A Theory of Political Polarization*

John W. Patty†   Elizabeth Maggie Penn‡

August 22, 2017

Abstract

We present a simple theory of voters' preferences over representatives, assuming that a representative will vote on one's behalf. Few, if any elected representatives are capable of unilaterally implementing their platforms: rather, they choose between options generated by other actors and/or external events. When this is the case, voters' preferences over candidates' platforms will almost always be asymmetric even if voters' preferences over policy outcomes are symmetric. Furthermore, these induced preferences tend to prefer more extreme ("polarizing") representatives when the legislative agenda is independent of the status quo.

*Sections 3 & 4 are currently incomplete. Of course, comments on all parts are very welcome. Prepared for presentation at the Colloquium on Law Economics and Politics at New York University School of Law, September 5, 2017. We thank Alberto Alesina, Chris Berry, Thomas Bräuninger, John Brehm, Ethan Bueno de Mesquita, Peter Buisseret, Bill Clark, Tom Clark, Torun Dewan, Wioletta Dziuda, Scott DeMarchi, Mark Fey, Sean Gailmard, Bernard Grofman, Gabe Lenz, Thomas Leeper, Tom Mann, Gerard Padró i Miquel, Amy Pond, Jas Sekhon, Ken Shepsle, Ahmer Tarar, Stephane Wolton, Dan Wood, and audience members at the University of Chicago, University of California-Berkeley, University of California-Irvine, University of California-Merced, Harvard University, London School of Economics, University of North Carolina-Chapel Hill, Stanford University, Texas A&M University, University of Warwick, and the 2016 annual meetings of the Southern and Midwest Political Science Associations for very helpful comments on, and conversations, about this project, which was previously circulated under the title "Does Representation Induce Polarization? A Theory of Choosing Representatives." All errors are our own.

†Professor of Political Science, University of Chicago. Email: *jwpatty@uchicago.edu*.

‡Professor of Political Science, University of Chicago. Email: *epenn@uchicago.edu*.

Introduction

Spatial theories of voting presume that voters reduce a candidate's platform to an ideological position.[1] This ideological position captures how the voter believes that the candidate will, if elected, affect policy outcomes.[2] In most theories of electoral competition, voters presume that the platform of the winning candidate will be the policy that is implemented after the election. This is a convenient and productive simplification but it is not an innocuous one, particularly when considering the role of most elected officials: few, if any, political offices in a democracy allow the officeholder to unilaterally impose his or her will by fiat. Rather, the official must work through an institutionalized process in order to have some effect on public policy.[3]

In this article, we focus on a near-ubiquitous characteristic of policymaking processes: the menu of choices from which a representative may choose is at least partially determined by actors and/or events beyond the representative's control. Put differently, most officials with decision-making authority spend most of their time making decisions about issues and between choices that were chosen by someone else. Both internal procedures and external events, such as disasters, force policymakers to choose between options other than their most-preferred policies.[4] Examples include a legislator who may implement a "platform" only through voting on bills that are not necessarily representative of the policies that he or she would implement if given unilateral authority; an executive who may only sign or veto legislation passed by a legislature; and a judge who may only make decisions on cases

[1]For discussions of spatial preferences, see Grofman (2004), Dewan and Shepsle (2011), Hinich and Munger (1992, 1996), Eguia (2013). In terms of electoral competition, this basic framework is sometimes extended to include a candidate-specific "valence" dimension (*e.g.*, Groseclose (2001), Schofield (2004), Ashworth and Bueno de Mesquita (2009), Carter and Patty (2015)).

[2]The question of the degree to which a candidate can choose (or, "commit to") a given platform has been considered in depth by many scholars (*e.g.*, Osborne and Slivinsky (1996), Besley and Coate (1997), and Dhillon and Lockwood (2002)). For our purposes, it is irrelevant where platforms "come from." Rather, we are interested in how voters should evaluate and compare various platforms when the agenda is at least partially beyond the control of the candidate.

[3]The institutionally imposed divergence between goals and actions is treated very generally in Penn, Patty and Gailmard (2011) and Gailmard, Patty and Penn (2008).

[4]We mention and set to the side for future work the fact that electoral incentives within a legislature could have similar effects, to the degree that some individuals seek to stake out positions on issues through dilatory tactics or other forms of obstruction (Patty (2016)).

brought by others.

We explore the implications of this reality in this article. Specifically, we model a simple setting in which a voter, with an ideological position ($v_i \in \mathbb{R}$), must choose a candidate with a known platform ($p_c \in \mathbb{R}$).[5] This candidate will then be faced with a pair of options to choose from, and will pick the option closest to his or her platform. We find that a voter's expectations about the options the representative will confront affect how the voter views ideological platforms distinct from his or her own position. When representatives are faced with an exogenous set of options to choose from, *voter preferences for candidate platforms do not look like voter preferences for policy*. In particular, in this environment a voter's strategic evaluation, or "induced preference," over candidate platforms becomes asymmetric about the voter's own ideal position. We show that, in a variety of choice environments, these expectations induce a strict preference on the part of the voter for more extreme candidates. This "extremity" is relative to the choice set: if the voter tends to be (say) to the right of the alternatives likely to be brought up as potential choices, then a more extreme candidate is one whose platform is to the right of the voter's ideological position.

A consequence of this is that when a voter feels that the legislative agenda is ideologically distant from his or her own position, the voter will have a taste for candidates farther from the legislative action than ones closer to it, even when considering candidate platforms that are equidistant from the voter. We demonstrate that this induced taste for extremism can be reversed under some circumstances. Perhaps unsurprisingly, one such circumstance is when the distribution of the options likely to arise has high variance, and places significant weight on policies that are extreme relative to the voter. In this case, voters will prefer moderate candidates: candidates whose platforms are closer to the alternatives likely to be brought up on the agenda.

Our goal in this article is not to offer a theory of electoral competition (although we do touch upon this issue later in the paper): we are not attempting to explain how candidates choose the platforms they offer to voters. It is also not a model of strategic delegation, in which voters strategically choose candidates that will engage in a larger collective choice problem, or candidates that have private information about their own types. It is well-

[5] As we will discuss later, there is nothing limiting our argument to a unidimensional policy space.

known that many agency problems generate incentives for a principal to choose an agent whose preferences differ from his or her own.[6] In strategic delegation settings these incentives arise from the fact that the agent's preferences will influence the strategic behavior of third-party actors in ways that benefit the principal, or from the fact that preference divergence between a principal and agent may induce beneficial behavior (*e.g.* effort) on the part of the agent. In contrast, we assume that the voter cares only about how his or her representative will sincerely choose from an exogenous slate of policy choices. Thus, the divergence in voter preferences for policies and preferences for platforms that we identify is based solely on unilateral and sincere behavior on the part of the representative. *When choosing someone to choose on one's behalf, a voter may be better represented by a candidate farther from their ideal position than one closer.*

To our knowledge, our paper is most closely related to Grofman (1985), which discusses the role of the status quo in spatial models of voting. Grofman identifies a problem similar to our starting concern: that candidates are limited in their ability to implement their platforms. He focuses on the role of the status quo as representing a starting point for policy change, and develops a model in which a voter does not care directly about party platforms but about the outcomes the voter thinks each party is capable of achieving. Capability is represented by a "performance weight" that dictates how far toward its ideal point a party is capable of shifting the status quo. Grofman argues that his model is closer in spirit to Downs's original work on the spatial model, quoting from Downs's *An Economic Theory of Democracy*,

> "...if [a voter] is rational, he knows that no party will be able to do everything that it says it will do. Hence he cannot merely compare platforms; instead he must estimate in his own mind what the parties would actually do were they in power."[7]

Moreover, Grofman's model generates two insights that we also identify, for related reasons. First, voter preferences over candidate platforms may change over time even when

[6] This is discussed in detail in Penn, Patty and Gailmard (2011).

[7] Quoted in (Grofman, 1985, p. 231).

those platforms remain unchanged. And second, it is not always advantageous for a candidate to be thought able to implement his or her platform; a very extreme but incapable candidate may be successful precisely because he or she moves the status quo in the right direction but not too far. In Grofman's model, both of these insights are consequences of the location of the status quo and how parties are assumed to shift policy from this point. In our model they stem from the voter's perception of the choices the candidate is likely to encounter. The main, and crucial, distinction between these two models is that we remain firmly grounded in the spatial model of choice. In our model there are only voter platforms, candidate platforms, and choices likely to be encountered by candidates. If two candidates will always vote the same way then the voter is indifferent between them, regardless of their platforms. In contrast, Grofman presents a directional model of party behavior in which the party is capable of implementing a weighted measure of its platform.

In the remainder of the paper we set up our simple model of voter choice, present a few general insights that fall out of the model, and walk the reader through an illustration of our argument that provides intuition for our finding that a taste for extremism will often emerge when voters choose a representative to choose on their behalf. We then focus in more detail on a natural setting in which the status quo is fixed and known to the voter, but a "bill" likely to be pitted against it is not, and the voter must choose a candidate to choose between this known status quo and unknown bill. Following this, we present three applications of our theory to well-known models of politics: Downsian competition, Osborne and Slivinsky's citizen-candidate model, and the Romer-Rosenthal model.[8] We conclude with a discussion of our model's contribution to current debates on the origins of polarization in American politics today.

1 The Model

We consider a model of voting between candidates with policy platforms in a unidimensional policy space $X \subseteq \mathbb{R}$ and denote the set of voters by $N = \{1, \dots, n\}$. We characterize each voter i by his or her *ideal point*, $v_i \in \mathbb{R}$, and his or her preferences are represented by

[8]These models are in Downs (1957), Osborne and Slivinsky (1996), and Romer and Rosenthal (1978).

a real-valued policy payoff function $u : \mathbf{R}^2 \to \mathbf{R}$ where $u(x, v_i)$ denote the payoff received by voter i if policy x is chosen.

For expository purposes (to highlight the asymmetric preferences our model induces) it is useful to assume that u is a strictly decreasing function of the distance between x and v_i, and thus that u is symmetric about the voter's ideal point.[9]

Assumption 1 *For any ideal point v and pair of policies x and y,*

$$u(x, v) > u(y, v) \Leftrightarrow |x - v| < |y - v|.$$

Suppose that a choice is to be made between two policies, a *bill*, b, and a *status quo*, q. With Assumption 1 in hand, Figure 1 displays the voter's optimal vote choice for all pairs of bills and status quos. This figure depicts every vote a candidate could take, with each vote represented by the pair (q, b), with q on the $x-$ axis denoting a preexisting status quo and b on the $y-$axis denoting a bill that has been proposed to replace the status quo.

Clearly when $b = q$ the voter is indifferent between the bill and status quo, because they are the same. The voter is also indifferent between b and q when $b = 2v - q$, because along this line $-|v - b| = -|v - (2v - q)| = -|v - q|$. These two lines along which the voter is indifferent between b and q are pictured in Figure 1; they intersect at the point (v, v). Above and below this point of intersection, in the shaded regions, the voter strictly prefers the status quo to the bill: $-|v - q| < -|v - b|$. To the left and right of the point of intersection the voter strictly prefers the bill to the status quo: $-|v - b| < -|v - q|$. This figure will be referred to later, in order to characterize the regions along which the voter agrees or disagrees with a candidate's vote choice.

Candidates' Platforms. Any candidate, c, is characterized by his or her platform, $p_c \in X$. If elected, candidate c's platform, p_c, will determine the candidate's subsequent voting behavior as follows. For any pair consisting of a bill and status quo, (q, b), a candidate

[9] Our general argument that induced preferences over candidates differ from preferences over policies does not require this assumption, and our equations that calculate these induced preferences do not require symmetry.

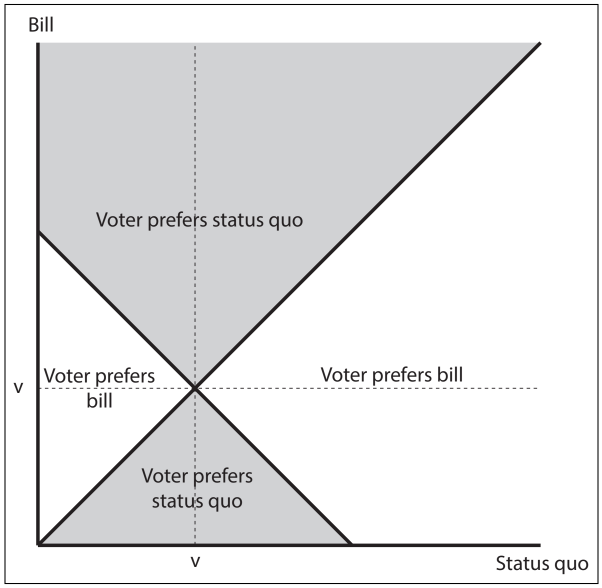


Figure 1: The Voter's Preferred Voting Behavior

with platform p_c will vote for the bill, b, if and only if $u(b, p_c) > u(q, p_c)$. That is, any candidate c would vote as the voter would vote if the voter had ideal point equal to p_c. For any platform p_c, this voting behavior is represented formally by the following function:

$$V(b, q, p_c) = \begin{cases} b & \text{if } u(b, p_c) > u(q, p_c), \\ q & \text{otherwise.} \end{cases}$$

The Agenda. We conceive of a legislative *agenda* that describes the likelihood that different (q, b) pairs will arise. To capture the idea that the voter is choosing a platform to represent his or her interests in the face of an exogenous agenda, we represent the agenda as a probability measure, α, over X^2. In later sections we consider more specific representations of the agenda. Prior to that, we summarize our theory of voter choice and establish a few general results.

Choosing Between Platforms. Viewed at its most general, our theory of voter choice is that the voter votes for (or selects) the candidate whose platform maximizes the following expected payoff function:

$$EU(p,v) = \int_{X^2} u(V(b,q,p),v)\mathrm{d}\alpha. \quad (1)$$

In words, a voter's expected payoff from a candidate with platform p is his or her expected utility from a policy the candidate will vote for. The assumption that the voter cares about the *vote* taken by the candidate can capture a number of different scenarios, depending on one's tastes. If we suppose that the chosen representative is decisive—that is, the alternative he or she chooses is actually implemented as policy—then our theory of voter preference over candidates is simply the voter's expected instrumental policy payoff induced by each candidate's platform. However, if the voter's chosen representative influences policy only indirectly (for example, as a member voting within a larger legislature), then these preferences neglect the details of that institution's decision-making processes. In this case the simplest interpretation is that these preferences reflect the "expressive" motivations of the voter.[10] Within an "expressive" interpretation, our theory interprets the voter's motivation as being simply to support candidates whose voting behavior would maximize the voter's payoff if those votes were converted into actual policy, regardless of whether or how those votes actually influence the implemented policy.

A different, *instrumental*, interpretation of our theory of voter preferences is available by supposing that there is a some positive probability that the elected representative's vote choice will be decisive. Such an interpretation can be derived from a larger model of (for example) probabilistic voting within a legislature.[11] A technical issue with such an interpretation regards whether this probability of the elected legislator being decisive is invariant to the elected legislator's platform. Unfortunately, a detailed consideration of this interesting issue is beyond the scope of this article. However, it is clear that if one does not assume that there is some positive probability of one's representative's vote being decisive,

[10]For example, Buchanan (1954), Tullock (1971), Brennan and Buchanan (1984), Brennan and Lomasky (1993), Brennan and Hamlin (1998, 2000), Schuessler (2000), Hillman (2010), and Hamlin and Jennings (2011).

[11]There are multiple possible sources for such probabilistic voting, including uncertainty about the platforms of the legislators elected from other districts and/or the preferences of the voters electing the legislators from other districts.

then an instrumental voter would be indifferent between all platforms, which is at odds with both intuition and the lengthy empirical literature on voter behavior.[12]

We can describe the voter's expected payoff function in some detail without specifying the agenda, α. To do so, we first introduce the notion of "disagreement sets." The disagreement set between a voter and any given candidate identifies the set of bill/status quo pairs on which the candidate's vote would differ from how the voter would vote.

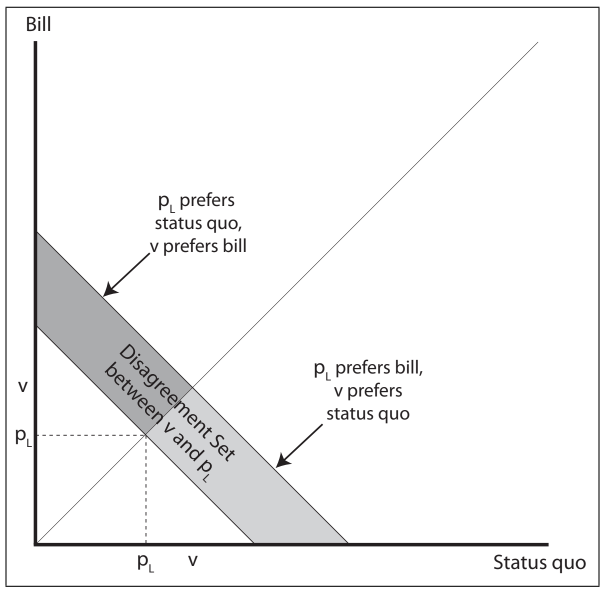


Figure 2: Voter-Candidate Disagreement Set

Disagreement Sets. For any voter with ideal point $v \in X$ and for any platform $p \in X$, let

$$D(p, v) \equiv \{(q, b) : V(b, q, p) \neq V(b, q, v)\}$$

denote the voter's *disagreement set* with respect to the platform p. This region is illustrated in Figure 2. In this figure a platform p_L is depicted that is to the left of the voter's ideal point v. The regions of disagreement between voter and candidate are generated by su-

[12]Among *many* others, see Tomz and Van Houweling (2008), Jessee (2012), and Montagnes and Rogowski (2015).

perimposing the voter's preferred voting behavior (shown in Figure 1) on the candidate's voting behavior (also Figure 1, but with (v, v) replaced by (p_L, p_L)). For (q, b) combinations in the darker region, the voter prefers b while the candidate prefers q. For (q, b) combinations in the lighter region, the voter prefers q while the candidate prefers b. For the rest of the possible votes that could occur, the voter and candidate agree. It is useful to use Figure 2 to explicitly define disagreement sets:

$$D(p, v) \equiv \begin{array}{ll} \{(q, b) : 2p - q < b < 2v - q\} & \text{if } p < v \\ \{(q, b) : 2p - q > b > 2v - q\} & \text{if } p > v. \end{array}$$

We start with an intuitive lemma that several of our results follow from. This lemma says that if the disagreement set for v and p_1 is a subset of the disagreement set for v and p_2, then a voter with ideal point v receives a weakly higher expected payoff from p_1 than p_2. All proofs are relegated to the appendix.

Lemma 1 *If $D(p_1, v) \subseteq D(p_2, v)$ then $EU(p_1, v) \geq EU(p_2, v)$.*

The following theorem is a direct corollary of Lemma 1, and is presented without proof. It states that a voter's expected payoff is maximized by a representative whose platform is equal to the voter's ideal point. This is because the disagreement region between a voter and platform $p = v$ is empty.

Theorem 1 *For any ideal point $v \in X$ and any agenda α, the function $EU(p, v)$ is maximized at $p = v$.*

The next theorem extends Theorem 1 by establishing that the expected payoff function is single-plateaued. Thus, as platforms move to either the left or right of the voter's ideal point, the voter's expected payoff weakly decreases. It is proved by showing that disagreement sets are nested as platforms move away from v in one direction.

Theorem 2 *For any ideal point $v \in X$ and any agenda α, the function $EU(p, v)$ is single plateaued: if $\underline{p} \leq p \leq v$ then $EU(\underline{p}, v) \leq EU(p, v)$ and, if $\overline{p} \geq p \geq v$ then $EU(\overline{p}, v) \leq$*

$EU(p, v)$.

Theorems 1 and 2 jointly establish a weak version of the "ally principle" (Bendor and Meirowitz (2004)) in this setting. Theorem 1 establishes that each voter should, if he or she can, appoint his or her ideological clone to vote on his or her behalf. Theorem 2 goes a step farther and implies that, when choosing among candidates whose platforms are all on the same side of the voter's ideal point then the voter should appoint the candidate whose platform is closest to his or her ideal point. These establish a "weak version" of the ally principle because they do not imply that "all else equal, a rational boss should choose her closest ally as an agent."[13] Specifically, the theorems do not address situations in which at least one candidate is offering a platform strictly higher, and another candidate is offering a platform that is strictly lower, than the voter's ideal point.

Our final result in this section establishes that the family of voter expected utility functions satisfies a single-crossing condition. This result is important because it establishes that there exists an individual whose voting behavior is equivalent to the majority preference relation for any binary vote that could occur. This enables us to assume the existence of a "representative voter." It is well-known that the symmetry imposed by the standard spatial model of preferences is sufficient to guarantee that the median voter is a representative voter. However, the expected utility functions we characterize are asymmetric, and it is also well-known that while the assumption of single-peaked preferences yields a median ideal point that is in the core, single-peakedness without symmetry is not sufficient to guarantee that the median voter is representative.[14]

Theorem 3 *There exists a representative voter over platforms, and it is the median voter.*

From this point forward we drop all subscripts on the voter's ideal point, and consider only the preferences of the representative voter.

[13]Bendor and Meirowitz (2004), p. 300.

[14]See Enelow and Hinich (1984) and Rothstein (1991) for expositions of these respective topics.

Taste for extremism

We conclude this introductory section with a detailed discussion of the notions of "extremism" and "moderation" that we consider throughout the remainder of this paper. We start with an illustration of what induced voter preferences over platforms look like in a simple agenda environment. This is followed by our definition of extremism and a general result about voter taste for extremism that the example hints at.

Example 1 Let agenda $\alpha = \{(b,q) \in \mathbb{R} \times \mathbb{R}\}$ such that bills and status quos are independently and identically distributed $N[0,1]$. Thus, α is the bivariate normal distribution centered at $(0,0)$ with a variance-covariance matrix equal to the identity matrix. The following figures depict the preferences of a voter with ideal point $v = -0.3$. On the right is the voter's preference for *policy* and on the left is his or her preference for *platforms*.

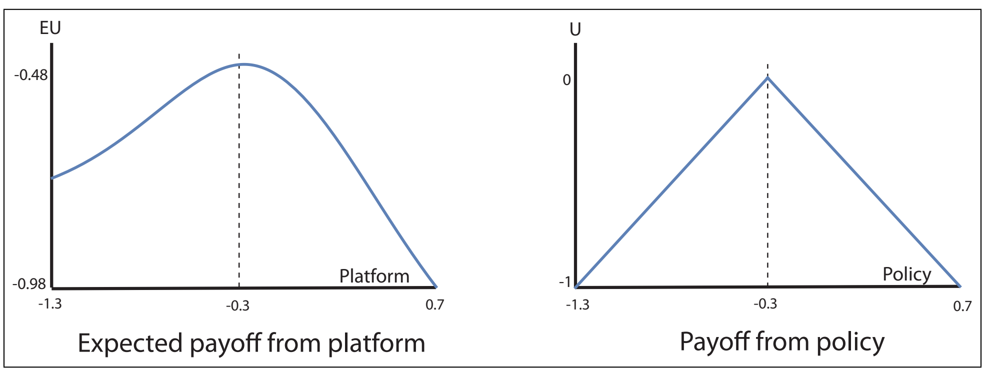


Figure 3: Differences in preferences for policies versus platforms

Figure 3 shows that a voter whose ideal policy v is to the left of the "center" of the legislative agenda has preferences that favor leftward deviations from v. Intuition for this finding comes from our definition of disagreement sets. Consider the voter's choice between two platforms that are equidistant from v to the left and right (p_L and p_R, respectively). Figure 4 shows the disagreement sets between v and p_L (the darker trapezoid) and v and p_R (the lighter one). A contour plot of an agenda that draws pairs centered to the right of (v,v) is shown, representing the bivariate normal distribution we have assumed. Since p_R is closer to the center of legislative activity, it is more likely that a vote will be drawn on which the

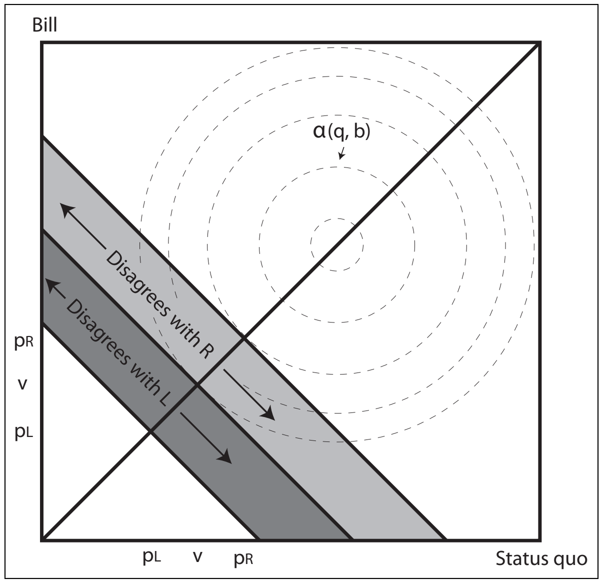


Figure 4: Intuition for voter preference for extremism

voter and R disagree. As the figure shows, agenda α assigns greater mass to policy pairs on the voter and R's disagreement set than on the voter and L's. Put differently, L is a "safer" candidate for the voter precisely because he or she is more extreme relative to the agenda. $\triangle$

Before returning to a generalization of the induced "taste for extremism" that this example provides, we formally define our notion of extremism. As noted earlier, our definition of extremism is relative to the agenda. In order to tighten up our definition, for the remainder of the paper we make a second assumption regarding the independence and single-peakedness of the bill and status quo distributions.[15]

Assumption 2 *Agenda α is a product measure of two strictly quasi concave distributions, f_q and f_b with modes μ_q and μ_b.*

We now define *candidate extremism* relative to agenda α, and voter *taste for extremism*.

[15]We relax this assumption when considering Romer-Rosenthal and uniformly distributed agendas in our extensions.

Definition 1 *Candidate* 1 *is* ***moderate*** *relative to Candidate* 2 *if* $|p_1 - \mu_q| \leq |p_2 - \mu_q|$ *and* $|p_1 - \mu_b| \leq |p_2 - \mu_b|$*, with at least one of the two inequalities strict. When this is the case, we refer to Candidate* 2 *as* ***extreme*** *relative to Candidate* 1*. Otherwise, we cannot rank the candidates in terms of moderation or extremism.*

The above definition says that a candidate is more moderate relative to another if his or her platform is closer to both the most likely bill and the most likely status quo to arise. Thus, a "moderate" candidate is closer to the center of legislative activity, with center conceived of as the modal vote likely to occur.

We define voter taste for extremism relative to a voter's baseline taste for policy. In our model, voter preferences for a candidate are determined solely by the candidate's platform and the legislative agenda. Consequently, our notion of extremism is indirect in that we define it solely in terms of platforms and agendas. In other words, we consider candidate extremism relative to the legislative agenda α, conceiving of the agenda as a proxy for political climate. The "middle" of the agenda represents something akin to moderation.

In keeping with this focus on the agenda as a indicator of extremism, we take a weak view of the concept of moderation. If, for example, the voter lies in the "center" of the agenda distribution, then it is unclear whether a vote to the left or right of the voter constitutes moderation or extremism. However, if the voter is to one side of this distribution–say, the right of both the status quo and the mode of the bill distribution–then a vote for a farther right candidate represents an extremist vote, and a vote for a farther left candidate represents a vote for moderation. These agendas for which the voter could exhibit a taste for extremism are defined formally below. They are *imbalanced*, in that the modal bill and status quo are both to one side of the voter. Figure 5 pictures these two types of agendas; on the left is a balanced agenda with v between μ_b and μ_q, and on the right is an imbalanced agenda with $\mu_b < \mu_q < v$.

Definition 2 *Agenda* α *is imbalanced if either* $v > \mu_b$ *and* $v > \mu_q$*, or if* $v < \mu_b$ *and* $v < \mu_q$*. Otherwise,* α *is balanced.*

The focus of this article is on asymmetries in voter preferences over representatives' platforms. Specifically, for any given divergence from the voter's ideal point, $\delta > 0$, when does

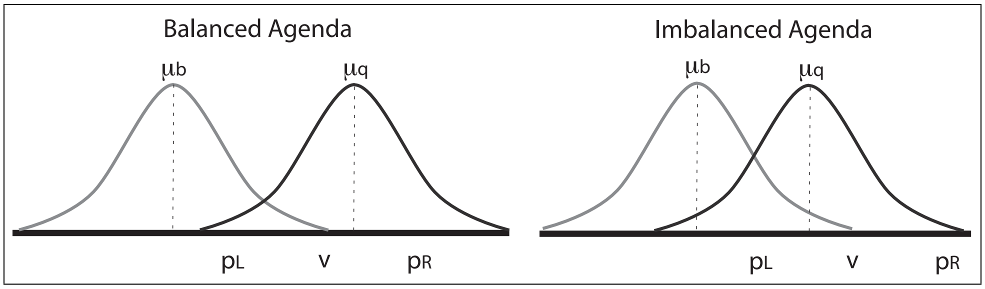


Figure 5: Agenda balance and extremism

the voter strictly prefer the extremist candidate, with platform $p_R = v + \delta$, to the moderate candidate, with platform $p_L = v - \delta$ (or vice versa)? Consider any two candidates with platforms that are equidistant from the voter's ideal point, v. If, whenever one of these candidates is more extreme than the other the voter strictly prefers the extreme candidate to the moderate candidate, then we say that the voter has a *taste for extremism.* Note that agenda imbalance is both necessary and sufficient for two candidates equidistant from the voter to be ranked in terms of extremism. This is again pictured in Figure 5. When platforms p_L and p_R are equidistant from v, a balanced agenda implies that neither platform can be more extreme than the other (in this case p_R is closer to μ_q and p_L closer to μ_b). An imbalanced agenda implies that one platform is *always* more extreme than the other (p_R in this case).

Definition 3 *Let $p_L = v - \delta$ and $p_R = v + \delta$ for $\delta > 0$. For an imbalanced agenda α, voter v has a taste for extremism if L more extreme than R implies that $EU(p_L, v) > EU(p_R, v)$ and if R more extreme than L implies that $EU(p_R, v) > EU(p_L, v)$.*

Returning to Example 1, we conclude this section with a result showing that voter taste for extremism will occur generally for certain agenda environments; namely, imbalanced agendas with circular isodensity curves, of which the particular bivariate Normal distribution considered in Example 1 is one example. We call agendas with circular isodensity curves *symmetric*. The following states one of the article's two main conclusions: voters have a strict taste for extremism whenever the agenda is symmetric and imbalanced.

Theorem 4 *For any symmetric and imbalanced agenda α the voter has a strict taste for extremism.*

The fact that Theorem 4 is based on the assumption that the agenda is symmetric does not undermine its generality: rather, symmetry as we define it in this setting simply removes the labels "bill" and "status quo" from the alternatives. Accordingly, to the degree that the voter is unaware of the more intricate details of public policy process, such as partisan gatekeeping (Cox and McCubbins (1993)) or institutionally induced "pivots" (Krehbiel (1998), Brady and Volden (2005)), the symmetric agenda model is arguably an appropriate approximation of his or her understanding of the legislative agenda. It is also arguably consistent with an approximation of how a voter might think about the distribution of policies likely to be brought up over the future. In any event, it serves as a useful baseline to gauge the impact of incorporating an exogenous agenda into how voters should evaluate different platforms. With that in hand, we now turn to consider a more detailed conception of the agenda process in which one policy, the "status quo," is known and the only uncertainty about the agenda is the location of the other (*i.e.*, the "bill").

2 Asymmetric agendas: When the status quo is known

It may be natural to consider the symmetric agenda environment described in Theorem 4 when choosing a delegate (such as a judge or bureaucrat) to adjudicate disputes that arise exogenously. In such cases we may wish to incorporate uncertainty about both options that the appointed representative will be choosing between. However, the standard spatial bargaining framework utilized by political economy scholars for the past 40 years typically assumes a status quo that is fixed and known to the voter.[16] In addition to comparability with existing scholarship, the assumption of an exogenous and known status quo is the most parsimonious way to consider "asymmetric" agendas.[17] We begin by considering a setting with a known status quo policy, which we normalize to zero, and a distribution of bills, f, satisfying our prior assumptions (namely strict quasi concavity—but not necessarily

[16]For example, Romer and Rosenthal (1978), Cox and McCubbins (1993), and Krehbiel (1998).

[17]Clearly, our framework is still very restrictive in the sense that we are constraining the representative to a binary choice, but we leave this extension for future work. However, it is worth noting that allowing the representative to choose from a set of more than two randomly drawn options will reduce the asymmetry in the voters' induced preferences.

symmetry— of f, with mode μ).[18] In this setting the expected payoff for a voter with ideal point v from a candidate with a platform equal to p is

$$EU(p, v) = \int_X u(V(b, 0, p), v) f(b) \mathrm{d}b.$$

This is simply the voter's expected utility for a vote between status quo $q = 0$ and a bill distributed according to f, taken by a candidate with platform p.

To begin, note that there is a single interval of bills on which the two candidates will vote differently from each other. When $b < 2p_L$ then both the left and right candidate (and the voter) prefer the status quo $q = 0$ to b. Similarly, when $b > 2p_R$ then both candidates and the voter prefer the status quo to the bill. We can therefore restrict attention to the interval $[2p_L, 2p_R]$; for any bill drawn from this interval the preferences of the two candidates diverge. As before, we call this interval the *disagreement set* for the known status quo setting. Figures 6, 7, and 8 show this region for the following three cases: Candidate divergence $\delta \le \frac{v}{2}$, $\delta \in [\frac{v}{2}, v]$, and $\delta \ge v$. Each figure also depicts the difference in the voter's utility calculation for R's vote over L's vote for each possible bill, or $u(V(b, 0, p_R), v) - u(V(b, 0, p_L), v)$ (note that we abuse notation slightly in the figures' captions).

When $p_L \ge 0$, or equivalently, $\delta \le v$, the left candidate will always vote for status quo $q = 0$ and the right candidate will always vote for bill b on the entire disagreement region. This is pictured in Figures 6 and 7. The voter prefers the bill on $[2p_L, 2v]$ and prefers the status quo on $[2v, 2p_R]$. What distinguishes these figures is the voter's expected utility calculation for the right candidate over the left. When the bill distribution extends below v, as in Figure 7, the difference between a vote for p_R over p_L begins to decrease. This difference is always positive on the interval $b \in [2p_L, 2v]$, but the magnitude of the difference gets smaller for smaller bills; when $b = 0$ the voter is indifferent between the two candidates (because the bill *equals* the status quo, so the candidates cannot be distinguished by their votes).

When $p_L < 0$ candidate behavior and voter preferences change slightly; in this case

[18] Indeed, somewhat interestingly, symmetry of f does not provide much additional purchase in this "known status quo" setting.

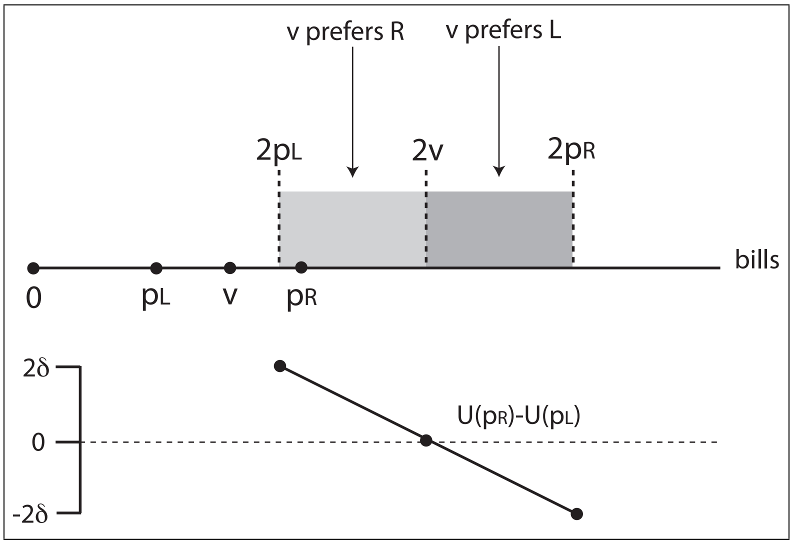


Figure 6: Disagreement region when $\delta \leq \frac{v}{2}$

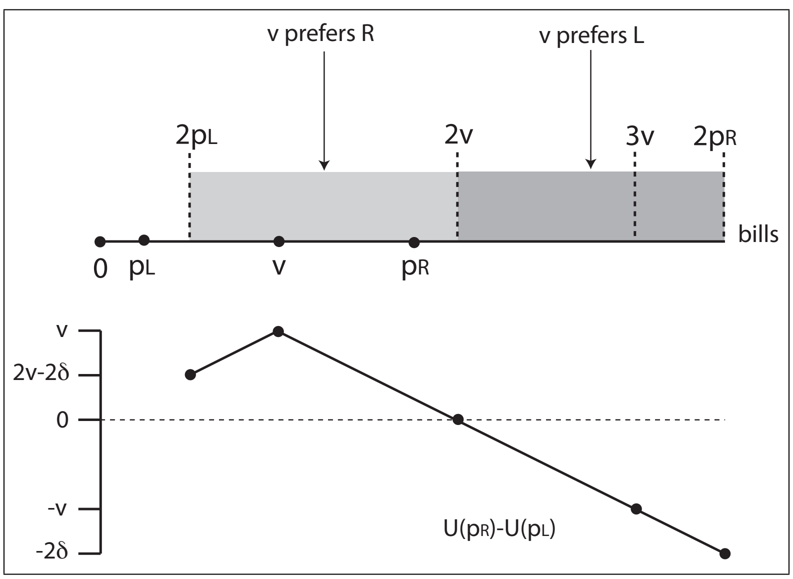


Figure 7: Disagreement region when $\delta \in [\frac{v}{2}, v]$

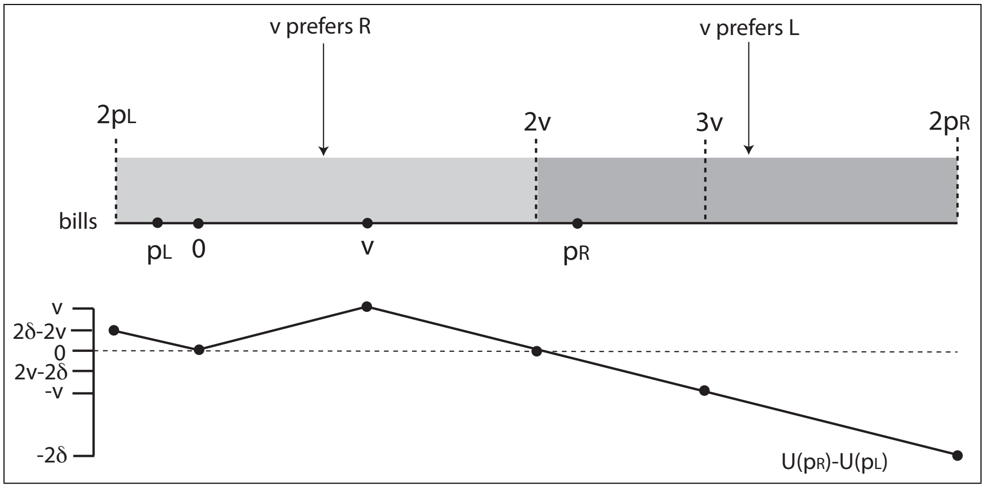


Figure 8: Disagreement region when $\delta \geq v$

the left candidate will vote for status quo $q = 0$ and the right candidate will vote for bill b when $b \in [0, 2p_R]$. When $b \in [2p_L, 0]$ then the left candidate will vote for the bill and the right candidate will vote for the status quo. On this disagreement region the voter prefers the bill on $[0, 2v]$ and prefers the status quo on $[2p_L, 0] \cup [2v, 2p_R]$. In this case, voter preference for the right candidate over the left gets smaller for smaller bills until $b = 0$; then this magnitude starts to rise as bills move left, past zero.

To evaluate voter taste for extremism over moderation, first note that when $v \geq p_L > \frac{v}{2}$ the disagreement region region lies to the right of v; in this case, $b \geq v$ for all b in the disagreement region. Thus,

$$\Delta(\delta, v) = \int_{2p_L}^{2p_R} \left(-(b - v) + v\right) f(b) db,$$

or

$$\Delta(\delta, v) = \int_{2p_L - q}^{2p_R - q} (2v - b) f(b) db. \tag{2}$$

If $0 \leq p_L \leq \frac{v}{2}$ then the voter's net expected payoff from the extremist candidate changes to:

$$\Delta(\delta, v) = \int_{2p_L}^{v} bf(b)db + \int_{v}^{2p_R} (2v - b)f(b)db. \tag{3}$$

Finally, if $p_L \leq 0$ it changes to:

$$\Delta(\delta, v) = \int_{2p_L}^{0} -bf(b)db + \int_{0}^{v} bf(b)db + \int_{v}^{2p_R} (2v - b)f(b)db. \tag{4}$$

When $p_L \in [\frac{v}{2}, v]$ (Figure 6) the voter always has a weak taste for extremism; he always prefers R to L. Intuitively, this is because the mode of the bill distribution is to the left of v, and so more likelihood is placed on bills arising from the left side of the disagreement region, where the voter prefers p_R. Additionally, $\Delta(\delta, v)$ is symmetric on the region overall, and so the increased likelihood of a bill being drawn from $[2p_L, 2v]$ leads to a preference for the right, extreme, candidate over the left.

Proposition 1 *When $\delta \leq \frac{v}{2}$ (i.e. for small enough policy divergence) the voter always prefers the extremist candidate to the moderate candidate.*

Proposition 1 can be rephrased as follows: Holding the divergence between two candidate platforms (δ) fixed, there always exists a voter ideal point v such that v is *indifferent between the policies* represented by the platforms, but *strictly prefers the extreme candidate to the moderate*. Such a voter has a sufficiently extreme ideal point, $v > \frac{\delta}{2}$. The following corollary formalizes this.

Corollary 1 *For any fixed level of candidate divergence, δ, any voter with $v \geq 2\delta$ always prefers the extremist candidate to the moderate candidate.*

Proposition 1 can also be used to derive sufficient conditions for extremism when $\delta > \frac{v}{2}$. These conditions are somewhat cumbersome, and are relegated to the Appendix. The relevant insight is that a voter's taste for extremism or moderation depends on the mass the distribution of bills f assigns to policies sufficiently to the right of the voter. Figure 8 illustrates this point. Since bills on the far right of the disagreement set yield higher disutility for the extremist candidate than bills on the far left do for the moderate candidate, the voter can have a strict taste for moderation if those bills are highly likely to arise. The most conservative case is if bills are drawn from a uniform distribution with support that spans the entire disagreement set: in this case the voter always has a strict taste for *moderation.* The uniform distribution provides a good illustration of the tradeoff between variance and extremity of the bill distribution in determining voter preferences for extremism. We present a full analysis of the uniform distribution in the appendix that explicitly characterizes this tradeoff. We conclude this section with an illustration of the fact that in the known status quo case voters may have a taste for either extremism or moderation, even when the distribution of bills is symmetric.

Example 2 Let the distribution of bills be $N[0.2, \sigma]$ and fix the status quo at $q = 0$. The Figures 9 and 10 depict the induced preferences of a voter with ideal point $v = 0.5$ for bill distributions with different variances (the distributions of bills are pictured on the right); in Figure 9 the variance of the bill distribution is $\sigma = 1$ and in Figure 10 the variance is $\sigma = 10$.

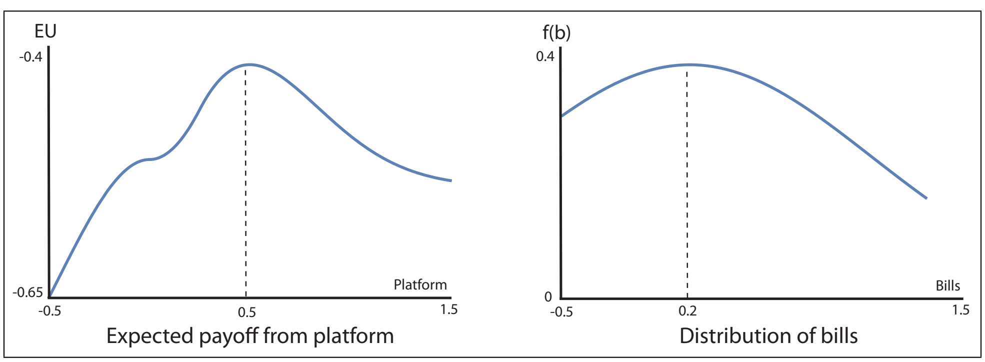


Figure 9: Preferences for platforms with low bill variance

In both figures the voter has a taste for the extreme, rightmost candidate for sufficiently small candidate divergence ($\delta \leq 0.25$, as Proposition 1 tells us). However, when the bill distribution is sufficiently dispersed, as in Figure 10, the voter has a taste for the moderate leftmost candidate as δ gets big. When $\delta = 1$ the voter prefers R to L (or platform $p_R = 1.5$ to $p_L = -.05$) when bill variance $\sigma = 1$, but prefers L to R when $\sigma = 10$. This is due to the

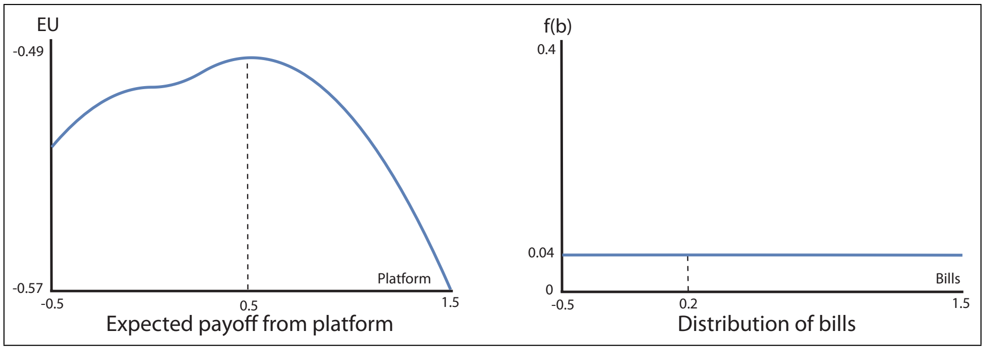


Figure 10: Preferences for platforms with high bill variance

high likelihood of an extreme right bill being drawn when the bill variance is high; such a bill would split the voter and the extremist candidate, and would yield a significant negative payoff to the voter. △

3 Implications in Three Canonical Settings [Incomplete]

In this section we illustrate some consequences of our framework for three well-known models of politics: a model of 2-candidate competition between vote-seeking candidates (*e.g.*, Downs (1957)), a "citizen-candidate" model with endogenous entry by policy-seeking candidates (*e.g.*, Osborne and Slivinsky (1996), Besley and Coate (1997)), and a model of voter preferences when the agenda is determined by a policy-motivated agenda setter (Romer and Rosenthal (1978)).

In the first two of these settings, a voter's taste for extremism alters well-known predictions about the platforms that will emerge in electoral competition. In the third, the voter's induced preference for platforms exhibits strong qualitative divergence from both symmetry and single-peakedness: the voter is indifferent with respect to platforms located between that of the "pivotal" member of the legislature and that of the "setter" who will choose what bill to propose after observing the status quo policy.

Two Vote-Seeking Candidates

Consider two vote-seeking candidates competing for the votes of the voters in N. By Theorem 3, this is equivalent to competing for the vote of the median voter, denoted by $m \in N$. Accordingly, the candidates' positions will be responsive to the median voter's ideal point, v_m.

Let $w \in \{L, R\}$ denote the winning candidate: central to our application of the framework presented above in this setting is an assumption that, prior to taking office, the winning candidate's platform, p_w, is perturbed to create a realized platform, $\hat{p}$, as follows:

$$\hat{p} = p + \pi,$$

where (for simplicity) we assume that π is distributed Uniform $[-\varepsilon, \varepsilon]$ where ε is an exogenous parameter. From a substantive standpoint, possible sources of perturbations (π) could include:

- External events with political implications (financial crises, wars, terrorist attacks)

- Actions / inducements of other political actors (party leaders, interest groups, donors, activists, the media)

⇒ Legislators susceptible to shifts in ideological leanings of their voting behavior

Given the uncertainty about each candidate's ultimate position, voter i's expected payoff from candidate with platform p is

$$\bar{U}(p, v_i; \varepsilon) \equiv \int_{-\varepsilon}^{\varepsilon} EU(p + \pi, v_i) d\pi. \tag{5}$$

In equilibrium, the candidates each seek to maximize $\bar{U}(p, v_m; \varepsilon)$, so that their positions converge to $p_L^* = p_R^* = p^*(\varepsilon)$, which is characterized in the following proposition.

Proposition 2 *Fix any $\varepsilon \geq 0$ and suppose that two purely vote-seeking candidates compete for the votes of N voters with ideal points $v =\equiv \{v_i\}_i$, with median ideal point v_m, who vote according to the payoff function $\bar{U}(p, v_i; \varepsilon)$ as defined in Equation* (5)*. The unique equilibrium involves both candidates choosing $p_L = p_R = p^*(\varepsilon)$, where*

$$p^*(\varepsilon) = \underset{p \in \mathbf{R}}{\operatorname{argmax}} \ \bar{U}(p, v_m; \varepsilon).$$

Furthermore, for any fixed median ideal point v_m, the function $p^* : \mathbf{R}_+ \to \mathbf{R}$ satisfies the following:

- *(Identifying First Order Condition.) The median voter is indifferent between the endpoints of the support of the distribution of π:*

$$EU(p^* - \varepsilon, v_m) = EU(p^* + \varepsilon, v_m)$$

- *(Uncertainty About π Required for Divergence.) The median voter is perfectly represented when there is no uncertainty:*

$$p^*(0) = v,$$

- *(Generically Not Representative.) If $EU(p, v_m)$ is asymmetric and $\varepsilon > 0$, then it is generically the case that $p^*(\varepsilon) \neq v$, and*

- *(Biased Toward Higher Side of Payoff Function.) If the utility function of the median voter is right asymmetric:*[19]

$$z > 0 \Rightarrow EU(v_m + z, v_m) > EU(v_m - z, v_m),$$

then $p^*(\varepsilon) \geq v$,

[19] An analogous symmetric conclusion follows for "left asymmetric" payoff functions.

Conclusion 2 of Proposition 2 is the principal conclusion with respect to a "taste for extremism" in this setting. For example, the following is an immediate corollary of Theorem 4.

Corollary 2 *If the agenda is symmetric and imbalanced, then the unique equilibrium position is more extreme than the voter.*

Similarly, though slightly less obvious, the following is a corollary of Proposition 1.

Corollary 3 *If the status quo, $q = 0$, is known and $\varepsilon \leq \frac{v_m}{2}$, then the unique equilibrium position, $p^*(\varepsilon)$, is more extreme than the voter.*

A Citizen-Candidate Model

Assume that each citizen chooses simultaneously whether to run in the election or not: those who run pay an exogenous and known cost $c > 0$. Following this, all citizens vote sincerely between the citizens (whose ideal points are assumed to be common knowledge) who chose to run. The winning candidate—determined by plurality rule with fair tie-breaking—chooses sincerely between b and q according to the exogenous agenda, α, based on his or her own ideal point. In other words: the game proceeds as follows:

- Timing is as follows

  1. Each citizen chooses whether to enter the race or stay out, entering costs $c > 0$,[20]
  2. Each citizen votes for one of the candidates; winner decided by plurality rule,
  3. Bill and status quo, (b, q), realized according to agenda, α,
  4. Winning citizen i votes according to $v_i \in X$, and

[20] If zero candidates enter, then a citizen is chosen at random to be the representative. *Note: there may be a technical issue with pure strategy equilibrium existence that we have to finesse with this assumption. Doing so is simple if we simply assign an exogenous reservation payoff that is common to all voters in the event that no candidates enter. However, that approach lacks a microfoundation. For reasons of time, we note this issue and leave it to the side: we are interested in two candidate equilibria, anyway.*

5. Denoting the policy chosen by the winning citizen by x, each citizen receives payoff of $u_i(x, v_i)$ if they did not run for election and $u_i(x, v_i) - c$ if they did run for reelection.[21]

Finally, we suppose that there is a continuum of voters whose ideal points are uniformly distributed between $v_{\min}$ and $v_{\max}$, with

$$\begin{aligned} v_{\min} &\geq 0, \\ v_{\max} &\leq 1, \text{ and} \\ v_m &\equiv \frac{v_{\min} + v_{\max}}{2}, \end{aligned}$$

and that the bill and status quo are independently distributed according to the Uniform$[0, 1]$ distribution.

Equilibrium number of candidates.

Equilibria can involve 0, 1, 2, or more citizens entering the race, and the number of candidates depends on

1. the cost of running (c),
2. the heterogeneity of voter ideal points ($v_{\min}$ and $v_{\max}$), and
3. the location of the median voter ($v_m = (v_{\min} + v_{\max})/2$) relative to the agenda, α.

The third of these conclusions—essentially that the number of candidates that enter in equilibrium depends on how far the median voter's ideal point is from $1/2$—qualitatively distinguishes this setting from that examined by Osborne and Slivinsky (1996) and Besley and Coate (1997).[22] In those models, of course, there is no agenda, and thus it is without loss of generality to (commonly) translate the voter's ideal points however one wants.

[21] We assume, for simplicity, that there are no "ego rents" from winning office, *per se*. Including such rents will complicate the analysis, but leave the comparative statics of interest unchanged *unless* one assumes that ego rents depend on ideal points in some way.

[22] Besley and Coate (1997) allow for strategic voting by the voters, which we do not. However, this is of interest only when considering entry by more than 2 candidtes, which is not our primary focus in this context.

Proposition 3 *Let $c^*(v_{\min}, v_{\max})$ denote the maximal cost at which there exists a pure strategy, 2 candidate equilibrium. The following properties characterize $c : [0,1] \times [0,1] \to \mathbf{R}$:*

1. *$c^*(v_{\min}, v_{\max})$ is increasing in $v_{\max} - v_{\min}$ and*
2. *$c^*(v_{\min}, v_{\max})$ is decreasing in $|v_m - 1/2|$.*

The logic behind these conclusions is that, as we come to again below, the median voters must be indifferent between both candidates in a pure strategy two-candidate equilibrium.

Interpreting Proposition 3. The first conclusion implies that one is more likely to observe uncontested elections in more homogenous districts. The second conclusion implies that one is more likely to observe uncontested elections in districts that are more extreme in the sense of their median voter is farther away from the center of the agenda $(1/2)$.

Polarization in 2 Candidate Elections

Theorem 3 implies that the election winner in a 2-candidate election will be that candidate whose platform is preferred by the median voter. Thus, because running for election costly and ultimately beneficial only if one wins the election, it follows that—as in existing citizen-candidate models—the median voter must be indifferent in a 2-candidate equilibrium. Given our assumptions about the agenda, Figure 11 illustrates the structure of a 2-candidate equilibrium.

Predictions about Equilibrium Platforms. We now suppose, without loss of generality, that $v_m \geq 1/2$. For any $(c, v_{\min}, v_{\max})$, let $p_L^*(c, v_{\min}, v_{\max})$ denote the maximal ideal point of a voter with ideal point no greater than the median voter's who will enter in a pure strategy 2-candidate equilibrium and let $p_R^*(c, v_{\min}, v_{\max})$ denote the minimal ideal point of a voter with ideal point no less than the median voter's who will enter in a pure strategy 2-candidate equilibrium. It is straightforward to show that both of the citizens will enter in

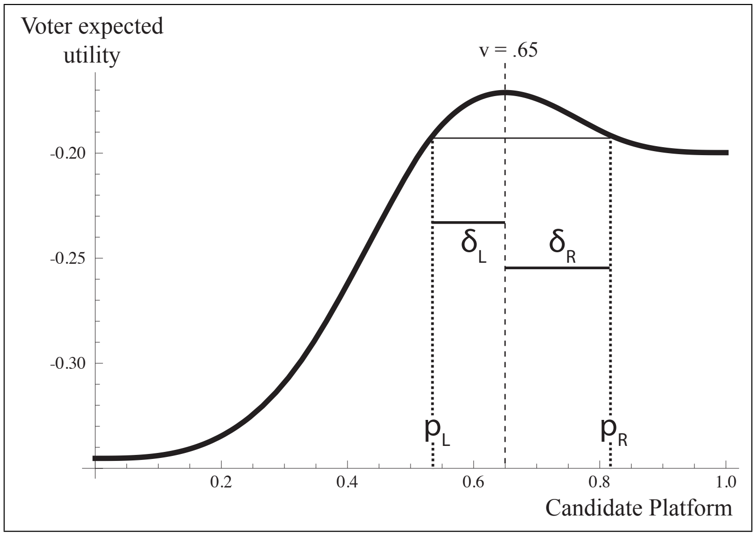


Figure 11: Citizen Candidate Model Example

the same equilibrium, given $(c, v_{\min}, v_{\max})$. Then let

$$\delta^*(c, v_{\min}, v_{\max}) = p_R^*(c, v_{\min}, v_{\max}) - p_L^*(c, v_{\min}, v_{\max})$$

denote the minimal divergence between the two platforms in a pure strategy, two-candidate race, given $(c, v_{\min}, v_{\max})$, and define the following "net extremism function":

$$\eta^*(c, v_{\min}, v_{\max}) = |p_R^*(c, v_{\min}, v_{\max}) - v_m| - |p_L^*(c, v_{\min}, v_{\max}) - v_m|,$$

which represents how much farther the extreme candidate (R) is from the median voter than is the moderate candidate (L). The next proposition characterizes δ^* and η^*.

Proposition 4 *The following characterize $\delta^*(c, v_{\min}, v_{\max})$ and $\eta^*(c, v_{\min}, v_{\max})$:*

- $\delta^*(c, v_{\min}, v_{\max})$ *is increasing in c,*
- $\delta^*(c, v_{\min}, v_{\max})$ *is increasing in $|v_m - 1/2|$,*
- $\eta^*(c, v_{\min}, v_{\max}) \geq 0$,

- $\eta^*(c, v_{\min}, v_{\max}) = 0 \Leftrightarrow v_m = 1/2$,

- $\eta^*(c, v_{\min}, v_{\max})$ *is increasing in* c,

- $\eta^*(c, v_{\min}, v_{\max})$ *is increasing in* $|v_m - 1/2|$.

Only conclusion 4 and (one direction of) conclusion 4 of Proposition 4 holds in standard citizen candidate models, in which $\delta^*(c, v_{\min}, v_{\max})$ is invariant to $|v_m - 1/2|$ and $\eta^*(c, v_{\min}, v_{\max}) = 0$ for all $(c, v_{\min}, v_{\max})$.

Figure 12 illustrates some of the comparative statics described above.

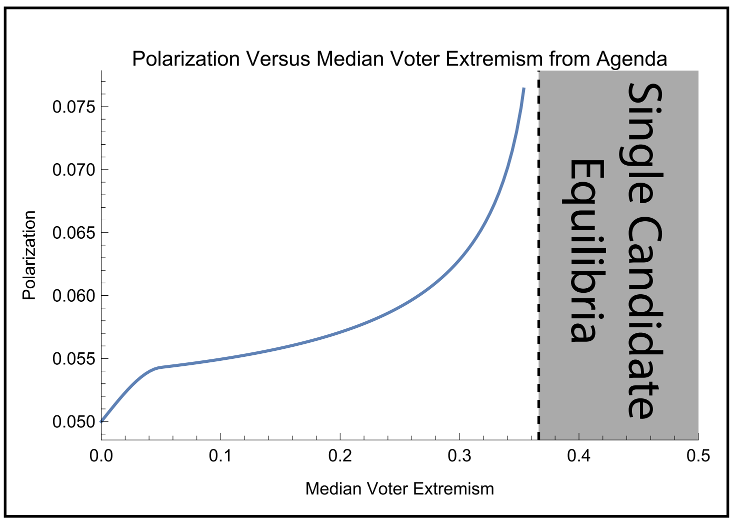
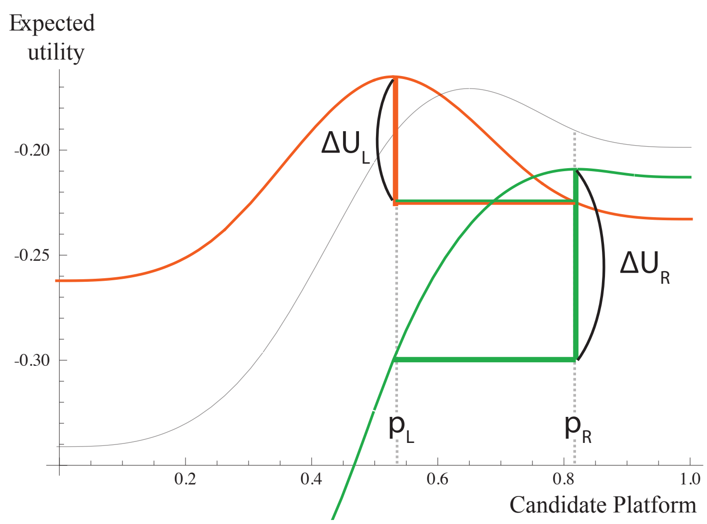


Figure 12: Some Comparative Statics of Net Extremism in Citizen Candidate Model

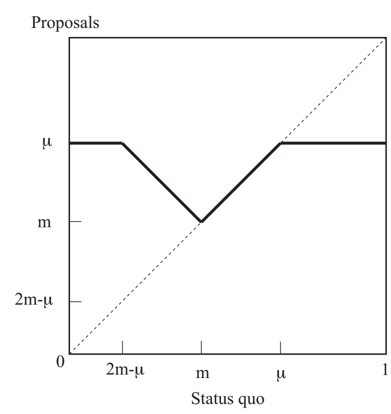
Another implication: extreme candidates would pay more to enter race



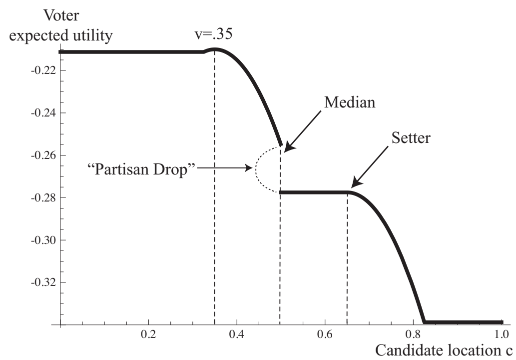
Voter Preferences When Bills Depends on Status Quo

(This is the least complete of the applications.)

Status quo drawn uniformly, bill chosen by strategic setter μ to be approved by receiver m

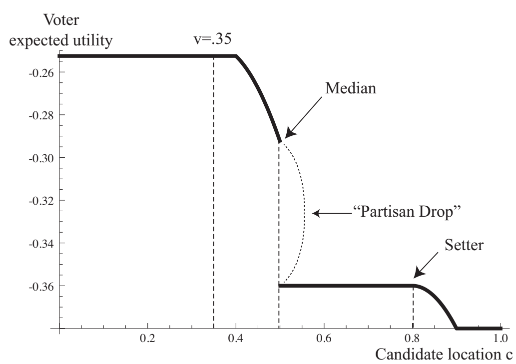


The setter model
(Setter at .65, receiver at .5)



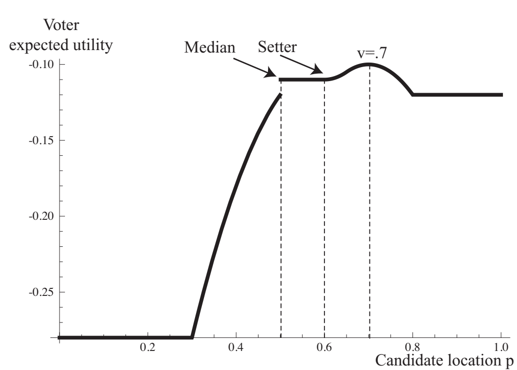
Example: A more extreme setter

(Setter at .8, receiver at .5)



Example:Setter model with majority party voter

(Setter at .6, receiver at .5, voter at .7)



Asymmetric / partisan gatekeeping procedures always induce a payoff discontinuity at legislative pivot (receiver)

- Discontinuity occurs because pivot is targeted to be indifferent between bill and status quo

- Empirically, legislators at pivot seem to be electorally disadvantaged

  - Out of step with party, etc.
  - Arlen Spector, Claire McCaskill

- Candidates perceived as slightly on wrong side take big hit

In the setter model, voters may have a taste for either extremism or moderation, depending on location of v and size of δ

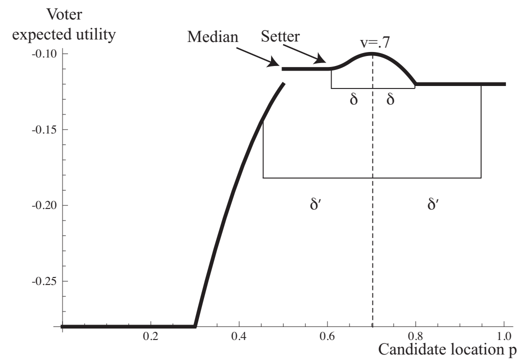
Let voter compare $v - \delta$ with $v + \delta$ (p_L with p_R)

- When δ is large, all voters prefer the candidate farthest from the legislative pivot (extremism)

- When δ is small *and voter is close to the interval* $[m, \mu]$, then voter may prefer candidate closer to the legislative pivot

- Taste for moderation when δ is small can happen regardless of whether voter is in majority party or not

Preferences for moderation or extremism
(Setter at .6, receiver at .5, voter at .7)



4 Conclusion (Incomplete/Outdated)

We have presented a theory of voting for representatives. The key finding is that, when candidates' platforms represent how they will vote over an exogenous agenda, the voter's preferences over these platforms will not generally be the same as his or her preferences over those platforms if they represented the policy that would be implemented by the candidate. Once recognized, this finding is intuitive but, to our knowledge, relatively unaccounted for in theoretical and empirical investigations of voting. We find that, when the uncertainty about the alternatives to be chosen from is single-peaked and strongly symmetric (in the sense that the isodensity curves of the joint distribution of bills and status quos are circular), then voters will *always* have a preference for a more extreme candidate over a more moderate one, holding the degree of divergence from the voter's ideal platform constant.

When the status quo is known *a priori*, then all voters will always have a weak preference for extreme candidates when the degree of divergence between the candidates is low or when the voter is sufficiently distant from both the center of the distribution of bills and the status quo policy. Otherwise, the voter in some cases may have a preference for the moderate candidate. Specifically and intuitively, when the voter is located sufficiently close to the center of the bill distribution, in that there is a sufficiently high probability of bills being proposed that fall on the opposite side of the voter's ideal point from the status quo, the voter may prefer a moderate candidate when comparing two candidates that are sufficiently distant from the voter's ideal point.

This somewhat convoluted conclusion is interesting when considered in substantive terms. In particular, the voters who, on the margin, might prefer moderate candidates, are those who are likely to be "in the middle" of, and face non-trivial trade-offs regarding, the comparisons that will confront their representative. In ongoing work, we are exploring how various models of strategic (and non-strategic) candidacy and partisan motivations interact with the voter incentives we have identified in this article.

References

Ashworth, Scott and Ethan Bueno de Mesquita. 2009. "Elections with Platform and Valence Competition." *Games and Economic Behavior* 67(1):191–216.

Bendor, Jonathan and Adam Meirowitz. 2004. "Spatial Models of Delegation." *American Political Science Review* 98(2):293–310.

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

Brady, David W and Craig Volden. 2005. *Revolving Gridlock: Politics and Policy from Jimmy Carter to George W. Bush*. Westview Press.

Brennan, Geoffrey and Alan Hamlin. 1998. "Expressive Voting and Electoral Equilibrium." *Public choice* 95(1-2):149–175.

Brennan, Geoffrey and Alan Hamlin. 2000. *Democratic Devices and Desires*. New York, NY: Cambridge University Press.

Brennan, Geoffrey and James Buchanan. 1984. "Voter Choice: Evaluating Political Alternatives." *American Behavioral Scientist* 28(2):185–201.

Brennan, H. Geoffrey and Loren E. Lomasky. 1993. *Democracy and Decision*. Cambridge: Cambridge University Press.

Buchanan, James M. 1954. "Individual Choice in Voting and the Market." *Journal of Political Economy* 62:334–343.

Carter, Jennifer and John W. Patty. 2015. "Valence and Campaigns." *American Journal of Political Science* 59(4):825–840.

Cox, Gary W. and Mathew D. McCubbins. 1993. *Legislative Leviathan: Party Government in the House*. Berkeley, CA: University of California Press.

Dewan, Torun and Kenneth A Shepsle. 2011. "Political Economy Models of Elections." *Annual Review of Political Science* 14:311–330.

Dhillon, Amrita and Ben Lockwood. 2002. "Multiple Equilibria in the Citizen-Candidate Model of Representative Democracy." *Journal of Public Economic Theory* 4(21):171–184.

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

Eguia, Jon X. 2013. Challenges to the Standard Euclidean Spatial Model. In *Advances in Political Economy: Institutions, Modelling and Empirical Analysis*, ed. Norman Schofield, Gonzalo Caballero and Daniel Kselman. Springer pp. 169–180.

Enelow, James M. and Melvin J. Hinich. 1984. *The Spatial Theory of Voting*. Cambridge: Cambridge University Press.

Gailmard, Sean, John W. Patty and Elizabeth Maggie Penn. 2008. Arrow's Theorem on Single-Peaked Domains. In *The Political Economy of Democracy*, ed. Enriquetta Aragones, Humberto Llavador and Norman Schofield. Barcelona: UAB.

Grofman, Bernard. 1985. "The Neglected Role of the Status Quo in Models of Issue Voting." *Journal of Politics* 47(1):230–237.

Grofman, Bernard. 2004. "Downs and Two-Party Convergence." *Annual Review of Political Science* 7:25–46.

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

Hamlin, Alan and Colin Jennings. 2011. "Expressive Political Behaviour: Foundations, Scope and Implications." *British Journal of Political Science* 41(3):645–670.

Hillman, Arye L. 2010. "Expressive Behavior in Economics and Politics." *European Journal of Political Economy* 26(4):403–418.

Hinich, Melvin J and Michael C Munger. 1992. "A spatial theory of ideology." *Journal of Theoretical Politics* 4(1):5–30.

Hinich, Melvin J and Michael C Munger. 1996. *Ideology and the Theory of Political Choice*. Ann Arbor, MI: University of Michigan Press.

Jessee, Stephen A. 2012. *Ideology and Spatial Voting in American Elections*. Cambridge University Press.

Krehbiel, Keith. 1998. *Pivotal Politics: A Theory of U.S. Lawmaking*. Chicago, IL: University of Chicago Press.

Montagnes, Brendan Pablo and Jon C Rogowski. 2015. "Testing Core Predictions of Spatial Models: Platform Moderation and Challenger Success." *Political Science Research and Methods* 3(3):619–640.

Osborne, Martin J. and Ashur Slivinsky. 1996. "A Model of Political Competition with Citizen-Candidates." *Quarterly Journal of Economics* 111(1):65–96.

Patty, John W. 2016. "Signaling through Obstruction." *American Journal of Political Science* 60(1):175–189.

Penn, Elizabeth Maggie, John W. Patty and Sean Gailmard. 2011. "Manipulation and Single-Peakedness: A General Result." *American Journal of Political Science* 55(2):436–449.

Romer, Thomas and Howard Rosenthal. 1978. “Political Resource Allocation, Controlled Agendas, and the Status Quo.” *Public Choice* 33:27–43.

Rothstein, Paul. 1991. “Representative Voter Theorems.” *Public Choice* 72:193–212.

Schofield, Norman. 2004. “Equilibrium in the Spatial ‘Valence’ Model of Politics.” *Journal of Theoretical Politics* 16(4):447–481.

Schuessler, Alexander A. 2000. *A Logic of Expressive Choice*. Princeton: Princeton University Press.

Tomz, Michael and Robert P Van Houweling. 2008. “Candidate Positioning and Voter Choice.” *American Political Science Review* 102(03):303–318.

Tullock, Gordon. 1971. “The Charity of the Uncharitable.” *Western Economic Journal* 9:379–392.

Lemma 1. *If $D(p_1, v) \subseteq D(p_2, v)$ then $EU(p_1, v) \geq EU(p_2, v)$.*

Proof: Fix any agenda α and an ideal point v. Consider two platforms p_1, p_2, with $D(p_1, v) \subseteq D(p_2, v)$. Letting $\Delta U(p_1, p_2) \equiv EU(p_1, v) - EU(p_2, v)$, we obtain

$$\begin{aligned}
\Delta U(p_1, p_2) &= \int_X \int_X u(V(b, q, p_1), v) f(b) f(q) \mathrm{d}\alpha - \int_X \int_X u(V(b, q, p_2), v) f(b) f(q) \mathrm{d}\alpha, \\
&= \int_{D(p_2,v) \setminus D(p_1,v)} \left[u(V(b, q, p_1), v) - u(V(b, q, p_2), v)\right] f(b) f(q) \mathrm{d}\alpha, \\
&\geq 0,
\end{aligned}$$

because for every $(b, q) \in D(p_2, v) \setminus D(p_1, v)$, $u(V(b, q, p_1), v) > u(V(b, q, p_2), v)$. Furthermore, the inequality is strict if $D(p_2, v) \setminus D(p_1, v)$ has positive measure under α. Thus, if the voter's disagreement set with p_2 contains the voter's disagreement set with p_1, the voter weakly prefers p_1 to p_2, as was to be shown. ∎

Theorem 2. *For any ideal point $v \in X$ and any agenda α, the function $EU(p, v)$ is single plateaued: if $\underline{p} \leq p \leq v$ then $EU(\underline{p}, v) \leq EU(p, v)$ and, if $\overline{p} \geq p \geq v$ then $EU(\overline{p}, v) \leq EU(p, v)$.*

Proof: The proof proceeds by showing that if $\underline{p} < p < v$ then $D(p, v) \subseteq D(\underline{p}, v)$ (with the case of $\underline{p} > p > v$ showed similarly). The result then follows immediately from Lemma 1.

Let $\underline{p} < p < v$. We know that $D(p, v) = \{(q, b) : 2p - q < b < 2v - q\}$ and that $D(\underline{p}, v) = \{(q, b) : 2\underline{p} - q < b < 2v - q\}$. As $p > \underline{p}$, it follows that $D(p, v) \subseteq D(\underline{p}, v)$, as was to be shown. ∎

Theorem 3. *There exists a representative voter over platforms, and it is the median voter.*

Proof: Let v_m be the ideal point of median voter m, and suppose that the median voter strictly prefers platform p_R to p_L, with $p_R > p_L$. We prove the result by showing that for any voter with ideal point $v > v_m$, it is the case that the voter also strictly prefers p_R to p_L. Moreover, if m is indifferent between p_R and p_L, then the voter weakly prefers p_R to p_L. By a symmetric argument applied to any voter with $v < v_m$, the result establishes m as a representative voter.

Candidates R and L always vote for the policy closest to their platform. Let $c_{xy} = \frac{x+y}{2}$ be the *cut point* of policy pair (x, y). Candidates with platforms higher than this point vote for the larger of the two policies, and those with platforms lower vote for the smaller. Thus, the set of policies on which candidates L and R vote differently is the set of policy pairs with cut points on the interval $[p_L, p_R]$. We characterize these pairs as $(x, y) \in X \times X$ with $x = c_{xy} + \delta$, $y = c_{xy} - \delta$, and $c_{xy} \in [p_L, p_R]$, for $\delta > 0$.

Consider any vote that the candidates differ on, with R voting for $x = c_{xy} + \delta$ and L voting for $y = c_{xy} - \delta$, which must always be the case since $p_R > p_L$. The following three cases characterize any such vote:

1. $v_m \geq x > y$. In this case, m prefers x to y, receiving $-(v_m - (c + \delta))$ from R's vote and $-(v_m - (c - \delta))$ from L's vote. The net utility difference from R's vote over L's is
$$-(v_m - (c + \delta)) + (v_m - (c - \delta)) = 2\delta.$$

2. $v_m \leq y < x$. In this case, m prefers y to x, receiving $-((c + \delta) - v_m)$ from R's vote and $-((c - \delta) - v_m)$ from L's vote. The net utility difference from R's vote over L's is
$$-((c + \delta) - v_m) + ((c - \delta) - v_m) = -2\delta.$$

3. $y \leq v_m \leq x$, with one inequality strict. In this case m is sometimes divided about which candidate he prefers, receiving $-((c+\delta)-v_m)$ from R's vote and $-(v_m-(c-\delta))$ from L's vote. The net utility difference from R's vote over L's is
$$-((c + \delta) - v_m) + (v_m - (c - \delta)) = 2v_m - 2c.$$

Note that the net utility difference possible from a vote by R versus L is bounded above by 2δ and below by -2δ. Now consider a voter with $v > v_m$. For any policy pair on which R and L disagree, the voter's net utility difference from R's vote over L's is weakly greater than the median voter's: If Case 1, then $v > x > y$, and, like m, the voter's net utility difference is 2δ. If Case 2, then either $v \leq y < x$, in which case the voter's net

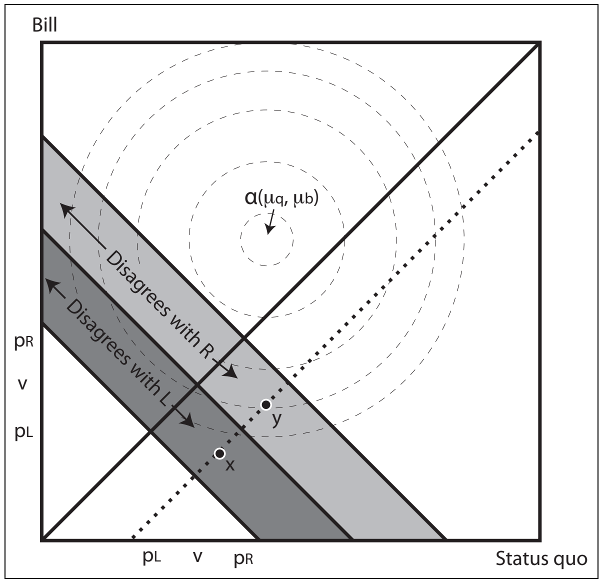


Figure 13: Illustration of Theorem 4

utility difference is equal to m's at -2δ, or $v > y$, in which case it is strictly higher (as the difference is bounded below by -2δ). Finally, if Case 3, then the voter's net utility difference is either $2v - 2c$ (if $v \leq x$), which is strictly greater than the median voter's net difference of $2v_m - 2c$, or it is 2δ, which again is strictly greater than the median voter's net difference. Thus, if m strictly prefers p_R to p_L, so does the voter.

The same argument can be used to show that if m strictly prefers p_L to p_R, then any voter with $v < v_m$ also strictly prefers p_L to p_R, and that if m is indifferent between p_L and p_R then voter above v_M weakly prefer p_R to p_L those below weakly prefer p_L to p_R. ∎

Theorem 4. *For any symmetric and imbalanced agenda α the voter has a strict taste for extremism.*

Proof: Let α be imbalanced and symmetric. Fix a value $(\mu_q, \mu_b) \in \mathbb{R}^2$ and let α be equal to a probability density function, f_{qb} with mode $\mu_{qb} = (\mu_q, \mu_b)$. Symmetry implies that for any two points $x, y \in \mathbb{R}^2$, if $||x - \mu_{qb}|| > ||y - \mu_{qb}||$ then $f_{qb}(x) < f_{qb}(y)$. Without loss of generality, assume that $v < \mu_q$ and $v < \mu_b$ (the agenda is imbalanced to the right of v).

Let $p_R = v + \delta$ and $p_L = v - \delta$. Given our definition of candidate extremism, L is more extreme than R. Take any point $y \in D(p_R, v)$, the disagreement set of the voter and R. Without loss of generality, assume that $y_1 > y_2$ so that the voter prefers bill y_2 and R prefers status quo y_1 (an identical argument holds for the other case). If we reflect this point on the 45^o line around the line $b = 2v - q$ we get the point $x = (2v - y_2, 2v - y_1)$, with $x \in D(p_L, v)$. This is pictured in Figure 13. The relevant insight is that at y R votes for y_1 whereas the voter prefers y_2; at x L votes for $x_2 = 2v - y_1$ whereas the voter prefers $x_1 = 2v - y_2$. In the former case the disutility the voter receives from R's incorrect vote for y_1 relative to L's correct vote for y_2 is

$$\Delta_y(R, L, v) = -|y_1 - v| + |y_2 - v|$$

whereas the disutility from L's incorrect vote for x_2 over R's correct vote for x_1 is

$$\begin{aligned} \Delta_x(L, R, v) &= -|2v - y_1 - v| + |2v - y_2 - v| \\ &= -|v - y_1| + |v - y_2| \\ &= \Delta_y(R, L, v). \end{aligned}$$

Thus, the voter receives identical disutility from the incorrect vote of R for y_1 over y_2 and L for x_2 over x_1. Symmetry and strict quasi concavity of f_{qb} implies that $f_{qb}(y) > f_{qb}(x)$. To see this, take the difference between the distance from μ_{qb} to x and to y:

$$||x - \mu_{qb}|| - ||y - \mu_{qb}|| = -2(\mu_q + \mu_b - 2v)(2v - y_1 - y_2).$$

Since $v \leq \mu_q, \mu_b$ with one inequality strict, this difference is strictly positive when

$$2v - y_1 - y_2 < 0.$$

By the definition of the disagreement set between the voter and R, we know that $y_2 > 2v - y_1$, or $2v - y_1 - y_2 < 0$. Thus, agenda α assigns strictly higher likelihood to vote y arising than to vote x. Integrating over all (x, y) pairs in $D(p_L, v) \times D(p_R, v)$ implies that moderate candidate R yields a lower expected payoff to the voter than extreme candidate L. Note that both strict single-peakedness *and* symmetry of f are key to this result, as they provide the stochastic dominance argument we require. ∎

Proposition 1. *When $\delta \leq \frac{v}{2}$ (i.e. for small enough policy divergence) the voter always prefers the extremist candidate to the moderate candidate.*

Proof: The voter prefers the right candidate over the left if and only if

$$-(E(b|b \in [2p_L, 2p_R] - v) \geq -v,$$

which we can rewrite as

$$2v \geq E(b|b \in [2p_L, 2p_R]).$$

By the definition of single-peakedness of f and the assumption that $\mu < v \leq 2p_L$, we have that $E[b|b \in [2p_L, 2p_R]] \leq \frac{2p_L+2p_R}{2} = 2v$. Thus, the voter prefers the right candidate over the left for small enough policy divergence between the two candidates (i.e. for $\delta \leq \frac{v}{2}$). ∎

Corollary 4 *When $\delta \in [\frac{v}{2}, v]$ a sufficient condition for the voter to prefer the right candidate to the left is*

$$2v \geq E(b|b \in [3v, 2p_R]) - E(b|b \in [2p_L, v]) \left(\frac{F(v) - F(2p_L)}{F(2p_R) - F(3v)} \right).$$

Proof: If $\delta \in [\frac{v}{2}, v]$ the voter's net expected payoff from R over L is given by Equation 3, which defines $\Delta(\delta, v) = \int_{2p_L}^{v} bf(b)db + \int_{v}^{2p_R} (2v - b)f(b)db$. We can rewrite

$$\begin{aligned}\Delta(\delta, v) = E(b|b \in [2p_L, v])(F(v) - F(2p_L)) + (2v - E(b|b \in [v, 3v]))(F(3v) - F(v)) \\ +(2v - E(b|b \in [3v, 2p_R]))(F(2p_R) - F(3v)).\end{aligned}$$

By the same logic as in Proposition 1, we know that $(2v - E(b|b \in [v, 3v]))(F(3v) - F(v)) \geq 0$, as $E(b|b \in [v, 3v]) \leq 2v$. The corollary follows immediately. □

For the case of $\delta \geq v$ we get the following corollary. The proof is similar to the proof of

Corollary 4, but utilizes Equation 4 instead of Equation 3. We omit the proof, though note that the corollary's condition differs from that in Corollary 4 due to the absolute value term on the right side of the equation.

Corollary 5 *When $\delta \geq v$ a sufficient condition for the voter to prefer the right candidate to the left is*

$$2v \geq E(b|b \in [3v, 2p_R]) - E(|b||b \in [2p_L, v])\left(\frac{F(v) - F(2p_L)}{F(2p_R) - F(3v)}\right).$$

Uniformly distributed bills. The uniform distribution on bills represents a conservative test for extremist tastes in our model; this is because we assume that the expected value of a bill is to the left of the voter, and because the voter prefers the moderate candidate to the extreme candidate for bills that fall to the right of $2v$. With a large enough support, the uniform distribution doesn't penalize extreme-right bills in terms of the likelihood they will arise. Thus, a uniform distribution of bills whose support contains the entire disagreement region represents a "best case" scenario for moderation, and we can show that in this case the voter always has a weak taste for moderation. In this best case scenario, when $\delta \leq \frac{v}{2}$ then $\Delta(\delta, v)$ equals zero; the voter is indifferent between the moderate and extreme candidates. The condition described in Corollary 4—the sufficient condition for extremism—is also a necessary condition for extremism in this scenario. This condition reduces to $2v \geq 2\delta + v$, which never holds when $\delta > \frac{v}{2}$. Similarly, the sufficient condition outlined in Corollary 5 is also a necessary condition when F is uniform and its support contains the entire disagreement region. In this case the condition requires $2v \geq \frac{4\delta^2}{2\delta - v}$, which never holds when $\delta > v$.

The best case scenario for moderation described above requires that the distribution of bills spans an interval with length at least $2v + 4\delta$, since the distribution of bills spans the entire disagreement region and $b < v$. The geometry of the disagreement region gives us a similar *best case scenario* for extremism; if the distribution of bills does not extend past the point $3v$ then the voter will always have a weak preference for extremism, with this preference being strict so long as the probability that a bill on the interval $[2p_L, v)$

has a strictly positive likelihood of being drawn (and, implicitly, so long as $2p_L < v$). The following proposition formalizes this result, and provides a necessary and sufficient condition for extremist preferences when F is uniform.

Proposition 5 *Let bills be drawn from a uniform distribution with support $[v-\kappa-\gamma, v-\kappa+\gamma]$. Thus, the distribution of bills is centered at $v-\kappa$ and the variance of the distribution is $\frac{1}{3}\gamma^2$. When when $v > \frac{\gamma-\kappa}{2}$, the voter has a weak preference for extremism. When*

$$2\delta + \gamma - \kappa > v > \gamma - \kappa \tag{6}$$

this preference for extremism is strict.

Proof: The result follows immediately from the definition of the disagreement region. Over the region $[v, 3v]$, the net utility for R over L (or $\Delta(\delta, v)$) cancels to zero. Thus, if a voter has a preference for the moderate candidate over the extreme, it is because the bills drawn from $[3v, 2p_R]$ (a region on which the voter prefers the moderate candidate to the extreme) outweigh the bills drawn from $[2p_L, v]$ (a region on which the voter prefers the extreme candidate). However, if $3v > v - \kappa + \gamma$, no bill on $[3v, 2p_R]$ can arise with positive probability. Thus, the voter has a weak preference for the extreme candidate in this case. □

Equation 6 leads to the following conclusions for the case where bills are drawn from a uniform distribution. First, if v is sufficiently extreme relative to the median of the bill distribution (in particular, if $v > \frac{\gamma-\kappa}{2}$,[23] or phrased differently, if $\mu + \gamma < 3v$) then the voter can never strictly prefer a moderate candidate to an extreme candidate. Moreover, in this case, increasing δ (the spread between the two candidates) will eventually lead the voter to strictly prefer the extreme candidate to the moderate candidate.

Second, when $\gamma < \kappa$ the voter always has a weak taste for extremism. This is because $v > 0$ by assumption, and so if $\gamma - \kappa < 0$ then $v > \frac{\gamma-\kappa}{2}$. This implies that when the support of the bill distribution $f(b)$ is sufficiently small, the voter will weakly prefer the extremist

[23] In the general symmetric uncertainty model, the mode and the median are identical by assumption.

candidate.

Finally, unlike previous results focusing on voter preference for extremism when the difference between the candidates is small (e.g. Proposition 1), large δ will always correspond to the voter preferring the extreme candidate when the voter himself is extreme relative to the bill mode (greater κ) and when the variance of the bill distribution is small (smaller γ). This is because the disagreement region is centered on $2v$ regardless of the size of δ. If the distribution of bills never crosses $3v$ then an increase in δ simply expands the region on which the voter disagrees with the moderate candidate.