Polarization and Pandering in Common Interest Elections

Joseph McMurray*

October 17, 2018

Abstract

This paper analyzes a spatial model of common interest elections, meaning that voter differences reflect private estimates of what is best for society, not idiosyncratic tastes. A moderate policy platform makes a candidate more likely to win, but can be socially detrimental. At the same time, an extreme candidate will win if truth is on her side. Because of this, candidates may be highly polarized in equilibrium, even when each cares much more about winning than about the policy outcome. This offers perspective into candidates' empirical reluctance toward political compromise.

JEL Classification Number D72, D82

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

1 Introduction

In one of the earliest applications of mathematics to social science, the French philosopher Condorcet (1785) promoted democracy on the grounds that elections can elicit collective wisdom, pooling information that is dispersed among voters. His "jury" theorem highlights conditions under which majority opinion can correctly identify which of two policies is best for society, even though individual voters cannot.[1]

*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, and Chris Ellis, for their interest and suggestions.

[1]Krishna and Morgan (2011) call this "the first welfare theorem of political economy."

Since then, however, scholars have largely overlooked or dismissed Condorcet's (1785) approach to elections, rejecting the premise that voters share a common interest in the welfare of the group. In particular, when analyzing the policy decisions of politicians, canonical literature relies on spatial election models in which voters hold idiosyncratic policy goals.

In McMurray (2017a) I point out that, though it is true that policies have idiosyncratic effects on voters, large elections can also substantially amplify small altruistic impulses, so that voters effectively approach policy decisions as social planners, and a common interest approach to elections is actually appropriate. That paper shows that Condorcet's (1785) binary decision structure extends naturally to a spatial environment, whether because any policy in an interval might be optimal, or because voter opinions on a binary issue range continuously from strong support for one side to strong support for the other. That paper also uses the spatial common interest model to shed light on a variety of otherwise puzzling voter behavior.[2] Like other papers on information aggregation, however, it specifies the menu of policies exogenously. This paper adds politicians to the analysis, exploring how common interest voting shapes incentives and policy choices of political candidates, and how this is similar to or different from the case of private interest voting.

In general, the incentives that voting behavior generates for candidates depends on candidates' intrinsic motivations. These are not obvious ex ante, so private interest literature considers various possibilities. For example, Hotelling (1929) and Downs (1957) suppose that candidates derive utility from winning office, regardless of any policy compromises that this requires, but Wittman (1977) and Calvert (1985) suppose instead that candidates derive utility from the policy outcome, regardless of who wins the election. In a common interest setting, candidates who care about policy outcomes also might share voters' objectives to do what is socially optimal, or might selfishly pursue some other policy goal at the expense of voters. This paper explores all of these possibilities in turn.

Incentives depend not only on candidates' underlying motivations, but also on their information or beliefs. Private interest models consider two possibilities, which

[2] The importance of information in elections is underscored by the tendency for voters to frequently switch sides on political issues. The common interest approach explains why voters often promote policies that do not favor themselves, expend effort trying to persuade others, and expect their side to win, and why voters who lack confidence in their own information tend to abstain from voting, in deference to others who know more. The spatial geometry explains why the latter group tend empirically to remain politically moderate.

are that candidates know voter preferences exactly, as in Hotelling (1929) and Downs (1957), and that candidates know only the distribution of possible voter preferences, as in the probabilistic voting models of Wittman (1977), Hinich (1978), Calvert (1985), and Lindbeck and Weibull (1987). In a common interest setting, where not even voters know their own policy preferences, information and beliefs play an even greater role. A candidate's beliefs about which policy is socially optimal are important, both because she might wish to implement this policy (depending on her motivation), and because the truth variable influences voter behavior.[3] In addition to any private information of her own, a candidate's behavior must optimally anticipate the private information that voters possess, which will guide their reactions to her policy choice. Similarly, voters should seek to infer candidates' private information from their platform choices, and candidates should infer information from one another. Explicitly modeling all of this would require a description of candidates' beliefs about voters' beliefs about candidates' beliefs, and so on. To keep things tractable, this paper instead models candidate beliefs in stylized ways that avoid such complexities, but illustrate the opposite extremes of putting too much or too little weight on candidates' initial private opinions, thus lending a sense of the range of possible equilibrium behavior.

Repeatedly, private interest literature has found that candidate motivations and information turn out not to matter: results like the *median voter theorem* predict that, in any case, candidates on the left and right should adopt similar or even identical platforms at the political center. This is problematic, however, because empirically, candidates in public elections seem to remain highly polarized. For example, legislative voting by members of the U.S. House, Senate, presidency, and state legislatures exhibits patterns similar to the survey responses of the most extremely liberal and conservative voters in the electorate.[4] 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 saw themselves as weakly more extreme than either candidate.[5] Numerous theories of polarization

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

[4]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).

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

have been explored, but as Section 2 explains, even the most prominent of these are problematic—especially to the extent that candidates value winning—to the point that Roemer (2004) refers to the "tyranny of the median voter theorem." This creates somewhat of a crisis, because in a private interest setting, as Davis and Hinich (1968) make clear, centrist policies maximize social welfare, minimizing the total disutility that voters experience from policies that are far from their ideal points. In that light, empirical polarization constitutes an inexplicable political failure.

Some of the theoretical forces generated by common interest turn out to be similar to those generated by private interest voting. Most notably, moving her policy platform toward her opponent's increases a candidate's vote share, by attracting voters who believe the optimum to lie between the two. This leads to a *median opinion theorem*, predicting that if candidates are sufficiently office motivated then their platforms will coincide. The welfare implication of this can be completely different from the private interest setting, however, because voters don't actually want a policy that matches their current opinions: they want a policy that matches the truth. Convergence can even produce policy outcomes that are known ex ante *not* to be optimal, and in that sense can be seen as a form of *pandering*—that is, doing what is popular instead of what is right. This is reminiscent of the binary models of Canes-Wrone, Herron, and Shotts (2001) and Maskin and Tirole (2004), but the spatial geometry here can explain why political compromise is sometimes viewed with disdain.

In large elections, the logic of the jury theorem guarantees that the candidate whose platform is truly superior will win the election with high probability. Even if she is more polarized than her opponent, therefore, a candidate who believes that truth is on her side expects to win the election. Candidates who over-weight their initial opinions can therefore remain highly polarized in equilibrium, relative to voters. Actually, it turns out that candidates who under-weight their initial opinions polarize substantially, as well, because a novel "pivotal" inference makes them endogenously confident. That candidates who over- and under-weight their own opinions polarize so similarly suggests that Bayesian candidates likely do so, as well, although a fully Bayesian model is not tractable.

Because voter opinions are correlated, a moderate candidate's electoral advantage is small in large elections, consistent with empirical evidence. Knowing this leads candidates to polarize even if they do not care what policy is optimal for voters.

The central message of this paper is therefore that, in contrast with the convergence that arises so robustly in private interest models, polarization emerges as the robust prediction in a common interest setting. In fact, candidates even polarize when they mostly just want to win, and care almost nothing about the policy outcome. Note that this is true polarization, not just non-convergence: for some specifications, candidates polarize to the far ends of the policy space. Such polarization can be good or bad for welfare, depending on the specification, but given the limited appeal of moderation, over-polarization seems to remain the greater danger.

2 Related Literature

With so many existing models, another theory of polarization may seem unnecessary. A complete review of this vast literature is beyond the scope of this paper, but this section explains a number of subtleties, often overlooked, such that what seem like obvious sources of polarization turn out not to be.[6] It also places this paper within the literature on information aggregation in common interest elections.

2.1 Electoral Convergence

The original convergence result of Hotelling (1929) and Downs (1957) assumes that candidates are willing to promise *any* policy outcome in order to win. This level of apathy about policy outcomes is unrealistic. Intuitively, it might seem that selfishly motivated candidates should be unwilling to commit to policies so distant from what they prefer. To exert any control over policy, however, a candidate must win first. Thus, if voting is deterministic, Calvert (1985) and Wittman (1977) show that candidates should cater to the median voter even if winning has no intrinsic benefit.

If the location of the median voter is unknown, as in the probabilistic voting literature,[7] a candidate might still win even when she is further from the political center than her opponent. Policy motivated candidates therefore no longer converge in

[6]Much of the literature is summarized by Duggan (2013).

[7]This includes the stochastic preference specification of Wittman (1983, 1990), Hansson and Stuart (1984), Calvert (1985), and Roemer (1994), the stochastic partisanship specification of Hinich (1978) and Lindbeck and Weibull (1987), and the stochastic valence specification of Bernhardt, Duggan, and Squintani (2009a)

equilibrium, each trading off some probability of winning for a better policy outcome conditional on winning. Political uncertainty seems inevitable, and this trade-off seems natural, so subsequent literature has attributed polarization to this combination of ingredients more frequently than to any other. This is problematic, however, because as Section 5.3 will make clear, probabilistic voting generates only negligible polarization unless uncertainty is severe.[8] Moreover, if candidates are office motivated, they simply converge to the estimated location of the median voter, as various authors have shown.[9]

Another prevalent conjecture is that a candidate polarizes to foster enthusiasm within her "base" of party insiders, extremist supporters, activists, and interest groups, who then finance her campaign, promote her candidacy, nominate her in the primary election, and turn out to vote in the general election. However, this type of theory has at least three weaknesses. First, it assumes that only extreme voters can engage in activism; if moderates can be activists as well, candidates need less support from extremists. Second, the extreme policies that motivate a candidate's own supporters should also make her *opponent's* activists more determined not to lose. Third, party activists should actually *favor* centrist policies, not penalize them: whether they care about winning or about the final policy outcome, they should favor a candidate who is more centrist than her opponent. In light of these considerations, Davis, Hinich, and Ordeshook (1970) show that convergence is robust to endogenous voter participation. In Aranson and Ordeshook's (1972) and Coleman's (1972) models of party primaries, candidates cater not to the median voters within their respective parties, but to the general election median voter, thus polarizing only to the extent of their uncertainty regarding the median voter's location, as in the probabilistic voting models above.[10] Aldrich (1983) and Glaeser, Ponzetto, and Shapiro (2005) show that party activists generate polarization only if they value polarization independent of electoral victory, or observe their own party's platform and not her opponent's; otherwise, candidates cater to the median voter (at least approximately, given probabilistic voting).

[8] Related to this, Calvert (1985), Roemer (1994), and Banks and Duggan (2005) show by continuity that small levels of uncertainty produce only a small degree of polarization.

[9] 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).

[10] Hirano, Snyder, Ansolabehere, and Hansen (2010) and McGhee et al. (2014) show empirically that primary elections seem not to influence polarization.

Another seemingly intuitive source of political extremism is voters' inability to enforce campaign promises: once elected, a candidate can be as extreme as she wishes. To the extent that candidates value reelection, however, Alesina (1988) shows that voters can incentivize moderation by only reelecting centrists. Even if they do not value reelection, of course, candidates might also intrinsically prefer moderate policies. Grosser and Palfrey (2014) point out that ideological extremists should have the greatest incentive to run for office (being the most harmed by adverse policy outcomes), but this assumes voters do not know candidates' true preferences: otherwise, intrinsic moderates have a competitive advantage, by the standard reasoning. With both of these forces at work, the entry models of Osborne and Slivinski (1996) and Besley and Coate (1997) exhibit equilibria with little or no polarization. Polarized equilibria exist as well, but require near-ties (Eguia, 2007), which often do not occur empirically.[11] In any case, if a lack of credibility explains why candidates who pretended to be moderate turn out not to be, it offers no explanation for candidates who openly advocate opposite extremes.

Other theories of polarization have been proposed, including minor party influence;[12] asymmetric ability, charisma, incumbency status, or media exposure;[13] informational asymmetries;[14] efforts to signal hidden types;[15] interactions across jurisdictions or between branches of government;[16] non-policy competition;[17] and convex voter utility.[18] Each of these has merits and weaknesses, but none has been as influential as the theories above, and the continued proliferation of theories attests that existing explanations are unsatisfactory. By comparison with these, the explanation below is also notable for its simplicity: if truth is on her side then a candidate

[11] There is nothing in this literature to systematically favor polarized equilibria. The same is true of the multiple-candidate positioning game studied by Myerson and Weber (1993) where, in equilibrium, strategic voters ignore all but two candidates, who may have any policy positions, polarized or not.

[12] See Palfrey (1984), Castanheira (2003), Callander and Wilson (2007), and Brusco and Roy (2011).

[13] See Bernhardt and Ingberman (1985), Ansolabehere and Snyder (2000), Groseclose, (2001), Aragonès and Palfrey (2002), Gul and Pesendorfer (2009), Soubeyran (2009), Krasa and Polborn (2010, 2012), and Matějka and Tabellini (2015).

[14] See Bernhardt, Duggan, and Squintani (2007), Asako (2014), and Aragonès and Xefteris (2017)

[15] See Banks (1990), Kartik and McAfee (2007), Callander and Wilkie (2007), Callander (2008), and Kartik, Squintani, and Tinn (2012).

[16] See Ortuño-Ortín (1997), Alesina and Rosenthal (2000), Eyster and Kittsteiner (2007), Krasa and Polborn (2015), and Polborn and Snyder (2016)

[17] See Ashworth and Bueno de Mesquita (2009) and Van Weelden (2013).

[18] See Kamada and Kojima (2014).

expects voter support, without the need to moderate. In particular, this does not require special electoral circumstances such as multiple elections, an incumbent, or third party pressure.

2.2 Information

There are three groups of papers that view elections as mechanisms for identifying truly optimal policies. Explicit extensions of Condorcet's (1785) model focus exclusively on voting: specifically, informational efficiency given informational impediments,[19] alternative voting rules,[20] or deviations from common value,[21] and strategic incentives to vote insincerely[22] or abstain.[23] 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, treating candidate behavior 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

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

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

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

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

[23] 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 positionining in a common interest spatial model.[24] 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

There are N voters in an electorate, where, as in Myerson (1998), N is drawn from a Poisson distribution with mean n. There is an interval $X = [-1, 1]$ of policy alternatives, and the electorate must implement one of these, which will then provide a common utility to each voter. The policy z provides the greatest utility possible, but its location is unknown; at the beginning of the game, nature draws z from the domain $Z \subseteq X$. If the state of the world is z but policy x is implemented then each voter receives the following utility,

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

[24] 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.

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

There are two important specifications of this model. The simpler of the two assumes *binary truth*, as in Condorcet's (1785) original model. That is, $Z = \{-1, 1\}$, 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 quite large or quite 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", across policy fields.

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.[25] 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.

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[26] of a voter's signal, drawn

[25] 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.

[26] 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

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. 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. Intuitively, the linearity of (3) does not seem important for any of the results below. However, it seems essential for keeping things tractable once a voter updates his private beliefs to condition on the event of a pivotal vote. 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),

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

and depends on q_i and s_i only through the product $\theta_i = q_i s_i$.[27] Once again, (4) can be interpreted either as 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*.[28] 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, and also point out how this so naturally produces

about the issues. That voters might miscalculate their own competency is an important possibility for future work to explore, but as Sunstein (2002) writes, "it is sensible to say that as a statistical matter, though not an invariable truth, people who are confident are more likely to be right".

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

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

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 voted to choose between two exogenously specified alternatives; here, the menu of policies is endogenous. Specifically, two candidates, A and B, 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]

The analysis below considers two possibile types of policy motivations. One is *welfare motivation*, meaning that candidates do whatever they honestly believe will maximize social welfare. Such public spirit could be intrinsic, or could reflect a more selfish desire to develop a favorable legacy or reputation among voters. Either way, it also makes the model parsimonious, in the sense that candidates are fundamentally no different from ordinary voters (like the "citizen candidates" of Osborne and Slivinski, 1996, and Besley and Coate, 1997). The second possibility is *selfish motivation*, meaning that candidates favor specific policies that privilege themselves (or favored interest groups), regardless of what is best for voters. In addition, candidates with either type of policy motivation might be *office motivated*, meaning that they value the prestige or other perks of winning office, regardless of the policy outcome. Relative to policy utility, the benefit $\beta \geq 0$ of winning could be large or small.

The behavior of candidates depends not only on their preferences, but also on their beliefs about what is optimal for voters—especially if they are welfare motivated. The most straightforward assumption would be that candidates are *Bayesian*, mean-

[29] Allowing abstention here would substantially complicate the model, presumably with results similar to those in McMurray (2017a).

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

ing that they start from the same prior belief as voters, receive private signals of their own, and update according to Bayes' rule. If platforms reflect candidates' private information, however, then the best response to an opponent's platform should try to infer the private information that likely prompted that platform choice. Similarly, voters should infer both candidates' signals in formulating their best-response voting strategies. In anticipating how voters will respond to their platforms, candidates should also infer the private signals that determine voter behavior. Voters should also infer information from one another. In equilibrium, then, a candidate must anticipate what her opponent believes that voters believe that other voters believe that she herself believes, and so on. Such higher order beliefs seem hopelessly intractable, so the analysis below instead considers candidates who are *overconfident* or *underconfident*, meaning that they put too much or too little weight on their private signals. This is modeled in an extreme way that, while less realistic, avoids (some of) the complexities described above. It also deviates from the Bayesian case in opposite directions, in hopes of identifying the range of the possible equilibrium behavior of Bayesian candidates.

For any of the above versions of the model, 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 *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 BNE in the voting subgame associated with every platform pair $(x_A, x_B) \in X^2$, and candidates' platform choices x_A^* and x_B^* maximize the appropriate objectives given their preferences and beliefs, taking the opposing platform and the voting strategy as given. Given the symmetry of the model, it is natural to focus further on equilibria that are *platform-symmetric*, meaning that $x_A^* = -x_B^*$.

4 Voting

4.1 Voting Equilibrium

This section analyzes equilibrium voting in the subgame associated with an arbitrary pair $x_A \leq x_B$ of candidate platforms, slightly extending a similar result from McMurray (2017a). 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,

$$\phi(j|z) = \int_Q \int_S 1_{v(q,s)=j} h(s|q,z)\, g(q)\, ds dq \tag{5}$$

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.

Austen-Smith and Banks (1996) point out that a voter should adopt a strategy that will be optimal in the event that his vote turns out to be *pivotal* (event P), meaning that it reverses the election outcome, even though such an event is unlikely, because otherwise his behavior does not influence his utility. Instead of merely supporting the candidate closest to $E(z|s)$, therefore, a voter optimally supports the candidate who is closest to $E(z|P,s)$. It is this pivotal updating that makes a general model intractable, which is why Section 3 employs such specific functional forms.[31] Given these simplifications, 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, Lemma 1 of McMurray (2017a) states that the best response to any subgame voting strategy is *ideological*, meaning that there is an *ideology threshold* $\tau \in X$ such that voters with ideology left of τ vote A and those with ideology right of τ vote B. Proposition 1 of that paper states the existence and uniqueness of an equilibrium strategy, characterized by the ideology threshold τ^*. Proposition 1 of this paper 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.

Proposition 1 *There exists a unique function $\tau^* : X \to X$ such that for any*

[31] In addition to inferring information from other voters, a citizen should infer whatever he can from candidates' platform choices. Below, however, candidates' information is specified below in a stylized way that conveys nothing useful to voters.

$x_A, x_B \in X$ with midpoint $\bar{x}$ the ideological strategy $v^*(\bar{x})$ characterized by the ideology threshold $\tau^*(\bar{x})$ constitutes a BNE in the voting subgame. For $x_A < x_B$, $v^*(\bar{x})$ is the unique BNE. Moreover, $\frac{d\tau^*(\bar{x})}{d\bar{x}} > 0$ and $\tau^*(-\bar{x}) = -\tau^*(\bar{x})$.

The last part of Proposition 1 states that τ^* is symmetric for symmetric values of $\bar{x}$. If candidates are equidistant from the center so that $\bar{x} = 0$, for example, as they often are in the equilibria analyzed in Section 5, 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, the pivotal voting logic seems entirely foreign to most voters (Esponda and Vespa, 2014). In other settings, this would seem to undermine the theoretical prediction that voters should make inference from the event of a pivotal vote; here, however, voting on the basis of s_i alone is exactly the same as voting on the basis of both P and s_i, at least on the equilibrium path.

4.2 Large Elections

Proposition 1 characterizes equilibrium voting for a fixed population parameter n. Since real-world electorates tend to be very large, the rest of this section analyzes voting behavior in the limit. To this end, first note that the number of votes that each candidate receives depends not only on the voting strategy, but on the realizations of voters' many private signals, which in turn depend on 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 $z_\tau < \bar{x}$, 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 ideology strategy of his peers. It must therefore be the case that, as n grows large, the equilibrium threshold $\tau^*(\bar{x})$ adjusts so that the implied state of the world $z_{\tau^*(\bar{x})}$ leaves voters indifferent between A and B, and therefore willing to follow their signals, as Lemma 1 now states.

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

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

In McMurray (2017a) I show that Condorcet's (1785) binary jury theorem extends in a natural way to this spatial environment: specifically, Proposition 3 of that paper states that the candidate whose platform is closest to the policy that is truly optimal almost surely wins a large election. That normative result applies here, as well, and is relevant in its own right, but also suggests an alternative intuition for the equilibrium balancing predicted above. If $x_A = .9$ and $x_B = 1$, for example, then it is highly likely that x_A is superior to x_B. If voters all merely voted naively for the candidate who seems superior, however, then none would ever vote for candidate B, and A would win *even* in the few states of the world where a B victory is optimal (namely, any state $z > .95$). When $z = .95$ exactly, policies at .9 and at 1 generate equivalent utility. In that case, the median signal realization is approximately .4. For maximal efficiency, therefore, a social planner instructs voters with signals lower than .4 to vote A and instructs voters with signals above .4 to vote B. In this way, A's vote share exceeds 50% precisely when $z < .95$, and B's vote share exceeds 50% precisely when $z > .95$. With common values, as McLennan (1998) points out, the planner's recommendation cannot be improved upon by any individual voter, so this behavior constitutes an equilibrium voting strategy.

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.[32] 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 also expects other voters, after weighing the evidence, to reach the same conclusion. The latter reasoning operates in the analysis below, as well, such that a candidate who believes she is on the side of truth expects to be rewarded with votes.

5 Candidates

Having characterized voters' equilibrium response to any platform pair, this section proceeds to analyze what incentives this creates for candidates in choosing platforms. Let σ_{τ^*} denote the strategy in Σ that induces ideological voting in every subgame, with ideology thresholds given by the function $\tau^*(\bar{x})$ identified in Proposition 1. An equilibrium $(x_A^*, x_B^*, \sigma^*) \in X^2 \times \Sigma$ in the complete game consists of platform positions for both candidates and a voting strategy, and Proposition 1 above implies that a necessary condition for equilibrium is that σ^* corresponds to σ_{τ^*} in every subgame for which $x_A \neq x_B$. Sections 5.1 and 5.1.2 analyze polarization for welfare motivation and selfish policy motivation, respectively, and Section 5.3 illustrates these results with a series of numerical examples. Section 5.4 then considers the case of office motivation.

5.1 Welfare Motivation

Welfare motivated candidates maximize the same objective function (1) as voters. Naturally, their behavior depends on their beliefs about z. The most straightforward

[32] 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.

assumption would be that candidates start from the same prior as voters, receive private signals of their own, and incorporate any additional information using Bayes' rule. However, it is also natural to worry that candidates might be overconfident in their private policy opinions.[33] Section 5.1.1 begins with the latter possibility, and, both to keep the analysis tractable and to provide a stark benchmark, overconfidence is assumed to be extreme. Specifically, candidates are *overconfident*, meaning that each assigns probability one to a particular realization of z, as if her own private opinion were infallible.[34]

Following the analysis of overconfident candidates, Section 5.1.2 considers candidates who are *underconfident*, meaning that they hold correct beliefs and incorporate new information using Bayes' rule. Informally, it is useful to think of candidates as each receiving signals of their own, regarding the location of the optimal policy. Formally adding such signals would tremendously complicate the model and sacrifice tractability, however, so candidate signals are actually not modeled at all. Candidates are thus identical ex ante, although substantially differences arise endogenously in equilibrium, as candidates infer opposite information from voters.

Clearly, it is unrealistic to assume that voters know more about policy decisions than politicians. Candidates are voters themselves, after all, and have career incentives to learn what consequences various policy alternatives will have. One interpretation of the model, however, is that candidates do observe private signals, but place too little (namely, zero) weight on their own opinions, and too much weight on the prior, in contrast with the overconfident candidates who place too little (namely, zero) weight on the prior, and too much weight on their private signals. As polar opposites, however, underconfident and overconfident candidates provide useful benchmarks. Moreover, these candidate types behave surprisingly similarly in equilibrium, suggesting that candidates who place intermediate weight on their signals (including the correct amount of weight) should behave similarly, as well. Given the model's other assumptions, therefore, I conjecture below that explicitly adding can-

[33] Caplan (2007) and Ortoleva and Snowberg (2015) document evidence that overconfidence is rampant among voters. Entry is not modeled here, but it seems intuitive that voters who are the most confident in their policy opinions should be the most inclined to run for office.

[34] In McMurray (2017b) I relax the commitment assumption, and show that welfare-motivated candidates can productively adjust their policy positions ex post, to incorporate information inferred from voting outcomes. The assumption of binding platform commitments may therefore seem inappropriate, as voters have no reason to insist that candidates fulfill their campaign promises. However, a culture of holding candidates to their promises could be a useful safeguard against rogue candidates who do not share voters' preferences.

didate signals to the model would have virtually no impact on equilibrium behavior, even if candidate opinions are less noisy than voters'.

5.1.1 Overconfident Candidates

Overconfident candidates believe themselves to be fully informed about z. Candidate A believes (with probability one) that $z = \theta_A$, while B feels certain that $z = \theta_B$ (where $\theta_A < \theta_B$). The expected utility EU_j^D of candidate $j \in \{A, B\}$ can therefore be written as follows,

$$EU_j^D = \sum_{w=j,-j} u(x_w, \theta_j) \Pr(w|z = \theta_j) + \beta \Pr(w = j|z = \theta_j) \tag{6}$$

where the utility $u(x_w, z)$ associated with the winning candidate's platform and the probability $\Pr(w|z)$ of that candidate winning are both evaluated at $z = \theta_j$. The second term in (6) reflects the possibility of office motivation, which Section 5.4 considers, but for now let $\beta = 0$.

From (6) it is clear that the trade-off faced by overconfident candidates is fundamentally the same as in standard private-value probabilistic voting models, 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. In other words, by standard reasoning, 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; in fact, there is exactly one such equilibrium.

Theorem 1 *If candidates are overconfident with $\beta = 0$ 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 of PBE,* $\lim_{n\to\infty} x_j^* = \theta_j$ *for* $j = A, B$.

While the basic logic of Theorem 1 is quite standard, the extent of polarization is not, as Section 5.3 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 approximately to the median voter, therefore 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, proposing the policies θ_A and θ_B that they most prefer, and not moderating at all. This more dramatic polarization essentially just 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 truth is binary, so that $\theta_A = -1$ and $\theta_B = 1$, and candidates adopt positions at opposite extremes of the policy space.

5.1.2 underconfident Candidates

Underconfident candidates are assumed to start from the same prior belief as voters. With no private information of her own, the expected utility of a underconfident candidate $j \in \{A, B\}$ can be written as follows,

$$EU_j^B = \int_Z \left[\sum_{w=j,-j} u(x_w, z) \Pr(w|z) \right] f(z)\, dz + \beta \Pr(w = j) \tag{7}$$

which differs from (6) in that it now integrates over all possible realizations of z. As before, the second term in (7) reflects the possibility of office motivation, but for now let $\beta = 0$.

Starting with identical prior beliefs, and with no exogenous differences such as incumbency status, ability, or charisma, the basic inclination of underconfident candidates would be to adopt identical platforms at the center of the policy interval. As Theorem 2 states, however, this does not occur in equilibrium: candidates adopt policy positions with opposite signs and, at least in large elections, are highly polarized. With continuous truth, for example, platforms approach $E(z|z<0)=-\frac{1}{2}$ and $E(z|z>0)=\frac{1}{2}$, even though the most extreme voters only favor policies $-\frac{1}{3}$ and $\frac{1}{3}$ (see Footnote 28); with binary truth, candidates polarize to the far extremes $E(z|z<0)=-1$ and $E(z|z>0)=1$ 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, which by Proposition 1 induces symmetric voting; for any n, there is exactly one such equilibrium.[35]

Theorem 2 *If candidates are underconfident with $\beta=0$ 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, there is exactly one platform-symmetric PBE, and for the sequence of these equilibria, $\lim_{n\to\infty}\left(x_{A,n}^*, x_{B,n}^*\right) = (E(z|z<0), E(z|z>0))$.*

It may be surprising that candidates should polarize so substantially, given that candidates are ex ante identical, and given the concavity of (1), which makes candidates risk averse. The key observation underlying Theorem 2 is that, when voting is ideological, candidate A tends to win the election when z turns out to be low and candidate B tends to win when z is high. Upon winning the election, therefore, the two candidates develop opposite posterior beliefs. Intuitively, posterior beliefs might seem irrelevant for platform decisions, which must be made before voting takes place. However, a candidate's platform choice will only affect her utility if she wins; thus, each candidate behaves in a way that will be optimal in the "pivotal" event of winning the election, just as voters restrict attention to the pivotal event of breaking a tie. In a sense, candidates infer information in equilibrium from voters, and incorporate that information into their platform decisions.[36]

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

[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. Candidates would then learn nothing useful about z from the event of winning the election, 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.

As noted above, the model of Section 3 would be more realistic if candidates observed private signals of their own. Unfortunately, explicitly adding candidate signals would severely impede tractability. According to Theorems 1 and 2, however, underconfident candidates and overconfident candidates are similarly polarized, especially if truth is binary. Since these specifications represent the opposite extremes of over- and under-weighting a candidate's private information, it seems reasonable to conjecture that intermediate weighting (including correct weighting) should produce similar behavior, as well. In that sense, omitting candidate signals from the model above actually seems rather inconsequential. A related observation that suggests the same conclusion is that inference based on the event $w = j$ already incorporates the private signals of all N voters. Even if her own signal s_j is much more reliable than a typical voter's, therefore, the expectation $E(z|w = j, s_j)$ based on $N + 1$ pieces of information (including her own) is similar to the expectation $E(z|w = j)$ based on the N voters alone, when N is large.[37]

At least since Austen-Smith and Banks (1996) and Feddersen and Pesendorfer (1996, 1998, 1999), it has been recognized that voters who ignore pivotal considerations will forego utility, and that such considerations can dramatically alter voting incentives. Theorem 2 makes clear that pivotal considerations are equally relevant for candidates, with just as dramatic implications. Underconfident candidates have no private reason for deviating from the center, but nevertheless are just as polarized in equilibrium as candidates who are supremely overconfident: with continuous truth in large elections, for example, underconfident candidates are as polarized as the median overconfident type, and either is more polarized than the most polarized voters; with binary truth, underconfident and overconfident candidates are indistinguishable.

Empirically, of course, it is not clear that candidates react so dramatically to pivotal considerations, any more than it is clear that pivotal considerations are important for voters. It is certainly rare for candidates to describe learning from voters

[37]With explicit candidate signals, a candidate would need to anticipate her opponent's signal s_{-j} in addition to incorporating her own signal s_j into her platform decision, but $E(z|w = j, s_j, s_{-j})$ would of course be similar to $E(z|w = j)$ and $E(z|w = j, s_j)$ when N is large, especially if s_j and s_{-j} have opposite signs, effectively canceling each other out. Formal analysis would be extremely complicated, however, because voters would need to infer both signals as well, in addition to the $N - 1$ peer signals that they already infer in the analysis above, from the event of a pivotal vote. A candidate would then need to discount her inference from voters, recognizing that voting no longer reflects voters' private information alone, but now also reflects voter perceptions of the information that she already possesses. Voters must similarly discount in making inference from one another and from candidates, to avoid duplicating information.

or, for that matter, to acknowledge any possibility that their policy opinions might ultimately prove to be wrong. One reason for this might be that voter information is more complex in the real world than in the model above, so that majority opinion is 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.[38] Another possibility is that candidates are overly confident, and therefore less willing to update their beliefs than they should be. In any case, the analysis above makes clear that, just like voters, candidates who ignore pivotal events will forego utility. Models that take voter and candidate information both seriously should therefore take both pivotal events into account.

5.2 Selfish Motivation

Section 5.1 assumes that candidates want whatever is best for voters. The other possibility, of course, is that they want what is best for themselves, or for some privileged interest group. Accordingly, this section considers candidates who are *selfishly motivated*, meaning that they prefer policies as close as possible to $\hat{x}_A$ and $\hat{x}_B$, regardless of the truth about z. Specifically, utility is still quadratic as in (1), but with z replaced by $\hat{x}_j$, so that expected utility can be written as follows.

$$EU_j^P = \sum_{w=j,-j} u\left(x_w, \hat{x}_j\right) \Pr\left(w\right) + \beta \Pr\left(w = j\right) \tag{8}$$

This is similar to the expression (6) for overconfident candidates, merely evaluating disutility in terms of the distance to $\hat{x}_j$ rather than θ_j, but differs in that a candidate's perceived probability $\Pr(w = j)$ of winning the election no longer depends on her preferred policy position. As before, the second term in (8) reflects the additional possibility of office motivation, which Section 5.4 considers, but for now let $\beta = 0$.

Unlike welfare motivated candidates who are overconfident or underconfident, a private interest candidate has no reason to expect voters to be on her side; to have any influence over policy, she must cater to voters' beliefs. As in the probabilistic

[38] 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 that are not modeled here.

voting literature, however, candidates cannot converge completely in equilibrium: if platforms were identical, a candidate could deviate toward her preferred policy, making herself no worse off if she loses but better off if she wins. Thus, as Theorem 3 now states, $\hat{x}_A < x_A^* < x_B^* < \hat{x}_B$ in equilibrium. If policy preferences are *symmetric*, meaning that $\hat{x}_A = -\hat{x}_B$, then there exists exactly one platform-symmetric equilibrium.

Theorem 3 *If candidates are selfishly motivated with $\hat{x}_A < \hat{x}_B$ and $\beta = 0$ then $(x_A^*, x_B^*, \sigma^*) \in X^2 \times \Sigma$ is a PBE only if $\hat{x}_A < x_A^* < x_B^* < \hat{x}_B$. If policy preferences are symmetric then, for any n, there is exactly one platform-symmetric PBE, and for this sequence of equilibria, $\lim_{n\to\infty} x_j^* = \begin{cases} \frac{\hat{x}_j}{1+|\hat{x}_j|} & \text{if } Z = [-1,1] \\ \hat{x}_j & \text{if } Z = \{-1,1\} \end{cases}$ for $j = A, B$.*

Like the case of welfare motivation, Theorem 3 predicts a high degree of polarization. In large elections, equilibrium platforms are closer to $\hat{x}_A$ and $\hat{x}_B$ than they are to the center. With binary truth, in fact, candidates do not moderate *at all* from their preferred positions. It is straightforward to adapt the proofs of Theorems 1, 2, and 3 to show that candidates polarize not only when they are welfare motivated or selfishly motivated, but also when they have a mixture of these motivations.

5.3 Numerical Examples

On the surface, the logic of Theorems 1 through 3 may seem closely related to the familiar logic of non-convergence in existing private interest probabilistic voting models. The important difference, however, is in degree. In probabilistic voting models, polarization decreases as the number of voters grows large. Asymptotically, platforms convergence completely. This contrasts with Theorems 1 through 3, where candidates polarize even in the limit. The purpose of this section is to demonstrate via numerical examples that, even in small elections, polarization is much more pronounced in a common interest setting.

Table 1 displays equilibrium policy positions for common and private interest elections with variable numbers of voters.[39] Columns 1 through 6 assume common

[39] To facilitate computation, these examples assume that the number $2n+1$ of voters is known and odd, rather than following a Poisson distribution. Thus, the probability of a voter's peers casting a votes for A and b votes for B is given by $\psi(a,b|z) = \frac{2n!}{a!b!}\phi(A|z)^a \phi(B|z)^b$, instead of the function (10) given in the proof of Proposition 1.

values with uniform expertise, for both binary and continuous truth and for candidates who are overconfident (with $\theta_A = -1$ and $\theta_B = 1$), underconfident, and selfishly motivated (with $\hat{x}_A = -1$ and $\hat{x}_B = 1$). There are two canonical specifications of private interest probabilistic voting (Duggan, 2013), which are displayed in columns 7 and 8. In both cases, voter utility $u_i(x) = -(x - \hat{x}_i)^2$ and candidate utility $u_j(x) = -(x - \hat{x}_j)^2$ are quadratic in the distance between the policy outcome and a privately preferred policy $\hat{x}_i$ or $\hat{x}_j$, and candidate ideal points $\hat{x}_A = -1$ and $\hat{x}_B = 1$ are assumed to lie at the far extremes of the policy space. In Column 7, candidates view each voter ideal point $\hat{x}_i$ as an i.i.d. draw from a uniform distribution on $[-1, 1]$, as in the *stochastic preference* model of Wittman (1983) and Calvert (1985). Column 8 assumes the same, but adds a benefit α_i if candidate A is elected (where a negative α_i reflects a preference for B), so that a voter votes A if and only if $u(x_A, \theta_i) + \alpha_i > u(x_B, \theta_i)$, as in the *stochastic partisanship* model of Hinich (1978) and Lindbeck and Weibull (1987). The computations below assume that α_i is an i.i.d. draw from a uniform distribution. For the sake of emphasis, its domain is the interval $[-4, 4]$, making biases so large that even the most extreme voters might vote for a candidate at the opposite end of the policy spectrum.[40]

[40] In general, Duggan (2013) points out that pure strategy equilibria need not exist in private interest probabilistic voting models. However, the specifications adopted here guarantee a unique signal-symmetric equilibrium for any n. Existence follows because expected utility

$$EU_B = u_B(x_A) \Pr(w = A) + u_B(x_B) \Pr(w = B) \tag{9}$$

is continuously differentiable, and if $|x_A| = |x_B| = x$ then $\frac{\partial EU_B}{\partial x_B}$ is positive for $x = 0$ but negative for $x = 1$, and is therefore downward-sloping through zero for some $x \in (0, 1)$. Uniqueness follows because $\frac{\partial}{\partial x} \frac{\partial EU_B}{\partial x_B}$ is negative for any x, implying that if $\frac{\partial EU_B}{\partial x_B}$ is zero for some pair $(-x^*, x^*)$ then it is non-zero for all other $(-x, x)$.

Equilibrium Candidate Positions

Model:	Binary Truth			Continuous Truth			Private Value	
	Dogmatic	Bayesian	Policy Motivated	Dogmatic	Bayesian	Policy Motivated	Stochastic Preference	Stochastic Partisanship
Voters	1	2	3	4	5	6	7	8
1	0.600	0.500	0.500	-	0.083	-	0.500	0.618
3	0.778	0.688	0.675	0.142	0.123	0.104	0.400	0.549
5	0.862	0.793	0.777	0.227	0.152	0.157	0.348	0.511
11	0.958	0.931	0.923	0.395	0.212	0.245	0.270	0.451
15	0.980	0.965	0.961	0.474	0.241	0.282	0.241	0.427
25	0.996	0.993	0.993	-	-	-	0.199	0.389
51	1.000	1.000	1.000	-	-	-	0.149	0.340
101	1.000	1.000	1.000	-	-	-	0.111	0.296
1,001	1.000	1.000	1.000	-	-	-	0.038	0.180
10,001	1.000	1.000	1.000	-	-	-	0.012	0.106
100,001	1.000	1.000	1.000	-	-	-	0.004	0.061
limit	*1.000*	*1.000*	*1.000*	*1.000*	*0.500*	*0.500*	*0.000*	*0.000*

Note: Table entries list the absolute value of candidates' equilibrium platform positions, for various specifications of the model. Omitted cells reflect computational limitations or, in the case of a single voter, a lack of platform-symmetric pure strategy equilibrium.

Table 1

Table 1 makes clear that, across specifications, polarization in common interest models is substantial for any n, and increases as n grows large.[41] For binary truth, this increase is especially rapid: even with only 15 voters, candidates are more extreme than 96% of the electorate. Private interest specifications exhibit some polarization as well, especially with stochastic partisanship, but only in small elections: with at least 100,000 voters, candidates are only more extreme than 6% of voters. Equilibrium platforms never coincide exactly, but converge asymptotically as n grows large.[42] Intuitively, this is because the realized location of the median voter converges to its expectation, so electoral uncertainty vanishes.

It is tempting to dismiss this asymptotic convergence as an artifact of the canonical probabilistic voting specification, which makes the unrealistic assumption that candidates already know the population distributions from which idiosyncratic $\hat{x}_i$ and α_i are drawn. More realistically, candidates must learn about voter preferences from public opinion polls, as in Bernhardt, Duggan, and Squintani (2009b). From that perspective, however, candidates should be eager to learn the median voter's location as precisely as possible, and so should survey voters widely.[43] Similarly, voters should

[41] Computational difficulties limit the size of electorates for which examples can be computed. Limiting policy positions for Poisson N are taken from Theorems 1 through 3.

[42] This can be shown formally by noting that, for any x, $\Pr(w = B)$ equals $\frac{1}{2}$ for $(x_A, x_B) = (-x, x)$ but tends to 1 as n grows large for $(x_A, x_B) = (-x, 0)$. Thus, for n large enough, deviating from $(-x, x)$ to $(-x, 0)$ increases expected utility (9) for candidate B.

[43] Polling has value in common interest settings as well, but less so, because voter opinion will continue to evolve after polling takes place, so outcomes may well deviate from polling predic-

be eager to broadcast their interests as clearly as possible to candidates. Both of these forces should ameliorate uncertainty, making it clear to a candidate that extreme policy positions will only sacrifice the election to her opponent.

Of course, polls will never deliver perfect information, and intuitively, it may seem that slight uncertainty about the location of the median voter should be enough to sustain polarization: after all, if candidates both adopt extreme platforms and the median voter's location is not known exactly then each may still perceive a 50% chance of winning, just as if both had converged. However, that reasoning is incomplete: with only a small amount of uncertainty about the location of the median voter, one of two polarized candidates could secure victory almost certainly, by deviating to the center. With concave utility, winning with high probability at the center should be better than the lottery that yields $-x$ and x with equal probability.[44] Polarization *per se* can only be sustained if uncertainty is almost as severe for a candidate who deviates to the center as for a candidate who remains polarized. This result is natural in a common interest setting, because if truth is against a candidate then moderating will not help, while if truth is on her side then she can win from either policy position.

Some probabilistic voting models sustain polarization by assuming that the *median voter's* ideal point is drawn uniformly from $[-1, 1]$. Within a private interest paradigm, however, such severe uncertainty about voter preferences is implausible: to believe that the median voter might lie at either extreme of the policy space, a candidate must place non-trivial probability on *half* of the electorate being bunched at that extreme. A common interpretation of a random median voter is that the entire distribution of policy preferences shifts with an unknown "valence" advantage for one candidate, reflecting commonly-valued characteristics such as ability or charisma. As the name suggests, however, valence considerations are typically thought to be of secondary importance to voters' policy preferences; if so, they should have limited impact on the median voter's location, and thus limited impact on voter uncertainty.[45] Even if uncertainty were so severe as to warrant the assumption of a uniform median

tions, as in the 2016 U.S. presidential election (www.nytimes.com/2016/11/09/us/politics/hillary-clinton-donald-trump-president.html) and U.K. "Brexit" referendom to leave the European Union (www.nytimes.com/2016/06/25/world/europe/britain-brexit-european-union-referendum.html).

[44] If the median voter is known to lie between $-\varepsilon$ and ε, for instance, symmetrically polarized candidates could each win with 50% probability, but equilibrium positions could not be more polarized than -2ε and 2ε, lest one candidate deviate to the center and win with certainty.

[45] If valence considerations are more important than policy considerations, of course, then the election is essentially common interest, as in the model above.

voter, the scope for polarization would be limited. To see this, reinterpret the lone voter in line 1 of column 7 as the median voter. This produces equilibrium candidate platforms at $-.5$ and $.5$, which are as polarized as the large electorates in columns 5 and 6, but only half as polarized as the electorates of columns 1 through 4.

Another assumption that would exacerbate uncertainty in a private value setting is that $\hat{x}_i$ are *correlated*, rather than i.i.d. As in the common interest specification, this would increase polarization by reducing the benefit of moderation: if one voter is on her side, others likely will be as well, so moderation is unnecessary; alternatively, all might be against her, even if she moderates. On the surface, this might seem to imply that common interest is unnecessary for polarization. However, assuming correlated preferences amounts to *defining* a common interest. Formally, de Finetti's (1990) theorem states that symmetrically correlated (or *exchangeable*) random variables can always 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.

A final observation regarding the comparisons of this section is that the specifications in Table 1 assume no intrinsic benefit from holding office: a candidate desires office only for the sake of achieving desirable policies, so winning on an undesirable policy platform is no better than losing. In contrast, many observers of elections see candidates as highly motivated to win, and willing to make large policy concessions to do so. In a private interest environment, however, 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) variously show that, when candidates are office- or vote-motivated, electoral uncertainty no longer sustains polarization—even in small elections. In contrast, Section 5.4 now shows that polarization in a common interest setting is highly robust to office motivation.

5.4 Office Motivation

The various results above all 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. To allow that possibility, this section specifies candidates' policy preferences as in the previous sections, but assumes that each candidate is also *office* *motivated*, receiving an additional benefit $\beta > 0$ if she wins office. This benefit could be small relative to policy considerations, or could be many times larger than the utility difference between the best and worst policy alternatives. Entry decisions are not modeled here, but it seems reasonable to conjecture that, for candidates who decide to run for office, β might be substantial.

A recurring theme from private interest literature is that office motivation decreases polarization, as candidates increasingly move toward each other to attract voters whose ideal points lie between their platforms.[46] As Theorem 4 now states, the same logic holds in a common interest setting, as well: moving toward each other attracts voters who tentatively believe z to lie between x_A and x_B. In fact, if β is large enough then full convergence occurs. For smaller values of β, full convergence does not occur, but there is a unique platform-symmetric equilibrium (as long as candidates are *symmetric*, meaning that $\theta_A = -\theta_B$ or $\hat{x}_A = -\hat{x}_B$, if applicable), in which polarization decreases with β. Theorem 4 is labeled as the median *opinion* theorem to emphasize that voter ideologies, which are all-important in determining political behavior, are in this model actually only approximations of a more fundamental preference. As Section 6 emphasizes below, this familiar behavior therefore has new implications for social welfare.

Theorem 4 (Median Opinion Theorem) *If candidates are overconfident, underconfident, or selfishly motivated then 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 symmetric then there is a unique platform-symmetric PBE $(x_A^*, x_B^*, \sigma^*) \in X^2 \times \Sigma$ and, in this equilibrium, polarization $|x_B^* - x_A^*|$ strictly decreases in β.*

On its surface, Theorem 4 might seem to suggest that the polarization predicted in Theorems 1 through 3 is not robust: candidates who are not office motivated polarize to varying degrees depending on their specific motivations, but all types of candidates converge when the benefit of winning is sufficiently high. However, this conclusion is premature. Theorem 4 fixes the population size n and considers arbitrarily large values of β; instead, Theorem 5 now fixes β and, for the various candidate types, analyzes behavior in the limit as n grows large. The precise consequences of this depend on the precise specification of candidate incentives, and on whether truth is

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

binary or continuous, but the common theme is that, to varying degrees, candidates polarize substantially even when the benefit of winning office is quite large—in some cases, arbitrarily large.

Theorem 5 *If truth is binary or candidates are overconfident then* $\lim_{n\to\infty}\left(x^*_{A,n}, x^*_{B,n}\right)$ *is the same for all* $\beta \geq 0$. *If truth is continuous then (for the unique sequence of platform-symmetric equilibria)* $\left|x^*_{j,n}\right|$ *approaches* $\min\left\{\frac{1}{2}-\frac{\beta}{4}, 0\right\}$ *if candidates are underconfident and* $\min\left\{\frac{\hat{x}_B-\frac{\beta}{4}}{1+|\hat{x}_B|}, 0\right\}$ *if candidates are symmetrically selfishly motivated.*

Theorem 5 makes clear that, for some specifications of the model, full candidate convergence is still possible, even with large electorates. However, the levels of office motivation that this requires are implausibly large. For example, full convergence by underconfident candidates requires $\beta \geq 2$, meaning that winning the election compensates for a policy outcome that is a distance of $\sqrt{2} \approx 1.4$—which is 70% of the length of the policy interval—from what is optimal. Full convergence by selfishly motivated candidates with ideal points at -1 and 1 requires $\beta \geq 4$, meaning that winning compensates for *any* policy concession. A more plausible value for β might be $\frac{1}{4}$, meaning that the perks of office compensate a candidate for a policy platform that is a distance of $\sqrt{\frac{1}{4}} = \frac{1}{2}$—or 25% of the length of the policy interval—from what she prefers. In that case, underconfident and selfishly motivated candidates adopt policy positions at $\pm\frac{7}{16} \approx .44$ and $\pm\frac{15}{32} \approx .47$, respectively, which are only barely less polarized than the positions $\pm.5$ that (by Theorems 2 and 3) would prevail if β were zero.

As noted above, Harrington (1993) has suggested that the truth underlying voters' ideologies might be fundamentally binary. If this is the case then, according to Theorem 5, platforms in large elections are just as polarized when β is large as they are in Theorems 1 through 3, where $\beta = 0$. The same is true if candidates are overconfident. In particular, if $\beta > 4$ then a candidate is willing to promise *any* policy in order to get elected. Nevertheless, in equilibrium, the candidates take up the same polarized positions they would have chosen if they didn't value winning at all. For overconfident candidates, the logic for this result is fundamentally the same as before: a candidate who is certain that truth is on her side finds policy concessions unnecessary: proposing what she believes to be truly optimal is a strategy that, in large elections, guarantees victory. When truth is binary, candidates of all types come to a similar conclusion, because there are no intermediate states of the world where policy concessions will make a difference.

Taken together, Theorems 4 and 5 make clear that the familiar logic that stronger office motivation should push candidates toward the political center does hold in a common interest environment, but that in large elections the competitive advantage of moderation is minimal: of far greater importance is the "luck" of having public opinion on one's side. As a consequence, candidates remain highly polarized. In fact, for several of the specifications, candidates who desperately want to win are just as polarized as those who do not care at all about winning.

A variety of empirical evidence supports the general findings above. For example, Hall (2015) presents causal evidence that extremism does indeed hurt a candidate in a general election. Nevertheless, as Ansolabehere, Snyder, and Stewart (2001), Canes-Wrone, Brady, and Cogan (2002), and Cohen, McGrath, Aronow, and Zaller (2016) point out, extremists often do prevail over moderates. Primary voters seem to recognize the disadvantage of extremism, but nominate extremists anyway, seeking a bold champion of the truth of their cause over a candidate who timidly appeases the other side. In the 2012 U.S. presidential primaries, for example, many Republicans 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."[47] According to Brady, Han, and Pope (2007) and Hall and Snyder (2013), this pattern is typical: primary elections tend to favor extremists. From a private interest perspective, such self-defeating behavior is inexplicable. Primaries are not formally modeled above, but primary voters are among the most confident voters in the electorate, as I show in McMurray (2017a), and so may well be prone to the same feelings of invincibility that overconfident candidates exhibit in the analysis above. In any case, available evidence seems to support the general view that candidates polarize *because* of the way voters behave, not just in spite of it.

[47] For example, see http://elections.nytimes.com/2012/primaries/states/ohio/exit-polls. In 2016, Bernie Sanders and Ted Cruz were widely viewed as more ideologically extreme than Democrat and Republican front-runners Hillary Clinton and Donald Trump, but still attracted substantial vote shares.

6 Welfare

The analysis above has focused on characterizing candidates' equilibrium behavior. This section analyzes the implications of such behavior for social welfare. Since voters and welfare-motivated candidates share a common objective and since elections are zero-sum for other types of candidates (at least if their preference or belief biases are symmetric), it is uncontroversial to measure welfare simply by the expected utility $E_{N,q,s,w,z}\left[u\left(x^*_{w,n},z\right)\right]$ of an individual voter. This averages over the many realizations of the population size N, the vectors $q=(q_i)_{i=1}^N$ and $s=(s_i)_{i=1}^N$ of private information, the identity w of the election winner, and the truth variable z, where candidates' policy positions $x^*_{j,n}$ depend on the expected number of voters n in combination with the equilibrium voting strategy.

In private interest settings, Davis and Hinich (1968) show that, quite generally, the policy that maximizes social welfare lies near the center of the policy interval.[48] Compromising between the competing interests of the left and right minimizes the total disutility that voters suffer from a policy that is far from their bliss points. In that light, the theoretical prediction that competition for office should drive candidates—who might otherwise prefer extreme policies—toward the political center is sometimes viewed as a sort of "invisible hand" of politics. Puzzled by the level of polarization in the U.S. and elsewhere, however, Ansolabehere, Snyder, and Stewart (2001) concede that, whatever its source, polarization must be interpreted as evidence of some kind of political failure. To the extent possible, this might warrant efforts to increase the benefit β that candidates perceive from winning an election, for example by adjusting office holder salaries. Augmenting the standard private interest model with a common valence shock that shifts the entire distribution of bliss points, Bernhardt, Duggan, and Squintani (2009a) show that voters are best served when candidate platforms do not coincide exactly, so that the policy outcome can be chosen by the median voter after the shock has been realized. Still, if policy preferences are much more important than valence, as is typically assumed, then the optimal level of polarization is low.

In contrast with these private interest models, Proposition 2 now states that, in a common interest environment, the optimal level of polarization can be substantial.

[48] Specifically, if voters' loss functions are linear or quadratic 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.

Specifically, the policy positions that maximize voter welfare are those adopted in equilibrium by underconfident candidates for whom $\beta = 0$. In large elections, as Section 5.1.2 emphasizes, these can be even more extreme than the policies favored by the most extreme voters. The last part of the proposition strengthens the conclusion of Condorcet's (1785) jury theorem: not only does majority voting select the better of two exogenous policies with probability approaching one, but when truth is binary, platforms adjust as n grows large, so that the final policy outcome converges precisely to z.

Proposition 2 *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 is also a PBE. If truth is binary then, for the optimal strategy vector, $|x^*_w - z| \to_{a.s.} 0$.*

The proof of Proposition 2 is a straightforward application of McLennan's (1998) insight that, in common interest environments such as this, socially optimal behavior is also individually optimal. When $\beta = 0$, the incentives of underconfident candidates align perfectly with voters', so the behavior that occurs in equilibrium for that specification of the model is optimal from voters' perspective. Intuitively, the value of polarization here is that it allows voters to tailor their policy choice more closely to the (perceived) state of the world. The optimal level of polarization is much higher than in a private interest model, even with a valence shock, because centrist policies no longer have the same utilitarian appeal, even for voters with moderate opinions: truth might lie in the center, but if it doesn't, voters across the ideological spectrum actually benefit from policies that are extreme.

Regardless of the truth variable, Theorem 4 predicts that competition for office should push candidates toward the center. Remarkably, this is identical to the behavior that arises in private-interest environments, but has opposite implications for welfare. The contrast is most stark for the case of binary truth, where moderate policies are always inferior. When candidates disagree whether economic stimulus should be large or small, for example, competition for votes might produce a compromise policy of moderate stimulus. However, this is known ex ante *not* to be optimal. That candidates should "pander" to voters' mistaken opinions in pursuit of votes is reminiscent of the models of Canes-Wrone, Herron, and Shotts (2001) and Maskin and Tirole (2004), where incumbents adopt bad but popular policies, to ensure reelection. However, those models are binary. With the richer geometry of this spatial environment, pandering takes the specific form of moderation—foregoing the policies that might dramatically improve welfare for the safety of the political center.

The geometry of pandering seems not to have been noted in existing academic literature, but popular political commentary often criticizes politicians for compromising their party's ideals. In the U.S., for example, moderates are sometimes pejoratively labeled DINO or RINO (i.e. Democrats- or Republicans-in-name-only). Green party presidential candidate Ralph Nader famously criticized Republicans and Democrats for being "look alike parties", "Tweedledum and Tweedledee".[49] Recognizing the electoral pressure on candidates to converge, 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 welfare- over office-motivation. Commentators frequently decry extremism, but usually only by the opposite side of the political spectrum: politicians on their own side are more often criticized for being too timid.

Theorem 5 suggests that pandering is actually of limited concern, as candidates polarize substantially even when β is high. In fact, the bigger danger may be *over*-polarization: underconfident candidates adopt optimal policy positions, but overconfident candidates are more polarized than this, which according to Proposition 2 is undesirable.[50] Moreover, this does not depend on β, and so cannot easily be eliminated by making winning more attractive. If truth is binary, as Harrington (1993) suggests it might be for the most fundamental ideological questions, then over-polarization may be negligible: as columns 1-3 of Table 1 make clear, candidates of all types adopt similar policy positions in that case. With continuous truth, however, the distance between underconfident and overconfident candidates (columns 4 and 5 of Table 1) may be substantial.

So far, this discussion has interpreted z as the exact optimal policy. As I explain in McMurray (2017a), another possibility is that there is *aggregate uncertainty*, meaning that $z = E(z^*)$ is only an approximation of an optimal policy z^* that re-

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

[50] According to Theorem 3, policy motivated candidates tend to under-polarize; if θ_j are sufficiently moderate then dogmatic candidates under-polarize, as well.

mains uncertain even after all private information is pooled. In particular, as that paper explains, z may be continuous when z^* is binary, with a moderate z indicating that $z^* = -1$ and $z^* = 1$ are equally likely, while z close to -1 or 1 indicates that $z^* = -1$ of $z^* = 1$ with high probability. A voter with a high-quality moderate signal is then a *skeptic*, favoring moderate policies not because he believes them ultimately to be optimal, or because he lacks information generally, but because he finds the evidence on either side to be unconvincing. This is important because aggregate uncertainty warrants additional caution: the appropriate policy positions for underconfident candidates would then be those listed in column 5 of Table 1, not column 2, even though z^* is binary. Overconfident candidates may still follow the predictions of column 1, however, in essence treating inconclusive evidence as black and white. With aggregate uncertainty, then, overextremism may be a problem even if truth is fundamentally binary.

Empirically, most modern elections tend to be rather close. In close elections, neither an A victory nor a B victory provides strong evidence that the optimal policy is extreme on either side; either the optimal policy is actually moderate, or if there is aggregate uncertainty, the collective evidence remains inconclusive. Either way, candidates' optimal policy positions $E\left(z|w=A\right)$ and $E\left(z|w=B\right)$ should be near to each other and close to the center. Empirical evidence from Section 1 of substantial polarization is therefore again worrisome—not because moderate policy positions are inherently better or more utilitarian, but because they more accurately reflect the collective ambivalence of the electorate. This may vindicate efforts to curb polarization, but if overextremism stems from overconfidence, it may be difficult to curb: increasing office perks could strengthen the desire to attract votes, for example, but according to Theorem 5, this may have negligible effect on candidate behavior.

7 Conclusion

From a private interest perspective, the extreme political polarization observed empirically remains both perplexing and disturbing, reflecting some inexplicable political failure. In particular, probabilistic voting cannot explain polarization when the number of voters is large, or when candidates value winning. This paper has pointed out that, in a common interest setting, polarization arises quite naturally, even when candidates are highly motivated to win. Centrist candidates do have a competitive advantage, but it is only slight, consistent with empirical observation, because with truth on her side, a candidate can handily beat a centrist opponent. Political compromise no longer holds the same utilitarian appeal, and can even produce policies that are known *not* to be optimal, which may be why it is often eschewed.[51] Over-polarization is still a danger, but for a new reason, which is that it may reflect greater confidence than public opinion warrants.[52] Unfortunately, this seems a difficult problem to remedy.

In addition to shedding light on polarization, adding candidates to a model of common interest voting sets the stage for McMurray (2017b), where candidates infer additional information from the size of the margin of victory, matching the popular conception of electoral "mandates" and providing a rationale for protest abstention, and for supporting minor party candidates who are unlikely to win. In McMurray (2017c) I show that the present model also extends readily to multiple dimensions, which is a well known limitation of private interest models, and show how logical correlations across issues shape the endogenous bundling of policy positions, so that a single ideological dimension emerges.

It is ironic to argue that candidates are polarized because voters (and perhaps candidates, as well) have fundamentally unified interests. This is also problematic, in that once voting takes place and the majority decision is made known, minority voters (and perhaps losing candidates, as well) ought to learn that they were wrong, update their opinions, and join the majority. Clearly, this does not happen, which makes it tempting to discard the common interest paradigm and attribute political differences to fundamental, immutable tastes. As I argue in McMurray (2017a), however, this response is inappropriate: empirically, the failure to reach consensus is not limited to settings where interests are at stake; individuals make forecasts about all kinds of future events—political and non-political—and maintain many opinions that they know are unpopular. Whatever the reasons for this, they evidently apply even in settings where preferences are irrelevant.

In reality, of course, the structure of information is much more complex than in the model above, which might explain why the path to consensus is so excruciatingly slow:

[51] That the same behavior can have such opposite welfare consequences is remarkable, and underscores the importance of correctly identifying the right model of politics, since behavior can be observed empirically but welfare can only be assessed through the lens of a model.

[52] 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.

policy decisions depend on a myriad of factors, so resolving disagreements about the main policy question requires determining which of these many factors is the source of disagreement.[53] In contrast with the conditionally independent signals of the model above, voters who communicate with one another before an election may also make correlated errors in judgment. Of course, the updating process could also be plagued by behavioral biases or other cognitive limitations; for example, Ortoleva and Snowberg (2015) present evidence that voters are systematically overconfident. The simple information structure above is useful as a benchmark, but realistic extensions such as the above are important directions for future exploration.[54]

Rendering majority opinion less fallible may well have consequences for polarization. If votes are less strongly correlated with the truth, for example, then polarization may decrease, because overconfident candidates are less certain of winning, underconfident candidates infer less from the pivotal event of winning, and selfishly motivated candidates perceive votes to be less strongly correlated with each other. Alternatively, if public opinion swings less predictably than before, moderate platforms should be even less reliable at attracting voters, so all types of candidates may polarize more. Informational impediments might also interact with electoral institutions: for example, overconfident primary election voters may nominate overly extreme candidates—confident, like the overconfident candidates above, that their nominee will prevail in the general election.

A Appendix

Proof of Proposition 1. In terms of (5), the joint distribution of exactly a votes for candidate A and b votes for candidate B is simply the product

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

[53] 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) then points out 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 would lead an individual to oppose the policy as well, but an individual with access to more information about the individual claims will support the policy. In that sense, individuals with different information may draw opposite conclusions from observing the majority policy decision.

[54] Another valuable extension would be for candidates to learn partial information about public opinion before choosing their policy platforms.

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.

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

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

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

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

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

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 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, the same fixed point $\tau^*(\bar{x})$ characterizes the unique equilibrium response to any pairs of candidate platforms with the midpoint $\bar{x}$. (If $x_A = x_B$ then any voting strategy—including the ideological strategy characterized by $\tau^*(\bar{x})$—constitutes a BNE.) From (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, (5) can be rewritten as follows.

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

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

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 1. For any τ, differentiating (14) and (15) with respect to z yields the following.

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

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 $\tau^{br}(\tau,\bar{x}) \to \begin{cases} 1 \text{ if } z_\tau < \bar{x} \\ 0 \text{ if } z_\tau = \bar{x} \\ -1 \text{ if } z_\tau > \bar{x}\end{cases}$. 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 $\tau^{br}(\tau_{\bar{x}}-\varepsilon) > \tau_{\bar{x}}+\varepsilon$ and $\tau^{br}(\tau_{\bar{x}}+\varepsilon) < \tau_{\bar{x}}-\varepsilon$. Since $\tau^{br}(\tau)$ decreases in 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. Proposition 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 (6) for candidate B with respect to her own platform yields the following.

$$\frac{\partial EU_B^D}{\partial x_B} = -2(x_B - \theta_B)\Pr(w = B|z = \theta_B) + \sum_{j=A,B} u(x_j, \theta_B)\frac{\partial}{\partial x_B}\Pr(w = j|z = \theta_B)$$

$$= 2(\theta_B - x_B)\Pr(w = B|z = \theta_B) \quad (16)$$

$$+[u(x_B, \theta_B) - u(x_A, \theta_B)]\frac{\partial}{\partial x_B}\Pr(w = B|z = \theta_B) \quad (17)$$

The result that $\theta_A \le x_A^* \le x_B^* \le \theta_B$ 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 that neither $\Pr(w = j|z)$ nor $\frac{\partial}{\partial x_B}\Pr(w = j|z)$ changes with x. Substituting into (17) and differentiating with respect to x therefore yields

$$\frac{\partial}{\partial x}\left[\frac{\partial EU_B^D}{\partial x_B}(-x, x)\right] = -x\Pr(w = B|z = \theta_B) + [-2(x - \theta_B) + 2(x + \theta_B)]\frac{\partial}{\partial x_B}\Pr(w = B|z = \theta_B)$$

$$= -x\Pr(w = B|z = \theta_B) + 4\theta_B\frac{\partial}{\partial x_B}\Pr(w = B|z = \theta_B),$$

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

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

$$E[u(x, z)|z = \theta_B] = u(x_A^*, \theta_B)\Pr(w = A|z = \theta_B; x_A = x_A^*, x_B = \theta_B)$$

$$+u(\theta_B, \theta_B)\Pr(w = B|z = \theta_B; x_A = x_A^*, x_B = \theta_B) \quad (18)$$

Since $\lim_{n\to\infty} \Pr(w = B|z = \theta_B; x_A = x_A^*, x_B = \theta_B) = 1$ (by Proposition 3 of McMurray, 2017a), a sequence of such deviations would yield utility $u(\theta_B, \theta_B)$ in the limit. But for every n, $x_B = x_B^*$ is a best response to x_A^*, and so provides weakly greater utility than $x_B = \theta_B$. This implies that equilibrium utility approaches $u(\theta_B, \theta_B)$ as well. This is possible only if $\lim_{n\to\infty} x_B^* = \theta_B$. By symmetric arguments, $\lim_{n\to\infty} x_A^* = \theta_A$ as well. ■

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

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

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

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

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

That $\phi(A|z;\tau)$ and $\phi(B|z;\tau)$ increase and decrease in τ, respectively, for any z, is clear from (14) and (15). These expressions can also be rewritten as follows,

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

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

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

exceeds 1 and increases in τ, so by the first part of the lemma, $\frac{\Pr(w=B|z)}{\Pr(w=B|-z)}$ exceeds 1 and increases in τ as well. The latter implies that, in states with magnitude $|z|$, the expectation $E(z|w = B, |z| = \bar{z}) = \frac{\bar{z}\Pr(w=B|\bar{z}) - \bar{z}\Pr(w=B|-\bar{z})}{\Pr(w=B|\bar{z}) + \Pr(w=B|-\bar{z})}$ is positive and increases in τ. Integrating over $\bar{z}$, $E(z|w = B)$ is positive for any τ and also increases in τ.

Symmetric arguments establish that $E(z|w=A)$ also increases in τ as well, but is negative for any τ. ■

Proof of Theorem 2. Setting $\beta = 0$ and differentiating (7) for candidate B with respect to her own platform yields the following (as long as $x_A \neq x_B$, so that, by Proposition 1, voting behavior is uniquely characterized by the ideological strategy σ_{τ^*}),

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

where the final equality follows because, by Proposition 3 of McMurray (2017a), 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 (19) is positive and B prefers to move to the right, while if $x_B = 1$ then (19) is negative and B prefers to move to the left. Thus, a best response x_B^{br} to x_A (and to the equilibrium voting response σ^*) that satisfies $x_B^{br} \neq x_A$ requires that (19) equal zero, which is the case if and only if $x_B^{br} = E(z|w=B)$. Similarly, a best response $x_A^{br} \neq x_B$ requires that $x_A^{br} = E(z|w=A)$. If $x_A = x_B$ then voting need not be ideological, but non-ideological voting cannot produce higher utility. Thus, $x_j^{br} = E(z|w=j)$ is the best response for either candidate, and $x_j^* = E(z|w=j)$ for $j = A, B$ is a necessary condition for a PBE. With ideological voting, Lemma A1 implies that $E(z|w=A) < 0 < E(z|w=B)$.[55]

For any pair of symmetric platforms $x_A = -x_B$, the midpoint $\bar{x} = 0$ lies exactly at the center of the policy interval, so by Proposition 1, voters' equilibrium response is characterized by an ideological strategy with ideology threshold $\tau^*(0) = 0$. By the symmetry of the model, this implies that candidates form symmetric expectations $E(z|w=A) = -E(z|w=B)$, and therefore symmetric platforms $x_A^* = -x_B^*$. Together with the equilibrium voting strategy 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 (5) therefore reduces to $\phi(A|z) = \Pr(s<0|z)$ and $\phi(B|z) = \Pr(s>0|z)$. If $z < 0$, therefore, then $\phi(A|z) > \frac{1}{2} > \phi(B|z)$, so Proposition 3 of McMurray (2017a) implies that $\lim_{n\to\infty}\Pr(w=A|z) = 1$, while if $z > 0$ then these inequalities are reversed, so $\lim_{n\to\infty}\Pr(w=B|z) = 1$. Thus, $f(z|w=A)$ and

[55] These inequalities could also be reversed in equilibrium, but then candidates could be relabeled so that A is to the left of B, as the statement of the theorem presumes.

$f(z|w=B)$ converge to $f(z|z<0)$ and $f(z|z>0)$, respectively, and $E(z|w=A)$ and $E(z|w=B)$ therefore converge to $E(z|z<0)$ and $E(z|z>0)$. ■

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

Candidate B's expected utility can be rewritten from (8) as follows,

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

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

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

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

To see the uniqueness of symmetric equilibrium platforms when $\hat{x}_A = -\hat{x}_B$, suppose that $(x_A, x_B) = (-x, x)$ for some $x \geq 0$. As x changes, $\Pr(w=B)$ and $\frac{\partial \Pr(w=B)}{\partial \bar{x}}$ do not change, but $\frac{\partial u(x_B = x, \hat{x}_B)}{\partial x_B}$ and $u(x_B, \hat{x}_B) - u(x_A, \hat{x}_B)$ are both linear in x. This implies that the right-hand side of (20) is linear in x as well, and therefore equals zero for a unique value x^*. Since candidate A's incentives are perfectly symmetric, $(x_A^*, x_B^*) = (-x^*, x^*)$ constitutes the unique pair of symmetric equilibrium platforms.[56] According to Lemma 1, $\tau_n^*(\bar{x})$ approaches the solution $\tau_\infty^*(\bar{x})$ to $z_\tau = \bar{x}$. That is, in the limit as n grows large, if $z = \bar{x}$ then the expected vote shares $\phi[A|z=\bar{x}; \tau_\infty^*(\bar{x})] = \phi[B|z=\bar{x}; \tau_\infty^*(\bar{x})]$ for the two candidates will be exactly the

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

same; if z turns out to be less than $\bar{x}$ then $\phi[A|z;\tau^*_\infty(\bar{x})] < \phi[B|z;\tau^*_\infty(\bar{x})]$ and if z turns out to exceed $\bar{x}$ then $\phi[A|z;\tau^*_\infty(\bar{x})] > \phi[B|z;\tau^*_\infty(\bar{x})]$. In a large election, the candidate with the larger expected vote share almost surely wins (see Myerson, 2002). ■

Proof. 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 $\lim_{n\to\infty} \frac{\partial \Pr(w=B)}{\partial \bar{x}} = \frac{\partial}{\partial \bar{x}}\left(\frac{1-\bar{x}}{2}\right) = -\frac{1}{2}$.[57] In a platform-symmetric equilibrium $x_A = -x_B$ and $\bar{x} = 0$, so (20) therefore converges to the following,

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

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

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

Proof of Theorem 4. According to Proposition 1, the ideological strategy 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 (7) generalizes from (17) to the following if candidates are overconfident,

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

generalizes from (19) to the following if candidates are underconfident,

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

[57] $\Pr\left[w = B; \tau^{b*}_n(\bar{x})\right]$ is continuously differentiable in $\bar{x}$ and n because $\tau^{br}_n(\tau; \bar{x})$ is continuously differentiable in $\bar{x}$ and n, so the solution $\tau^*_n(\bar{x})$ to the fixed point problem $\tau = \tau^{br}_n(\tau; \bar{x})$ is continuously differentiable in $\bar{x}$ and n by the implicit function theorem.

and generalizes from (20) to the following if candidates are selfishly motivated.

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

In all three 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 pair $(x_A, x_B) \in X^2$ of platforms, the other terms in the expression for the derivative are finite, so there exists a threshold $\bar{\beta}_{x_A,x_B}$ sufficiently large that for all $\beta > \bar{\beta}_{x_A,x_B}$ the derivative is negative. The set of platform pairs is compact and $\bar{\beta}_{x_A,x_B}$ is continuous in the platform pair, so there exists a maximum $\bar{\beta} = \max_{(x_A,x_B)} \bar{\beta}_{x_A,x_B}$, and for any $\beta > \bar{\beta}$ the derivative is negative, meaning that candidate B does not want to move x_B away from x_A. Symmetrically, A does not want to move x_A away from x_B. Thus, if $\beta > \bar{\beta}$ then there is no PBE with distinct platforms $x_A < x_B$.

While neither candidate wishes to move away from her opponent, a candidate might have incentive to "leap frog" her opponent, to attract more votes: if the two candidates converge to a position in which B wins with probability lower than one-half then B can move x_B just to the left of x_A and win with greater than $\frac{1}{2}$ probability instead. For the cases of underconfident or selfishly motivated candidates, this implies that the unique PBE is $x_A^* = x_B^* = 0$ (together with the voting strategy $\sigma^* = \sigma_{\tau^*}$). For the case of overconfident candidates, there is a range of 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 the voter response threshold is $\tau^*(\bar{x}) = 0$, regardless of the magnitude of x, implying that $\Pr(w=B)$ and $\frac{\partial \Pr(w=B)}{\partial \tau^*}\frac{\partial \tau^*}{\partial \bar{x}}\frac{\partial \bar{x}}{\partial x_B}$ do not depend on the magnitude of x. The utility differences $u(x,\theta_B) - u(-x,\theta_B) = 4\theta_B x$ and $u(x,\hat{x}_B) - u(-x,\hat{x}_B) = 4\hat{x}_B x$ are linear in x, implying that (22) through (24) are linear in x. For each of these motivations, therefore, there is a unique $x^* \in [0,1]$ such that the best responses for A and B, respectively, to any pair $(-x, x)$ of symmetric platforms are $x_A^* = -x^*$ and $x_B^* = x^*$. Thus, $(x_A^*, x_B^*) = (-x^*, x^*)$ (together with $\sigma^* = \sigma_{\tau^*}$) constitute the unique PBE with symmetric platforms. Since $\frac{\partial \Pr(w=B|z)}{\partial \tau}$ is negative and $\frac{\partial \tau(\bar{x})}{\partial \bar{x}}$ and $\frac{\partial \bar{x}}{\partial x_B}$ are positive, (22) through (24) are all decreasing in 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 5. If candidates are overconfident then, for any n, the utility (18) derived by deviating to $x_B = \theta_B$ in response to candidate A's equilibrium platform $x^*_{A,n}$ generalizes to include an additional term.

$$\begin{aligned}\frac{\partial EU^D_B}{\partial x_B} &= u\left(x^*_{A,n}, \theta_B\right) \Pr\left(w = A | z = \theta_B; x_A = x^*_{A,n}, x_B = \theta_B\right) \\ &\quad + \left[u\left(\theta_B, \theta_B\right) + \beta\right] \Pr\left(w = B | z = \theta_B; x_A = x^*_{A,n}, x_B = \theta_B\right)\end{aligned}$$

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

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

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

Symmetric platforms imply that $\tau^*\left(\bar{x}\right) = 0$ and therefore that $\Pr\left(w = B\right) = \frac{1}{2}$. As the proof of Theorem 2 shows, $E\left(z | w = B\right)$ also approaches $E\left(z | z > 0\right)$ as n grows large. If truth is continuous then the proof of Theorem 3 shows that $\frac{\partial \Pr(w=B)}{\partial \bar{x}}$ approaches $-\frac{1}{2}$ as n grows large, so $\frac{\partial EU^B_B}{\partial x_B}$ approaches $E\left(z | z > 0\right) - x^*_{B,\infty} - \frac{1}{4}\beta$, which is zero if and only if $x^*_{B,n}$ approaches $x^*_{B,\infty} = E\left(z | z > 0\right) - \frac{1}{4}\beta$. If β is sufficiently large 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 $\frac{\partial EU^B_B}{\partial x_B}$ approaches $E\left(z | z > 0\right) - x^*_{B,\infty}$, which is zero if and only if $x^*_{B,n}$ approaches $x^*_{B,\infty} = E\left(z | z > 0\right)$.

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

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

If truth is binary then $\frac{\partial \Pr(w=B)}{\partial \bar{x}}$ converges to 0, so maintaining $\frac{\partial EU^P_B}{\partial x_B} = 0$ requires that $x^*_{B,n}$ approach $\hat{x}_B$. A similar derivation for candidate A implies that $x^*_{A,n}$ approaches $\hat{x}_A$. For any sequence of platform-symmetric equilibria, $\bar{x}_n = 0$ and $\tau^*\left(\bar{x}_n\right) = 0$, so $\Pr_n\left(w = B\right) = \frac{1}{2}$. If truth is continuous then $\frac{\partial \Pr(w=B)}{\partial \bar{x}}$ converges in large elections to $-\frac{1}{2}$, as above, so $\frac{\partial EU^P_B}{\partial x_B}$ 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 Proposition 2. Drawing on the common-value logic of McLennan (1998), Proposition 3 of McMurray (2017a) 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 Proposition 1 of this paper, therefore, v_n^* is given by the ideological strategy $\sigma_{\tau_n^*}(x_A, x_B)$, evaluated at the platform pair. The optimal combination of voter and candidate behavior can then be obtained by maximizing over the set X^2 of 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 $\sigma_{\tau_n^*}(x_A^*, x_B^*)$ 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 $\tau = 0$. In that case, if truth is binary, then, by Proposition 3 of McMurray (2017a), $\Pr_n(w = A|z = -1)$ 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. ■

References

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

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

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

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

[5] 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.

[6] 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).

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

[8] Ansolabehere, Stephen, James M. Snyder, Jr. and Charles Stewart, III. 2001. "Candidate Positioning in U.S. House Elections." *American Journal of Political Science*, 45(1): 136-159.

[9] Aragonès, Enriqueta and Thomas R. Palfrey. 2004. "Mixed Equilibrium in a Downsian Model with a Favored Candidate." *Journal of Economic Theory*, 103: 131-161.

[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] 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.

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

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

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

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

[23] 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.

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

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

[26] Callander, Steven. 2008. “Political Motivations.” *Review of Economic Studies*, 64: 671-697.

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

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

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

[30] 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.

[31] 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.

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

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

[34] 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.

[35] 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.

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

[37] 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.

[38] 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.

[39] de Fenetti, Bruno. 1990. *Theory of Probability*, vol. 2, New York: Wiley.

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

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

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

[43] 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.

[44] Duggan, John. 2013. "A Survey of Equilibrium Analysis in Spatial Models of Elections." Manuscript, University of Rochester.

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

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

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

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

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

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

[51] 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.

[52] Feddersen, Timothy J. and Wolfgang Pesendorfer. 1999. "Abstention in Elections with Asymmetric Information and Diverse Preferences." *The American Political Science Review*, 93(2): 381-398.

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

[54] 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.

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

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

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

[58] Groseclose, Tim. 2001. "A Model of Candidate Location When One Candidate Has a Valence Advantage." *American Journal of Political Science*, 45(4): 862-886.

[59] 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.

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

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

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

[63] 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.

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

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

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

[67] 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.

[68] Hirano, Shigeo, James M. Snyder, Jr., Stephen Ansolabehere, and John Mark Hansen. 2010. "Primary Elections and Partisan Polarization in the U.S. Congress." *Quarterly Journal of Political Science*, 5: 169-191.

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

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

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

[72] 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.

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

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

[75] Kartik, Navin, Francesco Squintani, and Katrin Tinn. 2013. "Information Revelation and (Anti-) Pandering in Elections". Manuscript, Columbia University.

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

[77] Kornhauser, Lewis A. and Lawrence G. Sager. 1986. "Unpacking the Court." *Yale Law Journal*, 96(1): 82-117.
[78] Krasa, Stefan, and Mattias Polborn. 2010. "Competition Between Specialized Candidates." *American Political Science Review*. 104(4): 745-765.
[79] Krasa, Stefan, and Mattias Polborn. 2012. "Political Competition Between Differentiated Candidates." *Games and Economic Behavior*, 76: 249-271.
[80] Krasa, Stefan, and Mattias Polborn. 2015. "Political Competition in Legislative Elections." Manuscript, University of Illinois.
[81] Krishna, Vijay, and John Morgan. 2011. "Overcoming Ideological Bias in Elections". *Journal of Political Economy*, 119(2): 183-211.
[82] Krishna, Vijay, and John Morgan. 2012. "On the Benefits of Costly Voting." *Journal of Economic Theory*, forthcoming.
[83] Ladha, Krishna K. 1992. "The Condorcet Jury Theorem, Free Speech, and Correlated Votes," *American Journal of Political Science*, 36(3): 617-634.
[84] 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.
[85] Laslier, Jean-François and Karine Van der Straeten. 2004. "Electoral Competition Under Imperfect Information." *Economic Theory*, 24: 419-446.
[86] Lindbeck, Assar and Jörgen W. Weibull. 1987. "Balanced-Budget Redistribution as the Outcome of Political Competition." *Public Choice*, 52(3): 273-297.
[87] List, Christian. 2012. "The Theory of Judgment Aggregation: An Introductory Review." *Synthese*, 187: 179-207.
[88] List, Christian and Robert E. Goodin. 2001. "Epistemic Democracy: Generalizing the Condorcet Jury Theorem." *The Journal of Political Philosophy*, 9(3): 277-306.
[89] Loertscher, Simon. "Location Choice and Information Transmission." Working paper, University of Melbourne.
[90] Martinelli, Cesar. 2001. ""Elections with Privately Informed Parties and Voters." *Public Choice*, 108: 147-167.
[91] Maskin, Eric and Jean Tirole. 2004. "The Politician and the Judge: Accountability in Government". *American Economic Review*, 94(4): 1034-1054.
[92] Matějka, Filip and Guido Tabellini. 2015. "Electoral Competition with Rationally Inattentive Voters." CERGE-EI Working Paper Series # 552.
[93] 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.
[94] McGhee, Eric, Seth Masket, Boris Shor, Steven Rogers, and Nolan McCarty. 2014. "A Primary Cause of Partisanship? Nomination Systems and Legislator Ideology." *American Journal of Political Science*, 58(2): 337-351.
[95] McLennan, Andrew. 1998. "Consequences of the Condorcet Jury theorem for Beneficial Information Aggregation by Rational Agents." *American Political Science Review*, 92(2): 413-418.

[96] 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.
[97] McMurray, Joseph C. 2017a. "Ideology as Opinion: A Spatial Model of Common-value Elections." *American Economic Journal: Microeconomics*, 9(4): 108-140.
[98] McMurray, Joseph C. 2017b. "Signaling in Elections: Mandates, Minor Parties, and the Signaling Voter's Curse." *Games and Economic Behavior*, 102: 199-223.
[99] McMurray, Joseph C. 2017c. "Why the Political World is Flat: An Endogenous Left-Right Spectrum in Multidimensional Political Conflict." Manuscript, Brigham Young University.
[100] Milgrom, Paul R. and Robert J. Weber. 1982. "A Theory of Auctions and Competitive Bidding." *Econometrica*, 50(5): 1089-1122.
[101] Myerson, Roger. 1998. "Population Uncertainty and Poisson Games." *International Journal of Game Theory*, 27: 375-392.
[102] Myerson, Roger. 2000. "Large Poisson Games." *Journal of Economic Theory*, 94: 7-45.
[103] Myerson, Roger. 2002. "Comparison of Scoring Rules in Poisson Voting Games." *Journal of Economic Theory*, 103: 219-251.
[104] Myerson, Roger B. and Robert J. Weber. 1993. "A Theory of Voting Equilibria." *American Political Science Review*, 87(1): 102-114.
[105] Nitzan, Shmuel and Jacob Paroush. 1982. "Optimal Decision Rules in Uncertain Dichotomous Choice Situations." *International Economic Review*, 23(2): 289-297.
[106] Ortoleva, Pietro and Erik Snowberg. 2015. "Overconfidence in Political Behavior". *American Economic Review*, 105(2): 504-535.
[107] Ortuño-Ortín, Ignacio. "A Spatial Model of Political Competition and Proportional Representation." *Social Choice and Welfare*, 14: 427-438.
[108] Osborne, Martin J., and Al Slivinski. 1996. "A Model of Political Competition with Citizen-Candidates." *Quarterly Journal of Economics*, 111(1): 65-96.
[109] Palfrey, Thomas R. 1984. "Spatial Equilibrium with Entry." *The Review of Economic Studies*, 51(1): 139-156.
[110] 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.
[111] Pivato, Marcus. 2016. "Epistemic Democracy with Correlated Voters." Manuscript, Université de Cergy-Pontoise.
[112] Polborn, Mattias and James M. Snyder. 2016. "Intra-Party Disagreement and Inter-Party Polarization." Manuscript, University of Illinois and Harvard University.
[113] 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.
[114] Prato, Carlo and Staphane Wolton. 2017. "Wisdom of the Crowd? Information Aggregation and Electoral Incentives." Working paper, Columbia University and London School of Economics.
[115] Razin, Ronny. 2003. "Signaling and Election Motivations in a Voting Model with Common Values and Responsive Candidates." *Econometrica*, 71(4): 1083-1119.

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

[117] Roemer, John E. 2004. "Modeling Party Competition in General Elections." Manuscript, Yale University.

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

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

[120] Shor, Boris. 2011. "All Together Now: Putting Congress, State Legislatures, and Individuals in a Common Ideological Space to Assess Representation at the Macro and Micro Levels." Manuscript, University of Chicago.

[121] Sobel, Joel. 2006. "Information Aggregation and Group Decisions." Manuscript, University of California, San Diego.

[122] Soubeyran, Raphaël. 2009. "Does a Disadvantaged Candidate Choose an Extremist Position?" *Annals of Economics and Statistics*, 93/94: 327-348.

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

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

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

[126] Van Weelden, Richard. "Candidates, Credibility, and Re-election Incentives." *Review of Economic Studies*, 80: 1622-1651.

[127] Wittman, Donald. 1977. "Candidates with Policy Preferences: A Dynamic Model." *Journal of Economic Theory*, 14: 180-89.

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

[129] 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.

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