

Polarization of What?

A Model of Elections with Endogenous Valence*

Gilles Serra[†]

March 30, 2010

Abstract

We connect three characteristics of political candidates: their preferences, their platforms and their valence. To do so, we define and study three types of elite polarization: *preference* polarization, *platform* polarization and *valence* polarization. In our model, policy is represented as a position in a unidimensional space; candidates are policy-motivated; valence is represented as a dimension orthogonal to policy; and candidates can increase their valence by paying a cost. We find that candidates will in general not converge to one another in the policy dimension or in the valence dimension. A candidate's preference is positively, but not perfectly, correlated with her platform. Strikingly, under some circumstances, candidates' platforms will diverge when their preferences converge. And there is an unfortunate trade-off regarding candidates' attributes: they will display high valence or low polarization, but not both. One of our empirical predictions is a positive correlation between campaign spending and platform polarization.

*This paper was presented at Harvard University, Oxford University, Tilburg University, and the annual meetings of the American Political Science Association, the Midwest Political Science Association, the Public Choice Society and the European Public Choice Society. I thank the participants in those seminars, as well as Christopher Avery, Robert Bates, Justin Buchler, Steven Callander, Christian Henning, Colin Jennings, David Myatt, Kenneth Shepsle, James Snyder, Bruno Strulovici and Mor Zahavi for their useful comments. In particular Ken Shepsle, Jim Snyder, Bob Bates and Chris Avery read several versions closely, and constantly offered insightful comments and thoughtful guidance. I gratefully acknowledge financial support by the Institute of Quantitative Social Sciences. All remaining errors are my responsibility.

[†]Research Fellow, Nuffield College at Oxford University, gilles.serra@nuffield.ox.ac.uk.

1 Introduction

What may candidates do to make themselves electable? Most obviously, candidates can try to design a policy platform that appeals to voters. Since the seminal work of Anthony Downs (1957), a major interest of the theoretical literature has been to study the platforms that candidates design to compete in elections. But policy is not the only factor influencing voters. Candidates are also evaluated on their non-policy attributes, as Donald Stokes (1963) pointed out early on. A candidate's competence, honesty and campaigning skills are a few examples of non-policy attributes that are valued by all voters; Stokes called such attributes "valence". Only recently has the theoretical literature begun to incorporate a valence dimension in the Downsian framework, but almost all the existing theories have assumed that candidates take their valence as fixed and exogenous.

Yet, candidates can try to improve their non-policy attributes; in other words, candidates often take their valence as endogenous. If the source of valence is competence, candidates can obtain additional education to increase their human capital. If the source of valence is experience, candidates can take jobs in the bureaucracy or the private sector before running for office. If the source of valence is campaign funding, candidates can exert effort to attract donors. If the source of valence is credibility, candidates can build a reputation for honoring their promises. If the source of valence is the "personal vote", candidates can spend the previous years bonding with constituents.

This paper develops a model in which candidates have a chance to increase their valence (at a cost) before designing their platforms. In other words, candidates can make an "investment" in valence that can be later cashed in for a preferred policy outcome. With this framework we can investigate important questions for our understanding of electoral politics, such as which candidates have the most incentives to improve their non-policy characteristics; and, how does competition over non-policy attributes affect competition over policy.

Another important phenomenon in modern politics is elite polarization. In particular, scholars of American politics have long worried about polarization between the two parties. Several authors have actually documented a growing gap between the policies advocated by Democrats and Republicans.¹ This paper contributes to the debate on elite polarization by distinguishing three concepts: (1) *platform polarization*, meaning the distance between candidates' announced platforms; (2) *preference polarization*, meaning the distance between candidates' true preferences irrespective of their

¹See for example Ansolabehere, Snyder and Stewart (2001a, 2001b), Burden (2004) and McCarty, Poole and Rosenthal (2006).

platforms; and (3) *valence polarization*, meaning the difference between candidates' levels of valence. The existing literature has not usually drawn a clear distinction between (1) and (2), and has not paid enough attention to (3).

On one hand, empirical studies concerned with elite policies have tended to use two types of measures: surveys,² and roll-call votes.³ Those two measures have often been thought of as interchangeable, and yet, they are conceptually different. We can probably assume that surveys are measuring preferences rather than platforms whereas roll-call votes are measuring platforms rather than preferences. The distinction matters, since a candidate's preferences and platforms can be different and far apart. The empirical analysis in Ansolabehere, Snyder and Stewart (2001b) is exceptional in using both types of measures—surveys and roll-call votes—in the same study. Interestingly, they find that legislators' preferences are predictors of roll-call behavior which is consistent with the results in our paper. In light of our theoretical results, we believe that more empirical work distinguishing the candidates' preferences and platforms is worth doing.⁴

On the other hand, some recent empirical papers have studied the relationship between valence and platforms.⁵ They have been largely motivated by the theoretical results in Groseclose (2001). But the empirical literature has not studied the effects of ideological preferences on valence accumulation, perhaps due to the absence of a theoretical framework with testable predictions.

In order to connect those three dimensions – preferences, platforms and valence – we develop a spatial voting model with the following characteristics. Following Downs (1957), policy is represented as a position in the unidimensional left-right spectrum. Following Calvert (1985), the candidates are policy-motivated. And following Groseclose (2001, 2007), valence is represented as a half-dimension orthogonal to policy. Unlike that previous research, our paper allows candidates to choose their location in the valence dimension by paying a cost.

The paper starts by studying the choice of policy platforms when two candidates have exogenously given levels of valence, and we derive several results on how valence influences policy. We ask the questions: how does a higher valence affect a candidate's platform; will policy be biased toward the high-valence or the low-valence candidate; and what is the relationship between platform

²See for example Burden (2004); Erikson and Wright (1980); Miller and Stokes (1963); Sullivan and Minns (1976).

³See for example Ansolabehere, Snyder and Stewart (2001a); Canes-Wrone, Brady and Cogan (2002); McCarty, Poole and Rosenthal (2006).

⁴A similar call was made by Krehbiel (1993), who convincingly advocated for separate empirical measures of legislator preferences and legislator behavior.

⁵See for example Adams, Merrill and Grofman (2005, chap. 12); Adams, Merrill, Stone and Simas (2009); Somer (2009); Stone and Simas (2007).

polarization and valence polarization? As a second step, we study the costly decision of those two candidates to increase their valence, and we derive results on how ideology influences that decision. We ask the questions: who acquires the most valence; how does the cost of acquiring valence affect the election outcome; and what is the relationship between preference polarization and valence polarization?

The model provides intuitive predictions regarding policy. The policy implemented will always be strictly more moderate than the ideologies of either candidate, meaning that the winning platform is always located between the candidates' ideal points. Moreover, the policy implemented will be biased toward the most extremist candidate, meaning that the winning platform will be right-of-center if an extremely right-wing candidate faces a moderate left-wing candidate, and the winning platform will be left-of-center if an extremely left-wing candidate faces a moderate right-wing candidate.

We also derive predictions about the levels of valence that candidates decide to acquire. We find that both candidates will acquire strictly positive levels of valence. But interestingly, out of the two candidates, it is the most extremist one who will exert the most effort to acquire valence.

The predictions regarding polarization are surprising. The candidates will usually adopt different platforms and different levels of valence, and thus will not converge to one another in either of those two dimensions. Therefore our results contradict the famous Downsian prediction of convergence to the median voter's ideal point.⁶ Obtaining a divergence result is noteworthy given that our model is particularly parsimonious, and it does not include any type of uncertainty. Indeed, the only condition in our model to obtain polarization is for candidates' ideal points to be non-symmetric – a most likely condition.

Another result is that *platform* polarization and *valence* polarization are positively correlated. If we interpret valence as campaign funds, then we are making a prediction that is consistent with an empirical fact in American politics: an increase in campaign spending should be correlated with an increase in polarization, which is exactly what has been observed in the United States during the past decades (McCarty et al. 2006).

One of the most striking results in our paper pertains to the effect of *preference* polarization on *platform* polarization: their relationship is non-monotonic. Indeed, for some parameter values, candidates' platforms will diverge if their preferences converge. This opens up the possibility

⁶See Fiorina (1999) for a survey of models predicting convergence or divergence.

that empirical studies may find a *negative* correlation between the polarization of candidates' true preferences (as measured for example by surveys), and the polarization of those candidates' platforms (as measured for example by roll-call votes).

Another set of results pertain to the cost of acquiring valence: when that cost decreases, both candidates will increase their valence but they will also increase the polarization of their platforms. This implies an unfortunate trade-off in the election: candidates will display high valence or low polarization, but not both. That has profound consequences for the welfare of voters who are affected both by valence and by polarization. As we discuss in the conclusions, the severity of that trade-off depends on the normative interpretation of valence. As a normative framework we borrow the conceptualization by Adams, Merrill, Stone and Simas (2009) and Stone and Simas (2007). As we discuss later, this newfound tension between valence and policy poses a problem for the appropriate representation of voters in democratic systems.

The rest of the paper proceeds as follows. First we summarize the related theoretical literature on valence. Then we lay out the basic setup of the model. As a first step toward solving the model, we treat valence as exogenous. We then use those results to solve the model when valence is endogenous. The following section connects the three types of polarization that we introduced. Finally, we conclude by discussing some empirical, normative and policy implications of the model. The appendix of this paper contains the proofs of all the results, as well as some extensions.

2 Related literature

This paper belongs to the class of models positing a *valence parameter*, i.e. a parameter in addition to policy that enters the utility function of all voters positively.⁷ That literature is growing rapidly, but most of it has treated valence as a fixed parameter that is exogenously determined.

A few papers, however, have allowed the agents in their models to affect the valence parameter through their actions. Ashworth and Bueno de Mesquita (2007) study how the endogenous adoption of platforms affects the endogenous adoption of valence. Their results have some interesting complementarities with our results. For example, we make predictions about the effect of

⁷Models with at valence dimension include Adams (1999); Adams and Merrill (2009); Adams, Merrill and Grofman (2005; chap. 11); Adams, Merrill, Simas and Stone (2009); Ansolabehere and Snyder (2000); Aragonés and Palfrey (2002); Ashworth and Bueno de Mesquita (2006; 2007); Callander (2008); Carrillo and Castanheira (2008); Feld and Grofman (1991); Groseclose (2001; 2007); Meirowitz (2006; 2008); Penn (2009); Schofield (2003; 2004); Schofield and Sened (2005); Wiseman (2006); Zakharov (2009).

valence on polarization, which turns out to be positive, whereas they make predictions about the effect of polarization on valence, which turns out to be negative.⁸ Meirowitz (2008) studies the amount of money that will be spent by an incumbent and a challenger to increase their respective valence when they both have different marginal costs. He finds that the disadvantaged candidate will exert more effort to increase her valence, which is reminiscent of our findings.⁹ Zakharov (2009) studies the endogenous spending by candidates on valence, which is interpreted as campaign advertising. One of his empirical predictions is consistent with our results, namely a positive correlation between campaign spending and party polarization. Schofield develops a multidimensional, multiparty model where political parties differ in location and valence. Parties strive to win over the centrist voters, but they also need to obtain the support of their non-centrist activists who could provide monetary and time contributions (Schofield 2003; 2004; 2005). Given that activists provide support as a function of the party’s policy location, the party can increase or decrease that support with its choice of policy (Schofield 2006; 2007). His results coincide with our paper in finding that parties will typically not locate at the center of the political spectrum. Schofield, Claassen, Ozdemir, Schmidman and Zakharov (2008) have an extension that compares proportional representation (where activists may form multiple political parties) with plurality rule (where small groups of activists may be forced to coalesce). In Callander (2008), valence is observed when the election is over: after getting elected, the candidate chooses a level of effort that is valued by voters. The rich description of candidates in Callander’s paper includes three features that we also include in our paper: platform, valence and preferences. The candidates in Carrillo and Castanheira (2008) need to select a policy platform, which is observable, and make an investment in quality, which is unobservable. In a result similar to ours (but reversing the causation), the candidate choosing the most extremist platform is also the one trying harder to acquire valence.

Our paper contributes to that literature by studying how competition in the valence dimension and competition in the policy dimension affect each other. Furthermore, our paper introduces and connects three types of divergence: platform polarization, valence polarization, and preference polarization.

3 The basic setup

⁸Those authors also make the latter prediction in Ashworth and Bueno de Mesquita (2006) with endogenous valence and exogenous platforms.

⁹See Meirowitz (2006) for an extension that adds a policy dimension.

In this section we introduce the elements of a basic model of electoral competition in two dimensions: a policy dimension, which represents the left-right political spectrum, and a valence dimension, which represents a non-policy attribute of candidates. First, political candidates are free to locate their platforms in any point on the left-right policy spectrum, but their location in the valence dimension is exogenously predetermined. In later sections we allow candidates to choose their location in the valence dimension as well. We denote policy by x , which is a real number, and we denote valence by v , which is a non-negative number. Thus the election takes place in the space $\mathbb{R} \times \mathbb{R}_+$. We now describe the voters and candidates involved in this election.

3.1 The electorate

The voters care about the positions of candidates in both dimensions; in other words they are both *policy-motivated* and *valence-motivated*. The electorate has a median voter, which we call M , whose preferences are decisive in the election. Regarding policy, voters have single-peaked utility functions around their ideal point. We normalize the ideal point of the median voter to zero. In addition to the policy implemented x , the electorate also cares about the valence v of the winning candidate. The utility function of M is given by

$$U_M(x, v) = -f(|x|) + v$$

where $f(\cdot)$ is a continuous and strictly increasing function with $f(0) = 0$ and $\lim_{x \rightarrow +\infty} f(x) = +\infty$. Throughout most of the paper we will assume $f(\cdot)$ to be linear, but the model is reasonably robust across different specifications.¹⁰

3.2 Candidates

There are two candidates competing in this election, labeled candidate R and candidate L , who are policy-motivated with ideal points X_R and X_L , respectively. We assume $X_L < 0 < X_R$ so that R and L have clearly distinct ideologies on opposite sides of the median voter. Both candidates have

¹⁰Indeed, we have studied a number of different specifications for the voters' utility function, and all the equilibria, whenever they exist, show similar properties to the ones reported here. In particular, with a quadratic function $f(\cdot)$ equilibria fail to exist for some parameter values. But they do exist for other parameter values; and whenever they exist, the equilibria show properties similar to the ones described in the text. In particular, we still find that candidates do not converge to one another; that valence polarization and platform polarization are positively correlated; that the most extremist candidate exerts the most effort to acquire valence; and that voters face a trade-off between high valence and low polarization.

single-peaked utility functions of the following form:

$$\begin{aligned}U_R(x) &= -g(|X_R - x|) \\U_L(x) &= -g(|X_L - x|)\end{aligned}$$

where $g(\cdot)$ is a continuous and strictly increasing function with $g(0) = 0$ and $\lim_{x \rightarrow +\infty} g(x) = +\infty$. In subsequent sections, when we study the endogenous investment in valence, we will focus on a specific form for $g(\cdot)$. For all the results regarding exogenous valence, however, $g(\cdot)$ can take any general form.

Candidate R and candidate L adopt policy platforms x_R and x_L , with $x_R, x_L \in \mathbb{R}$. Candidates are also characterized by a parameter v denoting each candidate's valence, where v is a non-negative number. We call v_R and v_L the valences of R and L , respectively. In this section, candidates' valences are exogenously fixed. We call δ the difference in valence between R and L , meaning that $\delta \equiv v_R - v_L$. Note that this valence advantage of R over L can take any real value, that is, $\delta \in \mathbb{R}$.

3.3 Timing, information and solution concept

The timing of this election is the following:

1. Candidates' valences v_R and v_L are observed.
2. Candidates simultaneously announce x_R and x_L .
3. The median voter elects R or L .

All that information is common knowledge. This election is thus a deterministic game of complete information. The game must be solved by backward induction. The solution concept is subgame-perfect equilibrium (SPE), which requires that strategies form a Nash equilibrium (NE) in every subgame.

4 The model with exogenous valence

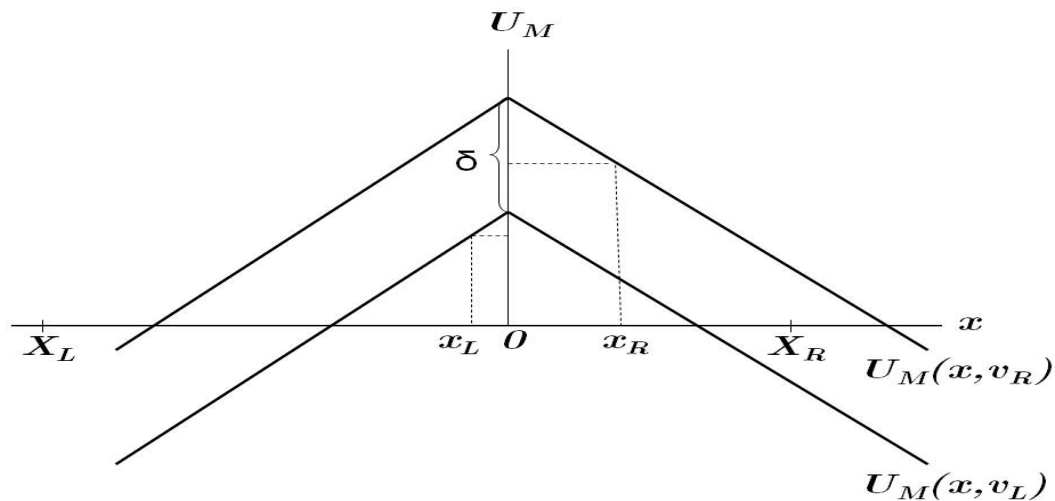
In this section we solve the basic exogenous-valence model laid out above. Let us first study how the median voter makes her decision at the ballot box. M will vote for the candidate who maximizes

her utility. If M is indifferent between the two candidates, we assume that she will vote for the one who has the highest valence – this assumption is important to keep in mind because, in equilibrium, that is exactly the situation that M will find herself in. If both candidates have the same valence, she will randomize equally between the two.¹¹ We assume that $f(\cdot)$ is linear, meaning that

$$U_M(x, v) = -|x| + v$$

M 's appreciation for a candidate decreases with the distance between her ideal point and that candidate's platform, and increases with the candidate's valence. In essence, the valence parameter v "shifts up" the utility function of M . An example of how M evaluates R and L is illustrated in Figure 1, where it is assumed that $v_L < v_R$ and $|x_L| < |x_R|$. In the case depicted in that graph, candidate R is strictly preferred to candidate L in spite of having a more extremist platform. Candidate R wins the election because her higher score in the valence dimension more than compensates her extremism in the policy dimension.

Figure 1: The effect of a valence advantage of R over L



4.1 Equilibrium results with exogenous valence

¹¹With other assumptions when M is indifferent an equilibrium might not exist. But the outcome would still converge arbitrarily close to the equilibria described in the text.

We now turn our attention to the behavior of candidates when they must formulate their policies in Stage 2 of the election. Given that candidates' valence levels are fixed when the game starts, the equilibrium platforms and equilibrium outcomes are contingent on those valences. In other words, there is a different subgame for each pair v_R, v_L .

Given the information they have, what platforms will candidates formulate? Our solution concept, subgame-perfect equilibrium (SPE), imposes that R and L must play a Nash equilibrium (NE) in every subgame. We call x_R^* and x_L^* such an equilibrium and x^* the ensuing winning platform. As we can see in the following theorem, an equilibrium exists for every pair of valence levels.¹²

Theorem 1 *The equilibrium strategies of candidates and the equilibrium outcome of this election, for all possible values of v_R and v_L , are given in the following table, where $\delta \equiv v_R - v_L$*

Value of δ	Equilibrium platforms	Winning platform	Winning candidate
	x_R^*, x_L^*	x^*	
$X_R \leq \delta$	$x_R^* = X_R$ $x_L^* \in \mathbb{R}$	X_R	R
$0 < \delta < X_R$	$x_R^* = \delta$ $x_L^* = 0$	δ	R
$\delta = 0$	$x_R^* = 0$ $x_L^* = 0$	0	R or L with equal probability
$X_L < \delta < 0$	$x_R^* = 0$ $x_L^* = \delta$	δ	L
$\delta \leq X_L$	$x_R^* \in \mathbb{R}$ $x_L^* = X_L$	X_L	L

There are several remarks to make about the results in this table. For starters note the results when $\delta = 0$, that is, when there is no valence difference between the candidates. This corresponds to the standard election between two candidates who are policy motivated and compete only in the policy dimension, such as studied in the canonical Calvert (1985). In fact the results correspond to those by Calvert: the centripetal forces in the election drive both candidates to converge completely to the median voter's ideal point.

¹²The proofs of all the results in this paper come in the appendix.

As soon as $\delta \neq 0$, however, the results depart from the standard outcome in notable ways. Most importantly, the candidate with higher valence has the ability to move the policy implemented toward her ideal point, and increasingly so as her valence advantage increases. For example when candidate R has the highest valence, meaning that $0 < \delta$, she is able to diverge away from the center and still win the election: she adopts a platform closer to her preferred policy, and increasingly closer the higher her valence advantage. As a matter of fact, when δ increases past a certain threshold (namely X_R), candidate R is actually able to propose exactly her ideal point and still win the election.

In the meantime, the candidate with lowest valence, say L , will converge to the center of the spectrum. The reason why L adopts the median voter's ideal point is to force R , who will win the election anyhow, to converge as much as possible. By adopting $x_L = 0$ she constrains R to diverge no further than $x = \delta$. If R 's valence advantage is too large however, meaning that $\delta \geq X_R$, then there is no location that L can adopt to prevent R from adopting her ideal point.

As a final remark, let us note an important implication regarding polarization. We define *platform polarization* as $|x_R - x_L|$, and *valence polarization* as $|v_R - v_L|$. Our model predicts that an increase in valence polarization will lead to an increase in platform polarization. To see this, imagine that δ increases from zero to positive values. Then candidate R will adopt an increasingly right-wing platform while candidate L will remain at the center of the political spectrum, and thus, the distance between x_R and x_L will increase.

5 The model with endogenous valence

In this section we allow the candidates' score in the valence dimension to be endogenously determined. Specifically, we assume that candidates can exert effort to increase their valence before the election takes place.

5.1 Utility functions and timing

R and L can now choose their own levels of valence, v_R and v_L , by paying a cost $C(v_R)$ and $C(v_L)$, respectively.¹³ The cost function $C(\cdot)$ is defined as $C(v) \equiv cv^2$ with $c > 0$. Smaller values of the

¹³So we assume that candidates do not have any valence to begin with, and they must pay a cost for any amount of valence that they wish to have. An alternative assumption would be that candidates come to the election with a *preexisting* amount of valence that is given exogenously, and then they pay a cost for any *additional* valence they

parameter c imply a smaller marginal cost of increasing v .

To provide explicit solutions, we assume that candidates have quadratic utility functions over policy, meaning that $g(\cdot)$ is quadratic.¹⁴ Thus, the payoffs for candidate R and candidate L are the following:

$$U_R(x) - C(v_R) = -(X_R - x)^2 - cv_R^2 \quad (1)$$

$$U_L(x) - C(v_L) = -(X_L - x)^2 - cv_L^2 \quad (2)$$

Candidates will choose the levels of valence that maximize their utility. If a candidate is indifferent between choosing two levels of valence, we assume that she chooses the highest one.

The timing of the election, now that valences are endogenous, is the following:

1. Candidates simultaneously choose v_R and v_L
2. Candidates' valences v_R and v_L are observed
3. Candidates simultaneously announce x_R and x_L
4. The median voter elects R or L

The game must be solved by backward induction, but Stages 2, 3 and 4 are identical to the game where valence is not endogenous, which has already been studied in the previous section. So we can take the results from that game as given (namely Theorem 1), and directly study the *reduced game* at Stage 1 of the timing above.

5.2 Optimal levels of valence

At this stage, we need to derive how each candidate chooses her optimal level of valence for any given valence of the other candidate. In game-theoretic parlance, each candidate's valence must be a *best-response* to all other parameters. We call v_R^* the best-response of candidate R , and v_L^*

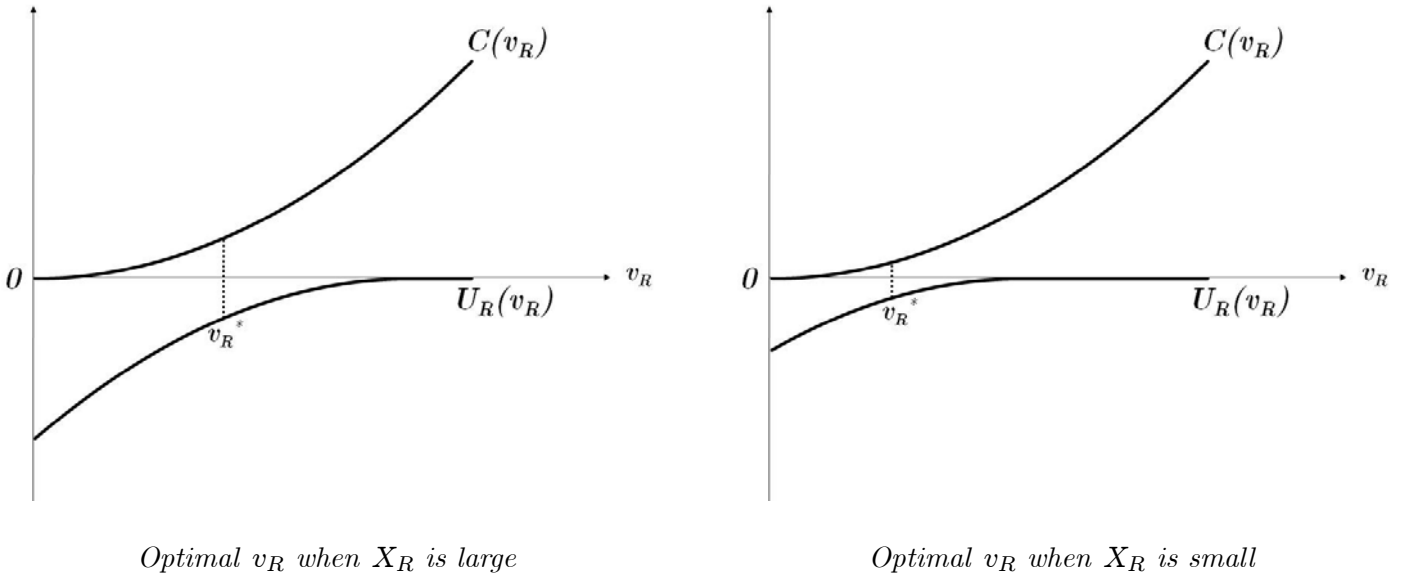
wish to acquire. Such initial levels of valence could be labeled v_{RI} and v_{LI} . We have solved the model with such initial valences, and we have found that all the results still hold. In particular, all the effects of X_R and X_L remain the same. In subsequent footnotes we report the results with preexisting levels of valence v_{RI} and v_{LI} .

¹⁴All the results would be qualitatively the same with any strictly concave $g(\cdot)$, and we use the quadratic only to provide explicit results. The utility functions could also be linear, as we show in the appendix. All the results with a linear $g(\cdot)$ remain the same, except regarding the ideal points of parties: indeed, the ideal points X_L and X_R lose their effect because they drop out of the equations.

the best-response of candidate L . Their exact expressions are derived in the appendix. Here we mention their most important properties.

To begin with, let us understand how each candidate calculates her optimal level of valence. Consider candidate R 's decision problem: she needs to maximize Equation 1 taking the other candidate's behavior as given. That is, for a fixed v_L , the candidate increases her valence v_R up to the point where its marginal cost equals its marginal benefit. As we can see in the figure below, the optimal v_R is found where the marginal increase in $U_R(v_R)$ is equal to the marginal increase in $C(v_R)$.

Figure 2: The optimal choice of valence by candidate R
 assuming that she is either extremist (left panel) or moderate (right panel)



The optimal levels of valence have three main properties. The first important property is that each candidate's optimal valence is increasing with her ideological extremism. In other words, v_L^* is increasing in $|X_L|$, and v_R^* is increasing in $|X_R|$. This can be seen in Figure 2. The figure depicts two scenarios. In one scenario R has an extremist ideology, meaning that X_R is large, and in the other scenario R has a moderate ideology, meaning that X_R is small. If X_R was small, meaning that R had moderate preferences, then v_R^* would also be small (see the right-hand panel). However,

if X_R was larger, implying more extremist preferences, then R would find it worthwhile to acquire additional valence in order to have a larger influence in the election (see the left-hand panel). The reason is that when X_R increases, the marginal benefit of v_R^* increases. That is a direct consequence of having a strictly concave utility function.

The second property is that each candidate's optimal valence is increasing with the opponents' valence. In other words, v_L^* is increasing in v_R , and v_R^* is increasing in v_L . This leads candidates to compete with each other on the valence dimension. To understand the intuition, let us think how R should respond if L increased her valence. We know, from Theorem 1, that the winning policy in the subsequent stages of the election would move closer to L 's ideal point and farther from R 's ideal point. Given that R has a strictly concave utility function, the marginal benefit of moving policy toward her ideal point increases, and therefore paying an extra cost of increasing her valence becomes worthwhile.

And the third property is that, not surprisingly, the optimal valence decreases with the cost of acquiring valence. In other words, v_R^* and v_L^* are decreasing with c .¹⁵ It should be noted that, in equilibrium, the candidates must take each other's behavior into account, and therefore v_R^* and v_L^* need to be solved as two simultaneous equations. That is our task in the following section.

5.3 Equilibrium results with endogenous valence

We know that any Nash equilibrium, if it exists, is such that L is best-responding to R , and R is best-responding to L . Therefore, to find the Nash equilibria we must intersect both best-response functions. In other words, v_R^* and v_L^* need to be solved simultaneously. We denote by v_R^{**} and v_L^{**} the solutions of that simultaneous problem. Thus, v_R^{**} and v_L^{**} are the equilibrium levels of valence in the endogenous case.

The next result provides the exact expressions for that equilibrium. For convenience, we assume that c is not too low.¹⁶

Theorem 2 *In this election, there exists a unique equilibrium in the candidates' choice of valence.*

¹⁵In the appendix we derive the optimal valence in the case where candidates have linear instead of strictly concave utility functions. As we show there, the first property (regarding X_R) does not hold in the linear case, but the other two properties (regarding v_R and c) do hold.

¹⁶Specifically we assume that $c > \max \left\{ \frac{X_L + \sqrt{5X_L^2 - 4X_L X_R}}{-2X_L}, \frac{-X_R + \sqrt{5X_R^2 - 4X_L X_R}}{2X_R} \right\}$. This assumption is mild as long as X_L and X_R are reasonably far from the median voter's ideal point. If c was below that threshold, candidates would overreact to each other's choice of valence so that an equilibrium might fail to exist.

The equilibrium is

$$\begin{aligned} v_R^{**} &= \frac{X_R(c+1) - X_L}{c(c+2)} \\ v_L^{**} &= \frac{X_R - X_L(c+1)}{c(c+2)} \end{aligned}$$

Note some interesting properties of this equilibrium.¹⁷ Both candidates choose a strictly positive level of valence, as can be seen by noting that all the values in those two formulas are positive. The level of valence that each candidate chooses is decreasing with c , as can easily be proved by differentiation of v_R^{**} and v_L^{**} . So the larger the marginal cost of acquiring in valence, the lower the effort that a candidate exerts to acquire it. Actually, as c tends to infinity each candidate's level of valence tends to zero.

Regarding ideology, as we had suggested before, each candidate's equilibrium valence increases with *her own* ideological extremism. Moreover, a candidate's equilibrium valence also increases when the *other* candidate becomes more extremist: R increases her valence when L 's ideal point becomes more leftist, and L increases her valence when R 's ideal point becomes more rightist.

Of particular interest is the question of which candidate will exert more effort to acquire valence. As the result below indicates, when one of the candidates has moderate preferences and the other has extremist preferences, the extremist one will exert more effort to increase her valence.

Corollary 1 *If $|X_L| < (>) (=) |X_R|$ then $v_L^{**} < (>) (=) v_R^{**}$.*

This result is intuitive, as it reflects the incentives of candidates. The most extremist candidate is the most sensitive to policy outcomes (due to concave utility functions). So she will find investing on valence more worthwhile.

The equilibrium characterized above has consequences down the road, in the subsequent stages of the election, as described in the following result. We denote by x_R^{**} and x_L^{**} the equilibrium platforms adopted by R and L when their choice of valence is endogenous, and by x^{**} the resulting winning platform.

Theorem 3 *In this election, the equilibrium platforms of candidates and the equilibrium strategy*

¹⁷As we mentioned in a previous footnote, it is possible that candidates may come to the election with some preexisting levels of valence. If we included in the model some initial amounts of valence for the candidates, v_{RI} and v_{LI} , then the equilibrium would be the following: $v_R^{**} = \frac{X_R(c+1) - X_L - cv_{RI} + cv_{LI}}{c(c+2)}$ and $v_L^{**} = \frac{X_R - X_L(c+1) + cv_{RI} - cv_{LI}}{c(c+2)}$. The important point is that all the effects of c , X_R , X_L would remain the same.

of the median voter are the following:

If $|X_R| < |X_L|$ candidate L adopts $x_L^{**} = \frac{1}{c+2}(X_L + X_R)$; candidate R adopts $x_R^{**} = 0$; and the median voter elects L .

If $|X_L| < |X_R|$ candidate L adopts $x_L^{**} = 0$; candidate R adopts $x_R^{**} = \frac{1}{c+2}(X_L + X_R)$; and the median voter elects R .

If $|X_L| = |X_R|$ candidate L adopts $x_L^{**} = 0$; candidate R adopts $x_R^{**} = 0$; and the median voter elects R or L with equal probability

One of the most relevant implications of this theorem is the non-negative relationship between a candidates' preferences and her platform. Take for example the relationship between R 's ideal point and her equilibrium platform. If $|X_R| < |X_L|$, the relationship is null: when X_R changes, x_R^{**} will not change at all (from zero). However, if $|X_L| < |X_R|$, the relationship is strictly positive: when X_R increases, x_R^{**} will also increase. From the empirical point of view, we are thus predicting a positive, but imperfect, correlation between a candidate's preferences and her platforms: a change in preferences should result in a change of platform in the same direction, or no change at all.

It is now possible to make predictions about the policy that we expect to be implemented in this election. Its equilibrium value comes directly from the strategies derived above.

Theorem 4 *The policy implemented in equilibrium is $x^{**} = \frac{1}{c+2}(X_L + X_R)$.*

Some of the properties of this equilibrium policy are worth spelling out as lemmas.¹⁸

Lemma 1 *The equilibrium policy, x^{**} , is always such that $X_L < x^{**} < X_R$.*

This result provides a good reality-check. It says that the implemented policy will always be located strictly between the preferred policies of either candidate. In particular, no candidate will be able to pull policy all the way to her ideal point.

Given how extremist the candidates are relative to each other, will the winning platform be leftist, rightist or centrist?

Lemma 2 *If $|X_L| < (>) (=) |X_R|$ then $x^{**} > (<) (=) 0$.*

¹⁸If we included some initial amounts of valence for the candidates, v_{RI} and v_{LI} , then the equilibrium would be the following: $x^{**} = \frac{1}{c+2}(X_L + X_R + cv_{RI} - cv_{LI})$. Note that all the effects of c , X_R , X_L would remain the same.

The result above indicates that the policy implemented will be biased toward the candidate with the most extremist ideal point. This prediction reflects the intensity of preferences among candidates. The most extremist one will exert a larger effort to acquire valence, which will skew policy in her preferred direction.

Some final properties of the policy implemented come from its comparative statics.

Lemma 3 *The equilibrium policy, x^{**} , is*

- *increasing with X_R*
- *increasing with X_L*
- *closer to zero when c increases*

An unexpected result is that c has a moderating effect: when c increases, the policy implemented is brought closer to the median voter's ideal point. From the normative point of view, this presents an unfortunate trade-off for the median voter with respect to the cost of acquiring valence: when c increases, on one hand the policy implemented becomes more moderate which is beneficial, but on the other hand both candidates reduce their valence which is detrimental.

6 Relationship between the three types of polarization

Finally, we turn to the connection between platform polarization defined as $|x_R - x_L|$, valence polarization defined as $|v_R - v_L|$, and preference polarization defined as $|X_R - X_L|$. All three types of elite polarization are connected by the following result.

Lemma 4 *In equilibrium we have $|x_R^{**} - x_L^{**}| = |v_R^{**} - v_L^{**}| = \frac{1}{c+2} |X_L + X_R|$*

Note that we are predicting a strong correlation between platform polarization and valence polarization. In equilibrium, they will exactly follow each other.

Now consider the candidates' preferences, which are the independent variables in this model. In the next theorem we study their effect on valence and platforms, which are the dependent variables.

Theorem 5 *If $|X_L| \neq |X_R|$ then platform polarization and valence polarization are strictly positive in equilibrium, and they increase when*

- c decreases
- the ideal point of the extremist candidate becomes more extremist
- the ideal point of the moderate candidate becomes more moderate

If $|X_L| = |X_R|$ then platform polarization and valence polarization are null in equilibrium.

The first thing to note is that platform polarization is expected to be strictly positive: In all instances, except the knife-edge case where $|X_L| = |X_R|$, there will be a distance between L 's platform and R 's platform. And such distance could be very large if $|X_L|$ and $|X_R|$ are very asymmetric. The model thus contradicts the famous results of full convergence contained in Downs (1957) and Calvert (1985).

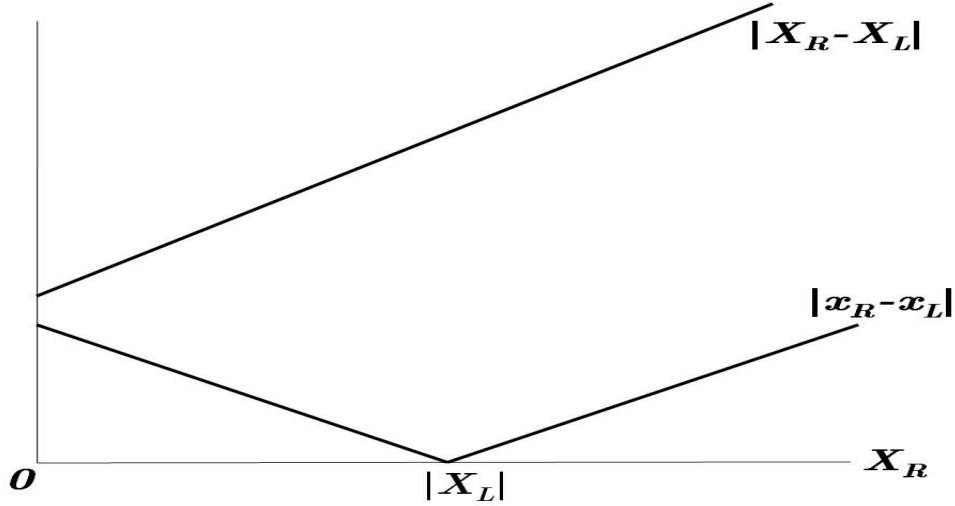
Another unexpected result is the cost of valence having a negative effect on platform polarization. When c decreases, both candidates will increase their valence but they will do so in different proportion, the extremist candidate acquiring more of it. As a consequence, valence polarization will increase and therefore platform polarization will increase. To summarize, increasing c will have three important effects: the *valence* of both candidates will increase; *valence polarization* will increase; and *platform polarization* will also increase.

But the most striking feature of this theorem is the non-monotonic effect of ideology on polarization. What happens to polarization when the preferences of a given candidate become more extremist, all else equal? It depends, it turns out, on whether that candidate is the most extremist or the most moderate of the two.

This relationship is depicted in Figure 3, where preference polarization, $|X_R - X_L|$, and platform polarization in equilibrium, $|x_R^{**} - x_L^{**}|$, are graphed as functions of R 's ideal point, X_R . Consider what would happen to polarization when the ideal point of candidate R gradually moves from moderate to extreme. For low values of X_R , namely the interval $X_R \in (0, |X_L|)$, we have a paradoxical result: if the preference polarization between R and L *increases*, their platform polarization will *decrease*. In other words, candidates' platforms will converge when their preferences diverge. The reason for this paradox is that, as R 's preferences become more extremist, R does not change her platform in equilibrium but she does increase her valence which forces L to moderate her platform. As a consequence, R 's platform and L 's platform become closer to each other. In

contrast, for high values of X_R , namely for the interval $X_R \in (|X_L|, \infty)$, we have the opposite and less surprising result: if preference polarization increases, platform polarization will also increase.

Figure 3: The non-monotonic relationship between platform polarization and preference polarization



7 Empirical, normative and policy implications

What distinguishes candidates from one another? This paper has emphasized three features. First, candidates are distinguished by their *policy platforms*, meaning the policies they promise to implement if they are elected. Second, candidates are distinguished by their *policy preferences*, meaning their ideal policy outcomes. Third, candidates are distinguished by their *valence*, meaning their non-policy assets. Seldom have those three characteristics been studied together, and yet, there is much to learn from an electoral theory that establishes causal connections between them. By including those three variables in a single theoretical framework, our paper has uncovered some surprising relationships. Several of our results have empirical, normative or policy implications.

7.1 Empirical support for the results

One prediction of our model is that a candidate's true preferences have a non-negative effect on the policies they actually adopt in office: for some intervals of the parameters the effect will be positive, and for other intervals the effect will be zero. This result is backed by the findings in Ansolabehere,

Snyder and Stuart (2001b) who find "a strong association between roll-call behavior in the 105th Congress and the electoral expression of preferences that preceded it" (p. 543). Therefore, in line with our predictions, those authors find that candidates' preferences are a strong, though imperfect, predictors of their subsequent votes in the legislature.

Another prediction is the positive correlation between valence and polarization. Evidence supporting this result comes from interpreting valence as candidates' campaign funds. We know that campaign contributions have increased dramatically in the last two decades. Soft money contributions, for example, went from \$100 million in 1992 to almost \$500 million in 2002. In that same period, the roll-call votes of the Democratic and Republican House representatives have been increasingly divergent, as pointed out by McCarty, Poole and Rosenthal (2006). Those authors concluded the following: "The past thirty years have been a period not only of political polarization but also of rapidly increasing campaign spending, which has led many to speculate that the two are linked" (McCarty, Poole and Rosenthal 2006, p.160). Our paper suggests a way in which the two might possibly be linked. The availability of donors – which decreases the marginal cost of acquiring campaign funds – should have two concurrent effects: increasing the overall campaign spending by candidates, and increasing the polarization of their platforms.

The model has other results that could be tested empirically. In particular, we predict the policy implemented to be biased toward the most extremist candidate. We predict the most extremist candidate to exert the most effort in acquiring valence. And we predict a non-monotonic relationship between platform polarization and preference polarization whereby it is possible to observe platforms diverging when preferences converge.

7.2 A policy consideration

These results have a relevant policy implication. Any new legislation that makes fund-raising more difficult, for example by limiting the influence of soft money as the McCain-Feingold act tried to do, will have an unintended benefit: it might decrease polarization. According to our theory, increasing the marginal cost of acquiring donations will decrease the candidates' disparity in campaign spending, which will force them to converge toward each other.

7.3 Normative evaluation

Finally, our analysis is informative for the normative evaluation of democratic elections. Theories of democracy have emphasized the importance of high levels of competitiveness on policy issues (for example in Downs 1957). Such competitiveness induces candidates to improve their offers to voters in terms of policy. Subsequent scholars have emphasized the existence of another dimension that is also valued by voters, namely valence (a concept originally introduced by Stokes 1963). What this paper wishes to emphasize is that candidates also compete in the valence dimension, and thus valence can be considered endogenous. Indeed, we believe that candidates are willing to pay the cost of investing in some valence before an election. On that basis, theories of democracy should recognize the importance of competition in non-policy issues as another way to improve the offers to voters.

This double competition, in policy and valence, is worth studying within the same theoretical framework. Indeed, those two dimensions might interact in non-trivial ways. As explained by Stone and Simas (2007), the question is important because it illuminates whether those two components of representation work at cross purposes or not. Those authors clearly identified the outcome that voters would prefer: "if constituent's interests in non-policy and policy concerns reinforce the quality of representation in both dimensions, there is cause for optimism about the electoral process (Stone and Simas 2007, p. 4)." As our model suggests, however, there are theoretical reasons to doubt such optimism. As it turns out, combining competition in the policy dimension with competition in the valence dimension reveals a surprising trade-off: candidates will display high valence or low polarization, but not both. This trade-off is unfortunate for the median voter, as she can expect only one of two possible outcomes: either candidates with high valence but divergent policies; or candidates with low valence but convergent policies.¹⁹

The severity of that trade-off depends on the normative interpretation of valence. To clarify this point, it is useful to borrow a dichotomy introduced by Adams, Merrill, Stone and Simas (2009). Those authors define two separate concepts. The first one, *character valence*, refers to factors that are intrinsically beneficial to voters after the election, such as competence, experience, credibility, honesty, and dedication to public service. This contrasts with the second one, *strategic valence*,

¹⁹The trade-off that we find in our paper is consistent with the theoretical results in Adams and Merrill (2009) and the empirical results in Somer (2009). But it seems to contradict the theoretical results in Groseclose (2001) and the empirical results in Stone and Simas (2007), who find that candidates can simultaneously have high valence and moderate platforms. More research is clearly needed to inform this budding debate.

which refers to factors that are only useful to gain votes during the electoral campaign, such as funding, name recognition, and campaigning skills.²⁰ If we interpret our parameter as character valence, then voters might be better-off on the net by suffering some polarization in exchange for a higher valence. However, if we interpret our parameter as strategic valence, then such additional polarization will represent a net loss given that voters will not be compensated by any significant post-electoral bonus. In future research we plan to study the effects of these results on voters' welfare. Our hope is that a formal analysis of these normative considerations will help clarify the tension between policy and valence, and will enrich our theories of representation.

References

- [1] Adams, James. 1999. "Policy Divergence in Multicandidate Probabilistic Spatial Voting." *Public Choice* 100: 103-122.
- [2] Adams, James and Samuel Merrill III. 2009. "Policy-Seeking Parties in a Parliamentary Democracy with Proportional Representation: A Valence-Uncertainty Model." *British Journal of Political Science* 39 (03) July: 539-558.
- [3] Adams, James, Samuel Merrill III and Bernard Grofman. 2005. *A Unified Theory of Party Competition: A Cross-National Analysis Integrating Spatial and Behavioral Factors*. New York: Cambridge University Press.
- [4] Adams, James, Samuel Merrill III, Walter J. Stone and Elizabeth N. Simas. 2009. "When Candidates Value Good Government: A Spatial Model with Applications to Congressional Elections." Typescript.
- [5] Ansolabehere, Stephen and James M. Snyder Jr. 2000. "Valence Politics and Equilibrium in Spatial Election Models." *Public Choice* 103 (3-4) June: 327-36.
- [6] Ansolabehere, Stephen, James M. Snyder, Jr. and Charles Stewart, III. 2001a. "Candidate Positioning in U.S. House Elections." *American Journal of Political Science* 45 (1) January: 136-159.

²⁰Stone and Simas (2007) introduced two equivalent concepts, which they called *personal quality* and *strategic quality*.

- [7] Ansolabehere, Stephen, James M. Snyder, Jr. and Charles Stewart, III. 2001b. "The Effects of Party and Preferences on Congressional Roll-Call Voting." *Legislative Studies Quarterly* XXVI (4) November: 533-572.
- [8] Aragonés, Enriqueta and Thomas R. Palfrey. 2002. "Mixed Equilibrium in a Downsian Model with a Favored Candidate." *Journal of Economic Theory* 103: 131-161.
- [9] Ashworth, Scott and Ethan Bueno de Mesquita. 2006. "Delivering the Goods: Legislative Particularism in different Electoral and Institutional Settings." *Journal of Politics* 68 (1) February: 168-179.
- [10] Ashworth, Scott and Ethan Bueno de Mesquita. 2007. "Valence Competition and Platform Divergence." Typescript.
- [11] Burden, Barry C. 2004. "Candidate Positioning in US Congressional Elections." *British Journal of Political Science* 34 (2): 211-227.
- [12] Callander, Steven. 2008. "Political Motivations." *Review of Economics Studies* 75: 671-697.
- [13] 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) March: 127-140.
- [14] Calvert, Randall L. 1985. "Robustness of the Multidimensional Voting Model: Candidate Motivations, Uncertainty, and Convergence." *American Journal of Political Science* 29: 69-95.
- [15] Carrillo, Juan and Micael Castanheira. 2008. "Information and Strategic Political Polarization." *Economic Journal* 118 July: 845-874.
- [16] Downs, Anthony. 1957. *An Economic Theory of Democracy*. New York: Harper and Brothers Publishers.
- [17] Erikson, Robert S. and Gerald C. Wright, Jr. 1980. "Policy Representation of Constituency Interests." *Political Behavior* 2 (1) 1980: 91-106.
- [18] Feld, Scott L. and Bernard Grofman. 1991. "Incumbency Advantage, Voter Loyalty and the Benefit of the Doubt." *Journal of Theoretical Politics* 3 (2): 115-137.

- [19] Fiorina, Morris P. 1999. "Whatever Happened to the Median Voter?." originally prepared for the MIT Conference on Parties and Congress. October 2. Cambridge, MA.
- [20] Groseclose, Tim. 2001. "A Model of Candidate Location When One Candidate Has a Valence Advantage." *American Journal of Political Science* 45 (4): 862-886.
- [21] Groseclose, Tim. 2007. "'One and a half dimensional' preferences and majority rule." *Social Choice and Welfare* 28: 321-335.
- [22] Krehbiel, Keith. 1993. "Where's the Party?." *British Journal of Political Science* 23 (2) April: 235-266.
- [23] McCarty, Nolan, Keith T. Poole, and Howard Rosenthal. 2006. *Polarized America: The Dance of Ideology and Unequal Riches*. Walras-Pareto Lectures. June 23. Cambridge MA: The MIT Press.
- [24] Meirowitz, Adam. 2006. "Electoral Contests." Typescript.
- [25] Meirowitz, Adam. 2008. "Electoral Contests, Incumbency Advantages, and Campaign Finance." *Journal of Politics* 70 (03) July: 681-699.
- [26] Miller, Warren E. and Donald E. Stokes. 1963. "Constituency Influence in Congress." *American Political Science Review* 57 (1) March: 45-56.
- [27] Penn, Elizabeth Maggie (2009), "From Many, One: State Representation and the Construction of an American Identity," *Journal of Theoretical Politics*, 21(3): 343-364.
- [28] Schofield, Norman. 2003. "Valence Competition in the Spatial Stochastic Model." *Journal of Theoretical Politics* 15 (4): 371-383.
- [29] Schofield, Norman. 2004. "Equilibrium in the Spatial 'Valence' Model of Politics." *Journal of Theoretical Politics* 16 (4): 447-481.
- [30] Schofield, Norman. 2005. "A Valence Model of Political Competition in Britain: 1992-1997." *Electoral Studies* 24: 347-370.
- [31] Schofield, Norman. 2006. "Equilibria in the Spatial Stochastic Model of Voting with Party Activists." *Review of Economic Design*. 10: 183-203.

- [32] Schofield, Norman. 2007. "The Mean Voter Theorem: Necessary and Sufficient Conditions for Convergent Equilibrium." *Review of Economic Studies*. 74: 965-980.
- [33] Schofield, Norman, Christopher Claassen, Ugur Ozdemir, Evan Schnidman and Alexei V. Zakharov. 2008. "Party Positioning under Proportional Representation and Plurality Rule." *Center in Political Economy*. Washington University. Typescript.
- [34] Schofield, Norman and Itai Sened. 2005. "Multiparty Competition in Israel, 1988-96." *British Journal of Political Science* 35: 635-663.
- [35] Somer, Zeynep. 2009. "Party Policy Strategies and Valence Issues: The Effect of Government Membership and Economic Performance on Parties' Policy Strategies in Post-Communist Europe." Typescript.
- [36] Stokes, Donald. 1963. "Spatial Models of candidate Competition." *American Political Science Review* 57: 368-377.
- [37] Stone, Walter J. and Elizabeth N. Simas. 2007. "Candidate Valence and Ideological Positioning in the 2006 House Elections." paper presented at the Annual Meeting of the Midwest Political Science Association. April 12. Chicago, IL.
- [38] Sullivan, John L. and Daniel Richard Minns. 1976. "Ideological Distance between Candidates: An Empirical Examination." *American Journal of Political Science* 20 (3) August: 439-468.
- [39] Wiseman, Alan E. 2006. "A Theory of Partisan Support and Entry Deterrence in Electoral Competition." *Journal of Theoretical Politics* 18 (2): 123-158.
- [40] Zakharov, Alexei V. 2009. "A Model of Candidate Location with Endogenous Valence." *Public Choice* 138: 347-366.

A Appendix

A.1 An extension: Candidates with linear preferences

In this section we study the case where $g(\cdot)$ is linear, that is, the utility functions of candidates are given by $U_L(x) = -|X_L - x|$ and $U_R(x) = -|X_R - x|$.

In order to find the Nash equilibria at this stage of the game, we need to derive the best-response functions of candidates, that is, the optimal valence the they would choose taking the valence of the other candidate as given. We call v_L^* the best-response function of candidate L taking v_R as given, and v_R^* the best-response function of candidate of candidate R taking v_L as given. Their values are given in the following lemma.

Lemma 5 *The best-response functions of L and R , respectively, are*

$$v_L^* = \begin{cases} \frac{1}{2c} & \text{if } v_R \in [0, X_R + \frac{1}{4c}] \\ 0 & \text{if } v_R \in (X_R + \frac{1}{4c}, +\infty) \end{cases}$$

and

$$v_R^* = \begin{cases} \frac{1}{2c} & \text{if } v_L \in [0, -X_L + \frac{1}{4c}] \\ 0 & \text{if } v_L \in (-X_L + \frac{1}{4c}, +\infty) \end{cases}$$

The most noteworthy feature of this lemma is that for small values of the opponent's valence, a given candidate will find it worth to have a strictly positive valence: the benefit from pulling policy towards its ideal point justifies a certain investment in valence. For large values of the opponent's valence, however, a given candidate will find it too costly to offset that valence (due to a convex cost function), and would rather not invest in valence at all, letting the opponent pull policy all the way to its ideal point. Another (very intuitive) feature of the best-response functions is that they are decreasing with c : all things equal, the larger the marginal cost of investing in valence, the lower the desired level of valence.

We should finally note that for small values of the opponent's valence, the best-response function has a unique value, \bar{v} . This is due to the linearity of the candidates' utility function which makes those candidates insensitive to the specific location of its ideal point.

We know that any NE, if it exists, is such that L is best-responding to R , and R is best-responding to L . This lemma therefore allows us to compute the Nash equilibria, if any exists, by

intersecting both best-response functions. To ensure that an equilibrium exists we must assume that c is not too low. (Relaxing that assumption could prevent the existence of equilibria because, costs being so low, candidates would overreact to to each other.) Technically we assume that $c > \max \left\{ \frac{1}{-2X^L}, \frac{1}{2X^R} \right\}$. This assumption is not too restrictive as long as X_L and X_R are reasonably far from the median voter's ideal point.

We are now equipped to state the main result of this section, which determines the equilibrium in the election. We denote by v_L^{**} and v_R^{**} any valence choices that form a NE of the reduced game at this stage.

Theorem 6 *In this election there exists a unique equilibrium in the candidates' choice of valences.*

*The equilibrium is $v_L^{**} = \frac{1}{2c}$ and $v_R^{**} = \frac{1}{2c}$.*

Note some interesting properties of this equilibrium. First, both candidates choose a strictly positive level of valence. Second, the level of valence that each candidate chooses is decreasing with c : the larger the marginal cost of investing in valence, the lower the equilibrium investment of valence by both candidates. Actually, as c tends to infinity each candidate's level of valence tends to zero, and thus the election converges to the standard Calvert framework where candidates are purely office motivated and there is no valence dimension. All these features will prove robust to the concavity of the candidates' utility function: they will hold as well when they are strictly concave.

Now, note that in equilibrium both candidates choose the exact same level of valence, that is, $v_L^{**} = v_R^{**}$. This is in spite of candidates being asymmetric in other respects; in particular we must remember that we did *not* assume that L and R were symmetrically located around the median voter; on the contrary we assumed that $|X_R| \neq |-X_L|$ in general. And lastly, note that this equilibrium level of valence is insensitive to the candidates' ideologies: the candidates' ideal points X_L and X_R do not influence the equilibrium; only the marginal cost c does. As we will see in the subsequent section, these last two features do not hold when the candidates' utility functions are strictly concave. We can therefore conclude that it is the *linearity* of candidates' preferences that leads them to adopt identical valences.

This equilibrium has consequences down the road, in the subsequent stages of the election, as described in the following corollary. We denote by x_L^{**} and x_R^{**} the equilibrium platforms adopted by L and R when their choice of valence is endogenous, and by x^{**} the resulting winning platform.

Corollary 2 *In this election, the equilibrium platforms of candidates and the equilibrium strategy of the median voter are for candidate L to adopt $x_L^{**} = 0$, for candidate R to adopt $x_R^{**} = 0$, and for the median voter to choose between L and R with equal probability. The winning platform is $x^{**} = 0$.*

Note that the winning platform is exactly the ideal point of the median voter.

A.2 Proofs of all the results

A.2.1 Theorem 1

Proof. First we need to derive how the median voter M chooses who to vote for. If $U_M(x_L, v_L) < U_M(x_R, v_R)$ then M will vote for R . If $U_M(x_R, v_R) < U_M(x_L, v_L)$ then M will vote for L . If $U_M(x_R, v_R) = U_M(x_L, v_L)$ then M will vote according to the following indifference assumptions: if candidates have different valence then M will vote for the highest-valence one; if candidates have the same valence then M will randomize equally between the two.

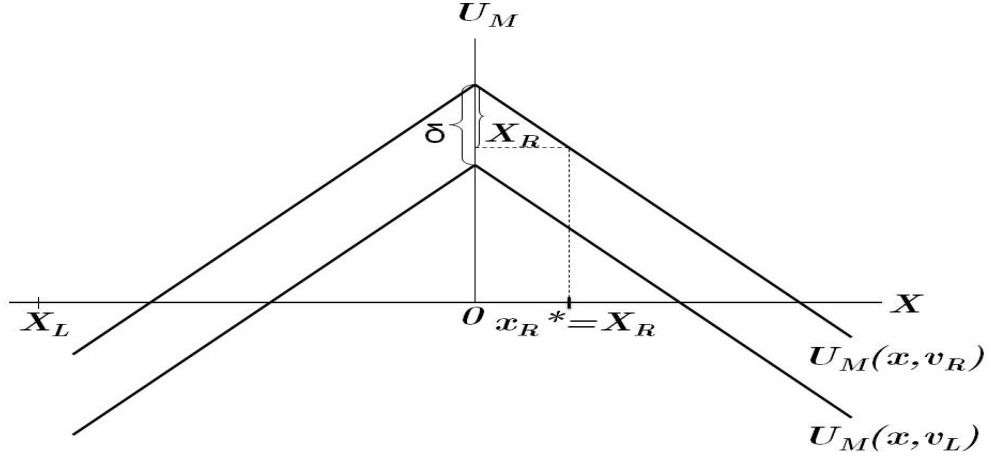
Now we need to derive how the candidates will choose their platforms, given this behavior of the median voter. We will need to divide δ in five possible intervals listed below. The following calculations assume that $f(\cdot)$ is linear, so that M 's utility function is given by $U_M(x, v) = -|x| + v$.

- Case $X_R \leq \delta$:

First note that in this case the valence advantage of candidate R is so large that R could propose her ideal point and still win the election, irrespective of the policy announced by candidate L . That is because $X_R \leq \delta \Rightarrow v_L \leq -X_R + v_R \Rightarrow U_M(0, v_L) \leq U_M(X_R, v_R) \Rightarrow \max_{x_L} U_M(x_L, v_L) \leq U_M(X_R, v_R)$ which implies that M will vote for R for any platform x_L (these equations hold with equality because $\delta > 0$ and due to the indifference assumptions we made). So if R announced $x_R = X_R$, L could announce any platform $x_L \in \mathbb{R}$ and still lose the election; and thus neither of the two candidates would have an incentive to deviate. Therefore those are Nash equilibria. This NE is depicted in the figure below.

Note that no other platform x_R can be sustained in equilibrium because R can always deviate to X_R and obtain her maximum payoff. There is one other way for candidate R to obtain her maximum payoff, which is for candidate L to adopt $x_L = X_R$, and for R to adopt a very extreme

platform that would allow L to win; but in that case L could benefit from deviating closer to her ideal point, so this cannot be a NE.



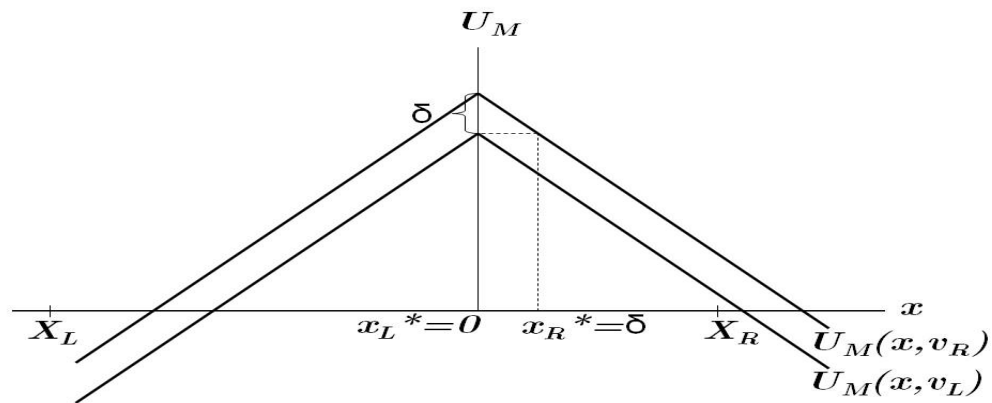
- Case $0 < \delta < X_R$:

Let us study all the possible locations of x_L to see which ones can be sustained in a NE. If x_L is too extreme, namely $|x_L| \geq X_R - \delta$, then R can propose $x_R = X_R$ and win the election. That is because $|x_L| \geq X_R - \delta \Rightarrow -|x_L| + v_L \leq -X_R + v_R \Rightarrow U_M(x_L, v_L) \leq U_M(X_R, v_R)$. But then L could deviate unilaterally to a more moderate platform and win the election, and therefore this cannot be an equilibrium.

On the other hand, if x_L is moderate enough, namely $|x_L| < X_R - \delta$, then R 's best response is to propose the rightmost platform that allows it to win the election, which is $x_R = |x_L| + \delta$. That is because $U_M(x_L, v_L) = U_M(|x_L| + \delta, v_R)$, so R could adopt $x_R = |x_L| + \delta$ and win the election (due to our indifference assumptions and the fact that $\delta > 0$). Note the crucial role that the indifference assumptions play in this last statement: without that assumption there would not exist an optimal x_R for R . Our results are robust to eliminating that assumption, however: An equilibrium would not exist but the optimal behavior of R would be nearly identical to the one described in the theorem – namely to increase x_R ever so closely to $|x_L| + \delta$ without reaching it.

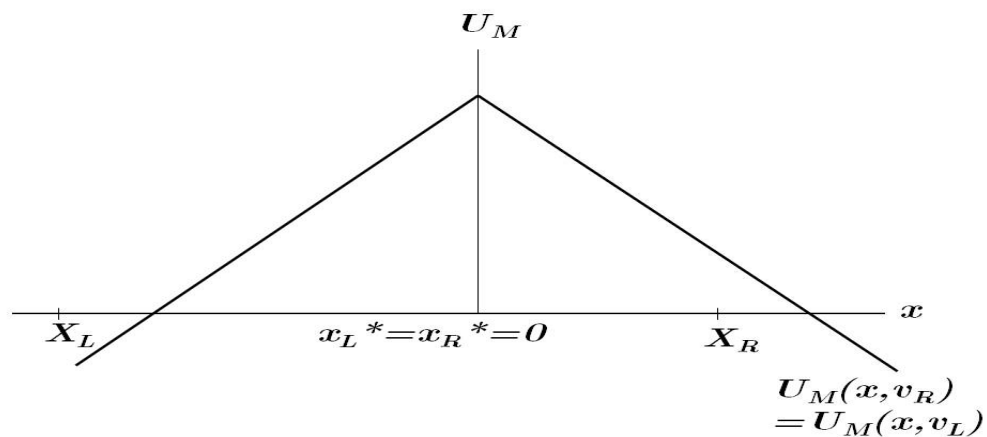
If $|x_L| > 0$ and $x_R = |x_L| + \delta$ then R is best responding to x_L , but L could adopt a more centrist platform and win the election, so this cannot be an equilibrium. Only if $x_L = 0$ and $x_R = \delta$ we have R best-responding to x_L and L best-responding to x_R , and that is because L cannot adopt a

more centrist platform. This is therefore the only NE, as illustrated in the figure below.



- Case $\delta = 0$:

In this case none of the candidates has a valence advantage. M will therefore vote for the candidate whose platform is closest to zero, or will randomize equally between the two if both platforms are equidistant from zero (as established by the indifference assumptions). It is well known in this setting that the unique Nash equilibrium is for both candidates to converge to the median voter's ideal point (see for example Calvert (1985)). Such an equilibrium is depicted in the figure below, where $x_L^* = x_R^* = 0$.



- Case $X_L < \delta < 0$:

This is the mirror image of the case $0 < \delta < X_R$. Hence the unique NE is for candidate R to converge to the median voter's ideal point, $x_R^* = 0$, and for candidate L to adopt the leftmost platform that can win the election, i.e. $x_L^* = \delta$.

- Case $\delta \leq X_L$:

This is the mirror image of the case $X_R \leq \delta$. Hence candidate L is able to adopt her ideal point, $x_L^* = X_L$, and win the election irrespective of the platform adopted by candidate R , $x_R^* \in \mathbb{R}$. ■

A.2.2 Theorem 2

This (long) proof has two steps. We must first state and prove a lemma with the exact values of the best-response functions of candidates, v_R^* and v_L^* . As a second step, we use that lemma to prove the theorem.

Lemma 6 *The best-response functions of R and L , respectively, are*

$$v_R^* = \begin{cases} \frac{1}{c+1} (v_L + X_R) & \text{if } v_L \in [0, \widehat{v}_L] \\ 0 & \text{if } v_L \in (\widehat{v}_L, +\infty) \end{cases}$$

$$\text{with } \widehat{v}_L \equiv -X_R + (X_R - X_L) \sqrt{\frac{c+1}{c}}$$

and

$$v_L^* = \begin{cases} \frac{1}{c+1} (v_R - X_L) & \text{if } v_R \in [0, \widehat{v}_R] \\ 0 & \text{if } v_R \in (\widehat{v}_R, +\infty) \end{cases}$$

$$\text{with } \widehat{v}_R \equiv X_L + (X_R - X_L) \sqrt{\frac{c+1}{c}}$$

Proof of Lemma 5. We derive only v_R^* given that the derivation of v_L^* follows the same steps. We start by noting from Theorem 1 that

$$x^* = \begin{cases} X_R & \text{if } X_R \leq \delta \\ \delta & \text{if } X_L \leq \delta \leq X_R \\ X_L & \text{if } \delta \leq X_L \end{cases}$$

We now separate all the possible values of v_L in a few intervals. First consider $v_L \in [0, -X_L)$, in which case $X_L < \delta$ for any possible value of v_R . What is the optimal v_R that R could adopt? R has the option of choosing a $v_R \in [0, X_R + v_L]$ such that $X_L \leq \delta \leq X_R$; or a $v_R \in [X_R + v_L, +\infty)$ such that $X_R \leq \delta$. If R was to adopt a v_R in the second interval, the optimum would clearly be $v_R = X_R + v_L$ given that all other values would give the same winning platform at a higher cost. If R was to adopt a v_R in the first interval, the optimum would be found by solving the problem

$$\begin{aligned} \max_{v_R} U_R(v_R - v_L, v_R) - C(v_R) &= -|X_R - (v_R - v_L)|^2 - cv_R^2 \\ \text{s.t. } v_R &\in [0, X_R + v_L) \end{aligned}$$

The first-order conditions give us the critical point $v_R = \frac{v_L + X_R}{c+1}$. Note that $\frac{v_L + X_R}{c+1} > 0$ and also $\frac{v_L + X_R}{c+1} < v_L + X_R$ (given our assumption that $c > 0$). Therefore $\frac{v_L + X_R}{c+1}$ satisfies the constraint and is thus an interior optimum (a quick look at the second order conditions reveals that it is indeed a maximum). So R compares the optima from each interval, $v_R = X_R + v_L$ and $v_R = \frac{v_L + X_R}{c+1}$; but given that $X_R + v_L$ belongs to both intervals, and yet is not the optimum in the first interval, it cannot be the global optimum. Therefore $\frac{v_L + X_R}{c+1}$ yields a higher utility in the union of both intervals and is the best response.

Now we consider $v_L \in [-X_L, +\infty)$. What is the optimal v_R that R could adopt? R has the option of choosing a $v_R \in [0, X_L + v_L]$ such that $\delta \leq X_L$; or a $v_R \in [X_L + v_L, X_R + v_L]$ such that $X_L \leq \delta \leq X_R$; or a $v_R \in [X_R + v_L, +\infty)$ such that $X_R \leq \delta$. If R was to adopt a v_R in the first interval, the optimal would clearly be $v_R = 0$ given that all other values would give the same winning platform at a higher cost. If R was to adopt a v_R in the third interval, the optimal would clearly be $v_R = X_R + v_L$ given that all other values would give the same winning platform at a higher cost. If R was to adopt a v_R in the second interval, the optimal would be found by solving the problem

$$\begin{aligned} \max_{v_R} U_R(v_R - v_L, v_R) - C(v_R) &= -|X_R - (v_R - v_L)|^2 - cv_R^2 \\ \text{s.t. } v_R &\in [X_L + v_L, X_R + v_L] \end{aligned}$$

As we saw above, The first-order conditions give us the maximum $v_R = \frac{v_L + X_R}{c+1}$. So the question is whether $\frac{v_L + X_R}{c+1}$ satisfies the constraint or not. We already saw that $\frac{v_L + X_R}{c+1} < v_L + X_R$, and we

will have $X_L + v_L \leq \frac{v_L + X_R}{c+1}$ if and only if $v_L \leq \frac{X_R - (c+1)X_L}{c}$. So we must further separate v_L in two intervals. If $v_L \in [-X_L, \frac{X_R - (c+1)X_L}{c}]$ then $\frac{v_L + X_R}{c+1}$ satisfies the constraint and is the optimal in the interval $[X_L + v_L, X_R + v_L]$. Then the optima from the three intervals that R compares are $v_R = 0$, $v_R = \frac{v_L + X_R}{c+1}$ and $v_R = X_R + v_L$; but we can ignore $X_R + v_L$ for the same argument as above; and R 's utility from $\frac{v_L + X_R}{c+1}$ will be higher-than-or-equal-to the utility from 0 if and only if $U_R(X_L, 0) \leq U_R\left(\frac{v_L + X_R}{c+1} - v_L, \frac{v_L + X_R}{c+1}\right)$. With some algebra we can see that this inequality is true if and only if $v_L \leq \widehat{v}_L$ with $\widehat{v}_L \equiv -X_R + (X_R - X_L)\sqrt{\frac{c+1}{c}}$. (Note that we can prove that $-X_L < \widehat{v}_L < \frac{X_R - (c+1)X_L}{c}$ using the fact that $c > 0$.) So for $v_L \in [-X_L, \widehat{v}_L]$, R 's best response is $v_R = \frac{v_L + X_R}{c+1}$, and for $v_L \in (\widehat{v}_L, \frac{X_R - (c+1)X_L}{c}]$, R 's best response is $v_R = 0$.

Finally we consider $v_L \in (\frac{X_R - (c+1)X_L}{c}, +\infty)$. Then $\frac{v_L + X_R}{c+1}$ does not satisfy the constraint in the interval $v_R \in [X_L + v_L, X_R + v_L]$ and the optimal is $X_L + v_L$. So the optima from the three intervals that R compares are $v_R = 0$, $v_R = X_L + v_L$ and $v_R = X_R + v_L$; but we can ignore $v_R = X_R + v_L$ for the same argument as above; and given that both $v_R = 0$ and $v_R = X_L + v_L$ give the same winning platform X_L , the optimum must be the smallest one, 0. Therefore $v_R = 0$ yields a higher utility in the union of all three intervals and is the best response.

In summary we have that R 's best response for $v_L \in [0, -X_L) \cup [-X_L, \widehat{v}_L]$ is $v_R = \frac{v_L + X_R}{c+1}$; whereas R 's best response for $v_L \in (\widehat{v}_L, \frac{X_R - (c+1)X_L}{c}] \cup (\frac{X_R - (c+1)X_L}{c}, +\infty)$ is $v_R = 0$, which is exactly what the lemma claims. The derivation of v_L^* , L 's best response as a function of v_R , follows the same logic. ■

We can now use the expressions for v_L^* and v_R^* in the previous lemma, to prove the theorem.

Proof of Theorem 2. We start by stating the following remark which can be proved using the fact that $c > 0$.

Remark 1 $-X_L < \widehat{v}_L$ and $X_R < \widehat{v}_R$

Now we note that the candidates for a pure-strategy Nash equilibrium (v_L^{**}, v_R^{**}) can only be of the following four types:

1. $v_L \in (\widehat{v}_L, +\infty)$ and $v_R \in (\widehat{v}_R, +\infty)$
2. $v_L \in (\widehat{v}_L, +\infty)$ and $v_R \in [0, \widehat{v}_R]$
3. $v_L \in [0, \widehat{v}_L]$ and $v_R \in (\widehat{v}_R, +\infty)$

4. $v_L \in [0, \widehat{v}_L]$ and $v_R \in [0, \widehat{v}_R]$

Looking at the best-response functions of R and L as given in Lemma 3 we can rule out the first type because R would prefer to deviate from the interval $v_R \in (\widehat{v}_R, +\infty)$ to $v_R = 0$. The second type can also be ruled out because R and L are not best-responding to one another: if $v_L \in (\widehat{v}_L, +\infty)$ then R 's best-response is $v_R = 0$, but L 's best-response to $v_R = 0$ is $v_L = \frac{-X_L}{c+1}$ which is strictly smaller than \widehat{v}_L by Remark 1, and therefore does not belong to $(\widehat{v}_L, +\infty)$. The third type can be ruled out with a similar logic.

For the fourth type to yield an equilibrium there must be a solution to the simultaneous equations: $v_L = \frac{1}{c+1}(v_R - X_L)$ and $v_R = \frac{1}{c+1}(v_L + X_R)$. The unique solution to these equations is $v_L^{**} \equiv \frac{X_R - (c+1)X_L}{c(c+2)}$ and $v_R^{**} \equiv \frac{(c+1)X_R - X_L}{c(c+2)}$. For that to be an equilibrium it must be that $v_L^{**} \in [0, \widehat{v}_L]$ and $v_R^{**} \in [0, \widehat{v}_R]$. We can clearly see that $v_L^{**} > 0$ and $v_R^{**} > 0$. To prove that $v_L^{**} < \widehat{v}_L$ and $v_R^{**} < \widehat{v}_R$ we state the following remark which can be proved using straightforward algebra:

Remark 2 If $c > \max \left\{ \frac{X_L + \sqrt{5X_L^2 - 4X_L X_R}}{-2X_L}, \frac{-X_R + \sqrt{5X_R^2 - 4X_L X_R}}{2X_R} \right\}$ then $v_L^{**} < -X_L$ and $v_R^{**} < X_R$.

Remarks 1 and 2, along with our assumption that $c > \max \left\{ \frac{X_L + \sqrt{5X_L^2 - 4X_L X_R}}{-2X_L}, \frac{-X_R + \sqrt{5X_R^2 - 4X_L X_R}}{2X_R} \right\}$, imply that $v_L^{**} < -X_L < \widehat{v}_L$ and $v_R^{**} < X_R < \widehat{v}_R$. This implies that v_R^{**} and v_L^{**} belong to the appropriate intervals. Therefore as long as c satisfies the constraint in Remark 2, v_R^{**} and v_L^{**} are a Nash equilibrium. In sum, v_R^{**} and v_L^{**} are the unique pure-strategy Nash equilibrium. ■

A.2.3 Corollary 1

Proof. We will only prove that $|X_L| < |X_R| \Rightarrow v_L^{**} < v_R^{**}$. The proofs for the cases where $|X_L| > |X_R|$ and $|X_L| = |X_R|$ follow the same steps.

Recalling that $X_L < 0 < X_R$ and $c > 0$, note that $|X_L| < |X_R| \Rightarrow -X_L < X_R \Rightarrow -cX_L < cX_R \Rightarrow \frac{1}{c(c+2)}(X_R - (c+1)X_L) < \frac{1}{c(c+2)}((c+1)X_R - X_L) \Rightarrow v_L^{**} < v_R^{**}$ ■

A.2.4 Theorem 3

Proof. Let us define $\delta^{**} \equiv v_R^{**} - v_L^{**}$. From Theorem 2 we can calculate that $\delta^{**} = \frac{1}{c+2}(X_L + X_R)$.

We then note the following straightforward remarks

Remark 3 If $|X_L| < (>)(=) |X_R|$ then $\delta^{**} > (<)(=) 0$

Remark 4 $X_L < \delta^{**} < X_R$

The results then come directly from looking at the corresponding outcomes for δ^{**} in the table of Theorem 1. ■

A.2.5 Theorem 4

Proof. This follows immediately from the winning platforms indicated in Theorem 3. ■

A.2.6 Lemma 1

Proof. We first prove that $x^{**} < X_R$. The following remark follows from $X_L < 0$.

Remark 5 $X_L + X_R < X_R$

The following remark follows from $c > 0$.

Remark 6 $\frac{1}{c+2} < 1$

From these two remarks it is evident that $\frac{1}{c+2}(X_L + X_R) < X_R$. With similar steps we can prove that $X_L < \frac{1}{c+2}(X_L + X_R)$, which completes the proof. ■

A.2.7 Lemma 2

Proof. This follows directly from the expression of x^{**} given in Theorem 4, and noting that if $|X_L| < (>)(=) |X_R|$ then $(X_L + X_R) > (<)(=) 0$. ■

A.2.8 Lemma 3

Proof. Straightforward differentiation of x^{**} shows that $\frac{\partial x^{**}}{\partial X_R} > 0$, $\frac{\partial x^{**}}{\partial X_L} > 0$ and $\frac{\partial |x^{**}|}{\partial c} < 0$, which is what the theorem claims. ■

A.2.9 Lemma 4

Proof. From Theorem 2 we can calculate that $|v_R^{**} - v_L^{**}| = \frac{1}{c+2} |X_L + X_R|$. From Theorem 3 we can calculate, for each of the three cases, that $|x_R^{**} - x_L^{**}| = \frac{1}{c+2} |X_L + X_R|$. ■

Theorem 5 Proof. We study the case where $|X_L| < |X_R|$, meaning that L is the moderate candidate and R is the extremist candidate. In that case $\frac{1}{c+2}(X_L + X_R)$ is strictly positive, so platform polarization and valence polarization are strictly positive. Straightforward differentiation shows that $\frac{\partial}{\partial c} \frac{1}{c+2}(X_L + X_R) < 0$, $\frac{\partial}{\partial X_R} \frac{1}{c+2}(X_L + X_R) > 0$, $\frac{\partial}{\partial X_L} \frac{1}{c+2}(X_L + X_R) > 0$. The case where $|X_R| < |X_L|$ is studied in the same way.

If $|X_L| = |X_R|$, implying that $-X_L = X_R$, we have that $\frac{1}{c+2}|X_L + X_R| = 0$ and polarization is null. ■

A.2.10 Lemma 5

Proof. We derive only v_R^* given that the derivation of v_L^* follows the same steps. We start by defining a variable that will play an important role. We call \bar{v} the value such that $C'(\bar{v}) = 1$. It can readily be calculated that $\bar{v} = \frac{1}{2c}$.

Now note from Theorem 1 that

$$x^* = \begin{cases} X_R & \text{if } X_R \leq \delta \\ \delta & \text{if } X_L \leq \delta \leq X_R \\ X_L & \text{if } \delta \leq X_L \end{cases}$$

We now separate all the possible values of v_L in a few intervals. First consider $v_L \in [0, -X_L)$, in which case $X_L < \delta$ for any possible value of v_R . What is the optimal v_R that R could adopt? R has the option of choosing a $v_R \in [0, X_R + v_L]$ such that $X_L \leq \delta \leq X_R$; or a $v_R \in [X_R + v_L, +\infty)$ such that $X_R \leq \delta$. If R was to adopt a v_R in the second interval, the optimum would clearly be $v_R = X_R + v_L$ given that all other values would give the same winning platform at a higher cost. If R was to adopt a v_R in the first interval, the optimum would be found by solving the problem

$$\begin{aligned} \max_{v_R} U_R(v_R - v_L, v_R) - C(v_R) &= -|X_R - (v_R - v_L)| - cv_R^2 \\ \text{s.t. } v_R &\in [0, X_R + v_L) \end{aligned}$$

The first-order conditions give us the critical point $\bar{v} = \frac{1}{2c}$. Note that $\bar{v} > 0$ and also $\bar{v} < X_R$ (given our assumption that $c > \frac{1}{2X_R}$). Therefore \bar{v} satisfies the constraint and is thus an interior optimum (a quick look at the second order conditions reveals that it is indeed a maximum). So R compares the optima from each interval, $v_R = X_R + v_L$ and $v_R = \bar{v}$; but given that $X_R + v_L$ belongs to

both intervals, and yet is not the optimum in the first interval, it cannot be the global optimum. Therefore \bar{v} yields a higher utility in the union of both intervals and is the best response.

Now we consider $v_L \in [-X_L, +\infty)$. What is the optimal v_R that R could adopt? R has the option of choosing a $v_R \in [0, X_L + v_L]$ such that $\delta \leq X_L$; or a $v_R \in [X_L + v_L, X_R + v_L]$ such that $X_L \leq \delta \leq X_R$; or a $v_R \in [X_R + v_L, +\infty)$ such that $X_R \leq \delta$. If R was to adopt a v_R in the first interval, the optimal would clearly be $v_R = 0$ given that all other values would give the same winning platform at a higher cost. If R was to adopt a v_R in the third interval, the optimal would clearly be $v_R = X_R + v_L$ given that all other values would give the same winning platform at a higher cost. If R was to adopt a v_R in the second interval, the optimal would be found by solving the problem

$$\begin{aligned} \max_{v_R} U_R(v_R - v_L, v_R) - C(v_R) &= -|X_R - (v_R - v_L)| - cv_R^2 \\ \text{s.t. } v_R &\in [X_L + v_L, X_R + v_L] \end{aligned}$$

As we saw above, The first-order conditions give us the maximum $\bar{v} = \frac{1}{2c}$. So the question is whether \bar{v} satisfies the constraint or not. We already saw that $\bar{v} < X_R + v_L$, and we will have $X_L + v_L \leq \bar{v}$ if and only if $v_L \leq -X_L + \frac{1}{2c}$. So we must further separate v_L in two intervals. If $v_L \in [-X_L, -X_L + \frac{1}{2c}]$ then \bar{v} satisfies the constraint and is the optimal in the interval $[X_L + v_L, X_R + v_L]$. Then the optima from the three intervals that R compares are $v_R = 0$, $v_R = \bar{v}$ and $v_R = X_R + v_L$; but we can ignore $X_R + v_L$ for the same argument as above; and R 's utility from \bar{v} will be higher-than-or-equal-to the utility from 0 if and only if $U_R(X_L, 0) \leq U_R(\bar{v} - v_L, \bar{v})$, which implies $v_L \leq -X_L + \frac{1}{4c}$. (Note that we can prove that $-X_L < -X_L + \frac{1}{4c} < -X_L + \bar{v}$ using the fact that $C'(\bar{v}) = 1$ and that $C(\bar{v}) > 0$.) So for $v_L \in [-X_L, -X_L + \frac{1}{4c}]$, R 's best response is $v_R = \bar{v}$, and for $v_L \in (-X_L + \frac{1}{4c}, -X_L + \bar{v}]$, R 's best response is $v_R = 0$.

If $v_L \in (-X_L + \bar{v}, +\infty)$ then \bar{v} does not satisfy the constraint in the interval $v_R \in [X_L + v_L, X_R + v_L]$ and the optimal is $X_L + v_L$. Then the optima from the three intervals that R compares are $v_R = 0$, $v_R = X_L + v_L$ and $v_R = X_R + v_L$; but we can ignore $v_R = X_R + v_L$ for the same argument as above; and given that both $v_R = 0$ and $v_R = X_L + v_L$ give the same winning platform X_L , the optimum must be the smallest one, 0. Therefore $v_R = 0$ yields a higher utility in the union of all three intervals and is the best response.

In summary we have that R 's best response for $v_L \in [0, -X_L) \cup [-X_L, -X_L + \frac{1}{4c}]$ is $v_R = \bar{v}$; whereas R 's best response for $v_L \in (-X_L + \frac{1}{4c}, -X_L + \bar{v}] \cup (-X_L + \bar{v}, +\infty)$ is $v_R = 0$, which is exactly what the Lemma claims. The derivation of v_L^* , L 's best response as a function of v_R , follows the same logic. ■

A.2.11 Theorem 6

Proof. According to Lemma 2, the best-response functions of L and R have only two values each, so in essence the candidates are playing a two-by-two game. The only candidates for a pure-strategy Nash equilibrium (v_L^{**}, v_R^{**}) are $(0, 0)$, $(0, \bar{v})$, $(\bar{v}, 0)$, and (\bar{v}, \bar{v}) . Given that $v_L(0) = \bar{v}$ and $v_R(0) = \bar{v}$, the pair of valences $(0, 0)$ cannot be a NE. Given that $\bar{v} < X_R$ (given our assumption that $c > \frac{1}{2X_R}$) we have that $v_L(\bar{v}) = \bar{v}$, so $(0, \bar{v})$ cannot be an equilibrium. By a similar argument (given our assumption that $c > \frac{1}{-2X_L}$), we have that $\bar{v} < -X_L$, so $v_R(\bar{v}) = \bar{v}$ and $(\bar{v}, 0)$ cannot be an equilibrium. But given that $v_L(\bar{v}) = \bar{v}$ and $v_R(\bar{v}) = \bar{v}$, the pair of valences (\bar{v}, \bar{v}) is indeed a NE. Given that there is a unique equilibrium in pure strategies, we know that an equilibrium in mixed strategies does not exist, and (\bar{v}, \bar{v}) is indeed the unique NE. ■

A.2.12 Corollary 2

Proof. Define δ^{**} as the equilibrium valence difference between candidates, in other words $\delta^{**} \equiv v_R^{**} - v_L^{**}$. The results come directly from observing that $v_L^{**} = v_R^{**}$ which implies that $\delta^{**} = 0$, and looking at the corresponding outcomes in the table of Theorem 1. ■