Polarization and Pandering in Common-Interest Elections

Joseph McMurray*

October 1, 2020

Abstract

Adding candidates to a one-dimensional common-interest voting model, this paper shows that catering to centrist voters can come at the expense of social welfare. The electoral advantage of doing so is weak, however, resolving the long-standing puzzle of why candidates remain so polarized, empirically.

JEL Classification Number D72, D82

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

1 Introduction

A prominent empirical feature of elections is polarization: across twelve U.S. presidential elections from 1972 to 2016, for example, 89% of Americans rated both major candidates as weakly more extreme than they rated themselves on a seven-point liberal-conservative scale, while only 12% rated both candidates as weakly more moderate.[1] Statistical analyses of the U.S. House, Senate, presidency, and state legislatures reach similar conclusions: elected officials behave similarly to the most extremely liberal and conservative voters in the electorate.[2]

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

[1] Data source: American National Election Studies (ANES).

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

From the perspective of standard election models, polarization is puzzling. If an extremely liberal candidate moved to the right, for example, liberal voters should still favor her over her a conservative opponent, but moderate voters should join her side.[3] Based on this logic, the well known *median voter theorem* predicts that candidates in an election should both adopt centrist platforms—not only when they value winning, as in Downs (1957), but even when they have polarized policy preferences, as in Calvert (1985), because indulging policy preferences requires winning first. Extensions of these early models have proven the logic of the median voter theorem to be extremely robust.[4]

One thing that might make a candidate reluctant to compromise is a conviction that her policy proposals are in voters' best interest. If she can convince voters of this, she may win their votes without needing to moderate her policy position. In fact, compromising on the truth may only hurt her cause. In that spirit, this paper explores a *common interest* model of elections, after the tradition of Condorcet (1785): voters share a common desire to do whatever is truly best for the group, but disagree for informational reasons what policy would be best. I adapted that voting model to a spatial environment in McMurray (2017a), showing (for exogenous candidate platforms) that equilibrium voting behavior matches many empirical features of real-world elections. This paper proceeds by backward induction to analyze candidate platform selection.

The basic trade-off that candidates make is the same here as in standard private-interest models: even if extreme positions deliver better policy outcomes, they still sacrifice the votes of voters who believe differently. However, two new subtleties emerge. First, it is well known that the pivotal voting calculus leads voters to favor underdog candidates, and the analysis below makes clear that this includes candidates disadvantaged by their own extreme policy positions. This leads to a *jury theorem*, stating that the candidate whose platform is truly superior wins the election with high probability, whether it is extreme or moderate. If a candidate were convinced that an extreme policy best serves the electorate, for example, she could confidently adopt that position, trusting voters to recognize her position as superior, and elect her to office.

Both to keep the analysis tractable and to transparently isolate and give emphasis

[3] Throughout this paper, female and male pronouns refer to candidates and voters, respectively.

[4] This logic is so robust, in fact, that Roemer (2004) refers to the "tyranny of the median voter theorem."

to the underlying incentives at work, candidates in the model below are not endowed with exogenous confidence in their policy positions, or even with private signals: they have no information beyond the common prior, and so are ex ante identical. Nevertheless, candidates develop *endogenous* confidence in opposite directions, with a surprisingly strong polarizing effect. This is because of a second subtlety, which is that a candidate faces a pivotal calculus of her own, as her platform choice will matter if she wins, but not if she loses.

As in private-interest settings, moderate policies attract votes from moderate voters, without sacrificing the support of extremists. If candidates are sufficiently motivated to win, a *median opinion theorem* therefore states that candidates propose identical policy platforms. For reasonable levels of office motivation, however, the forces above lead candidates to polarize despite their desire to win. In fact, candidates who are determined to win at all costs can still polarize maximally, confident that moderation is unnecessary. These equilibrium predictions are consistent with empirical evidence that elections penalize extreme candidates, but only slightly.

In standard models, compromising between the interests of the liberal left and the conservative right maximizes social welfare, because centrist policies minimize the disutility imposed on voters who prefer something very different. Indeed, the prediction that candidates should moderate even when they hold intrinsically polarized preferences can be viewed in that context as the “invisible hand” of political competition. On the other hand, empirical polarization must then be interpreted as evidence of some inexplicable political failure.

In the common-interest setting below, convergence instead amounts to a spatial form of *pandering*—compromising the truth and prioritizing popularity over principle. The geometry of this pandering closely mirrors concerns that are expressed prominently in popular discourse, further corroborating that voters and candidates view elections through a common interest lens. Polarization may still be troubling, but for the new reason that extreme policies are premature until consensus can be reached.

2 Related Literature

Following the convergence results of Hotelling (1929) and Downs (1957), numerous theories have attempted to explain polarization. Notable among these is the *probabilistic voting* literature.[5] As Wittman (1983) points out, candidates who are policy motivated and cannot perfectly predict election outcomes never converge, because deviating from a convergent position makes a candidate better off if she wins and no worse off if she loses. However, candidates and voters share an interest in reducing electoral uncertainty, and Calvert (1985) clarifies that if uncertainty is mild or office motivation is strong then polarization is minimal. Polarization may be substantial if there is *aggregate* uncertainty, but as Section 7 explains, this amounts to common-interest voting, as in the model below.[6] That section also shows examples where polarization is slight with aggregate uncertainty but increases with the various other assumptions made below.

Existing literature on common-interest elections focuses on information aggregation and on abstention or other voter behavior.[7] For the most part, this work also retains Condorcet's binary structure or extends to a small number of alternatives and truth states. In McMurray (2017a) I treat a continuum of policy alternatives and truth states, but candidate platforms are exogenous. This paper is the first to explore polarization in a spatial common-interest setting.[8]

Another literature focuses on the policy choices of a single incumbent politician in an implicitly common interest environment characterized by a representative voter. Like this paper, the pandering models of Canes-Wrone, Herron, and Shotts (2001) and Maskin and Tirole (2004) highlight a possible trade-off to a politician between doing what is popular and doing what is right. In Harrington (1993), the politician resists the temptation to pander with the expectation that voters will learn the truth, and reward him accordingly. The policy choice in these papers is binary; below, I show that a spatial model offers richer insight into the geometry of pandering, which matches empirical evidence. An electoral setting with more than one candidate is also essential for analyzing polarization, which is the central topic of this paper.

See also Prato and Wolton (2017), where office motivated candidates converge on

[5]See Duggan (2013) for a review.

[6]Bernhardt, Duggan, and Squintani (2009a) consider a private-interest spatial model augmented by common-interest valence, with the result that candidate platforms do not converge, and some polarization may be socially optimal. These results are discussed in more detail in Section 5.1 below.

[7]Examples include Austen-Smith and Banks (1996), Feddersen and Pesendorfer (1996, 1997), and McMurray (2013).

[8]Razin (2003) and McMurray (2017b) analyze candidate behavior in common interest spatial models as well, but focus on the single policy choice of an election winner who chooses policy *after* voting takes place.

whatever is favored ex ante, thereby preventing information aggregation by delivering voters a degenerate policy menu.

3 The Model

A society consists of N voters and two political candidates, A and B. As in Myerson (1998), N is drawn from a Poisson distribution with mean n. Together, these voters and candidates implement one policy x from an interval $X = [-1, 1]$ of alternatives. One of these policy alternatives $z \in X$ maximizes voter welfare, and voters and candidates all prefer policies as close as possible to z.[9] However, the location of z is unknown: at the beginning of the game, z is drawn from a known distribution F with density f. This density is *log-concave*, meaning that $\ln f(z)$ is concave in z.[10] For ease of exposition, f is also *symmetric* around the origin, meaning that $f(-z) = f(z)$.

The structure of the election game is as follows. In a first stage, the two candidates simultaneously commit to policy platforms $x_A \in X$ and $x_B \in X$. Observing these platforms, voters each vote simultaneously for one of the two candidates. The candidate $w \in \{A, B\}$ who receives more votes wins the election and implements her proposed platform, breaking a tie if necessary by a fair coin toss. If policy x is implemented when z was optimal, each voter receives the following utility,

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

which declines quadratically with the distance between x and z.[11] Candidate utility (where 1_j is an indicator function that equals one if $w = j$ and zero otherwise)

$$u_j(x, z) = u(x, z) + \beta 1_j \tag{2}$$

[9] Voters and candidates can be said to be *welfare motivated*. Section 7 also considers *selfish* voters and candidates, who prefer specific policies regardless of z. In McMurray (2017a) I explain that basically selfish voters and candidates might act welfare motivated as large elections amplify altruism. Selfish candidates might also act welfare motivated to cultivate a favorable image or legacy.

[10] Log-concavity is exhibited by many of the most commonly used distribution functions (see Bagnoli and Bergstrom, 2005).

[11] Quadratic utility is not essential but is standard, and is convenient because expected utility is quadratic in x and maximized at the expectation of z, conditional on any available information.

is similar, but a candidate receives an additional benefit β if she wins office. In spite of this slight difference, voters and candidates are both said to be *welfare motivated*.

Before voting, each voter observes a private signal s_i, drawn from the interval $S = [-1, 1]$ according to a known distribution $G(s|z)$ with density $g(s|z)$. Conditional on z, these private signals are jointly independent. The family of conditional distributions $G(s|z)$ satisfy the *monotone likelihood ratio property (MLRP)*, meaning that $\frac{g(s'|z)}{g(s|z)}$ increases in z, for all $s' > s$. For ease of exposition, g is also *symmetric*, meaning that $g(-s|-z) = g(s|z)$ for any s and z. For any z, $g(s|z)$ also has full support on S. Though voters share a common interest, a spectrum of private signals translates by Bayes' rule into a spectrum of private opinions regarding which policy is optimal.[12] s_i can also be referred to as a voter's *ideology*, ranging from the liberal left to the conservative right.

It would be natural to assume that candidates observe private signals of z, just as voters do. If anything, candidate signals ought to be more informative than the typical voter's, as candidates have career incentives to stay informed and have privileged access to information sources. However, candidate signals would introduce technical complexities that are beyond the scope of this paper, and so are excluded from the model. One interpretation of this is that candidates do receive signals, but publicly (and truthfully) reveal their policy opinions before the election takes place, so that the prior distribution f already takes these into account. Whether candidates possess private information or not, they should also infer whatever they can about voters' private signals from their equilibrium voting behavior.[13] Adding candidate signals is an important direction for future work, but omitting these signals is also useful for clarifying the equilibrium inference and emphasizing how influential it can be: below, candidates polarize *even* when they start from identical beliefs about what policies would be best.

In the subgame associated with any pair $(x_A, x_B) \in X^2$ of candidate platforms, a voting strategy is a measurable function $v : S \to \Delta(\{A, B\})$ from the space of signals

[12] In McMurray (2017a) I emphasize how single-peaked preferences make voters risk averse, so that those with imprecise signals favor moderate policies, consistent with various empirical evidence. In line with those observations, Garz (2018) presents empirical evidence that information has a causal effect on voter polarization.

[13] Candidate signals complicate the model considerably because of higher order beliefs: if she tries to infer her opponent's signal or voters' signals from their behavior, a candidate must take into account that their behavior reflects a mixture of their own private signals and their guesses of her own signal (as well as her opponent's guess of voters' guess of her own signal, and so on).

into the unit simplex over the set of candidates, with $v_j(s)$ specifying the probability of voting for candidate $j \in \{A, B\}$ in response to signal $s \in S$. Let V denote the set of such strategies. Abusing notation, let $v(s) = j$ indicate the pure strategy with $v_j(s) = 1$ and $v_{-j}(s) = 0$, where $-j$ is candidate j's opponent, and when clear from context, let event j also denote the event $w = j$ of candidate j winning the election. When his peers all vote according to the strategy $v \in V$, a voter's best response $v^{br} \in V$ maximizes the expectation $E_{w,z}[u(x_w, z); v]$ of (1) for every private signal realization in S. A *(symmetric) Bayesian Nash equilibrium (BNE)* in the voting subgame is a strategy v^* that is its own best response.[14]

A voting strategy $\sigma : X^2 \to V$ in the complete game specifies voting behavior for every subgame. Let Σ denote the set of such strategies. $(x_A^*, x_B^*, \sigma^*) \in X^2 \times \Sigma$ is a *(symmetric) perfect Bayesian equilibrium (PBE)* if $\sigma^*(x_A, x_B)$ constitutes a (symmetric) BNE in the voting subgame associated with every platform pair $(x_A, x_B) \in X^2$, and the platform choice x_j^* of candidate $j \in \{A, B\}$ maximizes the expectation $E_{w,z}[u(x_w, z) + \beta 1_j; x_{-j}, \sigma]$ of (2), taking her opponent's platform x_{-j} and the voting strategy σ as given.

4 Voting

This section analyzes equilibrium voting in the subgame associated with an arbitrary pair $x_A \leq x_B$ of candidate platforms, first for finite n and then for the limit as n grows large.[15] 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.

$$\phi(j|z) = \int_S v_j(s)\, g(s|z)\, ds \tag{3}$$

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

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

[15] Lemmas 1 and 2 and Proposition 1 extend the parallel results for specific F and G in McMurray (2017a).

either total.

With quadratic utility, a voter's basic inclination would be to favor the candidate whose platform is closest to $E(z|s_i)$. However, a voter's behavior will have no impact on his utility unless his vote happens to be *pivotal* (event P), meaning that it reverses the election outcome. As Austen-Smith and Banks (1996) point out, therefore, a voter optimally conditions his behavior on that event, and supports the candidate who is closest to $E(z|P, s_i)$. Since $G(s|z)$ satisfies MLRP, signals further to the left or right indicate to a voter that z lies further to the left or right, respectively. Accordingly, Lemma 1 characterizes best response voting as an *ideological* strategy v_t with some *ideology threshold* $t \in S$, meaning that $v_t(s) = \begin{cases} A & \text{if } s < t \\ B & \text{if } s > t \end{cases}$. In particular, the best response to ideological voting is ideological, with an ideology threshold that increases in the midpoint $\bar{x} = \frac{x_A + x_B}{2}$ between the two candidates' platforms. A fixed point argument, together with this monotonicity, guarantees the existence of a unique equilibrium ideology threshold $t^*(\bar{x})$, that is symmetric and increases in $\bar{x}$.[16]

Lemma 1 *There exists a unique function $t^* : X \to S$ such that, for any $(x_A, x_B) \in X^2$ with midpoint $\bar{x} \in X$, the ideological strategy $v_{t^*(\bar{x})}$ characterized by $t^*(\bar{x})$ constitutes a BNE in the voting subgame. For $x_A \neq x_B$, $v_{t^*(\bar{x})}$ is the unique BNE. Moreover, $\frac{dt^*(\bar{x})}{d\bar{x}} > 0$ and $t^*(-\bar{x}) = -t^*(\bar{x})$.*

Since voters share a common interest, the observation of McLennan (1998) applies: if a voting strategy is socially optimal, it is also individually optimal, and therefore constitutes an equilibrium. Showing that a socially optimal strategy exists is not trivial, but Lemma 2 does so by using the MLRP of $G(s|z)$ to show that every non-ideological strategy is dominated by an ideological strategy, and then invoking the extreme value theorem on the compact set of possible ideology thresholds. The unique equilibrium identified in Lemma 1 is therefore the uniquely optimal voting strategy.

Lemma 2 *For any $x_A < x_B$ there exists a uniquely optimal voting strategy $v^{**} = \arg\max_v E[u(x,z); v]$, which also constitutes a BNE in the voting subgame.*

Low realizations of z generate more liberal signals than conservative signals, and when voters follow an ideological strategy with threshold t, this translates into more

[16] This generalizes Lemma 1 and Proposition 1 from McMurray (2017a), which assumed specific distributions f and g and did not characterize t^*.

A votes than B votes. Symmetrically, high realizations of z generate more B votes than A votes. Somewhere between these extremes, there is a unique state of the world $\bar{z}(t) = \arg\min_z |\phi(A|z;v_t) - \phi(B|z;v_t)|$ that equalizes candidates' vote shares more than any other. As Myerson (2000) explains, a pivotal vote becomes less likely in general as n grows large, but becomes exponentially less likely in states other than $\bar{z}(t)$. Whatever his private signal, therefore, a voter assigns probability close to one on state $\bar{z}(t)$, conditional on being pivotal. If candidates are positioned with $\bar{x} < \bar{z}(t)$ then he should respond by voting B; if $\bar{x} > \bar{z}(t)$ then he should vote A. But citizens with other signal realizations have the same incentive, so neither of these scenarios is compatible with equilibrium. Lemma 3 thus concludes that the equilibrium ideology threshold t_n^* adjusts as n grows large, so that $\bar{z}(t_n^*)$ approaches $\bar{x}$ exactly.

Lemma 3 *For any $\bar{x} \in X$, the limiting equilibrium threshold $t_\infty^*(\bar{x}) = \lim_{n\to\infty} t_n^*(\bar{x})$ solves $\bar{z}(t_\infty^*(\bar{x})) = \bar{x}$.*

A well known consequence of the pivotal voting calculus is the *underdog effect*: one additional vote for an expected winner is less likely to matter than one additional vote for the expected loser. In McMurray (2017a) I emphasize how this can make voters less extreme: when candidate platforms are symmetric, for example, a voter may privately believe the state to be quite high, so that B is far superior to A, but if he is correct then B should win the election handily; if his vote makes or breaks a tie, truth is likely less extreme than he had thought. In that paper, such a voter might simply abstain from voting.

When platforms are asymmetric, the pivotal calculus and underdog effect can also make voters *more* extreme. For example, a citizen with $s_i = 0$ holds perfectly moderate private beliefs. If $\bar{x} > 0$, so that A is closer to the origin, such a voter would feel inclined to vote for candidate A. If the optimal policy is truly centrist, however, then other voters should favor A as well, and a tie is unlikely. If his own vote makes or breaks a tie, it is likely because the optimal policy is actually to the right of the origin, and may even favor candidate B. In that sense, the pivotal voting calculus favors extreme candidates: unconditionally, an extreme candidate is less likely than her opponent to win the election, but a pivotal vote suggests that truth is on the extremist's side, so she is better than she otherwise would seem.

From a welfare perspective, it may seem undesirable that pivotal considerations lead voters to favor the candidate who is worse in expectation. Contrary to that intuition, however, Lemma 2 has already stated that equilibrium voting is socially optimal. Letting $j^* = \arg\min_j |x_j - z|$ denote the candidate whose platform is truly superior, candidate j^* wins with certainty in the limit as n grows large. This is simply the (slightly generalized) Jury Theorem, stated here as a proposition.

Proposition 1 (Jury Theorem) *For any* $(x_A, x_B) \in X^2$, $\lim_{n\to\infty} \Pr(w = j^*; v_n^*) = 1$.

Proposition 1 relies crucially on the equilibrium balancing highlighted in Lemma 3. Suppose, for example, that $x_A = .8$ and $x_B = 1$. Of these platforms, x_A maximizes welfare when $z \leq .9$ but x_B maximizes welfare when $z \geq .9$. Since x_A is superior for so many realizations of z, only voters with the most extreme realizations of s believe B to be the better candidate. Depending on the distributions F and G, this group might not constitute a majority, even when z truly exceeds .9. In equilibrium, however, the logic of Lemma 3 is that a pivotal vote is more likely when z is more extreme. Thus, voters favor candidate B (the underdog, in this case) more than they would if they ignored the pivotal voting calculus. The consequence is that a full *half* of the electorate votes for candidate B when $z = .9$ exactly (even if far fewer than half have private signals directly supporting such a vote), more than half vote for B when $z > .9$, and fewer than half vote for B when $z < .9$. In this way, pivotal considerations and the implied equilibirum balancing ensure that the proper candidate wins in every state of the world.

The analysis of this section has focused on voters, but the jury theorem sets the stage for candidate polarization, which is anayzed in the following section. To see this, briefly consider a model different from the one posited in Section 3, in which candidates are supremely overconfident, each favoring some policy s_j, convinced (erroneously) that this is socially optimal—that is, believing with probability one that $z = s_j$. In general, though she (believes that) she knows which policy is socially optimal, she dares not commit to this policy in an election, because controlling policy requires winning first, which requires a moderate platform (ideally, not too different from s_j) that will attract as many votes as possible. In large elections, however, the jury theorem implies that such a candidate need have no inhibitions: individual voter misperceptions will average out, and the electorate will correctly favor whichever candidate's platform is truly superior. She can thus be confident doing what (she believes) is truly best for welfare, as this is also a winning strategy. In fact, this conclusion does not change if candidates place high value on the perks of winning office.

5 Candidates

5.1 Polarization

Section 4 analyzes the subgame associated with an arbitrary platform pair, showing that the ideological voting strategy $v_{t^*(\bar{x})}$ characterizes the unique equilibrium response to platforms $x_A \neq x_B$. This section analyzes candidates' platform choice in the first stage of the game, taking voters' equilibrium response as given. Let σ_{t^*} denote the strategy in Σ that specifies $v_{t^*(\bar{x})}$ for all (x_A, x_B) pairs (including pairs with $x_A = x_B$, for which any v constitutes an equilibrium).

To begin this analysis, Lemma 4 states the existence of a candidate's best response to her opponent's strategy x_{-j}, taking σ_{t^*} as given. If candidates do not care about winning (i.e. if $\beta = 0$) then (2) is quadratic in the policy choice, so expected utility is maximized at the expectation of z. Since candidates are assumed not to observe private signals, and given the symmetry of F, the ex ante expectation $E(z) = 0$ lies at the center of the policy space. As Lemma 4 states, however, this is not the policy that a candidate chooses. Instead, she updates her expectation to reflect the event of winning the election. Since equilibrium voting is ideological, the candidate on the left wins when lots of voters received low signal realizations, which tends to happen when z is left of center; the candidate on the right tends to win when z is right of center. Left of x_{-j}, candidate j's best response is the expectation of z condtional on beating x_{-j} from the left. Right of x_{-j}, her best response is the expectation of z conditional on beating x_{-j} from the right. Whichever of these policy positions provides greater utility is the candidate's best response.

Lemma 4 *For any $x_{-j} \in X$, a best response $x_j^{br}(x_{-j}, \sigma_{t^*})$ exists for $j = A, B$. If $\beta = 0$ then x_j^{br} satisfies $x_j^{br}(x_{-j}, \sigma_{t^*}) = E\left(z | w = j; x_j^{br}, x_{-j}, \sigma_{t^*}\right)$.*

It may seem odd that a candidate should condition her beliefs on the event of winning the election, as the timing of her platform choice is before she knows whether she will win or not—and even before votes are cast. The logic for this amounts to a sort of pivotal calculus, analogous to that performed by voters: if a candidate loses the election, her policy choice will not affect her utility, so she restricts attention to situations in which she wins, and chooses a platform that will be optimal in that event, just as a voter votes in a way that will be optimal if his own vote turns out to make or breaks a tie.

Building on Lemma 4, Theorem 1 states that a perfect Bayesian equilibrium exists. Given the symmetry of the model, candidates' equilibrium platforms can be symmetric around the origin. This does not rule out the possibility of asymmetric equilibrium platforms, but in large elections, the jury theorem ensures that the candidate on the left wins with near certainty when $z < \bar{x}$ while the candidate on the right wins when $z > \bar{x}$, so candidates' expectations converge to $E(z|z < \bar{x})$ and $E(z|z > \bar{x})$. The game of Section 3 therefore reduces in large elections to a simpler game in which candidates adopt platforms and the platform closest to the true z is implemented. $\bar{x} = 0$ is the only solution to $\frac{E(z|z<\bar{x})+E(z|z>\bar{x})}{2} = \bar{x}$, so at least in the case of $\beta = 0$, asymmetric equilibria become symmetric in the limit (if they exist). For ease of exposition, relabel candidates so that A is on the left. Then, platforms converge to $x^*_{A,\infty} \equiv \lim_{n\to\infty} x^*_{A,n} = E(z|z < 0)$ and $x^*_{B,\infty} \equiv \lim_{n\to\infty} x^*_{B,n} = E(z|z > 0)$.[17]

Theorem 1 *For any n, there exists one platform-symmetric PBE. If $\beta = 0$ then, for any sequence $(x^*_{An}, x^*_{Bn}, \sigma^*_n)$ of PBE, $\lim_{n\to\infty} |x^*_{jn}| = E(z|z > 0)$.*

If candidates adopted centrist platforms $(0, 0)$, voters would be indifferent, and might vote randomly. Their pivotal calculus would then tell candidates nothing, leading to identical posteriors $E(z|w = A) = E(z|w = B) = 0$. This might seem to validate $(0, 0)$ as an equilibrium platform pair, but it is not: a candidate who deviates from $(0, 0)$ will trigger a different subgame where voters vote ideologically, thus revealing information that validates the deviation. Thus, at least when $\beta = 0$ and n is large, Theorem 1 establishes equilibrium polarization. In fact, this polarization is substantial: x^*_A and x^*_B lie at the means of opposite sides of a partition. Examples 1 through 3 offer stark examples of this. Such polarization is remarkable, given that candidates in this model are identical: their polarizing beliefs are purely endogenous.

Example 1 *If F is uniform on $[-1, 1]$ then $\left(x^*_{A,\infty}, x^*_{B,\infty}\right) = \left(-\frac{1}{2}, \frac{1}{2}\right)$. These are the 25th and 75th percentiles of F. If voter signals are drawn from the family of linear densities $g(s|z) = \frac{1}{2}(1 + sz)$ for $s \in [-1, 1]$, as in McMurray (2017a), then private expectations $E(z|s_i) = \frac{s_i}{3}$ range only from $-\frac{1}{3}$ to $\frac{1}{3}$, making candidates more extreme than any voter.*

[17] Note that existence and asymptotic uniqueness (up to a relabeling of candidates) do not require model symmetry: asymmetries in F, G, or u may produce $\bar{x} \neq 0$, but a compact X and continuous best-response functions guarantee equilibrium existence, and a log-concave f guarantees a unique solution to $\frac{E(z|z<\bar{x})+E(z|z>\bar{x})}{2} = \bar{x}$ (as $E(z|z < 0)$ and $E(z|z > 0)$ both increase with slopes less than one).

Example 2 *If* $X = \mathbb{R}$ *and* F *is standard normal then* $\left(x^*_{A,\infty}, x^*_{B,\infty}\right) \approx (-.80, .80)$, *which are the 21st and 79th percentiles of* F. *If voter signals* $s_i = z + \varepsilon_i$ *differ from* z *by an amount* $\varepsilon_i \sim N(0,1)$ *then a voter's expectation* $E(z|s_i) = \frac{s_i}{2}$ *is less extreme than either candidate's platform with probability* .74.

In McMurray (2017a) I point out that many applications are best characterized by binary truth, meaning that one of two extremes is ultimately optimal. To exit an economic recession, for example, competing economic theories recommend either substantial economic stimulus or no stimulus at all. Intermediate levels of stimulus could be useful in the face of uncertainty, but are clearly not optimal, per se. Similarly, public funding can be split between two programs with a similar purpose, such as increasing teacher salaries or reducing class sizes to improve education, but if truth were known that one of these programs is actually more effective, it should ideally receive all available funding. Harrington (1993) proposes that deep philosophical attitudes toward the proper role of government may be binary, as well, as voter favor either "extensive or minimal government intervention in the economy". With applications such as these in mind, Example 3 shows that candidates are maximally extreme when z is binary.

Example 3 *If* F *is uniform on* $\{-1, 1\}$ *then* $\left(x^*_{A,\infty}, x^*_{B,n}\right) = (-1, 1)$, *making candidates the most extreme members of society.*

5.2 Office Motivation

Section 5.1 focuses on the case of $\beta = 0$, meaning that candidates care about the policy outcome but do not care who wins the election. To many observers, however, candidates in the real world seem highly motivated by prestige and other perks of winning office.[18] Accordingly, this section analyzes the case of $\beta > 0$.

If candidates polarize, there will be voters who favor policies between the candidates' platforms. By moving her own platform toward her opponent's, a candidate can attract some of these voters to her side. This is true even if truth is binary, as in Example 3: a candidate who splits funding between increasing teacher salaries and reducing class sizes will attract more votes from voters who are unsure which program is better, than a candidate who funds just one or the other.

[18] Entry decisions are not modeled here, but presumably, individuals who value office most highly should be the most likely to run.

When β is low, a candidate largely foregoes the support of centrist voters to remain as close as possible to $E(z|w = j)$. As the importance of winning increases, polarization decreases, as Theorem 2 now states.[19] If β is sufficiently large, the two candidates' platforms coincide, just as in standard median voter theorems. Theorem 2 is labeled the median *opinion* theorem because the voters attracted by centrist platforms are not those with centrist bliss points, but those with centrist signals of the common bliss point. As Section 5.3 emphasizes below, this subtle difference has important consequences for social welfare.

Theorem 2 (Median Opinion Theorem) *There exists $\bar{\beta}$ such that, in the platform-symmetric BNE, $\beta \geq \bar{\beta}$ implies $x_B^*(\beta) = 0$, and otherwise $x_B^*(\beta)$ decreases in β.*

Given the importance of public policy decisions, arbitrarily large values of β may not be relevant empirically. With quadratic policy utility, for example, $\beta \geq 4$ makes a candidate willing to implement the *worst* policy in X, in order to win, which seems implausible. If winning compensates a candidate for policies that deviate from the optimum by $\frac{1}{2}$, which is 25% of the length of X, then $\beta = \frac{1}{4}$.[20] Moreover, in analyzing the case of large β, Theorem 2 fixes the electorate size n. This matters because Section 5.1 emphasizes the polarizing impact of the jury theorem, which is a limit result. Suppose, for example, that candidates were grossly overconfident in their policy analysis, so that each (wrongly) believed with probability one to have identified the truly optimal policy $\hat{z}_j$. Believing this to be the case, each would believe that, in large elections, a policy position $x_j = \hat{z}_j$ will prevail almost surely against any competing platform. Thus, in large elections, deviating from the policy that seems best would be unnecessary, and candidate platforms would approach $x_{j,\infty}^*(\beta) = \lim_{n\to\infty} x_{j,n}^*(\beta) = \hat{z}_j$, *regardless* of β. In light of these observations, Theorem 3 now fixes β and analyzes behavior as n grows large. The result is that polarization can be quite robust. To illustrate this, Examples 4 through 6 extend Examples 1 through 3.

Theorem 3 *If $x_{B,n}^*(\beta)$ is the sequence of platform-symmetric BNE platforms then $x_{B,\infty}^*(\beta) = \lim_{n\to\infty} x_{B,n}^*(\beta) = \max\left\{0, E(z|z > 0) - \frac{1}{2}\beta f(0)\right\}$.*

---

[19] Theorem 2 focuses on the platform-symmetric equilibrium, where uniqueness facilitates clear comparative statics, but if equilibria exist with asymmetric platforms, the same reasoning guarantees that platforms coincide (not necessarily at zero) for β above some $\bar{\beta}$.

[20] In general, β compensates for a policy that differs from z by a distance that, as a fraction of the total width of the policy space, is given by $\frac{1}{2}\sqrt{\beta}$.

Example 4 *If F is uniform on $[-1,1]$ then $\left(x^*_{A,\infty}\left(\frac{1}{4}\right), x^*_{B,\infty}\left(\frac{1}{4}\right)\right) \approx (-.44, .44)$, which are the 28th and 72nd percentiles of F. This is only slightly less polarized than Example 1 and, for the linear $g(s|z)$ of Example 1, is still more polarized than any voter. Platforms only coincide if $\beta \geq \bar{\beta}_\infty = 2$, which seems implausibly high.*[21]

Example 5 *If $X = \mathbb{R}$ and F is standard normal then $\left(x^*_{A,\infty}\left(\frac{1}{4}\right), x^*_{B,\infty}\left(\frac{1}{4}\right)\right) \approx (-.75, .75)$, which are the 31st and 69th percentiles of F. If voter signals $s_i = z + \varepsilon_i$ differ from z by an amount $\varepsilon_i \sim N(0,1)$ then this is more polarized than a random voter's expectation $E(z|s) = \frac{s_i}{2}$ with probability .74. Platforms only coincide if $\beta \geq \bar{\beta}_\infty = 4$, meaning that winning compensates a candidate for any policy outcome.*

If truth is binary, as in Example 3, polarization is perfectly robust. As Example 6 now states, candidates polarize to opposite extremes of the policy space for large values of β, just as they do for $\beta = 0$. It remains true that a moderate platform would win more votes, and when β is large, a candidate is willing to promise any policy in order to win, but in a large election, compromise is simply unnecessary: if truth is on a candidate's side then she will win even from a polarized position, and if truth is against her then she will lose even if she moderates.

Example 6 *If the domain of F is $\{-1,1\}$ then $\left(x^*_{A,\infty}(\beta), x^*_{B,\infty}(\beta)\right) = (-1,1)$ for any β, still making candidates the most extreme members of society.*

5.3 Welfare

In private interest settings, the policies that maximize social welfare lie at the political center. By compromising between the extreme left and extreme right, such policies minimize the total disutility that voters suffer from a policy that is far from their bliss points.[22] Policy preferences might tempt candidates toward the extremes, but the need to win acts as an "invisible hand", pushing candidates toward the center in pursuit of votes. Bernhardt, Duggan, and Squintani (2009) amend this logic slightly, pointing out that some difference between candidates can be desirable, if "shocks" to voter preferences occur after platform commitments have been made:

[21] Here, $\bar{\beta}_\infty$ denotes the limit of the sequence of thresholds $\bar{\beta}_n$ for which $x^*_{B,n}(\beta) = 0$ for all $\beta \geq \bar{\beta}_n$.

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

in that case, the median voter benefits from positions slightly left and right of center, for option value. Office motivation in candidates still benefits voters to a point, but can make them excessively centrist, limiting voters' range of choice.

In the setting of this paper, voters and candidates share common interests and the benefit to candidates of winning is zero-sum, so it is uncontroversial to measure welfare simply by expected policy utility. When $\beta = 0$, voter and candidate preferences are identical, so the logic of McLennan (1998) implies that socially optimal behavior is also individually optimal and can arise in equilibrium, as Proposition 2 now states. Section 5.2 therefore makes clear that when $\beta > 0$, candidates are overly centrist.

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 $E_{N,s,w,z}\left[u\left(x^*_{w,n}, z\right)\right]$. If $\beta = 0$ then this vector also constitutes a PBE, and $x^*_{B,\infty}(\beta) = E(z|z > 0)$.*

Proposition 2 can be understood in the same light as Bernhardt, Duggan, and Squintani (2009): if the candidates had to agree on a single policy platform, the best choice would be in the center, but distinct platforms provide option value, allowing voters to make a choice after determining what is best. However, the result here is more stark: according to Examples 1 through 3, the optimal policy positions can be more polarized than *any* of the voter opinions, and may even lie at the far extremes of X.

Favoring policies that please voters but make society worse off amounts to a form of *pandering*, reminiscent of Canes-Wrone, Herron, and Shotts (2001) and Maskin and Tirole (2004). A spatial model adds geometry, though, that these binary models cannot: to be popular, a candidate avoids extreme policies even when they would improve welfare. As Section 6 discusses, this mirrors concerns that are often expressed in real-world elections.

The problem of pandering is especially pronounced in the binary environment of Example 3, where office motivated candidates compromise on a centrist policy even when they *know* that the optimum lies at one of the two extremes. In the examples of Section 5.1, such compromise could take the form of splitting funding between superior and inferior programs, or issuing moderate sized economic stimulus that is certainly either much too large or much too small.

6 Evidence

Hall (2015) presents causal evidence, using regression discontinuities from primaries, that extreme candidates are penalized in general elections. By itself, this only deepens the puzzle of why candidates should polarize: by moderating, a candidate would win more votes. The reluctance to compromise is often attributed to the influence of extremist supporters who serve as donors, activists, and primary voters. In the most recent US presidential primaries, for example, many liberal voters favored Bernie Sanders even though Joe Biden was widely viewed as more moderate and therefore more likely to win the general election. According to Brady, Han, and Pope (2007) and Hall and Snyder (2013), primary elections actually favor extremists on average.

That extremist supporters penalize centrist candidates can of course explain why candidates polarize, but begs the new question of why voters or party elites should behave this way: if they truly have sway over candidates, then they themselves face the trade-off of a standard candidate positioning game: pushing their candidate to be more extreme improves policy outcomes conditional on winning, but makes winning less likely. In equilibrium, these voters should favor moderate candidates, by the original logic of the median voter theorem.

The analysis above reconciles these findings by clarifying that the penalty for polarization is small, so candidates still polarize. This is consistent with evidence in Ansolabehere, Snyder, and Stewart (2001), Canes-Wrone, Brady, and Cogan (2002), and Cohen, McGrath, Aronow, and Zaller (2016), that extreme candidates often beat centrists. Related to all of these observations, I show in McMurray (2017a) that voters who prefer one candidate over the other also tend to expect that candidate to win: 96% of ANES survey respondents who voted Democrat in the 2012 presidential election also predicted a Democratic victory, while 83% of Republican voters predicted a Republican victory. The model above does not formally include primary elections, but in McMurray (2017a) I show that primary voters are the most confident in their policy opinions; in light of the evidence summarized here, these also seem the most prone to view an extreme candidate as invincible.

Almost two centuries ago, 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," and calling for "responsible parties" who believe that "putting a particular candidate into office is not an end in itself". From a standard private-interest perspective, such recommendations are odd, as centrist policies are socially optimal. They fit neatly in the common-interest paradigm here, however, as a call for candidates to pursue whatever policies are truly best, rather than pandering to the center.

This perspective also reinforces candidates' reluctance to compromise. 1964 Republican presidential candidate Barry Goldwater famously defended his extreme policy positions, for example, by proclaiming that "extremism in the defense of liberty is no vice," and "moderation in the pursuit of justice is no virtue."[23] This can also explain why many Democrats and Republicans criticize moderates within their own parties for being overly timid, labeling them DINO or RINO (i.e. Democrats- or Republicans-in-name-only). 2000 Green party presidential candidate Ralph Nader famously disparaged Republicans and Democrats as "look alike parties", "Tweedledum and Tweedledee".[24]

7 Sensitivity Analysis

The model above assumes that voters and candidates are all welfare motivated and that voters make strategic inference from the event of a pivotal vote. This section explores how sensitive polarization is to relaxing these various assumptions, using a series of examples with the linear densities of Examples 1 and 4 and then the binary specification of Examples 3 and 6. This approach quantifies the importance of pivotal considerations of voters and candidates, albeit only for special cases of F and G. It also relates the model above to a standard private interest framework. Each example focuses on the limiting platform pair $\left(x^*_{A,\infty}(\beta), x^*_{B,\infty}(\beta)\right)$ in the unique sequence platform-symmetric equilibria, as a function of office motivation β. Derivations behind these examples are presented in the Appendix.

Define a *selfish* candidate to have utility $u_j(x; \hat{x}_j) = -(x - \hat{x}_j)^2 + \beta 1_j$ that is quadratic in the policy outcome, like (2), but with idiosyncratic bliss points $\hat{x}_j$ different from z. In contrast, the candidates defined in Section 3 are *welfare motivated.* In Example 1, welfare-motivated candidates endogenously favored policies $-\frac{1}{2}$ and $\frac{1}{2}$; Example 7 now shows that, if selfish candidates *exogenously* favor the same policies,

[23]See https://www.washingtonpost.com/wp-srv/politics/daily/may98/goldwaterspeech.htm.
[24]See http://www.cbsnews.com/news/nader-assails-major-parties.

they are still extreme relative to voters, but are only half as polarized as before.

Example 7 *If F is uniform on $[-1, 1]$, voters are welfare motivated and strategic, and candidates are selfish with* $(\hat{x}_A, \hat{x}_B) = \left(-\frac{1}{2}, \frac{1}{2}\right)$*, then* $\left(x^*_{A,\infty}(0), x^*_{B,\infty}(0)\right) = \left(-\frac{1}{4}, \frac{1}{4}\right)$*. If $g(s|z)$ is linear as in Example 1 then this is more polarized than a random voter with probability .75.* $\left(x^*_{A,\infty}\left(\frac{1}{4}\right), x^*_{B,\infty}\left(\frac{1}{4}\right)\right) \approx (-.22, .22)$*, which is more polarized than a random voter with probability* $\frac{21}{32} \approx .66$*. Platforms only coincide if* $\beta \geq \bar{\beta}_\infty = 2$*.*

Since selfish candidates want the same policy as welfare motivated candidates, it might seem that, for a given β, they should adopt identical platforms. That they do not can be understood in terms of candidates' pivotal calculus: since the candidate closest to z wins the election, slightly adjusting one platform only matters when $z = \bar{x}$. When that happens, a welfare motivated candidate is actually indifferent between the two policy platforms, but a selfish candidate still strongly prefers her own platform to her opponent's, and so moderates further to secure the win. For any β, then, selfish candidates are less polarized.

So far, all of the examples above assume that voters are *strategic*, making inference from the event of a pivotal vote. Example 8 assumes instead that voters are *naive*, and so fail to make use of this information. A naive voter votes A if $E(z|s_i) < \bar{x}$ and for B if $E(z|s_i) > \bar{x}$ (using $E(z|s_i)$ as strategic candidates use $E(z|s_i, P)$). If candidates are purely welfare motivated then this has no effect on polarization, because naive voting is just as informative in the limit as strategic voting. If candidates are also office motivated, however, then polarization decreases substantially, because voters no longer favor the underdog extremist, so vote totals are more sensitive to polarization, and the reward to moderation is higher. For $\beta = \frac{1}{4}$, for example, platforms are only 30% as polarized as before.

Example 8 *If F is uniform on $[-1, 1]$, candidates are welfare motivated, and voters are welfare motivated but naive, then* $\left(x^*_{A,\infty}(0), x^*_{B,\infty}(0)\right) = \left(-\frac{1}{2}, \frac{1}{2}\right)$*. If $g(s|z)$ is linear as in Example 1 then this is more polarized than any voter.* $\left(x^*_{A,\infty}\left(\frac{1}{4}\right), x^*_{B,\infty}\left(\frac{1}{4}\right)\right) \approx (-.13, .13)$*, and if $g(s|z)$ is linear then this is only more polarized than a random voter with probability .13. Platforms coincide if* $\beta \geq \bar{\beta}_\infty = \frac{1}{3}$*.*

Examples 7 and 8 make clear that polarization falls relative to the model of Section 3 if candidates are selfish or if voters are naive. Example 9 combines both of these ingredients, and polarization declines further. Even if candidates do not care at all about winning office, candidates are only 30% as polarized as when they are welfare motivated and voters are strategic, and are only more polarized than a random voter with probability .37. If $\beta = \frac{1}{4}$ then candidates are only 10% as polarized as before, exceeding a random voter only with probability .09.

Example 9 *If F is uniform on $[-1, 1]$, candidates are selfish with $(\hat{x}_A, \hat{x}_B) = \left(-\frac{1}{2}, \frac{1}{2}\right)$, and voters are welfare motivated but naive, then $\left(x^*_{A,\infty}(0), x^*_{B,\infty}(0)\right) \approx (-.13, .13)$ and $\left(x^*_{A,\infty}\left(\frac{1}{4}\right), x^*_{B,\infty}\left(\frac{1}{4}\right)\right) \approx (-.03, .03)$. Platforms coincide if $\beta \geq \bar{\beta}_\infty = \frac{1}{3}$.*

The naive voters in Examples 8 and 9 still form expectations of the optimal policy. An alternative specification of naivete could be for voters to simply *vote their signals*, voting A if s_i is closer to x_A than x_B and voting B otherwise. Example 10 shows that this has a similar impact on polarization. If $\beta = 0$ then candidates are only half as polarized as when candidates are welfare motivated and voters are strategic, and are only more polarized than a random voter's signal with probability .25. If $\beta = \frac{1}{4}$ then candidates are 38% as polarized as before, exceeding a random voter's signal with probability .19.

Example 10 *If F is uniform on $[-1, 1]$, candidates are selfish with $(\hat{x}_A, \hat{x}_B) = \left(-\frac{1}{2}, \frac{1}{2}\right)$, and voters are welfare motivated but vote their signals, then $\left(x^*_{A,\infty}(0), x^*_{B,\infty}(0)\right) = \left(-\frac{1}{4}, \frac{1}{4}\right)$ and $\left(x^*_{A,\infty}\left(\frac{1}{4}\right), x^*_{B,\infty}\left(\frac{1}{4}\right)\right) \approx (-.19, .19)$. Platforms coincide if $\beta \geq \bar{\beta}_\infty = 1$.*

In these several examples, s_i is uniform on $[-1, 1]$ and conditionally i.i.d. Example 11 now modifies Example 10 by assuming that s_i are i.i.d. uniform instead. If voter signals are unrelated to the state variable and voters merely vote their signals, however, then they are functionally equivalent to idiosyncratic ideal points $\hat{x}_i = s_i$. Thus, Example 11 constitutes a canonical specification of private interest elections with selfish voters. For the case of $\beta = 0$, Wittman (1983) points out that platforms cannot coincide in equilibrium, as a candidate who deviated toward her ideal point would then make herself better off if she wins, and no worse off if she loses. That logic remains true here for finite electorates, but as n grows large, platforms converge to $x^*_{A,\infty}(\beta) = x^*_{B,\infty}(\beta) = 0$, so that polarization vanishes in the limit.

Example 11 (Private Interest) *If candidates are selfish with $(\hat{x}_A, \hat{x}_B) = \left(-\frac{1}{2}, \frac{1}{2}\right)$ and voters are selfish with ideal points $\hat{x}_i$ uniform on $[-1, 1]$, then $\left(x^*_{A,\infty}(\beta), x^*_{B,\infty}(\beta)\right) = (0, 0)$ for any β.*

Polarization vanishes in Example 11 because, when $\hat{x}_i$ are i.i.d., the election is asymptotically deterministic: as n grows large, the distribution of realized voter ideal points converges to the known prior. When candidates polarize symmetrically, this makes the election outcome highly sensitive to either candidate's policy position, giving either candidate a strong incentive to moderate slightly, with the result that they converge in the limit, as in Calvert (1985). Winning no longer conveys information, so welfare motivated candidates would do the same.

In contrast, when $\hat{x}_i$ are only conditionally i.i.d., as in the preceding examples, uncertainty about z induces *aggregate uncertainty*, meaning that candidates do not know which of two platforms will win. Actually, this is quite general: De Finetti's (1980) theorem states that any joint distribution of ideal points that is symmetric across voters can be decomposed so that $\hat{x}_i$ are i.i.d., conditional on some latent variable z, and this latent variable can *define* an object of common value.

If $\hat{x}_i$ do not represent voters' true preferences, but only candidates' *perceptions* of voters' preferences, then another source of aggregate uncertainty could be systematic polling errors that lead both candidates to perceive the entire electorate as more liberal or conservative than it truly is. Voters and candidates both have incentive to minimize such errors, however, so this source of uncertainty should be small. When candidates polarize equally, slight uncertainty about the location of the median voter generates substantial uncertainty as to who will win, so intuitively this might seem sufficient to sustain polarization. However, polarization actually requires more than this: a polarized candidate must be uncertain who will win *even* if she moderates her policy position.[25]

Examples 7 through 10 assume a particular family of linear signal densities. Example 12 instead considers the binary truth setting of Examples 3 and 6. As noted above, this is appropriate for many specific applications, and Harrington (1993) even proposes this as an apt description of voters' broadest competing worldviews. With binary truth, there is no state of the world that makes voters indifferent between x_A and x_B: one of the two is always superior. Asymptotically, this removes the margin on which candidates can adjust to increase their vote shares. As a consequence, candidates polarize to the extremes of the policy space, regardless of candidate motivations or behavioral voting.

[25] If the median voter were symmetrically distributed between $-\varepsilon$ and ε, for instance, polarized platforms $-x$ and x would win with equal probability, for any x, but equilibrium platforms could be no more polarized than -2ε and 2ε, as otherwise deviating to 0 would win with certainty.

Example 12 *If the domain of F is $\{-1,1\}$ then $\left(x^*_{A,\infty}(\beta), x^*_{B,\infty}(\beta)\right) = (-1,1)$ for any β, whether candidates are welfare motivated or selfish, and whether voters are strategic, naive, or vote their signals.*

For linear $g(s|z)$, Examples 7 through 11 make clear that candidates polarize at least a little bit as long as candidates face aggregate uncertainty about the electoral outcome, but the extent of polarization is sensitive to assumptions about candidate motivations and voter sophistication. Example 12 makes clear that, if truth is binary, polarization is completely robust. Together, these examples make clear that polarization can be more or less sensitive to various theoretical forces but relies crucially on common interest voting.

8 Conclusion

From the perspective of standard private-interest election models, polarization is both vexing and perplexing, representing some inexplicable democratic failure. This paper casts polarization in new light, through the common-interest lens of Condorcet (1785). Ironically, polarization stems from unity: a candidate trusts voters to recognize her superior policy position and support her, even when her opponent is less extreme. Centrist policies do attract more votes, but novel equilibrium forces make the penalty for extremism small, consistent with empirical evidence, so that candidates can be as extreme as the most extreme voters, even when they want very badly to win. This polarizing confidence can arise endogenously, even when candidates start from identical preferences and beliefs.

Relaxing the assumption of binding platform commitments, I show in McMurray (2017b) that a common-interest model similar to the one here can explain why large margins of victory convey policy "mandates" from voters, which can be shaped even by votes for losing parties. Unlike standard private-interest models, the model above also extends readily to multiple dimensions, as I show in McMurray (2020), eplaining why logically related issue positions are bundled together so consistently, so that multidimensional political contests seem nearly one dimensional.

In the model above, voters are candidates' only source of information, and are collectively infallible. Future work should model candidate signals, and should also explore impediments to voters' and candidates' learning. Polarization might decline if candidates lose confidence that voters will fully learn the truth. Alternatively, polarization might increase if election outcomes are determined by swings in public opinion that policy positioning cannot control.

Policy recommendations from standard private-interest models might include raising office holder salaries or otherwise strengthening office motivation, to incentivize moderation. In the model above, this can indeed decrease polarization, but only by increasing pandering, which reduces welfare.[26] Generalizing the information structure seems especially important in that regard: if candidates are somehow overconfident, for example, polarization might be excessive. Empirically, candidates do seem to over-polarize, given that elections tend to be close, suggesting that voter consensus is weak, and the optimal policy positions $E(z|w=A)$ and $E(z|w=B)$ should be similar. Certainly, under-polarization does not seem an immediate danger. The most unassailable policy priority may simply be to ensure that information is as widespread and accurate as possible.

A Appendix

Proof of Lemma 1. In terms of (3), the numbers N_A and N_B of A votes and B votes are Poisson random variables with means $n\phi(A|z)$ and $n\phi(B|z)$, so the joint distribution of vote totals a and b is simply the product

$$\psi(a,b|z) = \frac{e^{-n\phi(A|z)-n\phi(B|z)}}{a!b!}[n\phi(A|z)]^a[n\phi(B|z)]^b \tag{4}$$

of Poisson probabilities. Candidates A and B win the election with the probabilities $\Pr(A|z) = \Pr(N_A > N_B|z) + \frac{1}{2}\Pr(N_A = N_B|z)$ and $\Pr(B|z) = \Pr(N_A < N_B) + \frac{1}{2}\Pr(N_A = N_B|z)$, where the second term in both expressions reflects the possibility of winning a tie-breaking coin toss.

Because of the environmental equivalence property of Poisson games (Myerson, 1998), a voter reinterprets N_A and N_B as the number of votes cast by his peers: his own vote will add one to either total. His vote can be pivotal (event P) if the candidates otherwise tie or if one candidate trails by exactly one vote but would win

[26] That private- and common-interest models that are otherwise so similar can deliver such opposite welfare perspectives underscores the danger of taking the standard paradigm for granted.

the tie-breaking coin toss; in terms of (4), this occurs with the following probability.

$$\begin{aligned}\Pr(P|z) &= \Pr(N_A = N_B|z) + \frac{1}{2}\Pr(N_A = N_B + 1|z) + \frac{1}{2}\Pr(N_B = N_A + 1|z) \\ &= \sum_{k=0}^{\infty}\left[\psi(k,k|z) + \frac{1}{2}\psi(k,k+1|z) + \frac{1}{2}\psi(k+1,k|z)\right] \end{aligned} \quad (5)$$

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

$$t^{br} = \frac{\bar{x} - E(z|P)}{E(z^2|P) - \bar{x}E(z|P)} = \frac{\bar{x} - E(z|P)}{V(z|P) + E(z|P)^2 - \bar{x}E(z|P)} \quad (6)$$

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

$$E(z|P) = \frac{\int_Z z\Pr(P|z)f(z)\,dz}{\int_Z \Pr(P|z)f(z)\,dz} \quad (7)$$

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 $t^{br}(t)$ to an ideological strategy with ideology threshold t decreases with t, and using that fact shows that if $x_A < x_B$ then there exists a unique fixed point $t^* = t^{br}(t^*)$ that characterizes an ideological strategy that is its own best response. From (6) it can be seen that, for any $t \in X$, $t^{br}(t)$ depends on x_A and x_B only through the midpoint $\bar{x}$; accordingly, an ideological strategy v_{t^*} with ideology threshold t^* characterizes the unique equilibrium response to any pair of candidate platforms with the midpoint $\bar{x}$. From (6) it is clear that $t^{br}(t)$ also increases in $\bar{x}$, for any t; since $t^{br}(t)$ decreases in t but increases in $\bar{x}$ for any t, the solution $t^*(\bar{x})$ to $t^* = t^{br}(t^*;\bar{x})$ increases in $\bar{x}$, as claimed.

For an ideological strategy, (3) can be rewritten as follows.

$$\phi(A|z;v_t) = \int_{-\infty}^{t} g(s|z) = G(t|z) \quad (8)$$

$$\phi(B|z;v_t) = \int_{t}^{\infty} g(s|z) = 1 - G(t|z) \quad (9)$$

From these it is straightforward to show that $\phi(A|-z;v_{-t}) = \phi(B|z;v_t)$, which by (4) through (7) translates into symmetric pivot probabilities (i.e. $\Pr(P|-z;v_{-t}) = \Pr(P|-z;v_t)$) and therefore symmetric expectations $E(z|P;v_{-t}) = -E(z|P;v_t)$ and $E(z^2|P;v_{-t}) = E(z^2|P;v_t)$. If $t^{br}(t^*;\bar{x}) = t^*$, therefore, then from (6) it is clear that

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

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

Proof of Lemma 2. The first step is to show that if $v \in V$ is non-ideological then there exists an ideological strategy v_t such with $E[u(x,v);v_t] > E[u(x,v);v]$. To that end, define v_t such that $\phi(B;v_t) = \int_Z \int_t^1 g(s|z) f(z)\, ds dz = \int_Z \int_S v_B(s) g(s|z) f(z) = \phi(B;v)$. This is well defined because the left-hand side decreases in t, from 1 (for $t=-1$) to 0 (for $t=1$). Say that $\phi(B;v) < \frac{1}{2}$, so that $t>0$; otherwise, symmetric arguments apply. That $G(s|z)$ satisfies MLRP implies that $g(s'|z') g(s|z) > g(s'|z) g(s|z')$ for any $s'>s$ and $z'>z$. Writing $\phi(B|z;v) = \int_S v_B(s) g(s|z) = \int_{-1}^t v_B(s) g(s|z)\, ds + \int_t^1 v_B(s') g(s|z)\, ds'$ in terms of t then makes clear that $\frac{\phi(B|z;\sigma_t)}{\phi(B|z;\sigma)}$ increases in z, as the following holds for any $z'>z$.

$$
\begin{aligned}
& \phi(B|z';v_t)\,\phi(B|z;v) - \phi(B|z';v)\,\phi(B|z;v_t) \\
= & \int_t^1 g(s'|z')\, ds' \left[\int_{-1}^t v_B(s) g(s|z)\, ds + \int_t^1 v_B(s') g(s'|z)\, ds \right] \\
& - \left[\int_{-1}^t v_B(s) g(s|z')\, ds + \int_t^1 v_B(s') g(s'|z')\, ds' \right] \int_t^1 g(s'|z)\, ds' \\
= & \int_t^1 \int_{-1}^t v_B(s) \left[g(s'|z') g(s|z) - g(s|z') g(s'|z) \right] ds ds' \\
> & \ 0
\end{aligned}
$$

The monotonicity of $\frac{\phi(B|z;\sigma_t)}{\phi(B|z;\sigma)}$ implies that $\phi(B|z;v_t) > \phi(B|z;v)$ if and only if z exceeds some $\bar{z}$. $\Pr(w=B)$ (or, more succinctly, $\Pr(B)$) is monotonic in $\phi(B)$, so $\Pr(B|z;v_t) > \Pr(B|z;v)$ and

$$
f(z|B;v_t) = \frac{f(z)\Pr(B|z;v_t)}{\Pr(B;v_t)} > \frac{f(z)\Pr(B|z;v)}{\Pr(B;v)} = f(z|B;v)
$$

for any $z>\bar{z}$, as well.[27] Since $f(z|B;v_t)$ and $f(z|B;v)$ are single-crossing, $F(z|B;v_t)$ first-order stochastically dominates $F(z|B;v)$, implying that $E(z|B;v_t) > E(z|B;v)$.

The utility difference $\Delta u(x,z) = u(x_B,z) - u(x_A,z)$ between the two candidates' platforms reduces as follows,

$$
\begin{aligned}
\Delta u(x,z) &= -(z-x_B)^2 + (z-x_A)^2 \\
&= 2(x_B - x_A)(z-\bar{x})
\end{aligned}
$$

so that the welfare difference $\Delta W = E_{j,z}[u(x_j,z);v_t] - E_{j,z}[u(x_j,z);v]$ between the

[27] Increasing $\phi(B)$ causes a first-order stochastic dominance increase in the distribution of N_B, thus increasing $\Pr(N_B > N_A)$ (for any realization of N_A, and therefore unconditionally). $\Pr(N_A < N_B)$ similarly decreases in ϕ, so $\Pr(B) - \Pr(A) = \Pr(N_A < N_B) - \Pr(N_A > N_B)$ and therefore $\Pr(B)$ both increase in ϕ.

two voting strategies can be seen to be positive.

$$\begin{aligned}\Delta W &= E_z\left\{u\left(x_A,z\right)\left[\Pr\left(A|z;v_t\right)-\Pr\left(A|z;v\right)\right]+u\left(x_B,z\right)\left[\Pr\left(B|z;v_t\right)-\Pr\left(B|z;v\right)\right]\right\}\\ &= E_z\left\{\Delta u\left(x,z\right)\left[\Pr\left(B|z;v_t\right)-\Pr\left(B|z;v\right)\right]\right\}\\ &= 2\left(x_B-x_A\right)E_z\left\{\left(z-\bar{x}\right)\left[\Pr\left(B|z;\sigma_t\right)-\Pr\left(B|z;\sigma\right)\right]\right\}\\ &= 2\left(x_B-x_A\right)\left\{\Pr\left(B|z;v_t\right)\left[E\left(z|B;v_t\right)-\bar{x}\right]-\Pr\left(B;v\right)\left[E\left(z|B;v\right)-\bar{x}\right]\right\}\\ &= 2\left(x_B-x_A\right)\Pr\left(B;v\right)\left[E\left(z|B;v_t\right)-E\left(z|B;v\right)\right]\\ &> 0\end{aligned}$$

Here, the final equality holds because $\Pr(B;v_t)=\Pr(B;v)$ by the definition of v_t.

Welfare $W=E\left[u\left(x,v\right);v_t\right]$ is continuous in t over the compact set S of possible thresholds, so an optimal t^{**} exists by the extreme value theorem. Since every non-ideological strategy is dominated by an ideological strategy and an optimal ideological strategy exists, this is the optimal voting strategy for the game. In a common interest game such as this, McLennan (1998) shows that any behavior that is socially optimal is also individually optimal, and thus constitutes an equilibrium. By Lemma 1, there is only one such strategy, denoted here by $v^{**}=v_{t^{**}}$. ■

Proof of Lemma 3. As the proof of Lemma 2 notes, $\frac{\phi(B|z;v_t)}{\phi(A|z;v_t)}$ increases in z, implying that a unique state $\bar{z}(t)$ minimizes $|\phi(A|\bar{z}(t);v_t)-\phi(B|\bar{z}(t);v_t)|$, where $\bar{z}(t)=-1$ for t sufficiently low and $\bar{z}(t)=1$ for t sufficiently high. $\phi(A|z;v_t)$ and $\phi(B|z;v_t)$ also increase and decrease in t (see (8) and (9)), so $\bar{z}(t)$ increases in t. As n grows large, $\Pr_n(P|z)$ decreases to zero for any z, but as Myerson (2000) shows, the magnitude of $\Pr_n(P|z)$ shrinks at rate $\frac{1}{\sqrt{n}}$ for $z=\bar{z}(t)$ and at rate e^{-n} for all other z. Thus, $f(z|P)$ converges to a degenerate distribution with unit mass on $\bar{z}(t)$, implying that $\lim_{n\to\infty}E(z|P)=\bar{z}(t)$ and $\lim_{n\to\infty}V(z|P)=0$. Note that $\bar{z}(t)$ has the same sign as t, because $\phi(A|z=0;v_0)=\phi(B|z=0;v_0)=\frac{1}{2}$. Therefore, the right-hand side of (6) converges to $-\frac{1}{\bar{z}(t)}$, implying that $\lim_{n\to\infty}t_n^{br}(t;\bar{x})=\begin{cases}1\text{ if }\bar{z}(t)<\bar{x}\\ 0\text{ if }\bar{z}(t)=\bar{x}\\ -1\text{ if }\bar{z}(t)>\bar{x}\end{cases}$. Let $t_{\bar{x}}$ denote the solution to $\bar{z}(t)=\bar{x}$, which is unique since $\bar{z}(t)$ increases in t. For any ε there is an n large enough that $t_n^{br}(t_{\bar{x}}-\varepsilon)>t_{\bar{x}}+\varepsilon$ and $t_n^{br}(t_{\bar{x}}+\varepsilon)<t_{\bar{x}}-\varepsilon$. Since $t_n^{br}(t)$ decreases in t, this implies that $t_n^*\in(t_{\bar{x}}-\varepsilon,t_{\bar{x}}+\varepsilon)$. In other words, $t_\infty^*(\bar{x})\equiv\lim_{n\to\infty}t_n^*(\bar{x})=t_{\bar{x}}$. Equivalently, $\bar{z}(t_\infty^*(\bar{x}))=\bar{x}$. ■

Proof of Theorem 1. Since equilibirum voting is ideological (by Lemma 1), $\phi(j|z;v_n^*)$ are given by (8) and (9). Since $G(s|z)$ satisfy MLRP, these decrease and increase in z, respectively. Lemma 3 implies that, for a sequence (v_n^*) of equilibrium voting strategies, $\lim_{n\to\infty}\phi(A|z=\bar{x};v_n^*)=\lim_{n\to\infty}\phi(B|z=\bar{x};v_n^*)=\frac{1}{2}$, implying that $\lim_{n\to\infty}\phi(A|z;v_n^*)>\frac{1}{2}>\lim_{n\to\infty}\phi(B|z;v_n^*)$ for $z<\bar{x}$ and $\lim_{n\to\infty}\phi(A|z;v_n^*)<\frac{1}{2}<$

$\lim_{n\to\infty} \phi(B|z;v_n^*)$ for $z > \bar{x}$. By the law of large numbers, $\frac{N_B}{N_A+N_B}$ converges in probability to $\lim_{n\to\infty} \phi(B|z;v_n^*)$, so $\lim_{n\to\infty} \Pr(A|z;v_n^*) = 1$ if $z < \bar{x}$ and $\lim_{n\to\infty} \Pr(B|z;v_n^*) = 1$ if $z > \bar{x}$. In either of these cases, $\lim_{n\to\infty} \Pr(w = j^*|z) = 1$. This equality holds for $z = \bar{x}$, as well, since A and B are equidistant from z in that case, so $w = A$ and $w = B$ both imply $w = j^*$. Since $\lim_{n\to\infty} \Pr(w = j^*|z;v_n^*) = 1$ for all z, $\lim_{n\to\infty} \Pr(w = j^*;v_n^*) = 1$ unconditionally. ■

Proof of Lemma 4. The expectation of $u_j = u(x_w, z) + \beta 1_{w=j}$ can be written as follows,

$$E_{w,z}(u_j;\sigma_{t^*}) = E_z\left[\sum_{w=j,-j} u(x_w,z)\Pr(w|z;x_A,x_B,\sigma_{t^*})\right] + \beta\Pr(w = j;x_A,x_B,\sigma_{t^*}) \tag{10}$$

and is continuous with respect to x_B over the compact set X of platforms, so an optimum x_B^{br} exists by the extreme value theorem. Differentiating (10) with respect to x_B yields the following,

$$\begin{aligned}\frac{\partial E(u_B)}{\partial x_B} &= E_z\left[\frac{\partial u(x_B,z)}{\partial x_B}\Pr(B|z;\sigma_{t^*})\right] + \frac{\partial E[u(x,z);\sigma_{t^*}]}{\partial t^*(\bar{x})}\frac{\partial t^*(\bar{x})}{\partial \bar{x}}\frac{\partial \bar{x}}{\partial x_B} + \beta\frac{\partial\Pr(B;\sigma_{t^*})}{\partial x_B} \\ &= E_z[2(z - x_B)\Pr(B|z;\sigma_{t^*})] + \beta\frac{\partial\Pr(B;\sigma_{t^*})}{\partial\phi(B;\sigma_{t^*})}\frac{\partial\phi(B;\sigma_{t^*})}{\partial\bar{x}}\frac{\partial\bar{x}}{\partial x_B} \\ &= 2\Pr(B)[E(z|B;\sigma_{t^*}) - x_B] + \frac{1}{2}\beta\frac{\partial\Pr(B;\sigma_{t^*})}{\partial\phi(B;\sigma_{t^*})}\frac{\partial\phi(B;\sigma_{t^*})}{\partial t^*(\bar{x})}\frac{\partial t^*(\bar{x})}{\partial\bar{x}}\end{aligned} \tag{11}$$

which reduces as it does because σ_{t^*} characterizes equilibrium voting behavior for all platform pairs, by Lemma 1, and t^* is socially optimal, by Proposition 1, so $\frac{\partial E[u(x,z);\sigma_{t^*}]}{\partial t^*(\bar{x})} = 0$.

When $\beta = 0$, $\frac{\partial E(u_B)}{\partial x_B}$ has the same sign as $E(z|w = B) - x_B$. This is positive for $x_B = -1$, negative for $x_B = x_A - \varepsilon$ with ε sufficiently small, positive for $x_B = x_A + \varepsilon$, and negative for $x_B = 1$, implying the existence of a local optimum in $(-1, x_A)$ and another in $(x_A, 1)$, both of which satisfy $x_B^{br} = E\left(z|B;x_A,x_B^{br},\sigma\right)$. Whichever of these generates higher expected utility is the best response to x_A. Symmetrically, the best response to x_B exists and satisfies $x_A^{br} = E\left(z|A;x_A^{br},x_B,\sigma\right)$. ■

Proof of Theorem 1. If $(x_A, x_B) = (-x, x)$ then $\bar{x} = 0$, so by Lemma 1, voters respond in equilibrium with an ideological voting strategy with $\tau^*(0) = 0$. This generates symmetric voting outcomes: $\phi(A|z) = \phi(B|-z)$ and therefore $\Pr(A|z) = \Pr(B|-z)$, implying that $\Pr(A) = \Pr(B) = \frac{1}{2}$ and $E(z|A) = -E(z|B)$. From (11), noting that $\frac{\partial\Pr(A;\sigma_t)}{\partial t} = -\frac{\partial\Pr(B;\sigma_t)}{\partial t^*}$, this also implies that $\frac{\partial E(u_A;x_A=-x,x_B=x)}{\partial x_A} = -\frac{\partial E(u_B;x_A=-x,x_B=x)}{\partial x_B}$. These derivatives are (symmetrically) linear in x, so there exists

x^* such that $\frac{\partial E(u_A;x_A=-x^*,x_B=x^*)}{\partial x_A} = \frac{\partial E(u_B;x_A=-x^*,x_B=x^*)}{\partial x_B} = 0$ but $\frac{\partial E(u_A;x_A=-x,x_B=x)}{\partial x_A} > 0 > \frac{\partial E(u_B;x_A=-x,x_B=x)}{\partial x_B}$ for $x < x^*$ and inequalities are reversed for $x > x^*$, implying that $(-x^*, x^*, \sigma^*)$ constitutes a PBE.

Proposition 1 implies that $\lim_{n\to\infty} \Pr(B|z) = \begin{cases} 0 \text{ if } z < \bar{x} \\ 1 \text{ if } z > \bar{x} \end{cases}$, so $\lim_{n\to\infty} \Pr(B) = 1 - F(\bar{x})$, $\lim_{n\to\infty} E(z|B) = E(z|z > \bar{x})$, and $\lim_{n\to\infty} \frac{\partial \Pr(B)}{\partial \bar{x}} = -f(\bar{x})$. Accordingly, (11) reduces to the following.

$$\lim_{n\to\infty} \frac{\partial E(u_B)}{\partial x_B} = 2[1 - F(\bar{x})][E(z|z > \bar{x}) - x_B] - \frac{1}{2}\beta f(\bar{x}) \tag{12}$$

When $\beta = 0$, this expression is zero if and only if $\lim_{n\to\infty} x_{B,n}^{br} = E(z|z > \bar{x})$. The corresponding best response for A approaches $\lim_{n\to\infty} x_{A,n}^{br} = E(z|z < \bar{x})$. Equilibrium requires that $\frac{x_A^* + x_B^*}{2} = \bar{x}$, so the limit of equilibrium platforms $x_{A,\infty}^* = \lim_{n\to\infty} x_{A,n}^*$ and $x_{B,\infty}^* = \lim_{n\to\infty} x_{B,n}^*$ must satisfy $\frac{E(z|z<\bar{x}_\infty^*)+E(z|z>\bar{x}_\infty^*)}{2} = \bar{x}_\infty^* = \frac{x_{A,\infty}^* + x_{B,\infty}^*}{2}$. This is satisfied for $\bar{x}_\infty^* = 0$, given the symmetry of F, but no other solution exists, because the log-concavity of F implies that $E(z|z < \bar{x})$ and $E(z|z > \bar{x})$, and therefore $\frac{E(z|z<\bar{x})+E(z|z>\bar{x})}{2}$, all increase in $\bar{x}$ with slope smaller than one. ■

Proof of Theorem 2. For any symmetric platform pair, $\bar{x} = t^*(\bar{x}) = 0$. Other than x_B, the terms in (11) are the same for any such pair. Since $\frac{\partial \Pr(B;\sigma_{t^*})}{\partial \phi(B;\sigma_{t^*})} > 0$ (as explained in Footnote 27), $\frac{\partial \phi(B;\sigma_{t^*})}{\partial t^*(\bar{x})} < 0$ (as can be seen in (9)), and $\frac{\partial t^*(\bar{x})}{\partial \bar{x}} > 0$ (by Lemma 1), this expression decreases linearly in β, and is negative for any $\beta > \bar{\beta} = \frac{4\Pr(B)[x_B - E(z|B)]}{\frac{\partial \Pr(B;\sigma_{t^*})}{\partial \phi(B;\sigma_{t^*})} \frac{\partial \phi(B;\sigma_{t^*})}{\partial t^*(\bar{x})} \frac{\partial t^*(\bar{x})}{\partial \bar{x}}}$. In that case, the only platform-symmetric BNE is $(x_A, x_B) = (0, 0)$. For $\beta < \bar{\beta}$, (11) equals zero if and only if $x_B = x_B^{br}(\beta) \equiv E(z|B) + \frac{1}{2}\beta \frac{\partial \Pr(B;\sigma_{t^*})}{\partial \phi(B;\sigma_{t^*})} \frac{\partial \phi(B;\sigma_{t^*})}{\partial t^*(\bar{x})} \frac{\partial t^*(\bar{x})}{\partial \bar{x}}$. This final product is negative, so $x_B^{br}(\beta)$ decreases in β in that case. ■

Proof of Theorem 3. As n grows large, $\frac{\partial E(u_B)}{\partial x_B}$ approaches (12). For $\bar{x} = 0$, this reduces to the following,

$$\lim_{n\to\infty} \frac{\partial E(u_B)}{\partial x_B} = [E(z|z > 0) - x_B] - \frac{1}{2}\beta f(0)$$

which is negative for all x_B if $E(z|z > 0) - \frac{1}{2}\beta f(0) \leq 0$, implying that candidate B (and, symmetrically, candidate A) cannot improve on the platform pair $(0, 0)$. Otherwise, $x_B = E(z|z > 0) - \frac{1}{2}\beta f(0)$ (together with $x_A = -x_B$) satisfies the limiting first-order condition for equilibrium. ■

Proof of Theorem 2. Lemma 2 states that the equilibrium voting strategy σ_{t^*}

optimally responds to any pair $(x_A, x_B) \in X^2$ of candidate platforms, so a search for an optimal strategy vector reduces to a search for optimal candidate platforms $\left(x^{**}_{A,n}, x^{**}_{B,n}\right)$. Since expected utility is continuous over the compact set X^2 of such platform pairs, an optimal pair exists by the extreme value theorem. When $\beta = 0$, candidates share a common preference with each other and with voters. Since all players of a game share a common interest in that case, a strategy vector that is socially optimal is also individually optimal, as McLennan (1998) points out, and so constitutes a PBE. The limiting equilibrium platform is then given by Theorem 1. ■

Proof of Example 7. For candidate B, expected utility can then be written as follows,

$$E(u_B) = \int_X \left[\sum_{w=A,B} u(x_w, \hat{x}_B) \Pr(w|z) + \beta \Pr(B|z) \right] f(z)\, dz$$

with the following derivative,

$$\begin{aligned} \frac{\partial E(u_B)}{\partial x_B} &= \int_X \left[-2(x_B - \hat{x}_B)\Pr(B|z) + [2(x_B - x_A)(\hat{x}_B - \bar{x}) + \beta] \frac{\partial \Pr(B|z)}{\partial \bar{x}} \frac{\partial \bar{x}}{\partial x_B} \right] f(z)\, dz \\ &= -2(x_B - \hat{x}_B)\Pr(B) + \frac{1}{2}[2(x_B - x_A)(\hat{x}_B - \bar{x}) + \beta] \frac{\partial \Pr(B)}{\partial \bar{x}} \\ &= -(x_B - \hat{x}_B) + \frac{1}{2}(4x_B\hat{x}_B + \beta)\frac{\partial \Pr(B)}{\partial \bar{x}} \end{aligned} \quad (13)$$

where the final equality imposes $x_A = -x_B$, $\bar{x} = 0$, and $\Pr(B) = \frac{1}{2}$. With strategic voting, the jury theorem implies that $\lim_{n\to\infty} \Pr(B|z) = \begin{cases} 0 \text{ if } z < \bar{x} \\ 1 \text{ if } z > \bar{x} \end{cases}$, so $\lim_{n\to\infty} \Pr(B) = 1 - F(\bar{x})$ and $\lim_{n\to\infty} \frac{\partial \Pr(B)}{\partial \bar{x}} = -f(\bar{x}) = -\frac{1}{2}$, and the solution to $\frac{\partial E(u_B)}{\partial x_B} = 0$ approaches $x^*_{B,\infty} = \frac{\hat{x}_B - \frac{1}{4}\beta}{1+2\hat{x}_B}$, from which the result is immediate. ■

Proof of Example 8. Since $E(z|s_i) = \frac{1}{3}s_i$, a naive voter votes B if $s_i > 3\bar{x}$. Thus, $\phi(B|z) = 1 - G(3\bar{x}|z)$, where $G(3\bar{x}|z) = \int_{-1}^{3\bar{x}} \frac{1}{2}(1 + sz)\, ds = \frac{1}{2}(3\bar{x} + 1) + \frac{1}{4}z(9\bar{x}^2 - 1)$. Candidate B wins in a large election if $\phi(B|z)$ exceeds $\frac{1}{2}$, so $\lim_{n\to\infty} \Pr(B) = 1 - F\left(\frac{6\bar{x}}{1-9\bar{x}^2}\right)$ and $\lim_{n\to\infty} \frac{\partial \Pr(B)}{\partial \bar{x}} = -f\left(\frac{6\bar{x}}{1-9\bar{x}^2}\right) \frac{6\left(1-9\bar{x}^2\right) - 6\bar{x}(-18\bar{x})}{(1-9\bar{x}^2)^2} = -6f(0) = -3$. With welfare motivated candidates, expected utility can then be written as follows,

$$E(u_B) = \int_X \left[\sum_{w=A,B} u(x_w, z) \Pr(w|z) + \beta \Pr(B|z) \right] f(z)\, dz$$

with the following derivative,

$$\begin{aligned}
\frac{\partial E(u_B)}{\partial x_B} &= \int_X \left[-2(x_B - z)\Pr(B|z) + [2(x_B - x_A)(z - \bar{x}) + \beta] \frac{\partial \Pr(B|z)}{\partial \bar{x}} \frac{\partial \bar{x}}{\partial x_B}\right] f(z)\, dz \\
&= -2[x_B - E(z|B)]\Pr(B) \\
&\quad + \frac{1}{2}\left\{2(x_B - x_A)\int_X z\frac{\partial \Pr(B|z)}{\partial \bar{x}} f(z)\, dz + \beta \int_X \frac{\partial \Pr(B|z)}{\partial \bar{x}} f(z)\, dz\right\} \\
&= -[x_B - E(z|z>0)] + \frac{1}{2}\beta\frac{\partial \Pr(B)}{\partial \bar{x}} \qquad (14)
\end{aligned}$$

where the third equality follows because $\frac{\partial \Pr(B|-z)}{\partial \bar{x}} = -\frac{\partial \Pr(A|-z)}{\partial \bar{x}} = -\frac{\partial \Pr(B|z)}{\partial \bar{x}}$ and $f(-z) = f(z)$, so $\int_X z\frac{\partial \Pr(B|z)}{\partial \bar{x}} f(z)\, dz = 0$. This approaches $\frac{1}{2} - x_B - \frac{3}{2}\beta$ as n grows large, so the solution to $\frac{\partial E(u_B)}{\partial x_B} = 0$ approaches $x^*_{B,\infty} = \frac{1}{2} - \frac{3}{2}\beta$. ■

Proof of Example 9. With naive voters, $\frac{\partial \Pr(B)}{\partial \bar{x}}$ approaches -3, as in Example 8. With selfish candidates, $\frac{\partial E(u_B)}{\partial x_B}$ is given by (13) from Example 7. The latter approaches $\lim_{n\to\infty} \frac{\partial E(u_B)}{\partial x_B} = -\left(x_B - \frac{1}{2}\right) - \frac{3}{2}(2x_B + \beta)$, which is zero if $x^*_{B,\infty} = \frac{1}{8} - \frac{3}{8}\beta$, from which the claim is immediate. ■

Proof of Example 10. If $\hat{x}_A = -\frac{1}{2}$ and $\hat{x}_B = \frac{1}{2}$ then (13) reduces to $\frac{\partial E(u_B)}{\partial x_B} = -\left(x_B - \frac{1}{2}\right) + \frac{1}{2}(2x_B + \beta)\frac{\partial \Pr(B)}{\partial \bar{x}}$. When voters vote their signals, a simple modification of the proof of Example 8 shows that $\phi(B|z) > \frac{1}{2}$ if $z > \frac{2\bar{x}}{1-\bar{x}^2}$, so $\lim_{n\to\infty} \frac{\partial \Pr(B)}{\partial \bar{x}} = -1$, and the solution to $\frac{\partial E(u_B)}{\partial x_B} = 0$ approaches $x^*_{B,\infty} = \frac{1}{4}(1-\beta)$, from which the claim is immediate. ■

Proof of Example 11. For finite n, candidate B wins the election if and only if the median realized ideal point, $\hat{x}_m$, exceeds $\bar{x}$. Since $\hat{x}_m$ approaches the median of G which is zero, $\lim_{n\to\infty} \Pr(B) = \begin{cases} 0 \text{ if } \bar{x} < 0 \\ 1 \text{ if } \bar{x} > 0 \end{cases}$. Thus, $\lim_{n\to\infty} \frac{\partial \Pr(B)}{\partial \bar{x}} = -\infty$ for $\bar{x} = 0$, implying for any β that the limit of (13) is negative for any $x_A = -x_B$ pair, including for $x_A = x_B = 0$. Thus, any sequence of solution to $\frac{\partial E(u_B)}{\partial x_B} = 0$ approaches $x^*_{B,\infty} = 0$. ■

Proof of Example 12. Consider first the case of welfare motivated candidates and strategic voters. Proposition 1 ensures that $\lim_{n\to\infty} \Pr(B) = \begin{cases} 0 \text{ if } z < 0 \\ 1 \text{ if } z > 0 \end{cases}$ in that case, so $\lim_{n\to\infty} \frac{\partial \Pr(B)}{\partial \bar{x}} = 0$, and (14) reduces to $\frac{\partial E(u_B)}{\partial x_B} = -(x_B - 1)$, which is positive for any $x_B < 1$ and zero if $x_B = 1$, yielding the desired result. If voters vote their signals, $\phi(B|z) = 1 - G(\bar{x}|z)$ exceeds $\frac{1}{2}$ either if $z = -1$ and $\bar{x} > G^{-1}\left(\frac{1}{2}|z=-1\right)$ or if $z = 1$ and $\bar{x} > G^{-1}\left(\frac{1}{2}|z=1\right)$. Therefore, $\lim_{n\to\infty} \Pr(B) = \begin{cases} 0 \text{ if } \bar{x} > G^{-1}\left(\frac{1}{2}|z=1\right) \\ 1 \text{ if } \bar{x} < G^{-1}\left(\frac{1}{2}|z=-1\right) \\ \Pr(z=1) \text{ otherwise} \end{cases}$

and $\lim_{n\to\infty} \frac{\partial \Pr(B)}{\partial \bar{x}} = \begin{cases} -\infty \text{ if } \bar{x} \in \left\{ G^{-1}\left(\frac{1}{2}|z=-1\right), G^{-1}\left(\frac{1}{2}|z=1\right) \right\} \\ 0 \text{ otherwise} \end{cases}$. Again, (14) is positive unless $x_B < 1$ (except the two points at which it is negative). A similar conclusion holds if voters vote their signals, in which case $\phi(B|z) = 1 - G(3\bar{x}|z)$, so $\lim_{n\to\infty} \frac{\partial \Pr(B)}{\partial \bar{x}} = \begin{cases} -\infty \text{ if } \bar{x} \in \left\{ \frac{1}{3}G^{-1}\left(\frac{1}{2}|z=-1\right), \frac{1}{3}G^{-1}\left(\frac{1}{2}|z=1\right) \right\} \\ 0 \text{ otherwise} \end{cases}$. If candidates are selfish with $\hat{x}_A = -1$ and $\hat{x}_B = 1$ then (13) reduces to $\frac{\partial E(u_B)}{\partial x_B} = -(x_B - 1) + \frac{1}{2}(4x_B + \beta)\frac{\partial \Pr(B)}{\partial \bar{x}}$, producing the same conclusion. ■

References

[1] 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.

[2] 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.

[3] 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.

[4] 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.

[5] Bagnoli, Mark, and Ted Bergstrom. 2005. "Log-Concave Probability and Its Applications." *Economic Theory*, 26(2): 445-469.

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

[7] 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.

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

[9] 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.

[10] 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.

[11] Cohen, Marty, Mary C. McGrath, Peter Aronow, and John Zaller. 2016. "Ideologically Extreme Candidates in U.S. Presidential Elections, 1948-2012." *Annals of the American Academy of Political and Social Science*, 667(1): 126-142.

[12] 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.

[13] 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.

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

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

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

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

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

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

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

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

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

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

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

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

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

[27] 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.

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

[29] 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.

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

[31] 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.

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

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

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

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

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

[37] 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.

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

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

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

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

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

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