On the Benefits of Party Competition*

Dan Bernhardt
Department of Economics
University of Illinois
Urbana Champaign

Larissa Campuzano
Department of Economics
University of Rochester

Francesco Squintani
Department of Economics
University College, London

June, 2005

Abstract

We study the role of parties in a citizen-candidate repeated-elections model where voters have incomplete information. We identify a novel "party competition effect." Compared with "at large" selection of candidates, party selection makes office-holders more willing to avoid extreme ideological stands. Politicians follow party discipline, even in absence of a party-controlled reward mechanism. Voters of all ideologies benefit from the party-competition effect, which thus provides a novel rationale for political parties. When politicians have an (imperfect) informational advantage over voters, we additionally find a "party screening effect." Parties select moderate candidates, because they anticipate that their candidate's ideological record can be verifiably disclosed through campaigning. Under reasonable functional assumptions, all voters benefit from party screening.

*We thank the audiences of seminars at the Wallis Institute, Rochester University, at Southampton University, of the 2004 Workshop in Political Economy at Stony Brook, the 2004 Political Economy Workshop at Vienna, the 2005 Society for Advancement of Economic Theory Conference in Vigo, and especially Jim Snyder, for their comments. Nate Anthony helped significantly in this research.

1 Introduction

This paper studies the role of party competition in a citizen-candidate repeated-elections model where voters have limited information about candidates ideologies, and where incumbents' past actions help voters predict their future choices if re-elected. We contrast party selection of candidates with at-large selection in absence of parties. For politicians, we ask: Would politicians be more representative of their constituency with party or at-large selection? Would they be more or less willing to take ideological stands? For voters, we ask: Would party selection induce voters to set more demanding standards of representation on incumbents for them to win re-election? Would more incumbents lose with party or at-large selection of challengers? Would voters be better off with party selection than with at-large selection? For parties we ask: Can parties play an active role in screening candidates? What type of candidates would they select? Is such screening beneficial to voters?

We first abstract from the role of parties in screening their candidates, and derive general predictions about how the voters' choices and the office-holders' choices are affected by parties. We identify a novel "party competition effect": Party competition makes office-holders more reluctant to take extreme ideological stands, and induces voters to select more demanding standards for re-election. Second, we refine our predictions when politicians have an (imperfect) informational advantage over voters, identifying a "party screening effect": parties select moderate candidates. Third, we derive the implications for welfare. We find that the 'party-competition' effect is beneficial to voters of all ideologies. Under reasonable functional assumptions, we find that also party screening is beneficial to all voters. Hence, our analysis provides a new rationale for political parties.

Our model is as follows. Rational, forward-looking citizens, be they voters, electoral candidates or elected representatives, care about the valence and the policy choices of the representatives in office. Elected individuals may also receive 'ego rents' from being in office. Each period, an incumbent faces an electoral challenge from an opposing candidate. Voters' have limited information about the challenger's ideology. In contrast, voters can use the incumbent's past actions in office to make inferences about her ideology, and hence, her actions if re-elected.[1] We maintain the fundamental tenet of citizen-candidate models that candidates for office cannot make credible promises. We contrast outcomes when the challenger is drawn at random from the entire spectrum of ideologies "at large", with those that obtain when the challenger is drawn from a party representing the wing of the ideological space "opposite" from the incumbent.

Given standard assumptions the median voter is decisive and sets a simple retrospective voting rule: vote for the incumbent candidate if and only if the incumbent's policy is sufficiently moderate. An incumbent with sufficiently moderate ideology can represent his own ideology and win re-election. All other winning incumbents take positions as close to their ideologies as they possibly can and still be re-elected. Voters use this information to update their beliefs about an incumbent's ideology: The location of the marginal re-elected incumbent leaves the median voter indifferent between re-electing the incumbent and selecting an untried challenger.

The 'party competition effect' materializes because an incumbent fears being replaced by a challenger selected by the opposing party by more than he fears a challenger from at large. This reflects that the likely ideology and location of a challenger selected by the opposing party is further from the incumbent's own ideology. Thus, ceteris paribus, with party selection, an incumbent is more willing to adopt positions closer to the median voter's preferred position in order to win re-election. At the same time, the median voter sets more demanding standards for re-election with party selection because the random value of an untried politician becomes larger. Our "party competition effect" provides a fully endogenous theory of party discipline: office holders are willing to follow party lines, even though there is no party-controlled reward mechanism. Incumbents avoid extreme ideological stands and take the responsibility not to lose the elections. Because of this self-induced "party discipline", in equilibrium, incumbents are more likely to be re-elected with party selection than with at-large selection. That is, with party selection, there is less turnover, even though voters select more demanding standards for re-election.

[1] Most models of candidate location ignore differences between incumbents and challengers. Yet, we know that these differences must be important, if only because, in practice, incumbents often win re-election. Our model highlights one key difference between incumbents and challengers: Voters typically know much more about incumbents since they can observe their performances in office. Challengers, in contrast, are usually untried in the office for which they are running.

Our welfare analysis finds that the "party competition effect" benefits all voters. This is both because with party selection office-holders are forced to adopt positions closer to the median policy in order to win re-election; and, at the same time, candidates value re-election more and are more willing to compromise their ideologies by adopting more moderate positions. The median voter values this greater willingness to compromise. Ex-ante, all other voters benefit even more by party selection, because parties reduce both the per-period variance and the variability over time of the policies adopted by elected representatives.

We then extend the analysis to consider the possibility that politicians (e.g., candidates and parties) have an (imperfect) informational advantage over voters. We find that parties select moderate candidates. This 'party screening effect' reinforces the 'party competition effect'; it does not substitute for it. For simplicity, we assume that information is coarse: politicians can only assess and reveal whether a candidate is moderate or extremist. Such information about a candidate's ideology can be verifiably disclosed through campaigning, reflecting that it may be revealed through past public stands on policy issues. In equilibrium, parties anticipate that if they select extremist challengers, the incumbent will successfully campaign against them and win re-election.Under the natural model specification that ideologies are uniformly distributed and loss functions are homogenous, all voters benefit from party screening, that identifies moderate candidates. However, not all voters may benefit from the "screening effect".

Before proceeding with the formal analysis, we would like to emphasize that our model of parties is parsimonious. In our model, parties do not pool funds or financial resources. They do not exercise party discipline, nor dictate party lines. Parties have no control whatsoever on elected candidates' policy choices. There is no partisanship: citizens do not care about party identity per se, but only about the policy adopted by elected representatives and about their valence. We obtain the "party competition effect" by imputing to parties only the ability to aggregate citizens with like-minded political views. The "screening effect" only requires that politicians have some insider information on potential representatives, and that parties can endorse a candidate. It is reasonable to expect that with a broader theory of parties, that enriches the set of abilities and functions of parties in the political arena, one would uncover even broader roles for parties.

This paper is presented as follows. After the literature review, section 3 analyzes our base citizen-candidate repeated election model. Section 4 enriches the model to allow politicians to have better information on candidates' ideology than voters. Section 5 concludes. As is customary, most of the formal proofs are in the appendix.

2 Literature Review

Our analysis builds on an idea, which dates back at least to Downs (1957), that party labels provide voters with information about candidates. There is substantial evidence that voters learn about the policy positions of candidates from party labels. Fiorina (1981) and Lindbeck and Weibull (1993) are among the first to point out that party identification affects voting behavior. Snyder and Ting (2002) show that party dummies explain a significant part of the variation in the voters' placements of candidates on a left-right scale. Alvarez (1997) and Bartels (1986) show that voters are risk averse over the policy locations of candidates, so that the informational role of party labels is beneficial to the electorate. Cox and McCubbins (1993) highlight the role played by the legislative organization of the party in sustaining the electoral value of party labels. Snyder and Ting (2002) examine the relation between platform choices and the informational power of party labels. Ashworth and Bueno de Mesquita (2004) provide a formal account where party discipline, candidate affiliation, and ideological homogeneity are all determined endogenously within a strategic electoral-legislative setting. The informational role of parties is highlighted in our analysis by the 'screening effect'.In contrast to the papers cited, , we adopt a more minimalist approach to parties—our parties do not dictate party lines. Our 'party competition' effect is driven by the possibility of alternance in office of politicians from opposing parties. The role played by the information in party labels is minimal.

Another line of research related to our work examines how different mechanisms can be used to enforce party discipline. Huber (1996) and Diermeier and Feddersen (1998) demonstrate that party leaders can employ the vote of no confidence procedure to compel party members to follow the party line. Similarly, Calvert and Fox (2000) argue that agenda control can serve as a punishment mechanism in a repeated game to enforce intra-party policy compromise. In contrast in those papers, we have a fully endogenous theory of party discipline. The mere existence of a competing parties makes elected representative willing to follow party lines. This effect does not require the existence of any party-enforced mechanism to ensure adherence to party discipline.

Methodologically, our paper is most closely related to Duggan (2000) and to Bernhardt et al. (2004). Duggan (2000) develops the basic theory of repeated citizen-candidate election with incomplete information about the candidates' policy preferences. Bernhardt et al. (2004) study related issues when politicians face term limits, and more senior politicians can obtain "pork transfers" for their districts from districts with less senior politicians.[2] Banks and Duggan (1999) extend Duggan's (2000) analysis to allow for multiple ideological dimensions. Bernhardt and Ingberman (1985) is the first paper to consider the consequences of informational differences between incumbents and challengers. Most of the past literature that focuses on informational differences between incumbents and challengers focuses on legislative ability rather than ideology. Banks and Sundaram (1993) develop a dynamic model in which representatives exert effort to represent their constituency. Over time, voters learn which representatives are lazy, and vote them out, so a smaller fraction of more senior incumbents lose. In a similar vein, Austen-Smith and Banks (1991) and Ferejohn (1986) consider dynamic games in which representatives dislike exerting effort.

We conclude this review by briefly discussing the blossoming literature that formally models elections with endogenous party structures. Morelli (2004) studies a model in which parties help voters coordinate in the election and allow candidates to commit to policies. He finds conditions under which there are more parties under proportional representation than under plurality rule. Levy (2002) also focuses on the role of parties in insuring the credibility of policy commitments. Osborne and Tourky (2003) study increasing returns to party size, and their implications for the number of parties. Persson, Roland and Tabellini (2003) study a model in which electoral institutions endogenously determine party fragmentation. Unlike these papers, we do not ask ourselves how parties are formed. Rather, we take the existence of two parties as given, and show that all voters benefit when the pool of politicians partitions in two parties of opposite ideology.

[2] Relatedly, Reed (1994) considers an example featuring moral hazard and adverse selection in the provision of a public good where there is a term limit of two terms. He solves for parameter values that determine when voters re-elect incumbents. Banks and Sundaram (1998) and Reed (1989) consider related adverse selection models.

3 The 'Party Competition Effect'

There is an interval $[-a, +a]$ of citizen candidates, each indexed by her private ideology $x \in [-a, +a]$. Ideologies are private information to the candidates. Ideologies are distributed across society according to the c.d.f. F, with an associated single-peaked density f, that is symmetric about the median voter's ideology, $x = 0$. At time 0, an office holder is randomly determined. In any period t, an office holder with ideology x selects a policy $p(x) \equiv y \in [-a, +a]$. The time-$t$ utility of a citizen x depends on the implemented policy y, according to a symmetric, single-peaked loss function $L_x(y) = l(|x - y|)$, where l is $\mathcal{C}^2$, and $l' < 0$, $l'' \leq 0$. We normalize $l(0) = 0$ without loss of generality. Note that L satisfies the single-crossing property: L'_x is increasing in x. Period utilities are discounted with factor $\delta < 1$.[3]

We focus on symmetric, stationary and stage-undominated perfect Bayesian equilibrium (PBE). A stationary policy strategy p prescribes that at any time t, the policy y selected by a office holder depends only on his ideology y, regardless of t and of the history of play at time t. The policy strategy is symmetric if $p(x) = -p(-x)$. If office holders adopt symmetric stationary strategies, stage-undominated PBE voters' strategies are as follows. If the time-t incumbent adopts platform y, then each voter x votes to re-elect the incumbent if and only if $L_x(y) \geq U_x$, where U_x is the equilibrium expected continuation utility for selecting a new office holder at random. In a PBE, the median voter is said to be decisive whenever an office holder who adopts policy y is re-elected if and only if $L_0(y) \geq U_0$, i.e., the incumbent office holder is re-elected if and only if the median voter prefers the incumbent to the challenger.

3.1 At-large selection of challengers

With at-large selection of challengers, at the beginning of any period $t \geq 1$, the incumbent runs election against a challenger drawn at random from $f(\cdot)$. We show in the Appendix that as long as the loss functions do not display too much risk aversion, there is a unique symmetric, stage-

[3] For notational simplicity, we abstract from ego rents for holding office. Our results extend qualitatively if we allow for ego rents.

undominated, stationary perfect Bayesian equilibrium. The equilibrium is completely summarized by thresholds $\{c, w\}$, where $0 < w < c < a$. Candidates with centrist ideology $x \in [0, w]$ and extremist candidates $x \in [c, a]$ adopt their preferred policy $y = x$ when in office. Centrists are re-elected and extremists are ousted from office. Moderate candidates $x \in [w, c]$ do not adopt their preferred policy, as they would then lose office. They compromise and adopt the most extreme ideology that allows them to win re-election, i.e., they locate at w. The characterization is symmetric for $x < 0$.

The equilibrium obeys the following Bellman equations:

$$L_0(w) = U_0(c, w), \tag{1}$$

$$L_c(w) = (1-\delta) L_c(c) + \delta U_c = \delta U_c(c, w). \tag{2}$$

The median voter is decisive: she is indifferent between re-electing a candidate that implements policy w and electing the random challenger. The candidate c is indifferent between implementing policy w forever, or policy c once and then be replaced by a random challenger.

For any citizen x, the PBE continuation expected value for electing the challenger is:

$$\begin{aligned} U_x(w, c) = & \int_{-a}^{-c} (L_x(y)(1-\delta) + \delta U_x) \, dF(y) + \int_{-c}^{-w} L_x(-w) dF(y) \\ & + \int_{-w}^{w} L_x(y) dF(y) + \int_{w}^{c} L_x(w) dF(y) + \int_{c}^{a} (L_x(y)(1-\delta) + \delta U_c) \, dF(y). \end{aligned} \tag{3}$$

Throughout our analysis, we will assume that the parameters characterizing the economy are such that the median voter is decisive, and that candidates with more extreme ideologies are less willing to compromise, so that equilibrium is described by the set $\{c, w\}$. This amounts to assuming that citizens are not too risk averse.Theorem A1 in the appendix provides sufficient conditions for this to hold, extending Theorem 1 in Duggan (2000), which proves the result for linear loss $L_x(y) = -|x-y|$, with $a = 1/2$ and uniform distribution F. When l is strictly concave, the median voter may not be decisive in the stationary symmetric PBE. It cannot be the case that the median voter votes to re-elect an incumbent y and the incumbent y is defeated in the election. However, if l is sufficiently concave, an incumbent y may be re-elected by a coalition of voters with opposite extreme ideologies, despite the contrary vote by the median voter. This is because very risk-averse extremists may be so afraid that an extremist of the opposite side is selected, that they prefer to stay with the incumbent y, the lesser evil. The median voter faces less risk, and sets more demanding standards for re-election. Formally, the condition that may be violated if l is strictly concave is that $L_x(0) - U_x(c, w)$ be weakly decreasing in x, for any w, c. This assumption is not necessary for our results to hold, but it simplifies the analysis. This is because the equilibrium still has the feature that incumbents are retained in office if and only if their adopted policy y belongs to an interval symmetric $[-w, w]$ around the median policy 0.

More important is the fact that if l is sufficiently concave, the office holder's choice of location may not be determined by a cutoff c. An extremist policy holder may dislike the possibility of being replaced by an extremist of the opposite side by so much, that they compromise and adopt the most extreme possible policy in the set $[-w, w]$. Formally, the condition that may violated if l is strictly concave is that $L_x(w) - \delta U_x(x, w)$ be weakly decreasing in x,for any w.

3.2 Party selection of candidates

We contrast outcomes in the repeated election model with at-large selection of candidates—the standard modeling approach—with those that obtain when challenging candidates are chosen by opposing parties, A and B. We initially assume that party A includes all citizen-candidates with ideology $x < 0$, and party B all possible candidates with ideology $x > 0$. In every t election, the incumbent faces a challenger that comes from the opposite party. That is, incumbents are always endorsed by their parties. Equivalently, we could assume that if the party does not endorse "its" incumbent, then voters who are indifferent between untried challengers from the two parties select the candidate chosen by the opposing party. In contrast to at-large selection, the median voter is always decisive; intuitively, a challenger cannot be re-elected by a coalition of extremists, because the extremists that belong to the challenger's party prefer to vote for the challenger.

Again, we prove in the Appendix that, as long as voters are not too risk-averse, the symmetric, stage-undominated, stationary perfect Bayesian equilibrium is characterized by two thresholds: v and k. Centrist citizen $x \in [-v, v]$ and extremist candidates $x \in [-a, -k] \cup [k, a]$ adopt their preferred policy $y = x$ when in office. Moderates are re-elected and extremists are ousted. Moderate candidates $x \in [-k, -v]$, and $x \in [v, k]$, adopt policies $-w$ and w respectively, and are re-elected.

The Bellman equations characterizing the equilibrium are

$$L_0(v) = \overline{U}_0(v,k) = \underline{U}_0(v,k), \tag{4}$$

$$L_k(v) = (1-\delta) L_k(k) + \delta \underline{U}_k = \delta \underline{U}_k(v,k), \tag{5}$$

For each voter k, we need to distinguish the equilibrium expected continuation payoff when the next period's office holder is selected from k's own party, denoted by $\overline{U}_k$, from the continuation payoff when the next period's office holder is selected from the opposite party, $\underline{U}_k$.

$$\underline{U}_x(v,k) = 2\int_{-a}^{-k} \left(L_x(y)(1-\delta) + \delta \overline{U}_x(v,k)\right) dF(y) + 2\int_{-k}^{-v} L_x(-v) dF(y) + 2\int_0^v L_x(y) dF(y) \tag{6}$$

$$\overline{U}_x(v,k) = 2\int_0^v L_x(y) dF(y) + 2\int_v^k L_x(v) dF(y) + 2\int_k^a \left(L_x(y)(1-\delta) + \delta \underline{U}_x(v,k)\right) dF(y). \tag{7}$$

Unlike in the case of at-large, the median voter is always decisive in equilibrium. It cannot be the case the case that an incumbent from (say) party A, is re-elected against the vote of the median citizen by coalition of voters from the opposite extremes of the ideological spectrum. All right-wing extremists, in fact, vote for the challenger, who is drawn from party B.

3.3 Equilibrium and Welfare Comparison

We now show that the introduction of parties makes candidates more willing to compromise. We proceed in separate Lemmata. The first one shows that an office holder's payoff for being replaced by a candidate from the opposing party is smaller than the payoff for being replaced by a candidate randomly selected at large.

Lemma 1 *For any voter x, and any equilibrium thresholds (w, c), the equilibrium payoffs are ranked as follows: $\underline{U}_x(w,c) < U_x(w,c) < \overline{U}_x(w,c)$.*

Proof. Suppose that $x > 0$; the case for $x < 0$ is analogous by symmetry. Subtracting equation (6) from equation (3), and using the symmetry of f, yields:

$$U_x - \underline{U}_x = \int_{-a}^{-c} \delta \left(U_x - \overline{U}_x \right) dF(y) - \int_{-a}^{-c} \left(L_x(y)(1-\delta) + \delta \overline{U}_x \right) dF(y) + \int_0^w L_x(y) dF(y)$$

$$+ \int_c^a \left(L_x(y)(1-\delta) + \delta U_x \right) dF(y) - \int_{-c}^{-w} L_x(-w) dF(y) + \int_w^c L_x(w) dF(y) - \int_{-w}^0 L_x(y) dF(y)$$

$$\begin{aligned} = & \int_c^a [L_x(y) - L_x(-y)](1-\delta) dF(y) + \int_w^c [L_x(w) - L_x(-w)] dF(y) \\ & + \int_0^w [L_x(y) - L_x(-y)] dF(y) + 2 \int_c^a \delta \left(U_x - \overline{U}_x \right) dF(y) \end{aligned} \tag{8}$$

Thus,

$$\left(U_x - \underline{U}_x \right) \left(1 - (2\delta[F(a) - F(c)])^2 \right) = \left[\int_c^a [L_x(y) - L_x(-y)](1-\delta) dF(y) \right.$$

$$\left. + \int_w^c [L_x(w) - L_x(-w)] dF(y) + \int_0^w [L_x(y) - L_x(-y)] dF(y) \right] (1 - 2\delta[F(a) - F(c)]).$$

The result then follows because $L_x(y) > L_x(-y)$ for any $y > 0$, because $\delta \leq 1$ and because $F(0) = 1/2$, and $c > 0$ implies that $[F(a) - F(c)] < 1/2$.

The proof that $U_x - \overline{U}_x > 0$ is analogous. ■

The second Lemma shows that when comparing compromise set under party competition $[v, k]$ and at-large selection $[w, c]$, it must be either that $v < w$ and $k > c$, or that $v > w$ and $k < c$. The compromise set is either enlarged or reduced at both extremes.

Lemma 2 *When comparing the at-large selection compromise set* (w, c) *and the party competition compromise set* (v, k) *it must be the case that* $(w - v)(c - k) < 0$.

Proof. Consider at-large selection first. By substituting the continuation utility (3) in the Belman equation (1) we obtain:

$$0 = -L_0(w) + 2 \int_c^a (L_0(y)(1-\delta) + \delta L_0(w)) dF(y) + 2 \int_w^c L_0(w) dF(y) + 2 \int_0^w L_0(y) dF(y).$$

With party competition, substituting the continuation utility (6) in the Belman equation (4) yields:

$$0 = -L_0(v) + 2 \int_k^a (L_0(y)(1-\delta) + \delta L_0(v)) dF(y) + 2 \int_v^k L_0(v) dF(y) + 2 \int_0^v L_0(y) dF(y).$$

Because the two equations have the same form, letting $\phi(w,c)$ equal the right-hand side, the result follows because:

$$\begin{aligned}\frac{dw}{dc} &= -\frac{\phi_2(w,c)}{\phi_1(w,c)} = -\frac{-2\left(L_0(c)\left(1-\delta\right)+\delta L_0(w)\right)f(c)+2L_0(w)f(c)}{-L_0'(w)+2\int_c^a \delta L_0'(w)dF(y)-2L_0(w)f(w)+2\int_w^c L_0'(w)dF(y)+2L_0(w)f(w)} \\ &= \left(-\frac{1}{L_0'(w)}\right)\frac{2\left(1-\delta\right)\left[L_0(w)-L_0(c)\right]f(c)}{2\delta\left[F(a)-F(c)\right]+2\left[F(c)-F(w)\right]-1} < 0 \text{ for } 0 \le w \le c,\end{aligned}$$

where the inequality follows because $L_0'(w) < 0$, $L_0(w) > L_0(c)$, $f(c) > 0$ and $2\left[F(a)-F(c)\right]\delta + 2\left[F(c)-F(w)\right]-1 < -2\left[F(a)-F(w)\right]-1 < 0$, as $F(a)=1$ and $F(w) > 1/2$ because $w > 0$. ■

We conclude the equilibrium comparison analysis by showing that the compromise interval is larger under party competition than under at-large selection. In light of Lemma 2, this means both that more office-holders are willing to compromise, and that when they compromise, they take more centrist positions.

Proposition 1 *Given that $L_x(v) - \delta\underline{U}_x(v,x)$ strictly decreases in x, the comparison of the compromise set under party competition $[v,k]$ and at-large selection $[w,c]$ is such that $v < w$ and $k > c$.*

Proof. Suppose that $v = w$ and $c = k$. By Lemma 1, $U_x(w,c) > \underline{U}_x(w,c)$ for any w, c. By substituting in the Belman equations (2), we obtain that $L_c(w) = l(|c-w|) = \delta U_c(w,c) > \delta\underline{U}_c(c,w)$. That is to say, for any $\varepsilon > 0$ small enough, the candidate $c+\varepsilon$ prefers to adopt policy w than $c+\varepsilon$. Holding $v = w$ fixed, the condition that the function $L_x(v) - \delta\underline{U}_x(v,x)$ of x crosses zero only once and from above then implies that $k > c$. Given this, Lemma 2 implies that $v < w$. ■

We now turn to welfare comparisons. We show that the introduction of parties makes all voters better off. Unlike most welfare analysis in this literature we do not consider only the effect on the median voter's welfare. Our welfare concept is Pareto efficiency.

Theorem 1 *All voters prefer party competition to at-large selection of candidates.*

The intuition for this result is simple. All citizens like insurance because they are weakly risk averse, and they discount utilities. Parties provide ex-ante insurance against extremist policies, because (i) $v < w$, i.e. there is less expected turn-over, and (ii) $k > c$, i.e. positions are more moderate over all.[4]

Remark (General Parties) The notion of parties can be extended as follows. Each party is identified by a distribution of candidates G, that admits a density g. Suppose that G_B first-order stochastically dominates G_A and that g_A and g_B are symmetric in the sense that $g_A(x) = g_B(-x)$, for all x. When parties overlap, policy holders are still retained in office if and only if their adopted policy y belongs to an interval symmetric $[-v, v]$ around the median policy 0. The median voter is decisive. Because g_A is symmetric to g_B around the median policy 0, the median voter's continuation value for electing the challenger does not depend on the identity of the party. However, there cannot be any equilibrium where each candidate x chooses a policy y if and only if $-x$ chooses $-y$. Whether an office-holder compromises or not depends on her party identity. Consider a party A office holder with ideology $x > 0$. If she is ousted from office, she will be replaced from a challenger from party B, such a candidate's ideology is likely to the right of the median voter. Hence, the party A office holder has a smaller incentive to compromise than under the at-large selection. In sum, we need to introduce the thresholds $\underline{k}_i < 0 < \overline{k}_i$, for A, B. Party-i office holders compromise and adopt policy $-v$ if and only if $x \in [\underline{k}_i, -v]$, and adopt policy v if and only if $x \in [v, \underline{k}_i]$. Because g_A is symmetric to g_B around zero, and G_B dominates G_A, we obtain that $\underline{k}_A < \underline{k}_B$ and $\overline{k}_A < \overline{k}_B$. Because the equilibrium is symmetric across parties, we have $\underline{k}_A = -\overline{k}_B$, $\underline{k}_B = -\overline{k}_A$, as well as $-\underline{k}_A = \overline{k}_B$, $-\underline{k}_B = \overline{k}_A$. We conjecture that $v < w$, and that $\underline{k}_A < -c < \underline{k}_B$ and $\overline{k}_A < c < \overline{k}_B$.

4 The 'Party Screening Effect'

In the previous section we assumed that when voters consider voting for the challenger, they have no information about her ideology. We now extend the analysis to allow for campaigning, through which candidates can disclose information about their ideologies before voting takes place. We shall assume that candidates come into the election with a imprecise, but verifiable, record of their ideological

[4]It is worth observing, however, that if there are term limits so that candidates can only hold office for two periods, then turnover may be higher with party selection. This higher turnover implies that with linear loss functions, extreme voters would prefer at large selection of candidates.

positions. This record is known to parties and to both candidates running for office, but is not known, publicized or advertised to voters. We expand the basic model to let politicians enjoy ego rents ρ for being in power. We assume that politicians can disclose their own, and their opponents' records at a cost C, which we assume is small relative to the benefit of holding office, i.e. $2C < \delta\rho$.

For simplicity, we assume that the information contained in a candidate's record is coarse. Campaigning can only communicate whether the candidate's ideology is leftist or conservative, and whether the candidates is moderate or extremist, i.e. whether the candidate's ideology belongs to one of the following sets partitioning the ideology space: $E_A = [-a, -m)$, $M_A = [-m, 0)$, $M_B = (0, m]$ or $E_B = (m, a]$, where E stands for extremist and M for moderate, and we assume that the partition is symmetric around the median ideology $x = 0$. Modelling imprecise verifiable signals as a partition of the ideology space simplifies the analysis.

We begin by studying party competition. Because party members know their potential candidates' records, we assume that they can endorse candidates on the basis of their records. We show that in equilibrium, parties endorse moderate candidates whenever these candidates have a chance to win. If they endorsed extremist candidates, they would be subject to the incumbent's negative advertisement and lose the election. As is usually the case, verifiable information is disclosed in equilibrium (see Milgrom 1981). Whenever the election matters for the equilibrium payoffs, party A chooses a candidate in M_A and B chooses a candidate in M_B. Extremists never hold office. The equilibrium characterization of the previous section, determined by the thresholds v, k, also apply to this problem, given that we restrict attention to candidates in M_A and M_B.

Proposition 2 *Under the conditions in Theorem A2, there is a unique stationary symmetric stage-undominated equilibrium of the campaign model with party selection of candidates. In equilibrium, the median voter is decisive and extremist candidates are never in power. Office-holders with ideology* $x \in [0, v] \cup (k, m]$ *adopt policy* $y = x$, *and those with* $x \in (v, k]$ *compromise with policy* $y = v$. *Office holders are re-elected unless* $x \in (k, m]$. *Symmetrically for* $x < 0$. *Letting* $F^*(y) = F(y|y \in [-m, m])$, *the thresholds v, k are determined by the Belman equations (4) and (5), and for any $x > 0$,*

$$\begin{aligned}
\underline{U}_x(v,k) &= 2\int_{-m}^{-k}\left(L_x(y)(1-\delta)+\delta\overline{U}_x(v,k)\right)dF^*(y)+2\int_{-k}^{-v}L_x(-v)dF^*(y)+2\int_0^v L_x(y)dF^*(y)\\
\overline{U}_x(v,k) &= 2\int_0^v L_x(y)dF^*(y)+2\int_v^k L_x(v)dF^*(y)+2\int_k^m\left(L_x(y)(1-\delta)+\delta\underline{U}_x(v,k)\right)dF^*(y).
\end{aligned}$$

The key feature of the result is the 'party screening effect' that we have discussed. Under party competition, only moderate candidates are selected by parties and matter for the equilibrium characterization and welfare. For the sake of comparison, we first point out that an immediate consequence of Proposition 1 is that all citizens prefer party selection to at-large selection in the case that nature only selects moderate politicians. That is, the 'screening effect' complements and does not substitute for the 'party-competition effect'.

Corollary 1 *Suppose that with at-large selection, nature would select only moderate candidates. Then, in the campaign model, all voters prefer party selection to at-large selection of candidates.*

The role of parties in our citizen-candidate repeated election model cannot be reduced to only providing information about the endorsed candidates, nor just to select moderate candidates. Party competition is essential in restraining policy choices by office holders. Such result depends on election repetition and cannot be derived in simpler static models. The office holder is more willing to compromise (thus favoring the median voter and all citizens alike) under party competition, because she knows that if she loses the election she will be replaced by a candidate with the opposite ideology.

Note that the above corollary does not prove that restricting selection to moderate candidates necessarily benefits all voters; it only shows that if choice *is* restricted to moderate candidates, then voters prefer party to at-large selection of candidates. Obviously, the "direct" effect of restricting attention to moderate candidates (i.e., holding (c, w) fixed) is to raise the welfare of all voters. But, for example, the restricting selection to moderate candidates could make candidates less willing to compromise, raising the likelihood of turnover, and hurting extreme voters. We now provide sufficient conditions with at-large selection for this not to occur. It immediately follows that all voters prefer nature to select more moderate candidates, which, in turn, implies that voters prefer party selection to at-large selection in the campaign model.

Proposition 3 *Suppose that F is uniform and the loss function is homogenenous, i.e. $l(k|x-y|) = g(k)\,l(|x-y|)$, where $g(k) > 0$ for all k. Then the threshold function is linearly homogeneous $w(m) = mw(a)$ and $c(m) = mc(a)$. Hence the turnover probability is constant in m, whereas the welfare U_x strictly decreases in m for every x. Hence, in the campaign model, all voters prefer party selection to at-large of candidates.*

5 Conclusion

We have studied the role of parties in a citizen-candidate repeated-elections model where voters have incomplete information. We have identified a novel "party competition effect." Compared with "at large" selection of candidates, party selection makes office-holders more willing to avoid extreme ideological stands. Incumbents would like to minimize the chances to be replaced by a challenger from the opposite party. Politicians follow party discipline, even in absence of a party-controlled reward mechanism. Voters of all ideologies benefit from the party-competition effect. Hence, our analysis provides a novel rationale for political parties: one of the main benefits of party competition is that it provides choice tied to clear ideological positions. The mere existence of a left-wing party, prevents right-wing elected officials from drifting into extremism. When politicians have an (imperfect) informational advantage over voters, we have additionally found a "party screening effect." Parties select moderate candidates, because they anticipate that their candidate's ideological record can be verifiably disclosed through campaigning. This party screening effect complements and does not substitute the party competition effect.

Our analysis may be extended in several directions. As well as screening candidates according to their ideology, parties may also screen candidates according to their valence, i.e. competence ability and so on. When valence is included in our model, preliminary analysis shows that the equilibrium is described by valence dependent thresholds. The median voter is willing to retain higher valence office holders even if they adopt more extreme policies. At the same time, high-valence candidates are more willing to compromise. In equilibrium, party members prefer to endorse high-valence candidates, because their utility is directly influenced by the office-holder valence, and because high-valence candidates are less likely to be ousted from office and replaced by challengers from the opposite party. As for the case of party selection of moderate candidates, the screening of high-valence candidates effect is complementary to the party competition effect, and does not substitute it.

Among further extensions, costly entry could be considered. The candidates need to pay a cost after being nominated, and cannot borrow against future benefits for holding office (i.e. there is no lobbying). Preliminary calculations indicate that the equilibrium characterization gains one or more further thresholds. In a quadratic loss model, moderate candidates have the smallest incentives to enter the race. This is likely to make the presentation of results unbearably complex, while we do not expect any major qualitative changes. Following the insights of Bernhardt *et al.* (2004), a further possibility is to consider the impact of term limits. Preliminary calculations with term limits of length two find an increment in turnover with party selection of candidates that tempers the benefits of the party competition effect.

A Appendix: Omitted Proofs

Theorem A1. *There is a uniform bound M, such that if $M < l'' \leq 0$, then $L_x(0) - U_x(c,w)$ weakly decreases in x, and $L_x(w) - \delta U_x(c,w)$ weakly decreases in x, for any $0 < w < c < a$. As a result, there is a unique symmetric, stationary, stage-undominated equilibrium, and this equilibrium is determined by the thresholds $0 < w < c < a$.*

Proof. We first assume that the elected officials policy choice is described by the thresholds $0 < w < c < a$, and show that the median voter is decisive, and sets the standard for re-election w.

Let the retrospective set of voter x be defined as the positions y implemented by an incumbent that x will re-elect over a random challenger, hence $R_x = \{y | L_x(y) - U_x(c,w) \geq 0\}$. By proceeding in subsequent Lemmata, we will show that the retrospective set of the median voter R_0 is contained in the win set $W = [-w, w]$.

Lemma A1. *For any $x \in [-a, a]$, $0 \in R_x$, i.e. zero belongs to the retrospective set of all agents.*

Proof. Let $x \in [0, a]$, we need to show that $L_x(0) - U_x(c,w) \geq 0$. Note that $L_x(0) - U_x(c,w) \geq 0$ if and only if $(L_x(0) - U_x(c,w))(1 - 2\delta(F(a) - F(c))) \geq 0$,

$$\begin{aligned}
& (L_x(0) - U_x(c,w))(1 - 2\delta(F(a) - F(c))) \\
= & \int_0^w -L_x(y) - L_x(-y) + 2L_x(0) dF(y) \\
& + \int_w^c -L_x(w) - L_x(-w) + 2L_x(0) dF(y) + \int_c^a (-L_x(y) - L_x(-y) + 2L_x(0))(1-\delta) dF(y)
\end{aligned}$$

We have that $L_x(y) = l(|x-y|)$, $l'(\cdot) \leq 0$ and $l''(\cdot) \leq 0$. Since $|x - y| \leq |x + y|$ for all $y \in [0, a]$, then $l(x) - l(|x-y|) \leq l(x) - l(|x+y|)$ the left side of the inequality is positive, the right side can be negative or positive, but in any case $-L_x(y) - L_x(-y) + 2L_x(0) = l(x) - l(|x-y|) + l(x) - l(|x+y|) \geq 0$. Therefore, the inequality holds and $0 \in R_x$ for all $x \in [0, a]$. Analogously, we can show that $0 \in R_x$ for all $x \in [-a, 0]$. ■

Lemma A2. *For $x \in [-a, 0]$ we have that $\frac{\partial U_x(w,c)}{\partial x} \geq 0$ and for $x \in [0, a]$ $\frac{\partial U_x(w,c)}{\partial x} \leq 0$.*

Proof. Note that $\frac{\partial U_x(w,c)}{\partial x} \leq 0$ if and only if $\frac{\partial U_x(w,c)}{\partial x}(1 - 2\delta(F(a) - F(c))) \leq 0$. Let $x > 0$, then

$$\begin{aligned}
& \frac{\partial U_x(w,c)}{\partial x}(1 - 2\delta(F(a) - F(c))) \\
= & \int_0^w l'(|x-y|)\frac{\partial(|x-y|)}{\partial x} + l'(x+y) dF(y) + \int_w^c l'(|x-w|)\frac{\partial(|x-w|)}{\partial x} \\
& + l'(x+w) dF(y) + \int_c^a (l'(|x-y|)\frac{\partial(|x-y|)}{\partial x} + l'(x+y))(1-\delta) dF(y).
\end{aligned}$$

Concavity of l implies that $0 \geq l'(|x-y|) \geq l'(x+y)$ for all $y \in [0,a]$, and

$$\frac{\partial(|x-y|)}{\partial x} = \begin{cases} 1, & \text{if x>y;} \\ -1, & \text{if x<y;} \end{cases}$$

then $0 \geq l'(|x-y|)\frac{\partial(|x-y|)}{\partial x} + l'(x+y)$ for all $y \in [0,a]$. Therefore, for $x > 0$, $\frac{\partial U_x(w,c)}{\partial x} \leq 0$. Analogously, we can show that for $x < 0$, $\frac{\partial U_x(w,c)}{\partial x} \geq 0$. ■

Lemma A3. *The retrospective set of the median voter R_0 is contained in the win set $W = [-w, w]$.*

Proof. Let $x \in R_0$ and $x > 0$ by Lemma 1, $0 \in R_y$, then we must have that the lower extreme of the retrospective sets is less than zero, given that $y \in R_y$ for all y, then $x \in R_y$ for all $y \in [x, a]$. Note that the upper extreme of the retrospective set is given by $y + l^{-1}(U_y(w,c))$, where $l^{-1}(\cdot)$ denotes the inverse function of $l(\cdot)$; $l^{-1}(\cdot)$ is a decreasing function; and $l^{-1} : \Re^- \to \Re^+$. Lemma 2 implies that $U_y(w,c) \leq U_0(w,c)$, then $y + l^{-1}(U_y(w,c)) \geq 0 + l^{-1}(U_0(w,c))$, so $x \in R_y$ for all $y \in [0,x]$. Then, for any $x \in R_0$ and $x > 0$ all $y \in [0,a]$ will vote for an incumbent that implements x over a random challenger and therefore x will win at least half the votes and belong to the win set. ■

Lemma A4. *If $L_x(0) - U_x(w,c)$ decreases in x for any $x > 0$, then the win set W is contained in the retrospective set of the median voter R_0.*

Proof. We will show that if y is not in R_0, then y is not in the win set. Let y not in R_0 and $y < 0$, note that for any $x \in [0,a]$ the lower extreme of the retrospective set is given by $x - l^{-1}(U_x(w,c))$. Given that $L_x(0) - U_x(w,c)$ decreases in x, for all $0 < w < c < a$, it must be that $x - l^{-1}(U_x(w,c))$ increases in x, for any $x > 0$. Hence, we must have that $x - l^{-1}(U_x(w,c)) \geq 0 - l^{-1}(U_0(w,c)) \geq y$, so y is not in R_x for all $x \in [0,a]$. Then, at least half the voters will vote against y and y will not belong to the win set. We can show that the condition implies that $x + l^{-1}(U_x(w,c))$ increases for $x < 0$, so analogously we can show that any y not in R_0 and $y > 0$ will not belong to the win set. ■

Lemma A5. *There is a uniform bound M such that if $M < l'' \leq 0$, then $L_x(0) - U_x(w,c)$ decreases in x for any $x > 0$ and any $0 < w < c < a$.*

Proof. Because,

$$\begin{aligned} U_x(w,c) &= \frac{1}{1 - \delta 2\left[F(a) - F(c)\right]}\Big[\int_{-a}^{-c} L_x(y)(1-\delta)\, dF(y) + \int_{-c}^{-w} L_x(-w) dF(y) \\ &\quad + \int_{-w}^{w} L_x(y) dF(y) + \int_{w}^{c} L_x(w) dF(y) + \int_{c}^{a} L_x(y)(1-\delta)\, dF(y)\Big], \end{aligned}$$

it must be the case that:

$$\begin{aligned}
&\frac{\partial}{\partial x}L_x(0)-\frac{1}{1-2\delta\left[F(a)-F(c)\right]}[(1-\delta)\int_{-a}^{-c}\frac{\partial}{\partial x}L_x(y)dF(y)\\
&+\int_{-c}^{-w}\frac{\partial}{\partial x}L_x(-w)dF(y)+\int_{-w}^{w}\frac{\partial}{\partial x}L_x(y)dF(y)\\
&+\int_{w}^{c}\frac{\partial}{\partial x}L_x(w)dF(y)+(1-\delta)\int_{c}^{a}\frac{\partial}{\partial x}L_x(y)dF(y)]<0.
\end{aligned}$$

If $l''=0$, this quantity is indeed negative, because $\frac{\partial}{\partial x}L_x(y)$ is constant in x,y and negative for $y<x$, $\frac{\partial}{\partial x}L_x(y)>0$ for all $y>x$, and

$$\begin{aligned}
&\frac{1}{1-2\delta\left[F(a)-F(c)\right]}[(1-\delta)\int_{-a}^{-c}dF(y)+\int_{-c}^{-w}dF(y)+\int_{-w}^{w}dF(y)+\int_{w}^{c}dF(y)+(1-\delta)\int_{c}^{a}dF(y)]\\
=\;&\frac{2(1-\delta)\left[F(a)-F(c)\right]+2\left[F(c)-F(0)\right]}{1-2\delta\left[F(a)-F(c)\right]}=1,\ \text{when}\ \left[F(a)-F(c)\right]+\left[F(c)-F(0)\right]=1/2.
\end{aligned}$$

This implies that there is a uniform lower bound $M:0<M<l''$ guaranteeing that $L_x(0)-U_x(w,c)$ decreases in x, for all $0<w<c<a$. ■

We finally conclude the proof by showing that, given the standard for re-election w set by the median voter, the elected officials policy choice is described by the thresholds $0<w<c<a$.

Lemma A6. *There is a uniform bound M' such that if $M'<l''\leq 0$, then $L_x(w)-\delta U_x(c,w)$ decreases in x, for any $0<w<c<a$.*

Proof. We need that for all $w:0<w<a$, and all $x>w$, the following expression decreases in x:

$$\begin{aligned}
&L_x(w)-\delta U_x(c,w)\\
=\;&L_x(w)-\frac{\delta}{1-\delta 2\left[F(a)-F(c)\right]}[\int_{-a}^{-c}L_x(y)(1-\delta)dF(y)+\int_{-c}^{-w}L_x(-w)dF(y)\\
&+\int_{-w}^{w}L_x(y)dF(y)+\int_{w}^{c}L_x(w)dF(y)+\int_{c}^{a}L_x(y)(1-\delta)dF(y)].
\end{aligned}$$

$$\begin{aligned}
&\frac{\partial}{\partial x}L_x(w)-\frac{\delta}{1-\delta 2\left[F(a)-F(c)\right]}[(1-\delta)\int_{-a}^{-c}\frac{\partial}{\partial x}L_x(y)dF(y)+\int_{-c}^{-w}\frac{\partial}{\partial x}L_x(-w)dF(y)\\
&+\int_{-w}^{w}\frac{\partial}{\partial x}L_x(y)dF(y)+\int_{w}^{c}\frac{\partial}{\partial x}L_x(w)dF(y)+(1-\delta)\int_{c}^{a}\frac{\partial}{\partial x}L_x(y)dF(y)]
\end{aligned}$$

Proceeding as in the proof of Lemma A.5, we obtain that if $l''=0$, then this expression is decreasing in x. Hence there is a uniform lower bound $M':0<M'<l''$ guaranteeing that $L_x(w)-\delta U_x(c,w)$ decreases in x, for all $0<w<c<a$. ■ ■

Theorem A2. *There is a uniform bound M such that if $M < l'' \leq 0$, then $L_x(0) - \underline{U}_x(k,v)$ weakly decreases in x and $L_x(v) - \delta\underline{U}_x(k,v)$ weakly decreases in x, for any $0 < v < k < a$. As a result, there is a unique symmetric stationary stage-undominated equilibrium, and this equilibrium is determined by the thresholds $0 < v < k < a$.*

Proof. The proof that $L_x(v) - \delta\underline{U}_x(k,v)$ weakly decreases in x, for any v, and that as a result the elected politicians choice is described by thresholds $0 < v < k < a$ is unchanged with respect to the at-large selection case. The proof that the retrospective set of the median voter R_0 is contained in the win set W is unchanged with respect to the at-large selection case. The proof that the $W \subseteq R_0$ is as follows. Suppose that $y < 0$ and $L_0(y) < U_0(v,k)$. Pick any $x \geq 0$, we calculate:

$$\begin{aligned}
&\frac{\partial}{\partial x}\left[L_x(y) - \overline{U}_x(v,k)\right] \propto \frac{\partial}{\partial x}\left[2\int_0^v [L_x(y) - L_x(t)]\,dF(t) + 2\int_v^k [L_x(y) - L_x(v)]\,dF(t)\right. \\
&+2\int_k^a \left([L_x(y) - L_x(t)](1-\delta) + \delta\left[2\int_{-a}^{-k} ([L_x(y) - L_x(t)](1-\delta))\,dF(t)\right.\right. \\
&\left.\left.\left. +2\int_{-k}^{-v} [L_x(y) - L_x(-v)]\,dF(t) + 2\int_{-v}^{0} [L_x(y) - L_x(t)]\,dF(t)\right]\right)dF(t)\right] \\
= &\cdot\frac{\partial}{\partial x}\left[2\int_0^v [L_x(y) - L_x(t) + \delta[L_x(y) - L_x(-t)]\,2[F(a) - F(k)]]\,dF(t)\right. \\
&+2\int_v^k [L_x(y) - L_x(v) + \delta[L_x(y) - L_x(-v)]\,2[F(a) - F(k)]]\,dF(t) \\
&\left.+(1-\delta)\int_k^a [L_x(y) - L_x(t) + \delta[L_x(y) - L_x(-t)]\,2[F(a) - F(k)]]\,dF(t)\right].
\end{aligned}$$

Since $y < 0$, it follows that $\frac{\partial}{\partial x}[L_x(y) - L_x(t)] < 0$ for any $t > 0$ and that $\left|\frac{\partial}{\partial x}[L_x(y) - L_x(t)]\right| > \left|\frac{\partial}{\partial x}[L_x(y) - L_x(-t)]\right|$. This in turns, imply that the above quantity is positive. ■

Proof of Proposition 1. For any $x > 0$, the case for $x < 0$ is analogous, we need to compare the ex-ante at large welfare $U_x(c,w)$, with the party-competition ex-ante welfare

$$\begin{aligned}
U_x^*(k,v) = \frac{1}{2}[\underline{U}_x + \overline{U}_x] = &\int_{-a}^{-k} \left(L_x(y)(1-\delta) + \delta\overline{U}_x\right)dF(y) + \int_{-k}^{-v} L_x(-v)dF(y) \\
+ \int_{-v}^{0} L_x(y)dF(y) + &\int_0^v L_x(y)dF(y) + \int_v^k L_x(v)dF(y) + \int_k^a (L_x(y)(1-\delta) + \delta\underline{U}_x)\,dF(y).
\end{aligned}$$

Simple algebraic manipulations give:

$$U_x^* (k, v) - U_x (c, w) = \frac{1}{2}[\underline{U}_x + \overline{U}_x] - U_x = \int_k^a \delta \left(\overline{U}_x + \underline{U}_x - 2U_x\right) dF(y)$$
$$+ \int_c^k [L_x(v) - L_x(y)(1-\delta) - \delta U_x] dF(y) + \int_c^k [L_x(-v) - L_x(-y)(1-\delta) - \delta U_x] dF(y)$$
$$+ \int_w^c [L_x(-v) - L_x(-w)] dF(y) + \int_w^c [L_x(v) - L_x(w)] dF(y)$$
$$+ \int_v^w [L_x(-v) - L_x(-y)] dF(y) + \int_v^w [L_x(v) - L_x(y)] dF(y),$$

where we use the order $0 < v < w < c < k$, to simplify integral expressions. Thus,

$$U^* (k, v) - U_x (c, w) \propto \Phi \equiv \int_c^k [L_x(v) - L_x(y)(1-\delta) - \delta U_x] dF(y)$$
$$+ \int_c^k [L_x(-v) - L_x(-y)(1-\delta) - \delta U_x] dF(y) + \int_v^w [L_x(-v) - L_x(-y)] dF(y)$$
$$+ \int_v^w [L_x(v) - L_x(y)] dF(y) + \int_w^c [L_x(-v) - L_x(-w)] dF(y) + \int_w^c [L_x(v) - L_x(w)] dF(y).$$

For any $y : v < y \leq c$, we know that $L_x(-v) - L_x(-y) \geq 0$, by monotonicity of the function L_x for negative policies, and that $L_x(-v) - L_x(-y) \geq |L_x(v) - L_x(y)|$, because $x > 0$. Hence

$$\Phi \geq \delta[F(k) - F(c)]\Psi_x = \delta \int_c^k [L_x(v) - U_x] dF(y) + \delta \int_c^k [L_x(-v) - U_x] dF(y).$$

Let x momentarily be the median voter: Substituting in this equation the Belman equation (1), and using the symmetry of L_0 around zero, we obtain:

$$\Psi_x = L_0(v) - L_0(w) + L_0(v) - L_0(w) > 0,$$

because $v < w$ and hence $L_0(v) > L_0(w)$.

Now consider all other voters x. Note that:

$$U_x = \frac{\int_c^a [L_x(-y) + L_x(y)](1-\delta) dF(y) + \int_w^c [L_x(-w) + L_x(w)] dF(y) + \int_0^w [L_x(-y) + L_x(y)] dF(y)}{1 - 2\delta[F(a) - F(c)]}$$

Hence:

$$\Psi_x = L_x(v) + L_x(-v) - 2U_x$$
$$\propto \quad \Upsilon_x = (1 - 2\delta[F(a) - F(c)]) [L_x(v) + L_x(-v)] - (1 - 2\delta[F(a) - F(c)]) 2U_x$$

$$= \quad -2 \int_c^a [L_x(y) + L_x(-y)](1-\delta) dF(y) - 2\delta[F(a) - F(c)] [L_x(v) + L_x(-v)]$$
$$-2 \int_w^c [L_x(w) + L_x(-w)] dF(y) - \int_0^w 2 [L_x(y) + L_x(-y)] dF(y) + [L_x(v) + L_x(-v)]$$

$$
\begin{aligned}
= \quad & -2(1-\delta)\int_c^a [L_x(y) + L_x(-y) - [L_x(v) + L_x(-v)]]\, dF(y) \\
& -2[L_x(w) + L_x(-w) - [L_x(v) + L_x(-v)]][F(c) - F(w)] \\
& -2\int_0^w [L_x(y) + L_x(-y) - [L_x(v) + L_x(-v)]]\, dF(y).
\end{aligned}
$$

For any y, note that $\frac{d}{dx}[L_x(y) + L_x(-y)] \leq 0$, that $\frac{d^2}{dxdy}[L_x(y) + L_x(-y)] \geq 0$ if $L = -k|x-y|$ for any $k > 0$ and that for fixed $l'(|x-y|)$ and $l'(|x+y|)$, $\frac{d^2}{dxdy}[L_x(y) + L_x(-y)]$ is smaller the smaller are $l''(|x-y|)$ and $l''(|x+y|)$.

After inspecting the above equation, we conclude that we only need to show that $\Psi_x > 0$ for all x, for the case that $l'' = 0$. Indeed, we find:

$$
\Upsilon_x = \begin{cases}
-2\int_x^a [2x-2y](1-\delta)\, dF(y) > 0, & \text{for } 0 < v < w < c < x < a, \\
-2\int_c^a [2x-2y](1-\delta)\, dF(y) > 0, & \text{for } 0 < v < w < x < c < a, \\
-2\int_c^a [2x-2y](1-\delta)\, dF(y) - 2[2x-2w][F(c)-F(w)] + & \\
\quad -2\int_x^w [2x-2y]\, dF(y) > 0 & \text{for } v < x < w < c < a, \\
-2\int_c^a [2v-2y](1-\delta)\, dF(y) - 2[2v-2w][F(c)-F(w)] + & \\
\quad -2\int_v^w [2v-2y]\, dF(y) - 2\int_0^v [2v-2x]\, dF(y) > 0 & \text{for } 0 < x < v < w < c < a,
\end{cases}
$$

where the last expression is positive because increasing in x and because $\Upsilon_0 > 0$. ■

Proof of Proposition 2. Suppose that the incumbent is of party B and adopts policy $y > 0$. Let $U_{x'}(E_A)$, $U_{x'}(M_A)$, $U_{x'}(A)$ be respectively the equilibrium continuation value to a voter x' of electing a challenger with ideology $x \in E_A$, $x \in M_A$, and $x \in [-a, 0]$. Let v_E and v_M solve $L_0(v_E) = U_0(E_A)$ and $L_0(v_M) = U_0(M_A)$.

Lemma A7. *If challengers may be both moderate* $(x \in M_A)$ *and extremist* $(x \in E_A)$, *then the incumbent wins the election in equilibrium if* $0 \leq y \leq v_M$, *or if* $v_M < y \leq v_E$ *and* $x \in E_A$, *whereas she loses if* $v_E < y \leq a$, *or if* $v_M < y \leq v_E$ *and* $x \in M_A$.

Proof. Consider any stationary stage-undominated equilibrium where the median voter is decisive. When the voters know that $x \in E_A$, the incumbent wins the election if and only if $L_0(y) \geq U_0(E_A)$, i.e. $y \leq v_E$. If it is disclosed that $x \in M_A$, the incumbent wins whenever $L_0(y) \geq U_0(M_A)$, i.e. $y \leq v_M$. If the voters are uninformed, she retains office when $L_0(y) \geq U_0(A)$, i.e. $y \leq v_A$ where $L_0(v_A) = U_0(A)$. The record of the incumbent is immaterial.

In any stationary equilibrium, $U_0(E_A) < U_0(A) < U_0(M_A)$. Suppose in fact that the challenger A is elected. Her policy choice $y = p(x)$ depends only on her ideology $x < 0$, and not on whether her record has been revealed. By the single-crossing property of the loss functions $L_x(y) = l(|x-y|)$, the policy $p(x)$ is weakly increasing in x. Hence, the distribution $x|\{x \in M_A\}$ first-order stochastically dominates $x|\{x \in [-a,0]\}$ dominates $x|\{x \in E_A\}$. To complete the argument, note that the policy choices of future challengers do not depend on the challenger's record, and that $L_0(y)$ function is increasing in y, for $y < 0$

Because $U_0(E_A) < U_0(A) < U_0(M_A)$, it follows that $v_M < v_A < v_E$. When $0 \le y \le v_M$ the incumbent wins the election regardless of the challenger's record, when $y > v_E$, the challenger wins regardless of her record. When $y \in (v_A, v_E]$ the incumbent wins unless it is disclosed that $x \in M_A$. When $y \in (v_M, v_A]$ the challenger wins unless it is disclosed that $x \in E_A$. ■

Because the campaigning cost C is smaller than $\delta\rho$, it follows that (i) when $y \in (v_A, v_E]$ and $x \in M_A$, the challenger will disclose that $x \in M_A$, and (ii) when $y \in (v_M, v_A]$ and $x \in E_A$, the incumbent will disclose that $x \in E_A$. Hence, the incumbent wins the election if $0 \le y \le v_M$, or if $v_M < y \le v_E$ and $x \in E_A$, whereas she loses if $v_E < y \le a$, or if $v_M < y \le v_E$ and $x \in M_A$.

Anticipating this, party A unanimously endorses a candidate with record M_A whenever the incumbent's policy $y \in (v_M, v_E]$. To see that all citizens $x < 0$ prefer to have a candidate with record M_A in power, rather than retaining the incumbent y, note that because $y \le v_E$, the median voter prefers to elect the candidate with record M_A instead of the incumbent. A fortiori, this is also the preference of all voters $x < 0$. When $y \le v_M$ or $y \ge v_E$, the challenger record is not revealed in the election, and the party choice is irrelevant for our proof.

Because $U_0(E_A) < U_0(A) < U_0(M_A)$, in the time-0 election, both parties select a moderate candidate. In fact, if party (say) B choose a candidate $x \in E_B$, the best response of party A would be to select a candidate $x \in M_A$. Because $2C < \delta\rho$, the party-A candidate would reveal that $x \in E_B$ and that $x \in M_A$ and win the election.

This concludes that, in equilibrium, extremist candidates are never elected, and that the incumbent will lose the election to a moderate challenger if selecting policy $y > v_M$. This means that extremist candidates must be disregarded by the analysis. Hence, the equilibrium is characterized by the thresholds v_M and k_M such that $L_0(v_M) = U_0(M_A)$ and $L_{k_M}(v_M) = \delta U_{k_M}(M_A)$. ■

Proof of Proposition 3. Consider the equations characterizing equilibrium

$$L_0(mw) = U_0(mc, mw), \quad L_{mc}(mw) = \delta U_{mc}(mc, mw).$$

Hence,

$$L_{mc}(mw) = \frac{1}{2m\left[1-\delta\left[1-c\right]/2\right]} \begin{bmatrix} \int_0^{mw} \left[L_{mc}(y) + L_{mc}(-y)\right] dy \\ + \int_{mw}^{mc} \left[L_{mc}(mw) + L_{mc}(-mw)\right] dy \\ + (1-\delta) \int_{mc}^{am} \left[L_{mc}(y) + L_{mc}(-y)\right] dy \end{bmatrix},$$

$$l(|mw-mc|)=\frac{\delta}{2m\left[1-\delta\left[1-c\right]/2\right]}\left[\begin{array}{c}\int_0^{mw}\left[l\left(|mc-y|\right)+l\left(|mc+y|\right)\right]dy\\+\int_{mw}^{mc}\left[l\left(|mc-mw|\right)+l\left(|mc+mw|\right)\right]dy\\+\left(1-\delta\right)\int_{mc}^{am}\left[l(|mc-y|)+l(|mc+y|\right]dy\end{array}\right],$$

$$l(m\left|w-c\right|)=\frac{\delta}{2m\left[1-\delta\left[1-c\right]/2\right]}\left[\begin{array}{c}\int_0^{w}\left[l\left(m\left|c-y\right|\right)+l\left(m\left|c+y\right|\right)\right]mdy\\+\int_{w}^{c}\left[l\left(m\left|c-w\right|\right)+l\left(m\left|c+w\right|\right)\right]mdy\\+\left(1-\delta\right)\int_{c}^{1}\left[l(m\left|c-y\right|)+l(m\left|c+y\right|\right]mdy\end{array}\right],$$

$$l(m\left|w-c\right|)=\frac{\delta}{2\left[1-\delta\left[1-c\right]/2\right]}\left[\begin{array}{c}\int_0^{w}\left[l\left(m\left|c-y\right|\right)+l\left(m\left|c+y\right|\right)\right]dy\\+\int_{w}^{c}\left[l\left(m\left|c-w\right|\right)+l\left(m\left|c+w\right|\right)\right]dy\\+\left(1-\delta\right)\int_{c}^{a}\left[l(m\left|c-y\right|)+l\left(m\left|c+y\right|\right)\right]dy\end{array}\right],$$

Suppose that l is a homogeneous function: $l(m\left|w-c\right|)=g\left(m\right)l\left(|w-c|\right)$. Then we obtain:

$$g\left(m\right)l(|w-c|)=\frac{\delta}{2\left[1-\delta\left[1-c\right]/2\right]}\left[\begin{array}{c}g\left(m\right)\int_0^{w}\left[l\left(|c-y|\right)+l\left(|c+y|\right)\right]dy\\+g\left(m\right)\int_{w}^{c}\left[l\left(|c-w|\right)+l\left(|c+w|\right)\right]dy\\+\left(1-\delta\right)g\left(m\right)\int_{c}^{a}\left[l(|c-y|)+l\left(|c+y|\right)\right]dy\end{array}\right],$$

and hence

$$l(|w-c|)=\frac{\delta}{2\left[1-\delta\left[1-c\right]/2\right]}\left[\begin{array}{c}\int_0^{w}\left[l\left(|c-y|\right)+l\left(|c+y|\right)\right]dy\\+\int_{w}^{c}\left[l\left(|c-w|\right)+l\left(|c+w|\right)\right]dy\\+\left(1-\delta\right)\int_{c}^{a}\left[l(|c-y|)+l\left(|c+y|\right)\right]dy\end{array}\right],$$

analogously, we obtain:

$$l(|w|)=\frac{1}{2\left[1-\delta\left[1-c\right]/2\right]}\left[\begin{array}{c}\int_0^{w}\left[l\left(|-y|\right)+l\left(|y|\right)\right]dy\\+\int_{w}^{c}\left[l\left(|-w|\right)+l\left(|w|\right)\right]dy\\+\left(1-\delta\right)\int_{c}^{a}\left[l\left(|-y|\right)+l\left(|y|\right)\right]dy\end{array}\right],$$

and this verifies the result that $w\left(m\right)=mw$, and that $c\left(m\right)=cm$. Further, consider

$$U_x\left(mw,mc\right)=\frac{\delta}{2m\left[1-\delta\left[1-c\right]/2\right]}\left[\begin{array}{c}\int_0^{mw}\left[L_x(y)+L_x\left(-y\right)\right]dy+\\\int_{mw}^{mc}\left[L_x(mw)+L_x(-mw)\right]dy\\+\left(1-\delta\right)\int_{mc}^{am}\left[L_x(y)+L_x\left(-y\right)\right]dy\end{array}\right].$$

$$U_x\left(mw,mc\right)=\frac{\delta}{2\left[1-\delta\left[1-c\right]/2\right]}\left[\begin{array}{c}\int_0^{w}\left[l(|x-my|)+l(|x+my|)\right]dy+\\\int_{w}^{c}\left[l(|x-mv|+l(|x+mv|)\right]dy\\+\left(1-\delta\right)\int_{c}^{a}\left[l(|x-my|)+l(|x+my|)\right]dy\end{array}\right].$$

Because $l''\leq 0$, it follows that $\partial\left[l(|x-my|)+l(|x+my|)\right]/\partial m<0$ and hence that $\partial U_x/\partial m$. ■

References

[1] Alvarez, R. Michael (1997), *Information and Elections.* Ann Arbor: University of Michigan Press.

[2] Ashworth, Scott and Ethan Bueno de Mesquita (2004): "Party Discipline, Executive Power, and Ideological Balance" mimeo, Princeton University.

[3] Austen-Smith, David and Jeffrey S. Banks (1989): "Electoral Accountability and Incumbency," in *Models of Strategic Choice in Politics*, ed. by Peter Ordeshook. Ann Arbor: University of Michigan Press.

[4] Banks, Jeffrey S. and John Duggan (2001), "A Multidimensional Model of Repeated Elections," *mimeo*, University of Rochester.

[5] Banks, Jeffrey S. and Rangarajan K. Sundaram (1998), "Optimal Retention in Agency Problems, *Journal of Economic Theory*, 82, p.293-323.

[6] Banks Jeffrey S. and Rangarajan K. Sundaram (1993), "Moral Hazard and Adverse Selection in a Model of Repeated Elections", in *Political Economy: Institutions, Information, Competition and Representation*, W. Barnett *et al.* editors, New York, Cambridge University Press.

[7] Bartels, Larry M. (1986): "Issue Voting Under Uncertainty: An Empiricial Test." *American Journal of Political Science* 30: 709-728.

[8] Bernhardt, Dan, Sangita Dubey and Eric Hughson (2004): "Term Limits and Pork Barrel Politics", *Journal of Public Economics,* 2383-2422.

[9] Calvert, Randall and Justin Fox (2000): "Effective Parties in a Model of Repeated Legislative Bargaining." mimeo, University of Rochester.

[10] Cox, Gary W. and Matthew D. McCubbins (1993): *Legislative Leviathan: Party Government in the House.* Berkeley: University of California Press.

[11] Diermeier, Daniel and Timothy J. Feddersen (1998): "Cohesion in Legislatures and the Vote of Confidence Procedure." *American Political Science Review* 92: 611-621.

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

[13] Duggan, John (2000): "Repeated Elections with Asymmetric Information, *Economics and Politics*, 12, 109-135.

[14] Ferejohn, John (1986): "Incumbent Performance and Electoral Control," *Public Choice*, 43, p.5-25.

[15] Fiorina, Morris (1981): *Retrospective Voting in American National Elections,* Cambridge University Press, New Haven, CT.

[16] Huber, John (1996): "The Impact of Confidence Votes on Legislative Politics in Parliamentary Systems" *American Political Science Review* 90: 269-282.

[17] Levy, Gilat (2004): "A Model of Political Parties", *Journal of Economic Theory,* 115: 250-277.

[18] Lindbeck, Assar and Jorgen Weibull (1993): "A Model of Political Equilibrium in a Representative Democracy," *Journal of Public Economics* 51: 195-209.

[19] Milgrom, Paul R. (1981): "Good News and Bad News: Representation Theorems and Applications," *Bell Journal of Economics,* 12: 380-391.

[20] Morelli, Massimo (2004): "Party Formation and Policy Outcomes under Different Electoral Systems," *Review of Economic Studies* 71: 829-853.

[21] Osborne, Martin J. and Rabee Tourky (2003): "Party Formation in Collective Decision-Making." Unpublished paper.

[22] Persson, Torsten, Gerard Roland and Guido Tabellini (2003): "How do electoral rules shape party structures, government coalitions, and economic policies?," Unpublished Paper.

[23] Reed, Robert (1989): "Information in Political Markets: A Little Knowledge can be a Dangerous Thing," *Journal of Law, Economics, and Organization*, 5: 355-373.

[24] Reed, Robert (1994): "A Retrospective Voting Model with Heterogeneous Politicians ", *Economics and Politics,* 6: 39-58.

[25] Snyder, James M. and Michael M. Ting (2002): "An Informational Rationale for Political Parties," *American Journal of Political Science* 46: 90-110.