. 2007. "Party Platforms in Electoral Competition with Heterogeneous Constituencies." *Theoretical Economics, 2: 41-70.*

[53] Feddersen, Timothy J. and Wolfgang Pesendorfer. 1996. "The Swing Voter's Curse." *American Economic Review*, 86(3): 408-424.

[54] Feddersen, Timothy J. and Wolfgang Pesendorfer. 1997. "Voting Behavior and Information Aggregation in Elections with Private Information." *Econometrica*, 65(5): 1029-1058.

[55] 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.

[56] Fowler, Anthony and Andrew B. Hall. 2016. "The Elusive Quest for Convergence." *Quarterly Journal of Political Science*, 11:131-149.

[57] Galeotti, Andrea and Andrea Mattozzi. 2011. "'Personal Influence': Social Context and Political Competition." *American Economic Journal: Microeconomics*, 3: 307-327.

[58] Garz, Marcel. 2018. "Retirement, Consumption of Political Information, and Political Knowledge." *European Journal of Political Economy*, 53: 109-119.

[59] 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.

[60] Glaeser, Edward L. and Cass R. Sunstein. 2009. "Extremism and Social Learning." *Journal of Legal Analysis*, 1(1): 263-324.

[61] Goren, Paul. 1997. "Political Expertise and Issue Voting in Presidential Elections." *Political Research Quarterly*, 50: 387-412.

[62] Gratton, Gabriele. "Pandering and Electoral Competition." *Games and Economic Behavior* 84: 163-179.

[63] Grosser, Jens and Thomas R. Palfrey. 2014. "Candidate Entry and Political Polarization: An Anitmedian Voter Theorem." *The American Journal of Political Science*, 58(1): 127-143.

[64] Gul, Faruk and Wolfgang Pesendorfer. 2009. "Partisan Politics and Election Failure with Ignorant Voters." *Journal of Economic Theory*, 144: 146-174.

[65] Gul, Faruk and Wolfgang Pesendorfer. 2012. "Media and Policy." Working paper, Princeton University.

[66] Hall, Andrew B. 2015. "What Happens When Extremists Win Primaries?" *American Political Science Review*, 109(1): 18-42.

[67] Hall, Andrew B. and James M. Snyder, Jr. 2015. "Candidate Ideology and Electoral Success." Working paper, Stanford University.

[68] 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.

[69] Harrington, Joseph E., Jr. 1993. "Economic Policy, Economic Performance, and Elections." *The American Economic Review*, 83(1): 27-42.

[70] Heidhues, Paul and Johan Lagerlöf. 2003. "Hiding Information in Electoral Competition." *Games and Economic Behavior*, 42: 48-74.

[71] Hinich, Melvin J. 1977. "Equilibrium in Spatial Voting: The Median Voter Result is an Artifact", *Journal of Economic Theory*, 16(2): 208-219.

[72] 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.

[73] Hotelling, Harold. 1929. "Stability in Competition." *Economic Journal*, 39(153): 41-57.

[74] Jessee, Stephen A. 2009. "Spatial Voting in the 2004 Presidential Election." *American Political Science Review*, 103(1): 59-81.

[75] Jessee, Stephen A. 2010. "Voter Ideology and Candidate Positioning in the 2008 Presidential Election." *American Politics Research*, 38(2): 195-210.

[76] 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.

[77] Kamada, Yuichiro and Fuhito Kojima. 2014. "Voter Preferences, Polarization, and Electoral Policies." *American Economic Journal: Microeconomics*, 6(4): 203-236.

[78] Kartik, Navin, and R. Preston McAfee. 2007. "Signaling Character in Electoral Competition." *American Economic Review*, 97(3): 852-870.

[79] Kartik, Navin, Francesco Squintani, and Katrin Tinn. 2015. "Information Revelation and (Anti-) Pandering in Elections". Working paper, Columbia University.

[80] Kim, Jaehoon and Mark Fey. 2007. "The Swing Voter's Curse with Adversarial Preferences." *Journal of Economic Theory*, 135: 236-252.

[81] Kornhauser, Lewis A. and Lawrence G. Sager. 1986. "Unpacking the Court." *Yale Law Journal*, 96(1): 82-117.

[82] Krasa, Stefan, and Mattias Polborn. 2010. "Competition Between Specialized Candidates." *American Political Science Review*. 104(4): 745-765.

[83] Krasa, Stefan, and Mattias Polborn. 2012. "Political Competition Between Differentiated Candidates." *Games and Economic Behavior*, 76: 249-271.

[84] Krasa, Stefan, and Mattias Polborn. 2015. "Political Competition in Legislative Elections." Working paper, University of Illinois.

[85] Krishna, Vijay, and John Morgan. 2011. "Overcoming Ideological Bias in Elections". *Journal of Political Economy*, 119(2): 183-211.

[86] Krishna, Vijay, and John Morgan. 2012. "On the Benefits of Costly Voting." *Journal of Economic Theory*, forthcoming.

[87] Ladha, Krishna K. 1992. "The Condorcet Jury Theorem, Free Speech, and Correlated Votes," *American Journal of Political Science*, 36(3): 617-634.

[88] 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.

[89] Laslier, Jean-François and Karine Van der Straeten. 2004. "Electoral Competition Under Imperfect Information." *Economic Theory*, 24: 419-446.

[90] Lindbeck, Assar and Jörgen W. Weibull. 1987. "Balanced-Budget Redistribution as the Outcome of Political Competition." *Public Choice*, 52(3): 273-297.

[91] List, Christian. 2012. "The Theory of Judgment Aggregation: An Introductory Review." *Synthese*, 187: 179-207.

[92] List, Christian and Robert E. Goodin. 2001. "Epistemic Democracy: Generalizing the Condorcet Jury Theorem." *The Journal of Political Philosophy*, 9(3): 277-306.

[93] Loertscher, Simon. "Location Choice and Information Transmission." Working paper, University of Melbourne.

[94] Mailath, George J. and Larry Samuelson. 2018. "The Wisdom of a Confused Crowd: Model-Based Inference." Working paper, University of Pennsylvania and Yale University.

[95] Maravall-Rodriguez, Carlos. 2006. "A Spatial Election with Common Values." *Contributions to Theoretical Economics*, 6(1).

[96] Martinelli, Cesar. 2001. "Elections with Privately Informed Parties and Voters." *Public Choice*, 108: 147-167.

[97] Maskin, Eric and Jean Tirole. 2004. "The Politician and the Judge: Accountability in Government". *American Economic Review*, 94(4): 1034-1054.

[98] Matějka, Filip and Guido Tabellini. 2018. "Electoral Competition with Rationally Inattentive Voters." Working Paper, CERGE-EI and Bocconi University.

[99] 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.

[100] McLennan, Andrew. 1998. "Consequences of the Condorcet Jury theorem for Beneficial Information Aggregation by Rational Agents." *American Political Science Review*, 92(2): 413-418.

[101] 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.

[102] McMurray, Joseph C. 2017a. "Ideology as Opinion: A Spatial Model of Common-value Elections." *American Economic Journal: Microeconomics*, 9(4): 108-140.

[103] McMurray, Joseph C. 2017b. "Signaling in Elections: Mandates, Minor Parties, and the Signaling Voter's Curse." *Games and Economic Behavior*, 102: 199-223.

[104] McMurray, Joseph C. 2017c. "Why the Political World is Flat: An Endogenous Left-Right Spectrum in Multidimensional Political Conflict." Working paper, Brigham Young University.

[105] Milgrom, Paul R. and Robert J. Weber. 1982. "A Theory of Auctions and Competitive Bidding." *Econometrica*, 50(5): 1089-1122.

[106] Myerson, Roger. 1998. "Population Uncertainty and Poisson Games." *International Journal of Game Theory*, 27: 375-392.

[107] Myerson, Roger. 2000. "Large Poisson Games." *Journal of Economic Theory*, 94: 7-45.

[108] Myerson, Roger. 2002. "Comparison of Scoring Rules in Poisson Voting Games." *Journal of Economic Theory*, 103: 219-251.

[109] Myerson, Roger B. and Robert J. Weber. 1993. "A Theory of Voting Equilibria." *American Political Science Review*, 87(1): 102-114.

[110] Nitzan, Shmuel and Jacob Paroush. 1982. "Optimal Decision Rules in Uncertain Dichotomous Choice Situations." *International Economic Review*, 23(2): 289-297.

[111] Ortoleva, Pietro and Erik Snowberg. 2015. "Overconfidence in Political Behavior". *American Economic Review*, 105(2): 504-535.

[112] Ortuño-Ortín, Ignacio. "A Spatial Model of Political Competition and Proportional Representation." *Social Choice and Welfare*, 14: 427-438.

[113] Osborne, Martin J., and Al Slivinski. 1996. "A Model of Political Competition with Citizen-Candidates." *Quarterly Journal of Economics*, 111(1): 65-96.

[114] Palfrey, Thomas R. 1984. "Spatial Equilibrium with Entry." *The Review of Economic Studies*, 51(1): 139-156.

[115] 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.

[116] Pivato, Marcus. 2016. "Epistemic Democracy with Correlated Voters." Working paper, Université de Cergy-Pontoise.

[117] Polborn, Mattias and James M. Snyder. 2017. "Party Polarization in Legislatures with Office-Motivated Candidates." *Quarterly Journal of Economics*, 132(3): 1509-1550.

[118] 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.

[119] Prato, Carlo and Staphane Wolton. 2017. "Wisdom of the Crowd? Information Aggregation and Electoral Incentives." Working paper, Columbia University and London School of Economics.

[120] Razin, Ronny. 2003. "Signaling and Election Motivations in a Voting Model with Common Values and Responsive Candidates." *Econometrica*, 71(4): 1083-1119.

[121] Roemer, John E. 1994. "A Theory of Policy Differentiation in Single Issue Electoral Politics," *Social Choice and Welfare*, 11: 355–380.

[122] Roemer, John E. 2004. "Modeling Party Competition in General Elections." Working paper, Yale University.

[123] Romer, Thomas. 1975. "Individual Welfare, Majority Voting, and the Properties of a Linear Income Tax." *Journal of Public Economics*, 4: 163-185.

[124] Schultz, Christian. "Polarization and Inefficient Policies." *Review of Economic Studies* 63: 331-344.

[125] 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." Working paper, University of Chicago.

[126] Sobel, Joel. 2006. “Information Aggregation and Group Decisions.” Working paper, University of California, San Diego.

[127] Sunstein, Cass R. 2002. “The Law of Group Polarization.” *The Journal of Political Philosophy*, 10(2): 175-195.

[128] Tocqueville, Alexis de. 1835. *Democracy in America*. Trans. Henry Reeve. Ed. Phillips Bradley. New York: Alfred A. Knopf. 1945.

[129] Triossi, Matteo. 2013. “Costly Information Acquisition. Is it Better to Toss a Coin?” *Games and Economic Behavior*, 82: 169-191.

[130] Van Weelden, Richard. 2013. “Candidates, Credibility, and Re-election Incentives.” *Review of Economic Studies*, 80: 1622-1651.

[131] Wittman, Donald. 1983. “Candidate Motivation: A Synthesis of Alternative Theories.” *American Political Science Review*, 77(1): 142-157.

[132] 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.

[133] Young, H. Peyton. 1995. “Optimal Voting Rules.” *The Journal of Economic Perspectives*, 9(1): 51-64.

[134] Yuksel, Sevgi. 2018. “Specialized Learning and Political Polarization.” Working paper, UC Santa Barbara.