DISCUSSION PAPER SERIES

No. 3180

PLATFORM DIVERGENCE, POLITICAL EFFICIENCY AND THE MEDIAN VOTER THEOREM

Juan D Carrillo and Micael Castanheira

PUBLIC POLICY

Centre for Economic Policy Research

www.cepr.org

Available online at: **www.cepr.org/pubs/dps/DP3180.asp**

ISSN 0265-8003

PLATFORM DIVERGENCE, POLITICAL EFFICIENCY AND THE MEDIAN VOTER THEOREM

Juan D Carrillo, Columbia Business School and CEPR
Micael Castanheira, ECARES, Université Libre de Bruxelles and CEPR

Discussion Paper No. 3180
January 2002

Centre for Economic Policy Research
90–98 Goswell Rd, London EC1V 7RR, UK
Tel: (44 20) 7878 2900, Fax: (44 20) 7878 2999
Email: cepr@cepr.org, Website: www.cepr.org

CEPR Discussion Paper No. 3180

January 2002

ABSTRACT

Platform Divergence, Political Efficiency and the Median Voter Theorem*

The Paper analyses a standard Downsian model of election with two opportunistic parties. We assume that, after choosing their ideological position but before the election takes place, parties can affect the quality of their platforms by exerting some unobservable effort. When voters either (almost) always or (almost) never observe the resulting quality before the election, the standard Median Voter Theorem holds. For the more general case of imperfect observability of quality, however, we show that parties may optimally deviate from the median voter's bliss point as an implicit commitment to exert high effort (and therefore obtain a high-quality platform). The Paper thus argues that extremist parties are endogenously more committed to their ideas than moderate parties. Moreover, the extra quality implied by the divergence of parties will sometimes offset their worse ideology proposed, in which case the voters' welfare under divergence is greater than under convergence of platforms. Last, we endogenize the amount of information revealed to voters by assuming that a profit maximizing press collects the news about the quality of parties and sells it to the electorate. We show that the press may collect an amount of information that is excessively high from a social viewpoint.

JEL Classification: D72, D78 and L15
Keywords: median voter theorem, moral hazard and party competition

Juan D Carrillo
Columbia Business School
Uris Hall # 812
3022 Broadway
New York, NY 10027
USA
Tel: (1 212) 854 2455
Fax: (1 212) 662 8474
Email: jc2095@columbia.edu

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=131469

Micael Castanheira
ECARES
CP114
50 Avenue Roosevelt
1050 Bruxelles
BELGIUM
Tel: (32 2) 650 4467/3
Fax: (32 2) 650 4474
Email: mcasta@uni-bocconi.it

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=126467

* We thank Isabelle Brocas, Johan Lagerlöf, Howard Rosenthal, Loic Sadoulet, Raph Thomadsen and seminar participants at the CEPR workshop in Hydra and Columbia for comments and discussions. Micael Castanheira is 'Chargé de recherche aupres du FNRS' and gratefully acknowledges their financial support.

Submitted 04 January 2002

NON-TECHNICAL SUMMARY

What makes the political platform of a party sound ‘attractive’ or ‘convincing’? Why may proposing the median voter’s preferred platform sometimes be a self-defeating political choice? More generally, why, contrary to standard predictions of the theory, do we rarely observe convergence of parties to the median voter’s favourite platform?

In this Paper, we develop a theoretical model that explains some of these observations. We want to understand why some platforms may sound more credible or attractive than others, and how this appeal – which we call ‘quality’ – is linked to the ideological position of the platform. The novelty of our approach lies in the combination of two factors. First, the quality of the platform offered by each party is the result of some (costly) investment or ‘effort’. Second, ideological positions are perfectly observed, whereas quality is not. This captures the idea that, at the election time, voters are often uncertain about which program is more sensible. They have to rely on scarce information and decide to trust the party whose platform sounds more convincing. Using this approach, we identify interesting interactions between the quality of information and political competition.

We consider purely opportunistic parties and assume that voters’ preferences are public knowledge. Voters care both about the *ideology* and the *quality* of the platform offered by each party. Even if these two dimensions are orthogonal, our results show that the amount of information on the quality of platforms crucially affects the ideological position of the parties. Only for the extreme cases already analysed in the literature, namely when voters either (almost) always or (almost) never observe the quality of the platforms before the election, does the standard Median Voter Theorem hold.

For the more general case of imperfect observability of quality, the Median Voter Theorem does not hold anymore. Why? Because voters want parties to invest in the quality of their platforms. Parties cannot, however, commit to this. Thus, the party may need to ‘handicap’ itself (and adopt a non-centrist platform) in order to *implicitly* commit to provide a platform of higher quality than its rival. When this is the case, the extremist party will benefit from the ‘trust’ of the electorate whenever qualities are not observed. In other words, we show that some party may use its ideological position to manipulate its own incentives to exert higher effort. As a result, voters may find it optimal to reward the party which puts itself in a difficult position (a non-median platform) whenever they are uncertain about the quality level obtained by parties.

Two lessons can be drawn from these results. First, our theory explains why even purely opportunistic parties may appear ‘ideologically biased’. Second, our Paper sheds light on the effects of information on political competition. Insofar as our ‘accuracy of information’ parameter can reflect the result of an investment by the press to learn the quality of parties (and sell it to the electorate), our model measures the inefficiencies due to information being collected and provided by a privately interested press that does not maximize social welfare. Interestingly, we show that the press sometimes provides an excessively high amount of information, which results in parties choosing inefficient platforms and/or excessively low qualities.

1 Introduction

In 1994, during the national election campaign held in Belgium, the VLD (a traditionally right-wing party) promised to follow popular will. They organized a referendum at the national level (*het grote referendum*), and committed to defend the policy most favored by the majority of participants in the poll. The attempt failed spectacularly, and the VLD lost the election in favor of the left-wing coalition. Instead, in 1999, the VLD proposed what could be considered an excessively rightist platform from the median voter's viewpoint, and yet it became the main party in the governing coalition. The Belgium case is suggestive but by no means exceptional. In 1997, Tony Blair received an unprecedented support for a Labor Party, not because he was perceived as being more centrist than the Tories, but instead because he convinced the electorate that his "third way" policies were better than any of the other parties'. In Russia too, Vladimir Putin's agenda was by no means centrist. However, he persuaded voters that his strategy would bring quick economic recovery. These examples raise several questions. What makes the political platform of a party sound "attractive" or "convincing"? Why proposing the median voter's preferred platform may sometimes be a self-defeating political choice? More generally, why, contrary to standard predictions of the theory, do we rarely observe convergence of parties to the median voter's favorite platform?

In this paper, we develop a theoretical model that explains some of these observations. We want to understand why some platforms may sound more credible or attractive than others, and how this appeal –which we call "quality"– is linked to the ideological position of the platform. Introducing a quality dimension is a recurrent theme in the recent Political Economics literature (see e.g. Persson and Tabellini (2000, ch. 4) for a review). The novelty of our approach lies in the combination of two factors. First, parties are not exogenously different. Instead, the quality of the platform offered by each party is the result of some (costly) investment or "effort", under the anticipation that better platforms may help winning the election.[1] Second, ideological positions are perfectly observed whereas the quality of parties is only imperfectly observed by citizens. This captures the idea that, at the election time, voters are often uncertain about which program is more sensible. They have to rely on scarce information and decide to trust the party whose platform sounds more convincing. Using this approach, we identify some interesting interactions between observability of quality and the competition between parties.

We consider purely opportunistic (office-seeking) parties and assume that voters' preferences are public knowledge.[2] Voters care both about the ideology and the quality of the

[1] As usual in moral hazard contexts, "effort" to obtain "quality" should be interpreted in a broad sense. It can represent for example the resources spent by the party in order to form a suitable group, or the time dedicated to understand what are the needs of citizens and how to satisfy them, etc. A similar modelling has already been used in papers like Ferejohn (1986), Caillaud and Tirole (1999, 2001) or Persson, Roland and Tabellini (1997, 1998).

[2] In addition, to avoid possible non-transitivity of aggregate preferences, we assume that all voters share common preferences.

platform offered by each party. Even if these two dimensions are orthogonal, our results show that the amount of information revealed to the electorate about the second dimension before casting their ballot crucially affects the ideological position adopted by parties.

For the extreme cases already analyzed in the literature, namely when voters either (almost) always or (almost) never observe the quality of the platforms before the election, the standard Median Voter Theorem (Downs, 1956) holds: parties position themselves at the median voter's bliss point in the ideological dimension. Then, they compete in quality if this is perfectly observed by voters, or they neglect this dimension if voters are completely unable to discern the result of their effort. However, for the more general case of imperfect observability of the quality obtained by parties, the Median Voter Theorem does not hold anymore. Why would a party deliberately deviate from the median voter's preferred ideology? The point we make is that voters want parties to invest resources in increasing the quality of their platforms, but parties cannot commit to this. A party who offers an extremist ideology is handicapping himself: his only chance to win the election is to be sufficiently better than his moderate opponent (so as to compensate for his worse ideology). By adopting a non-centrist platform, he may increase his marginal return to effort and thus implicitly commit to provide a platform of higher quality than his rival. If this is the case, the extremist party will benefit from the "trust" of the electorate whenever qualities are not observed in equilibrium. In other words, we show that some party may use his ideological position to manipulate his own incentives to exert higher effort. As a result, voters may find it optimal to reward the party who puts himself in a difficult position (a non-median platform) whenever they are uncertain about the quality level obtained by parties. The mechanism works only if the information revealed prior to the election is neither "too small" (otherwise, the incentives to put effort are too weak in any ideological position) nor "too high" (otherwise, the extremist party never benefits from the trust of the electorate, which is his main reason for diverging in a first place). Convergence to the median voter's favorite platform is therefore a result that holds in extreme cases (no or full revelation of information) but not anymore in more general settings with imperfect revelation of information.

Two lessons can be drawn from these results. First, our theory explains why even purely opportunistic parties may appear "ideologically biased".[3] More importantly, the extra quality implied by the divergence of parties will sometimes offset their worse ideology proposed, in which case the voters' welfare under divergence is greater than under convergence of platforms. Second, our paper sheds light on the effects of information on political competition. Insofar as our "accuracy of information" parameter can reflect the result of an investment by the press to learn the quality of parties (and sell it to the electorate), our model measures the inefficiencies due to information being collected and provided by a privately interested press that does not maximize social welfare. Interestingly, we show that the press sometimes provides an

[3] See Poole and Rosenthal (1991,1997) for an empirical evaluation of parties' and candidates' political biases in the US since 1789.

excessively high amount of information, which results in parties choosing inefficient platforms and/or excessively low qualities.

Before proceeding to our formal analysis, we would like to mention some papers related to ours. There is an extensive literature on platform divergence and polarization. However, the literature has devoted surprisingly little attention to *understand* why the median voter theorem is rarely observed in practice. In most of the literature, divergence occurs because parties are assumed to have an exogenous "cost" of offering median platforms. This cost can be an intrinsic preference for some ideology (see Wittman (1983), Calvert (1985), Alesina (1988), Roemer (1997), Besley and Coate (1997), Caillaud and Tirole (1999,2001) and Rivière (2000) among others), or a reputation loss of changing the political discourse over time (Bernhardt and Ingberman, 1985).[4] However, under these circumstances, moderate parties should always dominate the political scene, with strategies like the one proposed by the VLD in Belgium being consistently successful. The papers most closely related to ours are Rogoff (1990) and Aragonés and Palfrey (2001). The first paper studies the optimal strategy of a high-ability incumbent with private information about his own capacity –namely his ability to finance a deficit with lower taxes. The paper shows that, in order to communicate his information to the electorate (and therefore increase the chances of reelection), the incumbent will use sub-optimal policies (excessive deficits), that a low-quality incumbent would not be able to implement. His model could also be reinterpreted in terms of a high-quality politician who handicaps himself by adopting an extremist platform in order to signal his capacity. The second paper shows that if one party enjoys an exogenous and known advantage over the other, then there is a unique mixed-strategy equilibrium in the positioning game with the favorite party adopting more moderate positions than his rival. In both cases, exogenous differences in quality (whether privately or publicly observed) are crucial to observe divergence. Moreover, the size of divergence is proportional to the quality differential.

The remainder of the paper is organized as follows. In Section 2, we present the model and some motivating examples. We also derive the equilibrium level of effort exerted, quality obtained and utility derived by parties conditional on their ideological choice. In Section 3, we characterize the equilibria of the game, including the optimal platform positions, as a function of the amount of information available to the public. In Section 4, we endogenize the provision information by assuming that it is collected by a profit-maximizing press. In Section 5 we discuss the robustness of our results and in Section 6 we conclude.

[4] Obviously, the goal of these papers is not to explain divergence only, but rather to use ideological preferences to understand other issues like the trade-off between ideology and probability of victory, the incumbency advantage, party formation, intra-party competition, or the evolution of a candidate reputation. See also Alesina and Rosenthal (1995, ch. 2) for a brief survey.

2 The model

2.1 Players, Information, and Platforms

There are two parties indexed by $i \in \{A, B\}$, and a homogeneous electorate with single peaked preferences. Voters care both about the *ideology* (or platform position) x_i of each party i and about the *quality* z_i of their proposed program. Denoting $V(x_i, z_i)$ the utility of voters if party i is elected, we have:

$$V(x_i, z_i) = \lambda f(x_i) + z_i \qquad \text{with } \lambda > 0, \tag{1}$$

where $f(x)$ is decreasing in $|x|$, so that the favorite platform of all voters is $x_i = 0$,[5] and λ represents the weight of pure ideology relative to platform quality. Note that a high-quality platform is always valuable for voters ($\partial V/\partial z_i > 0$), independently of the political platform of the party ($\partial^2 V/\partial x_i \, \partial z_i = 0$).

Parties on their side are purely opportunistic, which means that they do not have any ideology or intrinsic preference over platform positions.[6] We assume that the quality of the platform of each party can only take two levels: $z_i \in \{0, 1\}$. The probability that party i obtains a high-quality platform (i.e. $z_i = 1$), depends on the amount of his (unobservable and costly) effort e^i exerted. We formalize this as follows:

$$\Pr(z_i = 1) = z(e^i) \qquad \text{and} \qquad \Pr(z_i = 0) = 1 - z\left(e^i\right),$$

where $z'(e) > 0$ and $z''(e) \leq 0$ for all $e \geq 0$. This probability distribution is common knowledge, but neither parties nor voters know the realizations of platform quality before the election stage (which we define below). We denote by $U^i_{x_A x_B}(e^A, e^B)$ the expected utility of party i when the platforms selected are (x_A, x_B) and the efforts exerted are (e^A, e^B). We have:

$$U^i_{x_A x_B}(e^A, e^B) = \pi^i_{x_A x_B}(e^A, e^B) - \alpha(e^i), \tag{2}$$

where $\pi^i_{x_A x_B}(\cdot)$ is the (ex-ante) probability of party i being elected, and $\alpha(\cdot)$ is a function that represents the cost of effort, with $\alpha'(0) = 0$, $\alpha'(e) \geq 0$, and $\alpha''(e) > 0$ for all $e \geq 0$. This formalization accounts for the fact that (i) parties are only interested in winning the election, (ii) platform choices are costless, and (iii) quality choices are costly (as we will se below, the effort of parties influence their probability of being elected through their effect on the quality of platforms).

It is important to note right away that the model proposed here rules out any exogenous rationale for divergence. First, parties are a priori perfectly identical in the mind of voters,

[5] This framework is formally equivalent to the case in which voters have heterogeneous preferences but the position of the median voter is known and given by $x_i = 0$. The results also extend to the case of a median voter randomly situated around $x_i = 0$.

[6] This extreme assumption is debatable, to say the least. However, one of the goals of the paper is to show that ideology is not *necessary* to explain party divergence. Therefore, it is quite natural to adopt this most adverse situation.

who only value their ideological position and their quality (see (1)). As voters are also identical and with single peaked preferences, they have an unambiguous predilection for a moderate position and a high quality.[7] Second, we voluntarily assume that the quality of a party is valued by voters independently of the platform position. Third, and partly as a result of this, we rule out any exogenous interaction between quality and ideology on the side of parties: each party's probability of obtaining a high quality platform (the return to effort), as well as the cost of achieving a high quality (the cost of effort), are independent of both parties' platform positions and of the effort exerted by the rival.[8]

The quality of a platform can be evaluated for sure only once implemented. Nevertheless, voters may receive an informative signal about the quality of each party's platform before the election date (see below for the timing). More specifically, we assume that with some exogenous probability p and before the voting decision, they become perfectly informed about the quality of both platforms, while with probability $(1-p)$ they obtain no information at all.[9] Moreover, the value of p is known by all players of the game.

Let us insist on the fact that our model is not about multidimensional or non-transitive preferences.[10] As there is only one type of voter, aggregate preferences are always transitive, independently of the number of ideology dimensions in x_i.[11] For any given level of quality, centrist platforms are always Condorcet winners, and for any given ideological position, voters always prefer high quality platforms to low quality ones. The difference between the ideology and the quality dimensions is that the choice of the former is costless for parties whereas satisfying voters in the latter dimension is costly for parties. In the jargon of the Industrial Organization literature, parties can differentiate horizontally (ideology) and vertically (quality).

The introduction of vertical differentiation in political economy models is not new (see e.g. Berhardt and Ingberman (1985), Rogoff (1990), Feddersen and Pesendorfer (1997), Persson, Roland and Tabellini (1997), Polo (1998), Caillaud and Tirole (1999), Prat (2000), Aragonés and Palfrey (2001) among others). Instead, the novelty of our analysis lies in the *imperfect* observability of quality. In other words, our work embraces two extreme situations: the case in which the quality dimension is not considered (formally, $p = 0$), and the case in which quality and ideology cannot interact (formally, $p = 1$). Our contribution is to analyze what

[7] In Aragonés and Palfrey (2001), the incentives to diverge of both parties are triggered-off by the exogenous and known a priori disadvantage of one of them (and when the disadvantage goes to zero, parties converge).

[8] This avoids any exogenous trade-off from the parties' perspective between say, how desirable a given ideology is for voters vs. how costly it is for the party to reach a high-quality in this ideological position.

[9] This extreme signaling assumption (no information or perfect information) and the fact that voters cannot learn the quality of only one party is not necessary for the results. However, it greatly simplifies the analysis by reducing the number of cases to two. In Section 4, we endogenize p and show how it can be interpreted as the quality of the press.

[10] It is well-known that, under multidimensional preferences, an equilibrium in the choice of platform positions might not exist and, when it does, it may not imply full convergence (see e.g. Mueller (1989) and Ordeshook (1997) for a review). However, as the empirical estimates of Poole and Rosenthal (1991,1997) reveal, the dimensionality of the policy space is generally close to one, albeit for "exceptional times".

[11] If there are more than one ideological dimension, $|x_i|$ represents the norm of the policy vector x_i.

happens when endogenous interactions between the two dimensions are possible (which, given the assumptions of our game, occurs when $p \in (0,1)$).

In order to limit the number of cases, we assume that party A can only choose between a 'leftist' (L) and a 'centrist' (C) platform, whereas B chooses between a 'centrist' (C) and a 'rightist' (R) platform: $x_A \in \{L, C\}$ and $x_B \in \{C, R\}$.[12] Platform C is the most preferred by voters, and platforms L and R are positioned symmetrically around C. Therefore, $f(C) > f(L) = f(R)$. The relevant variable that we will use from now on is:

$$\Delta \equiv \lambda \Big[f(C) - f(L) \Big] \quad (> 0),$$

which corresponds to the differential in voters' utility between a party with a centrist platform and a party with an extremist platform (left or right), weighed by the importance attached to ideology relative to quality.

2.2 Examples

Before entering into the technical part of the model, we would like to sketch two stylized situations that fit into our framework. Obviously, these examples are only suggestive. Their main purpose is to clarify the meaning of an up to now abstract representation.

1. Taxes and expenditures. During their political campaign, two parties ($i \in \{A, B\}$) have to announce their policy x_i concerning how they plan to spend the revenues raised through taxation. For simplicity, there are three main possibilities. First, to increase the provision of public goods such as schools, hospitals, and libraries (L). Second, to focus on subsidies to agriculture (R). Last, to divide the budget equally between public goods and subsidies (C). All individuals value positively expenditures in both dimensions, however the median voter's favorite policy is the combination of expenses. Naturally, money needs to be raised in a first place. Since all expenditures are sensible, every individual strictly prefers a system where the cost of transferring public funds is as small as possible. This characteristic depends on how well the fiscal administration is designed, which in turn is affected by the amount of effort devoted by the party to this task (z_i in our model). Overall, voters are concerned on how money is distributed (the horizontal differentiation parameter x_i, where individuals have different preferences over the set of possibilities), and how efficiently the money is raised (the vertical differentiation parameter z_i, where everybody agrees on the importance of not wasting resources). Moreover, there is a priori no reason to think that voters with different preferences about optimal money redistribution will value differently the efficiency of the fiscal administration (formally, $\partial^2 V / \partial x_i \, \partial z_i = 0$).

2. Macroeconomic policy. Think of a country that suffers from macroeconomic problems. Both inflation and unemployment are high. One of the necessary reforms is the creation of an independent central bank, but its exact statuses are still not defined. The two parties

[12] Section 5 briefly discusses what happens when (i) we allow party A to adopt platform R and party B platform L and (ii) we consider a continuum of possible platforms.

competing for the election must decide which platform x_i to offer. The policy most desired by voters (C) is to have a central bank that pays equal attention to stabilization of output and stabilization of inflation. The left-wing alternative (L) is to concentrate on the former, and the right-wing alternative (R) is to concentrate on the latter. Independently of which alternative is proposed, lower inflation and lower unemployment can both be achieved simultaneously if the structural reform is of high quality (in economic terms, a good reform will affect the position of the Phillips curve). The quality of the reform thus depends on the effort exerted ($z(e^i)$), but not necessarily on the ideological bias of the central bank ($\partial^2 V/\partial x_i \, \partial z_i = 0$).

As the reader can notice, there are many other simple interpretations of our settings: type of foreign policy (x_i) and quality of diplomatic relations (z_i), type of reform for the Social Security or education system (x_i) and quality of the reform (z_i), etc.[13]

2.3 Timing

At this stage we can summarize the timing of the game played by parties and voters. First, parties simultaneously select the ideological position of their platform on the political spectrum (x_i), which becomes publicly observable. Second, parties simultaneously select their effort level, which determines their expected quality. Third, the electoral campaign takes place: the quality of the two platforms are realized and nature chooses whether both qualities are observed by voters or not.[14] Fourth, voters elect one candidate given the platform positions and the information they have on qualities (if any). This timing is depicted in Figure 1.

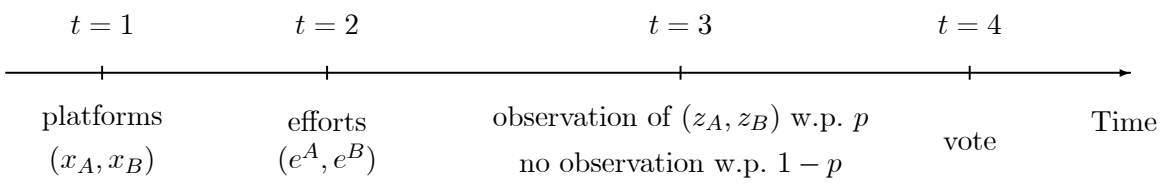

Figure 1. Timing.

Given that each party decides between two platforms, there are four potential pairs of positions: both parties located in the median voter's preferred platform (C, C), both parties in the extremes (L, R), and one extremist and one centrist party (L, C) and (C, R).[15]

[13] In some cases, a citizen may prefer a low-quality policy if the party in power has a different ideology (because, for example, this may undermine the party's reputation and therefore decrease its chances of reelection). We do not treat that case in the paper.

[14] By assumption, parties at this point have no instrument to credibly signal quality. Therefore, it makes no difference if we assume that parties always observe the quality of their own platform or that they learn their quality only if voters also do. Also, none of our qualitative results would be affected if we rather assumed that platforms and efforts were decided simultaneously (see point 1 in Section 5).

[15] We will often say "full convergence" when referring to (C, C), "full divergence" or "symmetric divergence" when referring to (L, R) and "partial divergence" or "asymmetric divergence" when referring to (L, C) and (C, R).

To focus on the interesting situation, we assume that the quality dimension is "sufficiently important", so that voters prefer a high quality party with an extremist platform rather than a low quality party with a centrist platform. This is summarized as follows.

Assumption 1 $\Delta < 1$.

Were this assumption violated, a party adopting an extremist position would automatically lose the election. Hence, platforms L or R could never be of potential interest and the unique possible equilibrium would imply full convergence to (C, C).[16]

Before studying the optimal behavior of parties and voters, note that if we define the problem in terms of the *expected quality* achieved for a given level of effort:

$$q^i = z\left(e^i\right),$$

then we can express the cost of achieving this level of quality as:

$$c\left(q^i\right) \equiv \alpha\left(z^{-1}\left(q^i\right)\right).$$

From the properties of $z(\cdot)$ and $\alpha(\cdot)$ defined in Section 2.1, we have $c'(0) = 0$, $c'(q) \geq 0$, and $c''(q) > 0$ for all $q \geq 0$. Using the above expression of costs, the expected utility of voters and parties given by (1) and (2) can be expressed in a simpler way as a function of platform positions and expected qualities q^i (instead of effort levels e^i):

$$V(x_i, q^i) = \lambda f(x_i) + q^i, \tag{3}$$

$$U^i_{x_A x_B}(q^A, q^B) = \pi^i_{x_A x_B}(q^A, q^B) - c(q^i). \tag{4}$$

As we will see below, using this indirect utility function considerably simplifies the exposition.

2.4 Parties' equilibrium effort, quality and payoff

As usual, we solve the model backwards. The voters' decision at $t = 4$ is simple. Given Assumption 1, if qualities are observed, then the median voter supports the party with highest-quality independently of his ideological position. If both parties achieved the same quality, then the most centrist party is elected. If both ideology and quality are equal, then voters are indifferent and elect each of them with probability 1/2. The same rule applies when qualities are not observed except that, in this case, voters rely on the expected quality of parties. Parties rationally anticipate this behavior of voters. Our next step is therefore to determine the parties' optimal choice of expected quality (date $t = 2$ of the game) given (i) the observable pair of platform positions (x_A, x_B) selected at $t = 1$, (ii) the existing uncertainty about the future revelation of information that will occur at $t = 3$ and the (iii) the anticipated behavior of voters at $t = 4$.

[16] Obviously, Assumption 1 is automatically satisfied if we consider a continuous policy space (see Section 5).

Denote by $q^i_{x_A x_B}$ the equilibrium choice of expected quality (we will often refer to "quality" for short) by party i when platform positions are (x_A, x_B). Similarly, $\hat{q}^i_{x_A x_B}$ denotes the quality of i *anticipated* by voters when the observed locations are (x_A, x_B) and the actual qualities remain unrevealed.[17] Last, $\kappa^i_{x_A x_B}(\hat{q}^A_{x_A x_B}, \hat{q}^B_{x_A x_B})$ denotes the probability that voters elect party i when qualities remain unobserved and platforms are located at x_A and x_B. Naturally, this probability depends on the anticipated qualities $\hat{q}^A_{x_A x_B}$ and $\hat{q}^B_{x_A x_B}$.[18] Given these definitions, we can compute the optimal quality chosen by parties at $t = 2$ for any pair of locations, as well as their corresponding payoffs.

Case 1. Symmetric positioning: (C, C) or (L, R).

As already noted, if both parties are located at the same distance from the median voter and their actual qualities are observed, the party with highest quality is elected with probability one. If both qualities are equal, each party is elected with probability one-half. This is true both under full convergence (C, C) and under full divergence (L, R). Therefore, using the subscript S (for "symmetric") to denote locations (C, C) or (L, R), we have the following expression for the parties' probabilities of being elected:

$$\pi^A_S = p\left[q^A\left(1-q^B\right) + \frac{q^A q^B + (1-q^A)(1-q^B)}{2}\right] + (1-p)\,\kappa^A_S(\hat{q}^A_S, \hat{q}^B_S) \quad (= 1 - \pi^B_S), \tag{5}$$

$$\pi^B_S = p\left[q^B\left(1-q^A\right) + \frac{q^A q^B + (1-q^A)(1-q^B)}{2}\right] + (1-p)\,\kappa^B_S(\hat{q}^A_S, \hat{q}^B_S) \quad (= 1 - \pi^A_S). \tag{6}$$

Given symmetric positioning, if qualities are not observed, voters prefer the party whose *anticipated quality* $\hat{q}^i$ is highest. Hence, $\kappa^A_S = 1$ if $\hat{q}^A_S > \hat{q}^B_S$, $\kappa^A_S = 1/2$ if $\hat{q}^A_S = \hat{q}^B_S$ and $\kappa^A_S = 0$ if $\hat{q}^A_S < \hat{q}^B_S$. From (4), (5) and (6), and taking first order conditions, we see that the optimal level of quality chosen by party i does not depend on voters' anticipation of quality $\hat{q}^i_S$ nor on the expected quality of the other party. Overall, if parties adopt symmetric platform positions at $t = 1$, then the *optimal* level of quality q_S that both parties will choose at $t = 2$ is unique and determined by:

$$c'(q_S) = \frac{p}{2}. \tag{C1}$$

where, given that $c(q) = \alpha(z^{-1}(q))$, we have $c'(q) = \alpha'(e)/z'(e)$. Note from **(C1)** that parties have incentives to look "intensively" for a good program (to get a high q_S) when effort has a substantial impact on the chances of obtaining a high quality platform (z' large), when the marginal cost of effort is low (α' small), and when the actual realization of qualities are frequently observed by voters (p large). Given rational anticipation of efforts, $\hat{q}^i_S = q_S$ and therefore, in equilibrium we have $\kappa^A_S = \kappa^B_S = 1/2$. The expected utility that, in equilibrium, both parties obtain under symmetric positioning is given by:

$$U^A_S(q_S, q_S) = U^B_S(q_S, q_S) = \frac{1}{2} - c(q_S). \tag{7}$$

[17] Obviously, in a rational expectation equilibrium, $\hat{q}^i_{x_A x_B} = q^i_{x_A x_B}$.

[18] Note that defining the equilibrium quality of both parties is redundant: given our symmetry assumption, we have for example that $q^B_{LC} = q^A_{CR}$. Besides, for any given effort levels and platform positions, we have $\pi^A_{x_A x_B} = 1 - \pi^B_{x_A x_B}$ and $\kappa^A_{x_A x_B} = 1 - \kappa^B_{x_A x_B}$.

As already mentioned, the optimal quality level selected by parties depends exclusively on the *relative* distance between each party's platform position and the median voter's ideology. Hence, q_S holds for *any* symmetric pair of platform positions. This in turn implies that symmetric divergence is unambiguously detrimental for the median voter: it implies a less desirable ideology and the same expected quality (we will elaborate on this point later on).

Case 2. Asymmetric positioning: (L, C) or (C, R).

Suppose that party A is more extreme than party B, i.e. $x_A = L$ and $x_B = C$ (the case $x_A = C$ and $x_B = R$ is identical). If the quality of both platforms is observed by voters then, given Assumption 1, the extremist party wins if and only if he obtains a high quality and his opponent a low one, i.e. if $z_A = 1$ and $z_B = 0$. Hence:

$$\pi^A_{LC} = p\, q^A(1-q^B) + (1-p)\, \kappa^A_{LC}(\hat{q}^A_{LC}, \hat{q}^B_{LC}) \quad (= 1 - \pi^B_{LC}), \tag{8}$$

$$\pi^B_{LC} = p\, \left[1 - q^A(1-q^B)\right] + (1-p)\, \kappa^B_{LC}(\hat{q}^A_{LC}, \hat{q}^B_{LC}) \quad (= 1 - \pi^A_{LC}). \tag{9}$$

In words, π^A_{LC} is the probability that party A (the "extremist") achieves a higher quality than his moderate opponent weighted by the probability that qualities become public, plus the probability that the quality of platforms remain unrevealed weighted by A's probability of being elected in that case. The probability π^B_{LC} that party B (the "moderate") wins is determined in a similar way, except that when qualities are observed and turn out to be equal, then B is preferred to his rival. From (4), (8), (9), and taking first order conditions, we deduce that if parties adopt asymmetric platform positions at $t = 1$, then the *optimal* level of expected quality chosen by parties at $t = 2$ is independent of the quality anticipated by voters in case of non-observability at $t = 3$. However, and contrary to the symmetric case, they now depend on the rival's choice of quality. Formally, expected qualities are given by:

$$c'(q_X) = p\,(1 - q_M)\,, \tag{C2}$$

$$c'(q_M) = p\, q_X, \tag{C3}$$

where subscripts X and M denote the equilibrium level of the variable for the "extremist" and the "moderate" party respectively (i.e. $q_X = q^A_{LC}$ and $q_M = q^B_{LC}$, $\pi_X = \pi^A_{LC}$, $\kappa_X = \kappa^A_{LC}$, and so on).[19] Again, note that if an equilibrium with asymmetric positioning exists, then as long as Assumption 1 holds, q_M and q_X will be selected by the moderate and extremist party independently of Δ, the relative distance between the two platform positions. Naturally, the equilibrium utility of parties will be also a function of κ_X and κ_M which themselves depend on the anticipated expected qualities. It is therefore necessary to determine whether q_X is greater or smaller than q_M.

Using **(C1)**, **(C2)**, and **(C3)**, we can rank the equilibrium levels of quality.

[19] By symmetry, we also have $q_X = q^B_{CR}$ and $q_M = q^A_{CR}$, $\pi_X = \pi^B_{CR}$, $\kappa_X = \kappa^B_{CR}$, and so on.

Lemma 1 *There exists a non-empty compact set $\mathcal{P}_X = (0, p_X)$ such that the extremist party has a strictly higher expected quality than the moderate party ($q_X > q_M$) if and only if $p \in \mathcal{P}_X$. More precisely:*

$$1/2 > q_X > q_S > q_M \quad \text{or} \quad q_X > 1/2 > q_M > q_S \qquad \text{if} \quad p \in \mathcal{P}_X$$
$$q_M > q_S > q_X > 1/2 \quad \text{or} \quad q_S > q_M > 1/2 > q_X \qquad \text{if} \quad p \notin \mathcal{P}_X$$

Proof. See Appendix A1. □

This result stresses the idea that there always exist some values of p for which an extremist *endogenously* chooses to offer a better platform than a moderate party. The intuition for the result is simple. Given that the median voter's favorite ideology is C, a party who deviates from that platform position is handicapping himself (i.e. becoming relatively less attractive to voters). When p is "small enough" ($p < p_X$), the moderate party has weak incentives to invest in quality (see **(C3)**). The extremist party also has relatively weak incentives to invest in quality. However, there is an asymmetry between the two parties: if actual qualities become public, the extremist is only elected if he has high quality and the moderate has low quality. Altogether, this implies that the marginal benefits of increasing quality are higher for the extremist than for the moderate (see **(C2)**). As a result, the former will work harder than the latter. Since voters do not integrate in their own welfare the parties' cost of effort, such a hard working attitude is always appreciated and, in equilibrium, compensated. Naturally, the opposite is true for p "high enough" ($p > p_X$).[20]

At this stage, it is useful to define the set of probabilities for which, in the absence of information on the actual qualities of platforms, voters strictly prefer an extremist party to a moderate one.

Definition 1 *For any Δ, we denote by $\mathcal{P}\,(\subset \mathcal{P}_X)$ the set of probabilities p satisfying $\Delta < q_X - q_M$. Formally, $p \in \mathcal{P} \;\Leftrightarrow\; \kappa_X = 1$ and $\kappa_M = 0$.*

$\mathcal{P}$ is thus the set of probabilities p such that if (i) an equilibrium with one centrist and one extremist party exists and (ii) voters do not observe the actual qualities of the platforms, then the extremist party is elected with probability one. In words, there is a trade-off between the higher expected *quality* of the extremist and the lower utility conferred by his proposed *ideology* ($q_X - \Delta$ vs. q_M). The set $\mathcal{P}$ thus corresponds to the case in which voters prefer to "trust" the extremist rather than the moderate party because the difference in expected qualities offsets the difference in ideologies. Given **(C2)** and **(C3)** one can immediately see that there always exists a threshold $\underline{p}$ (> 0) such that $[0, \underline{p}] \not\subset \mathcal{P}$. The idea is simply that if qualities are seldom revealed ($p < \underline{p}$), an extremist party is indeed more willing to provide a high-quality platform than a moderate one ($q_X > q_M$, as shown in Lemma 1). However, the incentives are very weak for both of them ($q_X \to 0$, $q_M \to 0$). Voters rationally anticipate this lack of incentives

[20] Note however that it may well be the case that $\mathcal{P}_X = (0, 1]$.

and therefore are not willing to bet on an extremist: his (slightly) higher expected quality does not compensate his less desirable ideology (formally, $p \in [0, \underline{p}] \Leftrightarrow 0 \leq q_X - q_M < \Delta$). Last, as q_X and q_M are independent of Δ, the set $\mathcal{P}$ will be smaller for larger values of Δ and $\mathcal{P} = \varnothing$ if Δ approaches 1.

Given Definition 1, we are finally in a position to determine the expected utility of parties in the asymmetric position case:

$$U_X(q_M, q_X) = \begin{cases} p\, q_X(1 - q_M) + (1 - p) - c(q_X) & \text{if} \quad p \in \mathcal{P} \\ p\, q_X(1 - q_M) - c(q_X) & \text{if} \quad p \notin \mathcal{P} \end{cases} \tag{10}$$

$$U_M(q_M, q_X) = \begin{cases} p\Big[1 - q_X(1 - q_M)\Big] - c(q_M) & \text{if} \quad p \in \mathcal{P} \\ p\Big[1 - q_X(1 - q_M)\Big] + (1 - p) - c(q_M) & \text{if} \quad p \notin \mathcal{P} \end{cases} \tag{11}$$

These will be used to determine the overall equilibrium of the game.

3 Characterization of the equilibria

3.1 The determinants of platform selection

In Section 2.4 we have determined the voting strategy of the electorate ($t = 4$) and the parties' optimal level of effort, quality and expected utility for every pair of platforms ($t = 2$). Working by backward induction, we can now determine the ideology optimally adopted by parties in the policy space ($t = 1$). This step will complete the analysis of the game.

Recall that (i) parties are purely opportunistic and dislike effort, (ii) voters dislike distance from the centrist platform, and (iii) the cost of achieving a given quality is independent of the platform position adopted. Therefore, it seems natural to expect that parties will always select the platform most preferred by the median voter, that is $x_A = x_B = C$. Yet, the existence of imperfect albeit symmetric information on the quality dimension affects the incentives for effort in the different platform positions, and therefore has an impact on the location strategy of parties. Our first result consists of determining a sufficient condition for the median voter theorem to hold.

Proposition 1 ***(Conditions for Median Voter Theorem)***
A sufficient condition for (C, C) to be the unique equilibrium in the choice of ideology by parties is $p \notin \mathcal{P}_X$. Hence, $q_X > q_M$ is a necessary condition for divergence.

Proof. See Appendix A2. □

According to Lemma 1, the reasons for which parties might be willing to offer non-centrist platforms can be of two different natures. If $p \in \mathcal{P}_X$, a deviation from (C, C) constitutes an implicit commitment to increase the quality of the platform (formally, $q_X > q_S$). Such deviation is thus costly on the effort side but will also end up increasing the probability of

getting elected (since, in equilibrium, $q_X > q_M$). We call it "deviation for quality". If $p \notin \mathcal{P}_X$, then a deviation from (C, C) induces the party to reduce his quality, with the corresponding effort-saving (formally, $q_X < q_S$). However, it also implies a smaller probability of being elected (since, in equilibrium, $q_X < q_M$). We call it "deviation for laziness". Proposition 1 shows that *only deviation for quality may occur*. Indeed, when a party deviates for laziness, the benefits of lower effort do not compensate for the decrease in his probability of being elected. This result stems from the fact that voters never directly take this cost of effort into account when they decide to elect one party or another. Only ideology and quality matter. Hence, our first conclusion is that if we ever observe one extreme and one moderate platform position, then we know for sure that the extremist party will look more intensively for a high-quality platform than his moderate rival.[21]

In light of this Proposition, it only remains to derive the platform positions adopted in equilibrium by parties when $p \notin \mathcal{P}_X$ (i.e. when $q_X > q_M$). In order to restrict the number of cases to analyze, we introduce the following technical Assumption that will be maintained throughout the rest of the paper.

Assumption 2 *The functions $z(e)$ and $\alpha(e)$ are such that $c'''(q) \geq 0$ for all $q > 0$.*[22]

Imposing convexity of the marginal cost of quality is just a convenient way of keeping all equilibrium qualities within a certain range. In other words, this assumption rules out a situation in which the qualities of parties are all very low except for that of an extremist party, which becomes very high when confronted to a moderate one.[23]

We are now in a position to complete the first stage of the game, which consists of a complete characterization of the optimal ideology adopted by parties in the policy space at $t = 1$ as a function of the probability p of qualities being revealed. Naturally, once the optimal policy platforms are determined, the qualities chosen by parties at $t = 2$ are simply given by **(C1)**, **(C2)** and **(C3),** and their corresponding utilities are determined by (7), (10) and (11).

[21] In other words, the fact that deviation for quality can occur but deviation for laziness cannot is a result of our model and not something we have imposed.

[22] Note that $z''' \leq 0$ and $\alpha''' \geq 0$ are sufficient conditions to ensure that $c''' > 0$. More generally, $c''' \geq 0$ requires that α''' be "not too small" relative to z'''. The exact formulation is:

$$c'''(q) = \frac{\alpha'''(e)}{z'(e)} - \frac{\alpha'(e)z'''(e)}{[z'(e)]^2} + \frac{2}{[z'(e)]^2}\left(\frac{\alpha'(e)[z''(e)]^2}{z'(e)} - \alpha''(e)z''(e)\right) > 0.$$

[23] Technically, Assumption 2 combined to **(C1)** and **(C2)** implies that q_X cannot be larger than $2q_S$. As informally explained in Section 5, most of the results still apply when Assumption 2 does not hold. Note that restrictions on the rate of convexity of cost functions, although difficult to interpret economically, are quite frequent in contract theory as technical devices to avoid non-convexities in the overall maximization problem (see e.g. the classical papers of Guesnerie and Laffont (1984) or Laffont and Tirole (1986)).

Proposition 2 ***(Complete characterization of equilibria)***
There exist two non-empty sets $\mathcal{P}_1$ and $\mathcal{P}_2$ such that:

- (C, C) *is an equilibrium if and only if* $p \in \mathcal{P}_1 \cap \mathcal{P}$ *or* $p \notin \mathcal{P}$*;*
- (L, R) *is an equilibrium if and only if* $p \in \mathcal{P}_2 \cap \mathcal{P}$*;*[24]
- (L, C) *and* (C, R) *are both equilibria if and only if* $p \in \mathcal{P} \setminus (\mathcal{P}_1 \cup \mathcal{P}_2)$*.*

Moreover, the equilibrium choice of platforms has also the following properties:

(i) There exist two values $\underline{p}$ (> 0) *and* $\overline{p}$ (< 1)*, such that* (C, C) *is the unique equilibrium if* $p \in [0, \underline{p}] \cup [\overline{p}, 1]$*.*

(ii) For Δ *sufficiently small, there always exist values of* p *such that* (L, R) *is the unique equilibrium.*

Proof. See Appendix A3. □

A first glance at Proposition 2 reveals that the median voter theorem generally does not hold under imperfect observability of quality. By contrast, convergence to (C, C) always occurs when either quality is almost an irrelevant issue (p close to zero) or quality cannot interact with the platform choice (p close to one). However, the reasons for convergence in these two extreme cases are of very different nature.

First, we have shown that for values of $p \in \mathcal{P}_X$, parties are willing to exert more effort the higher their *relative* distance to the median voter. Nevertheless, if p is sufficiently small (quality is rarely observed) the difference in incentives is very weak as witnessed by **(C2)** and **(C3)**. Voters anticipate this, and therefore are not willing to support an extremist candidate in case of not being informed about qualities. In other words, when p is small enough, voters prefer centrist parties because the loss in the ideology dimension associated to divergence always offsets the gain of a "slightly" higher expected quality. As a result, no party has incentives to diverge in a first place (technically, if p is close to 0, then $p \notin \mathcal{P}$). The case $p \to 0$ thus corresponds to the standard Hotelling model of voting in which the quality dimension is not taken into account (why would anyone exert a costly effort to increase quality if it can never be observed by the electorate?). In this case, voters have no reason to support extremism and then parties have no reason to diverge.

Second, recall from Proposition 1 that a party may only be willing to adopt an extreme position if it serves as a commitment to high quality. This "deviation for quality" yields benefits precisely when information about the realized quality of parties does not become public since, in that case, voters rely on anticipated efforts (i.e. on expected qualities) and optimally choose to support the extremist party. Therefore, if the probability p that information becomes public is sufficiently high, the likelihood that an extremist party benefits from the confidence of voters is so low that he does not have enough incentives to adopt this position (technically, if p is close to 1, then $p \in \mathcal{P}_1$). The case $p \to 1$ thus corresponds to another standard model studied in the literature, in which quality matters (and it is a choice variable)

[24] This implies that if $\mathcal{P}_1 \cap \mathcal{P}_2 \neq \varnothing$, then (C, C) and (L, R) can both be equilibria of the game.

but convergence still occurs because no endogenous interaction between incentives to effort and choice of platform is allowed; voters are willing to trust extremist parties, but parties do not have enough incentives to diverge in a first place.

To sum up, as p increases, two effects operate in opposite directions. First, if the quality of platforms is not revealed, voters are relatively more willing to support extremist parties. Second, parties are less willing to deviate from centrist positions because they have a lower probability of taking advantage of this trust. Overall, convergence occurs for p sufficiently small because voters do not benefit enough from the extra quality of extremist platforms, and it occurs for p sufficiently large because parties do not reap enough rents from the extra effort implied by divergence.

One contribution of this paper is to show that the median voter theorem always applies to these extreme cases ($p = 0$ and $p = 1$) but it does not necessarily hold when $p \in (0, 1)$. In fact, when the quality of parties is neither fixed nor perfectly observed, then there is scope for *endogenous* interactions between the ideological position and the incentives to increase the quality of the platform. In turn, these interactions affect the strategy of location by parties and, in particular, may induce them to optimally select extremist positions in a first place. Which situation prevails (convergence, partial divergence or full divergence) will crucially depend, not only on the degree of imperfect information p, but also on the shape of the cost function $c(\cdot)$. Note also the existence of multiple equilibria for some parameter configurations: whenever (L, C) is an equilibrium then (C, R) is another one, and if $\mathcal{P}_1 \cap \mathcal{P}_2 \neq \varnothing$ then (C, C) and (L, R) may also coexist. Last, divergence is not just a theoretical curiosity: as part (ii) of the proposition shows, if the amount of divergence can be sufficiently small or if the weight of platform quality relative to ideology is sufficiently strong (in both cases, if Δ is small enough), then divergence always occurs for some levels of information revelation.

Proposition 2 shows that even if parties are ex-ante identical, the one who decides to propose an extremist ideology will, in equilibrium, exert more effort and offer a political platform of higher expected quality than his rival. This contrasts with previous works on political divergence where parties (or candidates) have private information about their intrinsic quality or ability and the best one decides to signal his superiority by handicapping himself and offering a non-centrist ideology (see e.g. Rogoff, 1990). Although apparently similar, these two approaches are indeed quite different. According to the traditional signaling model, extremist leaders like Le Pen in France or Perot in the US are the most competent candidates in the political arena. They propose a non-centrist ideology only to prove to the public that they can afford such a wasteful choice. According to the moral hazard model proposed in this paper, Le Pen and Perot are not more competent than their moderate opponents. However, once they have announced an extremist ideology, they do work harder to convince the electorate about the value of their ideas.[25] The paper also offers an explanation of why the VLD's

[25] Needless to say that, in reality, there is always some component of intrinsic ideology in every extremist political leader. However, this will only reinforce the willingness to diverge (see point 6 in Section 5).

strategy of promising to blindly follow popular will did not receive the support of the Belgian electorate. In the absence of strong evidence of the contrary, voters may have perceived the party's absence of ideological commitment as a sign of their inability to offer sensible solutions to the political needs of the citizens. Last, our results have also some implications for the stylized examples presented in Section 2.2. Consider a party who announces a concentration of expenditures on public goods (example 1) and a macroeconomic policy with output stabilization as its main target (example 2). Propositions 1 and 2 suggest that this party will have greater incentives to raise money efficiently (example 1) and to find a coherent structural reform (example 2) than a party who offers the platform most desired by the median voter. This higher incentives will be reflected in the quality obtained in equilibrium.

Once we have studied the conditions for divergence, it is easy to determine the welfare impact of the different choices of political platforms.

Corollary 1 *(i) (C, C) always Pareto dominates (L, R).*
(ii) For some values of p, (L, C) and (C, R) Pareto dominate (C, C).

Recall that the expected quality incurred by parties depends exclusively on their *relative* degree of extremism ($q^A_{CC} = q^A_{LR} = q_S$). When their incentives to deviate are too strong, both of them offer extremist positions (L, R). This decreases the welfare of voters (as it implies a cost in terms of ideology and the same expected qualities than under full convergence) without altering the welfare of parties.[26] By contrast, if only one party adopts an extreme position, then both parties may be induced to increase their quality ($q_X > q_M > q_S$ for some $p \in \mathcal{P}$, see Lemma 1). Hence, when the gains of a higher expected quality offset the costs of a less desirable ideology, voters do benefit from partial divergence. Moreover, if an equilibrium with asymmetric positioning exists, then both parties are necessarily better-off under this pair of platforms than either under full convergence or under full divergence.[27]

Remark 1. The fact that when $p \in \mathcal{P} \setminus (\mathcal{P}_1 \cup \mathcal{P}_2)$, then both parties strictly prefer (L, C) or (C, R) rather than (C, C) or (L, R) has important consequences. It implies that the equilibrium with asymmetric positions is quite robust: even if parties could collude or communicate at the platform selection stage, they would still choose the partial divergence outcome.

Remark 2. Our model shares some features with the well-established literature on career-concerns (see the seminal paper by Holmström, 1999). As in Holmström's paper, the efforts of agents (here parties) stochastically affect output (here quality of platforms) which is noisily observed by the principal (here the median voter). As in a rat race, parties are trapped by the

[26] We do not want to excessively emphasize this result, since it hinges upon the restricted number of platform positions that can be offered. In Section 5 we argue that, without this restriction, greater divergence is always associated to higher effort and therefore higher expected quality.

[27] To understand the formal argument suppose that (C, R) is an equilibrium of the game. By construction $U^B_{CR}(q^A_{CR}, q^B_{CR}) > U^B_{CC}(q_S, q_S)$, i.e. B prefers (C, R) to (C, C) (otherwise he would deviate). Furthermore, also by construction, $U^A_{CR}(q^A_{CR}, q^B_{CR}) > U^A_{LR}(q_S, q_S)$, i.e. A prefers (C, R) to (L, R). Since $U^A_{LR}(q_S, q_S) = U^A_{CC}(q_S, q_S)$, this means that party A is indifferent between (C, C) and (L, R). Hence, A also prefers (C, R) to (C, C).

effort anticipated by voters. The key novelty of this paper is that the level of effort in which parties are trapped in, is a choice variable. Indeed, by selecting a particular (observable) policy platform at $t = 1$, parties are implicitly committing to exert a given amount of effort at $t = 2$, and therefore to offer a specific expected quality. This expected quality is rationally anticipated by voters and affects their election strategy at $t = 4$.

3.2 A numerical example

To illustrate these results, consider the following example. Let $c(q) = a\,q^2/2$.[28] In equilibrium and for values of a such that interior solutions exist, we have:

$$q_S = \frac{p}{2a}; \quad q_X = \frac{ap}{a^2 + p^2}; \quad q_M = \frac{p^2}{a^2 + p^2}.$$

Notice that $q_X > q_M \;\Leftrightarrow\; a > p$ and therefore $\mathcal{P}_X = (0, a)$. Defining $\Delta_{\max}(p) \equiv (q_X - q_M)$, the necessary condition for divergence $p \in \mathcal{P}$ is then:

$$\Delta < \Delta_{\max}(p) = \frac{p\,(a - p)}{a^2 + p^2}.$$

Figure 2 illustrates the equilibrium value of $\Delta_{\max}$ for $a = 1$ and $a = 2/3$ respectively.

[Insert Figure 2 about here]

The sets $\mathcal{P}_1$ and $\mathcal{P}_2$ as defined in Proposition 2 are such that $\mathcal{P}_1 = [p_1, 1]$ and $\mathcal{P}_2 = [0, p_2]$ with $p_1 < p_2$ for all a (see Appendix A2, Step 2 for the functional derivation of these values). Since $\mathcal{P}_1 \cup \mathcal{P}_2 = [0, 1]$, parties can never adopt asymmetric positions in equilibrium. Figure 3 depicts the boundaries of the sets $\mathcal{P}$, $\mathcal{P}_1$ and $\mathcal{P}_2$ for different pairs (p, a) and two values of Δ (0.05 and 0.15). The bounds of $\mathcal{P}$ are given by the left-most and right-most solid (resp. dotted) curves when $\Delta = 0.05$ (resp. $\Delta = 0.15$). One can immediately observe that $\mathcal{P}$ shrinks as Δ increases. The left and lower-right areas thus correspond to the equilibrium (C, C) when $\Delta = 0.05$ as, in these regions, $p \notin \mathcal{P}$. Next, the left (solid) curve inside $\mathcal{P}$ displays p_1 as a function of a and the right (solid) curve represents p_2 as a function of a. Hence, within $\mathcal{P}$ and according to Proposition 2, in the area to the right of p_2 the unique equilibrium is (C, C), and in the area to the left of p_1 the unique equilibrium is (L, R). Finally, inside the region between these two values, both (C, C) and (L, R) are equilibria.

[Insert Figure 3 about here]

Finally, Figure 4 displays the voters' expected utility as a function of p when $a = 1.5$ and $a = 3$ (the dotted lines account for the fact that there are two possible levels of welfare when (C, C) and (L, R) coexist). Clearly, for given platform positions, voters' welfare increases as the probability p of revealing information increases and as the cost of effort a decreases (both

[28] This could for example result from $z(e) = e$ and $\alpha(e) = a\,e^2/2$ or from $z(e) = \sqrt{e}$ and $\alpha(e) = a\,e/2$. This functional form satisfies Assumption 2.

measures affect positively the expected quality reached by parties). However, the endogeneity of platform positions implies that marginal increases in p or marginal reductions in a may trigger-off the decision of parties to diverge and end up being detrimental to voters.

[Insert Figure 4 about here]

4 The quality of information

4.1 Time inconsistency

The results of Section 3 shed light on two important aspects of the relationship between the welfare of voters and the probability that qualities be revealed. First, from **(C1)**, one can notice that, under symmetric positioning, efforts are increasing in p. Therefore, for given platforms, a higher probability of revelation of qualities is always beneficial to voters. However, these policy choices are endogenous. If p is close to $\underline{p}$, a marginal increase in p may trigger-off the incentives of parties to deviate from (C, C) to (L, R) which would reduce the welfare of the population (see for example the discontinuities of the voters' welfare function in Figure 4). In other words, the ability of parties to select their platforms may sometimes create a *time-inconsistency problem for voters*: before the announcement of platforms, individuals want to commit not to pay "too much" attention to information about qualities (p "low") in order to ensure the choice of centrist positions. But once platforms are fixed, this information will never be disregarded as it enables voters to make a better choice between the two contenders.

Perhaps more surprisingly, *parties may also face a time-inconsistency problem.* Consider for instance the case in which p is close to one (both parties adopt a centrist position) but the set $\mathcal{P} \setminus (\mathcal{P}_1 \cup \mathcal{P}_2)$ is not empty ((L, C) and (C, R) become the equilibria if p decreases sufficiently). For the initial value of p, both parties ex-ante strictly prefer p to be reduced. Indeed, as the probability of quality revelation decreases, either (C, C) persists as the unique equilibrium (in which case parties maintain a probability of winning of 1/2, but save on effort) or asymmetric positions (L, C) and (C, R) become optimal and, by revealed preferences, both parties also benefit from it (see Remark 1). However, once an asymmetric pair of platforms is adopted, parties have opposite incentives. The centrist party can only win when qualities are revealed, so he ex-post prefers p to be as high as possible, which creates his time-inconsistency problem. Conversely, the extremist party is elected for sure when qualities remain unknown, so he still prefers p to be as low as possible. It is interesting to note that a party adopts an extremist position only as a commitment to exert high effort, but then favors the situation in which the result of his endeavor does not become public.

4.2 Endogenous quality of information and the role of the press

One could think of p as representing the *quality of the press.* The results presented so far thus show how the press affects the political competition both in the ideology and the quality dimensions. The purpose of this section is to endogenize the role of the press. More precisely,

we study what type of inefficiencies are likely to emerge if an independent and privately interested press is in charge of collecting the information about the quality of parties.[29]

To analyze this issue, we extend the game represented in Figure 1 as follows. At date $t = 0$, that is before platforms are selected by parties, an independent and profit maximizing press decides how much to invest in learning the future quality of platforms. In order to keep the model as close as possible to that of Sections 2 and 3, we assume that by paying a cost $\gamma(p)$ (where $\gamma'(0) = 0$, $\gamma'(p) > 0$ and $\gamma''(p) > 0$ for all $p > 0$) the press will learn at $t = 3$ the quality of both parties with probability p and it will get no information with probability $1 - p$.[30] The press does not derive any direct value from such information. However, it can use it to extract resources from the electorate. One can think for instance of the number of newspapers sold as a proxy for the profits generated by the news collected. Let us assume that the press is able to extract the entire voters' surplus generated by the information. Naturally, this maximizes its incentives to learn the quality of parties.[31]

Two things must be noted upfront. First, the press is subject to the classical hold-up problem. Consider for example the case of symmetric positioning: voters benefit from a higher investment in information by the press at $t = 0$ because it encourages higher quality by parties at $t = 2$ (see **(C1)**). However, at $t = 3$ this investment is sunk so, at that point, the electorate will never remunerate the press for having increased the expected quality of parties. Instead, the press will only obtain rents from its ex-post ability to discern which party (if any) has succeeded in obtaining a high-quality. Second, due to this same hold-up problem, the press will not fully internalize the effects of p on the choice of platform positions by parties (for example the press will extract the same rents at $t = 3$ if the positions adopted are (C, C) or (L, R)). Hence, we can immediately see that a self-interested press is likely to induce a suboptimal platform choice from the voters' viewpoint. Since this inefficiency is straightforward and has already been discussed in Sections 3.2 and 4.1 we will leave it aside. Instead, we will concentrate on the effect that information collection has on the *quality* (not the positioning) of parties.

Suppose that at date $t = 1$ parties optimally adopt symmetric positions and denote by $\Pi_{sym}(p)$ the expected gross benefit that the press can extract from voters given that, at $t = 3$, it will learn the quality of parties with probability p. We have:

$$\Pi_{sym}(p) = p \times 2q_S\,(1 - q_S) \times \frac{1}{2}, \tag{12}$$

which reads as follows. The press can only make profits on the event that it observes qualities –which happens with probability p– and that exactly one party achieves high quality –which

[29] See Strömberg (2001a) for an empirical analysis of the links between political competition and the actions of a profit maximizing press.

[30] Once again, this extreme structure of information revelation is made only for expositional convenience.

[31] This assumption is not necessary. However, it provides a useful benchmark for comparison between private and social value of information. Alternative specifications could then be easily incorporated in this framework.

happens with probability $2q_S(1-q_S)$.[32] In that case, voters are ready to pay 1/2 to obtain this information, which is the utility difference between obtaining a quality $z_i = 1$ with probability one instead of one-half.

Quite naturally, this net profit of the press will differ from the social value of information. If transfers between voters and press are costless and parties again choose symmetric platforms, the social value of acquiring information with probability p is simply the utility derived by voters given this probability of being informed. Formally, the gross social benefit $G_{sym}(p)$ is:

$$G_{sym}(p) = p\left[1-(1-q_S)^2\right] + (1-p)\,q_S - \Delta \times \mathbb{1}_{(L,R)} \tag{13}$$

$$= \Pi_{sym}(p) + q_S - \Delta \times \mathbb{1}_{(L,R)} \tag{14}$$

where $\mathbb{1}_{(L,R)} = 1$ under platforms (L,R) and $\mathbb{1}_{(L,R)} = 0$ under platforms (C,C).

Assuming the existence of an interior solution, a cost $\gamma(p)$ of gathering news, and (as mentioned before) the same platform choice in both cases, then the first-order conditions of (12) and (14) determine the optimal amount of information collected by the press $\tilde{p}_{sym}$ and its socially optimal level p^*_{sym}:

$$\Pi'_{sym}(\tilde{p}_{sym}) = \gamma'(\tilde{p}_{sym}) \qquad \text{and} \qquad \Pi'_{sym}(p^*_{sym}) + \left.\frac{d\,q_S}{d\,p}\right|_{p=p^*_{sym}} = \gamma'(p^*_{sym}). \tag{15}$$

The analysis changes when the information collected induces an asymmetric choice of platforms by parties. Recall that, in this case, voters support the extremist party in the absence of evidence about qualities. Hence, information is valuable to voters every time the moderate party is of high quality or the extremist party is of low quality. This gives an expected gross benefit for the press $\Pi_{asy}(p)$ equal to:

$$\Pi_{asy}(p) = p\Big[q_M\,q_X\,\Delta + (1-q_M)(1-q_X)\,\Delta + q_M(1-q_X)\,(1+\Delta)\Big] \tag{16}$$

which reads as follows. Δ represents the increase in payoff of electing a moderate rather than an extremist party when both are of high quality (which occurs with probability $q_M\,q_X$) or both are of low quality (which occurs with probability $(1-q_M)(1-q_X)$). $1+\Delta$ represents the increase in payoff of electing a moderate rather than an extremist party when the former is of high quality and the latter is of low quality, which occurs with probability $q_M(1-q_X)$. Using the same principles as before, we can also determine $G_{asy}(p)$, the gross social utility under asymmetric positions:

$$G_{asy}(p) = p\left[q_M + (1-q_M)q_X\,(1-\Delta)\right] + (1-p)\Big[q_X - \Delta\Big] \tag{17}$$

$$= \Pi_{asy}(p) + q_X - \Delta \tag{18}$$

From (16) and (18), the optimal amount of information collected by the press and its socially optimal level in the asymmetric positioning case, $\tilde{p}_{asy}$ and p^*_{asy}, are given by:

$$\Pi'_{asy}(\tilde{p}_{asy}) = \gamma'(\tilde{p}_{asy}) \qquad \text{and} \qquad \Pi'_{asy}(p^*_{asy}) + \left.\frac{d\,q_X}{d\,p}\right|_{p=p^*_{asy}} = \gamma'(p^*_{asy}). \tag{19}$$

[32] Put it differently, if the two parties have the same quality, voters will get the same payoff whether they are informed at the voting stage or not.

Finally, using (15) and (19), we can compare the level of information disclosed and the corresponding qualities reached by parties under private $(\tilde{p}_{sym}, \tilde{p}_{asy})$ and socially optimal (p^*_{sym}, p^*_{asy}) collection of information, given the same platform choice of parties.

Proposition 3 ***(Role of the press)***

In the equilibrium where parties adopt symmetric positions, we have: $\tilde{p}_{sym} < p^*_{sym}$ *and* $q_S(\tilde{p}_{sym}) < q_S(p^*_{sym})$. *In the equilibrium where parties adopt asymmetric positions, we have:* $\tilde{p}_{asy} \lesseqgtr p^*_{asy}$, $q_X(\tilde{p}_{asy}) < q_X(p^*_{asy})$ *and* $q_M(\tilde{p}_{asy}) \lesseqgtr q_M(p^*_{asy})$.

Proof. See Appendix A4. □

The intuition behind this result is rather straightforward. A profit-maximizing press will not implement the social optimum, because markets do not internalize the entire social value of information. Among other things, recall that the mere *possibility* of information revelation influences the parties' investment in quality. This is valuable to voters be qualities actually revealed or not. However, the investment in information collection by the press is sunk at the time of claiming any benefits (i.e. at $t = 3$). Therefore, the press will only be compensated for their endeavor when qualities are revealed. As a result, it will not have any incentive to internalize the positive externality of information on the quality of parties.

Probably the most surprising result in Proposition 3 is the idea that, in this variation of the hold-up problem, the press may spend an *excessively high* amount of resources in acquiring information when parties adopt asymmetric positions. From (16) and (18), one can identify the crucial difference between the social and the private value of information: in the former case, an increase in the quality of the extremist is directly valued, whereas in the latter case it is not (this occurs because the extremist party is always elected whenever qualities remain unknown). Since the extremist's quality can be decreasing in p, the press may end up collecting such a high level of information that it eventually discourages the extremist to look intensively for a good platform, with the resulting adverse effect on voters' welfare.[33]

5 Robustness

In the analysis previously conducted we have kept the simplest possible structure, at the expense of some unrealistic hypotheses. The purpose of this section is to discuss the robustness of our results when we relax some of the most restrictive assumptions.

1. Timing. Often in the Industrial Organization literature, the timing of the actions taken by the different agents is crucial to determine the outcome of the game. It is therefore important to justify the sequencing adopted. We have assumed that parties choose how much effort to

[33] Even if the decision by the press to collect information takes place between the selection of platforms and the selection of effort (i.e. between $t = 1$ and $t = 2$), we still have an inefficiency in the parties' choice of both platforms and qualities. By contrast, if information is collected after the parties' selection of effort (i.e. between $t = 2$ and $t = 3$), then q_i is fixed at the time of choosing p, i.e. $dq_i/dp = 0$. As a result, and given (15) and (19), one can see that in this case the social inefficiency vanishes.

exert after having selected their ideological position (see the timing in Figure 1). However, one might think that parties choose their platforms and effort levels simultaneously. In a previous version of the paper (available upon request) we analyzed this possibility. We showed that considering a simultaneous choice of platforms and efforts only changes the set of probabilities that determine the regions of convergence, partial divergence and full divergence, without altering the main qualitative insights: for p sufficiently close to either 0 or 1, convergence is still the unique equilibrium, whereas for interior values of p, partial and full divergence can be sustainable and imply a commitment to effort.

2. Unrestricted choice of platforms. The incentives of parties to exert effort depend exclusively on their *relative* degree of extremism. In our model, this implies that (L, R) is strategically equivalent to (L, L) and (R, R). Hence, if we allowed party A to adopt platform R and/or party B to adopt platform L, then we could reach the unusual (and rather counter-intuitive) equilibrium situation in which the median voter has centrist preferences and, still, both parties adopt the same extremist platform. The reader might view this as a potential weakness for the predictive power of our model. However, we think it is not. In fact, it is easy to show that if we extend our basic model and assume that (i) the median voter is randomly and symmetrically located around C, and (ii) voters are risk-averse with respect to policy platforms ($f''(x) < 0$), then an equilibrium in which both parties share an extremist platform is no longer sustainable.[34]

3. Continuous platform space. Probably the most natural extension of our basic model would consist of allowing parties to select their platforms from a continuous ideological space: $x_A \leq C$ and $x_B \geq C$. A full characterization of the solution in this case is certainly interesting, but beyond the scope of this paper. Still, using a heuristic approach, one can see that this more comprehensive modelling of the political game would not substantially modify our current results. First, allowing arbitrarily small moves in the ideology space is formally equivalent to $\Delta \to 0$ in the discrete case. This means that, for any positive but sufficiently small p, an equilibrium with full convergence cannot exist (see Proposition 2) whereas for values of p sufficiently high (so that $p \in \mathcal{P} \cap \mathcal{P}_1 \backslash \mathcal{P}_2$) or for $p = 0$, (C, C) remains the unique equilibrium. Second and obviously, most of our results directly extend to a continuous policy space if we impose an exogenous and sufficiently tight upper bound on the level of admissible divergence. Third and most interestingly, if we do not impose such bound, one can check that there

[34]The idea is the following. Consider the equilibrium (L, R). If the median voter is sufficiently extremist, then only ideology matters (and not quality), which means that the effort of parties is wasted. In contrast, under locations (L, L) or (R, R) quality is always important no matter how extremist the median voter is. In other words, by selecting platforms (L, R) rather than (L, L) or (R, R), parties keep the same equilibrium probability of winning (1/2) but reduce their (costly) level of effort, since they anticipate that it can be useless with some probability.

This result holds even if there is an *arbitrarily small* probability that the median voter be extremist. In fact, the technical condition to avoid the equilibria (L, L) and (R, R) is that the median voter's preferred platform x^* must satisfy with positive probability the following inequality: $\lambda|f(x^* - L) - f(x^* - R)| > 1$. With random (and unbounded) position of the median voter, $f''(x) < 0$ is just a sufficient condition for this inequality to hold.

cannot exist a pure strategy equilibrium: parties will find it optimal to randomize between a continuous set of platforms, anticipating a higher expected level of effort the more extremist their platform selected. Moreover, in this case, the highest degree of extremism chosen in equilibrium becomes endogenous, but still remains bounded.[35]

4. Generalization of the cost and quality functional forms. Assumption 2 only plays a role to show that (L, R) cannot be an equilibrium when $p \to 1$. The idea is the following. Suppose that q_S and q_M are close to 0 and q_X is close to 1 (a situation that can only occur if Assumption 2 is violated). If parties expect (L, R) to be the equilibrium, then no party is willing to moderate the platform as it would imply a tremendous increase in effort by the opponent that would more than compensate his worse (extremist) ideology. Hence, in this specific situation, both (C, C) and (L, R) could be equilibria of the game. However, the result heavily draws on the binary structure of platform positions and qualities.

5. Increasing the number of parties. It is well known that the median voter theorem does not hold when three or more parties run for election and the position of the median voter is not fixed and known (see e.g., Palfrey (1984) or Castanheira (2000)). Hence, the contribution of our model with imperfectly observed qualities is more limited and less surprising in that context. However, adding a third party would still bring one interesting feature that we wish to stress here (since the formal proof of this argument is straightforward, it is omitted for the sake of brevity). Suppose that one party has an exogenously fixed, centrist position whereas the other two can decide which ideology they propose. If these two parties decide to diverge as a commitment for quality (i.e. offer to the electorate an equilibrium situation with platforms (L, C, R)), then they will exert more effort in equilibrium than if they all keep their moderate platforms (C, C, C). As a result, their probability of being elected if qualities remain unobserved increases from 1/3 to 1/2. The main lesson of this exercise is that focusing on a two-party situation may underestimate the welfare value of divergence: voters always prefer (C, C) to (L, R), but they may prefer (L, C, R) to (C, C, C).

6. Political ideology. We have assumed throughout the paper that parties have no ideology, which put ourselves in the most adverse situation for observing divergence. If, instead, some parties have an intrinsic preference over platforms, then divergence is more likely to occur in equilibrium. First and trivially, because parties with an extremist ideology will trade-off probability of election vs. utility of implementing a given platform. More interestingly, one ideological left-wing and one ideological right-wing party competing under platforms (L, R) will exert more effort than two (either opportunistic or ideological) parties located at C, because losing the election implies a bigger disutility (in terms of the policy implemented) for the former than for the latter. This in turn makes parties with extremist ideologies (and extremist positions) even more appealing to voters than the opportunistic ones, and therefore more likely to "survive" in a competitive environment.

[35] When one party diverges, the opponent has incentives to diverge even more. However, if the platform of one party becomes overly extreme, the rival has more incentives to become centrist than to keep increasing his own degree of extremism.

6 Conclusions

The main goal of this paper was to show that if parties can affect the quality of their platform (e.g. by exerting effort) and this is imperfectly observed by voters, then there is scope for an interesting endogenous interaction between the implicit incentives to exert effort and the political platform proposed to the electorate. More specifically, we show that opportunistic parties may rationally decide to adopt extremist positions only as a commitment for quality. Moreover, voters can end up benefitting from such divergence. Our explanation for divergence is thus based on a moral hazard problem and suggests that extremist politicians are not necessarily better than their moderate counterparts. However, they are more committed to their own ideas. In turn, it also shows why a rational electorate is reluctant to support a party who does not exhibit a strong commitment to a given ideology (like the VLD in Belgium in 1994).

This approach can explain some existing puzzles of the political economy literature. First, one may wonder what makes it so difficult for a new, moderate party to challenge the lead of existing (non-moderate) parties. In the absence of moral hazard considerations, the centrist party should have a substantial advantage over the extremist ones. Instead, we argue that its lower popularity will stem from its lower implicit incentives to exert effort (see also Caillaud and Tirole (1999) for an alternative explanation based on intra-party competition). Second, it is not clear either why a moderate electorate would shift from supporting overly left-wing to supporting overly right-wing parties, instead of electing systematically the party closest to the median voter. According to our model, such shifts are not exclusively motivated by changes in the mood or intrinsic preferences of the electorate. Instead, differences in the perception of the candidates' relative quality may trigger-off these radical changes in their voting behavior.[36] Last, to the best of our knowledge, our model is the first one to endogenize the role of the press in determining political divergence and platform quality.[37] Given the candidates' moral hazard problem, we show that the press may induce inefficient platform choices of parties and suboptimally low investments in quality.

[36] Our approach is consistent with the fact that, in order to moderate the implemented policy, voters may prefer to use split-ticket strategies (see Alesina and Rosenthal, 1995) rather than voting for a centrist party.

[37] Strömberg's (2001b) analysis focuses on a different type of information: in his framework, the press informs the voters about the *location* of the parties and he shows that, in this case, parties tend to move away from uninformed voters.

References

[1] Alesina, Alberto, 1988. Credibility and Policy Convergence in a Two–Party System with Rational Voters. American Economic Review, Vol. 78, Issue 4, pp. 796-805.

[2] Alesina, Alberto and Howard Rosenthal, 1995. Partisan Politics, Divided Government, and the Economy, Cambridge University Press.

[3] Aragonés, Enriqueta and Thomas Palfrey, 2001. Mixed Equilibrium in a Downsian Model with a Favored Candidate. Forthcoming in Journal of Economic Theory.

[4] Bernhardt, Daniel and Daniel Ingberman, 1985. Candidate Reputations and the 'Incumbency Effect'. Journal of Public Economics; 27(1): 47-67.

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

[6] Caillaud, Bernard and Jean Tirole, 1999. Party Governance and Ideological Bias. European Economic Review, Vol. 43, Issues 4-6: 779-789.

[7] Caillaud, Bernard and Jean Tirole, 2001. Parties as Political Intermediaries. CERAS and IDEI Mimeo.

[8] Calvert, R. 1985. Robustness of the Multidimensional Voting Model: Candidates's Motivations, Uncertainty and Convergence. American Political Science Review, Vol. 29, No1.

[9] Castanheira, Micael, 2001. Why Vote for Losers? ECARES mimeo.

[10] Downs, Anthony, 1957. An Economic theory of Democracy. Harper and Row.

[11] Feddersen Timothy and Wolfgang Pesendorfer, 1997. Voting Behavior And Information Aggregation in Elections with Private Information. Econometrica, Vol. 65, No5.

[12] Ferejohn, John, 1986. Incumbent Performance and Electoral Control. Public Choice, Vol. 50: 5-26.

[13] Guesnerie, Roger and Jean-Jacques Laffont, 1984. A Complete Solution to a Class of Principal-Agent Problems with an Application to the Control of a Self-Managed Firm. Journal of Public Economics, Vol. 25: 329-369.

[14] Holmström, Bengt, 1999. Managerial Incentive Problems: A Dynamic Perspective. Review of Economic Studies; 66(1), January: 169-82.

[15] Laffont, Jean-Jacques and Jean Tirole, 1986. Using Cost Observation to Regulate Firms. Journal of Political Economy; 94(3), Part 1, June: 614-41.

[16] Mueller, Dennis C, 1989. Public Choice II. Cambridge University Press.

[17] Ordeshook, Peter C. 1997. The spatial analysis of elections and committees: Four decades of research. In: Perspectives on Public Choice. A Handbook. Dennis C. Mueller editor. Cambridge University Press

[18] Palfrey, Thomas, 1984. Spatial Equilibrium with Entry. Review of Economic Studies: 139-156.

[19] Persson, Torsten, Gérard Roland, and Guido Tabellini, 1997. Separation of Powers and Political Accountability, Quarterly Journal of Economics, 112.4: 1163-1202.

[20] Persson, Torsten, Gérard Roland, and Guido Tabellini, 1998. Towards Micropolitical Foundations of Public Finance, European Economic Review, 42: 685-694.

[21] Persson, Torsten, and Guido Tabellini, 2000. Political Economics: Explaining Economic Policy, Cambridge Massachusetts, MIT Press.

[22] Polo, Michele. 1998. Electoral competition and political rents. IGIER Discussion Paper No 144.

[23] Poole, Keith and Howard Rosenthal, 1991. Patterns of Congressional Voting. American Journal of Political Science, 35: 228-78.

[24] Poole, Keith and Howard Rosenthal, 1997. Congress: A Political-Economic History of Roll Call Voting. New York: Oxford University Press.

[25] Prat, Andrea, 2000. Campaign Advertising and Voter Welfare. Tilburg Mimeo.

[26] Rivière, Anouk, 2000. Citizen Candidacy, Party Formation and Duverger's Law, mimeo.

[27] Roemer, John, 1997. Political–Economic Equilibrium When Parties Represent Constituents: The Unidimensional Case. Social Choice and Welfare 14(4), pp479-502.

[28] Rogoff, Kenneth, 1990. Equilibrium Political Budget Cycles, American Economic Review; 80(1), March: 21-36.

[29] Strömberg, David. 2001a. Radio's Impact on Public Spending, Stockholm university Mimeo.

[30] Strömberg, David. 2001b. Mass Media Competition, Political Competition, and Public Policy, Stockholm university Mimeo.

[31] Wittman, Donald. 1983. Candidate Motivation: A Synthesis of Alternatives. American Political Science Review, Vol. 77 pp142-157.

Appendix

A1. Proof of Lemma 1

The set of possible rankings among the different qualities is a direct result of **(C1)**, **(C2)** and **(C3)**. Showing that $\mathcal{P}_X$ is a compact set requires a more elaborated proof. Differentiating **(C2)** and **(C3)** with respect to p we get:

$$c''(q_X)\,\frac{dq_X}{dp} = (1-q_M) - p\,\frac{dq_M}{dp}, \tag{20}$$

$$c''(q_M)\,\frac{dq_M}{dp} = q_X + p\,\frac{dq_X}{dp}. \tag{21}$$

As $p = 0$ implies $q_M = q_X = 0$, we can infer that $\left.\frac{dq_X}{dp}\right|_{p=0} > \left.\frac{dq_M}{dp}\right|_{p=0} = 0$. Therefore, $q_X > q_M$ when $p \to 0$. As a result, it is sufficient to prove that q_X and q_M intersect at most once in order to conclude that $\mathcal{P}_X$ is a compact set. From **(C2)** and **(C3)**, we know that $q_X = q_M \Leftrightarrow q_X = q_M = 1/2$. Using this relationship, and computing both the sum and the difference between (20) and (21), it is straightforward to see that

$$\left.\frac{dq_X}{dp}\right|_{q_X=q_M} < \left.\frac{dq_M}{dp}\right|_{q_X=q_M},$$

which is sufficient to prove that q_X and q_M intersect at most once. □

A2. Proof of Proposition 1

From Lemma 1, we know that if $p \notin \mathcal{P}_X$ then $q_M > q_X$, and therefore $\kappa_M = 1$ and $\kappa_X = 0$. Hence, the utility of the extremist and moderate party when $p \notin \mathcal{P}_X$ are:

$$U_X = p\,q_X(1-q_M) - c(q_X) \quad \text{and} \quad U_M = p\Big[1 - q_X(1-q_M)\Big] + (1-p) - c(q_M).$$

Convergence to (C, C) is the unique equilibrium if $U_X < U_S = \frac{1}{2} - c(q_S)$ and $U_M > U_S = \frac{1}{2} - c(q_S)$, where the first inequality ensures that it is not profitable to deviate from (C, C) and the second one ensures that (L, R) is not sustainable. Using the above expressions of U_X and U_M and rearranging terms, we get:

$$p\,q_X\,(1-q_M) + c\,(q_S) - c\,(q_X) < \tfrac{1}{2}, \tag{B1}$$

$$p\,q_X\,(1-q_M) + c\,(q_M) - c\,(q_S) < \tfrac{1}{2}, \tag{B2}$$

as the necessary and sufficient conditions for (C, C) to be the unique equilibrium of the positioning stage when $p \notin \mathcal{P}$.

Given $q \geq 0$, $c(q) \geq 0$ for all q, and $c''(q) > 0$, we have that for any pair $(q, \tilde{q})$:

$$(q-\tilde{q}) \cdot c'(\tilde{q}) \;<\; c(q) - c(\tilde{q}) \;<\; (q-\tilde{q}) \cdot c'(q). \tag{22}$$

Using (22), **(C1)** and **(C3)**, we obtain the following inequalities:

$$c(q_S) - c(q_X) \;<\; (q_S - q_X) \cdot c'(q_S) = \tfrac{1}{2}\, p\, (q_S - q_X), \tag{23}$$

$$c(q_M) - c(q_S) \;<\; (q_M - q_S) \cdot c'(q_M) = p\, q_X\, (q_M - q_S). \tag{24}$$

Given (23) and (24), then sufficient conditions for **(B1)** and **(B2)** to hold are respectively:

$$p\, q_X\, (1 - q_M) + \tfrac{1}{2}\, p\, (q_S - q_X) < \tfrac{1}{2} \quad \Leftrightarrow \quad p\Big(q_X\, (1 - 2q_M) + q_S\Big) < 1, \tag{25}$$

$$p\, q_X\, (1 - q_M) + p\, q_X\, (q_M - q_S) < \tfrac{1}{2} \quad \Leftrightarrow \quad p\, q_X\, (1 - q_S) < \tfrac{1}{2}. \tag{26}$$

Last and again from Lemma 1, we know that when $p \notin \mathcal{P}$, then both $q_M > 1/2$ (which ensures that (25) holds) and $q_S > 1/2$ (which ensures that (26) holds). □

A3. Proof of Proposition 2

Recall from Definition 1 that $p \in \mathcal{P} \;\Leftrightarrow\; \kappa_X = 1$ (and $\kappa_M = 1 - \kappa_X = 0$). Let us define $\mathcal{P}_1$ as the set of probabilities such that, if $p \in \mathcal{P}$, then a deviation from (C, C) is dominated. Similarly, we define $\mathcal{P}_2$ as the set of probabilities such that, if $p \in \mathcal{P}$, then a deviation from (L, R) is dominated. Formally,

$$\begin{aligned} \mathcal{P}_1 &= \{p:\ \kappa_X = 1 \;\Rightarrow\; U_S > U_X\}, \text{ and} \\ \mathcal{P}_2 &= \{p:\ \kappa_X = 1 \;\Rightarrow\; U_S > U_M\}. \end{aligned}$$

The proof consists of the following steps. First, we show that (C, C) is the unique equilibrium of the game if $p \notin \mathcal{P}$. Second, we show that the sets $\mathcal{P}_1$ and $\mathcal{P}_2$ are not empty. Third, we show that if $p \in \mathcal{P} \backslash (\mathcal{P}_1 \cup \mathcal{P}_2)$, only asymmetric positions can be equilibria of the game. Fourth, we show that (C, C) is the unique equilibrium when $p \to 1$. Last, we show that (L, R) is the unique equilibrium when $p \to 0$ and $\Delta \to 0$.

Step 1. (C, C) is the unique positioning equilibrium for any $p \notin \mathcal{P}$.[38]

When $p \in \bar{\mathcal{P}}$, we must have $\kappa_X \leq 1/2$ which implies:

$$\pi_X \;\leq\; p\, q_X\, (1 - q_M) + \frac{1-p}{2}, \quad \text{and} \tag{27}$$

$$\pi_M \;\geq\; p\, [1 - q_X\, (1 - q_M)] + \frac{1-p}{2}. \tag{28}$$

Two situations must be considered separately: $p \in \bar{\mathcal{P}} \cap \bar{\mathcal{P}}_X$ and $p \in \bar{\mathcal{P}} \cap \mathcal{P}_X$. From Proposition 1 we already know that (C, C) is the only equilibrium if $p \in \bar{\mathcal{P}}_X$. Using Lemma 1, only two cases remain to be considered for $p \in \bar{\mathcal{P}} \cap \mathcal{P}_X$:

- $1/2 > q_X > q_S > q_M$. In this case, $\pi_X < 1/2$, $\pi_S = 1/2$ and $\pi_M > 1/2$. Noting that $c(q_X) > c(q_S) > c(q_M)$ immediately proves that both $U_X < U_S$ and $U_M > U_S$ hold and therefore that (C, C) is the unique equilibrium of the game.

[38] We will often use the notation $\bar{\mathcal{P}} = [0, 1] \setminus \mathcal{P}$.

- $q_X > 1/2 > q_M > q_S$. Using (27), the inequality $U_X < U_S$ becomes:

$$p\left[\frac{1}{2} - q_X\ (1 - q_M)\right] > c\,(q_S) - c\,(q_X)\,.$$

From (23), $c(q_S) - c(q_X) < \frac{1}{2}\,p\,(q_S - q_X)$. Therefore, $U_X < U_S$ is satisfied if:

$$1 - 2q_X\,(1 - q_M) > (q_S - q_X) \;\Leftrightarrow\; q_S + q_X(1 - 2q_M) < 1.$$

Given that $q_X < 1$ and $q_S < q_M$, we know that $q_S + q_X(1-2q_M) < 1 - q_M$, which means that the above inequality always hold. Last, using (28), $U_M > U_S$ becomes:

$$p\left[\frac{1}{2} - q_X\ (1 - q_M)\right] > c(q_M) - c(q_S).$$

From (24), $c(q_M) - c(q_S) < p\,q_X\,(q_M - q_S)$. Therefore, $U_M > U_S$ is satisfied if:

$$p\left[\frac{1}{2} - q_X\,(1 - q_M)\right] > p\,q_X\,(q_M - q_S) \;\Leftrightarrow\; \frac{1}{2} > q_X\,(1 - q_S)$$

which always holds given that $q_S < 1/2$ and, from Assumption 2, $q_X < 2q_S$.

Step 2. The sets $\mathcal{P}_1$ and $\mathcal{P}_2$ are not empty.

When $\kappa_X = 1$, we have:

$$U_X = 1 - p\,[1 - q_X\,(1 - q_M)] - c(q_X) \qquad \text{and} \qquad U_M = p\,[1 - q_X\,(1 - q_M)] - c(q_M)$$

Therefore, $\mathcal{P}_1$ is the set of p for which the following condition holds:

$$p > f_1\,(p) = \frac{\frac{1}{2} - c\,(q_X) + c\,(q_S)}{1 - q_X\,(1 - q_M)},$$

and $\mathcal{P}_2$ is the set of p for which the following condition holds:

$$p < f_2\,(p) = \frac{\frac{1}{2} + c\,(q_M) - c\,(q_S)}{1 - q_X\,(1 - q_M)}.$$

when $p \to 0$, then $q_M \to 0$, $q_S \to 0$ and $q_X \to 0$, so $f_2(p) \to 1/2$ and therefore $\mathcal{P}_2$ cannot be empty. Next, using (23), the inequality $p > f_1(p)$ necessarily holds if:

$$p > \frac{\frac{1}{2} - \frac{1}{2}\,p\,(q_X - q_S)}{1 - q_X\,(1 - q_M)} \quad \Leftrightarrow \quad p\left[1 - \frac{1}{2}\,q_X(1 - 2q_M) - \frac{1}{2}\,q_S\right] > \frac{1}{2}.$$

From Lemma 1, two cases are possible: (i) $q_S < q_M < 1/2$, and (ii) $q_M < q_S < q_X < 1/2$. In both of them, the term in brackets is smaller than $1/2$ and therefore the inequality always hold when $p \to 1$. Hence, $\mathcal{P}_1$ cannot be empty either.

Step 3. (C, R) and (L, C) are the only possible equilibria for all $p \in \mathcal{P} \setminus (\mathcal{P}_1 \cup \mathcal{P}_2)$.

This is immediate since, by definition, asymmetric positioning dominates (C, C) when $p \in \mathcal{P} \cap \bar{\mathcal{P}}_1$ and asymmetric positioning dominates (L, R) when $p \in \mathcal{P} \cap \bar{\mathcal{P}}_2$.

Step 4. (C, C) is the only possible equilibrium for $p \to 1$.

We know that if $p \to 1$, then $p \in \mathcal{P}_1$. Thus, it is sufficient to show that $p \notin \mathcal{P}_2$ when $p \to 1$ to conclude that (C, C) is the only equilibrium for $p \to 1$. Note that $p \in \mathcal{P}_2 \;\Leftrightarrow\; p < f_2(p)$. Using (24), we know that $c(q_M) - c(q_S) < p\, q_X\, (q_M - q_S)$. Hence, a necessary condition for $p \in \mathcal{P}_2$ is:

$$p < \frac{\frac{1}{2} + p\, q_X\, (q_M - q_S)}{1 - q_X(1 - q_M)} \quad \Leftrightarrow \quad p\Big[1 - q_X(1 - q_S)\Big] < \frac{1}{2}.$$

However, the above inequality cannot hold for $p = 1$ given that, by Assumption 2, $q_X < 2q_S$.

Step 5. (L, R) is the only possible equilibrium when $p \to 0$ and $\Delta \to 0$.

When $\Delta \to 0$, then $\mathcal{P} \to \mathcal{P}_X$. Hence, from Lemma 1 and Step 2 hereabove, values of $p \to 0$ are such that $p \in \mathcal{P} \cap \mathcal{P}_2$ and $p \notin \mathcal{P}_1$. □

A4. Proof of Proposition 3

Symmetric positions. By revealed preferences, and assuming fixed platforms, we deduce from (12) and (14) that:

$$\Pi_{sym}(p^*_{sym}) + q_S(p^*_{sym}) - \gamma(p^*_{sym}) > \Pi_{sym}(\tilde{p}_{sym}) + q_S(\tilde{p}_{sym}) - \gamma(\tilde{p}_{sym})$$

$$\Leftrightarrow \; q_S(p^*_{sym}) - q_S(\tilde{p}_{sym}) > \Big[\Pi_{sym}(\tilde{p}_{sym}) - \gamma(\tilde{p}_{sym})\Big] - \Big[\Pi_{sym}(p^*_{sym}) - \gamma(p^*_{sym})\Big]$$

Since the right-hand side of this inequality is positive by construction (the press maximizes profits in $\tilde{p}_{sym}$), we necessarily have $q_S(p^*_{sym}) - q_S(\tilde{p}_{sym}) > 0$. Given that $dq_S/dp > 0$ for all p (see **(C1)**), then $p^*_{sym} > \tilde{p}_{sym}$.

Asymmetric positions. First, notice that $dq_M/dp > 0$. To see this, suppose that $dq_M/dp < 0$. From (20), this would imply $dq_X/dp > 0$. But then, by (21), we would get $dq_M/dp > 0$ which is a contradiction. Second, suppose that $q_M \to 1$. By (20), $c''(q_X)\frac{dq_X}{dp} \simeq -p\,\frac{dq_M}{dp} < 0$. This means that $dq_X/dp \gtrless 0$. Now, given (16), (18) and following the same reasoning as in the previous case, we have:

$$q_X\left(p^*_{asy}\right) - q_X\left(\tilde{p}_{asy}\right) > \Big[\Pi_{asy}(\tilde{p}_{asy}) - \gamma(\tilde{p}_{asy})\Big] - \Big[\Pi_{asy}(p^*_{asy}) - \gamma(p^*_{asy})\Big],$$

and therefore $q_X\left(p^*_{asy}\right) > q_X\left(\tilde{p}_{asy}\right)$. However, since $dq_X/dp \gtrless 0$, then by (19) we have $\tilde{p}_{asy} \gtrless p^*_{asy}$. Last, given that $dq_M/dp > 0$, we get $q_M(\tilde{p}_{asy}) \gtrless q_M(p^*_{asy})$. □

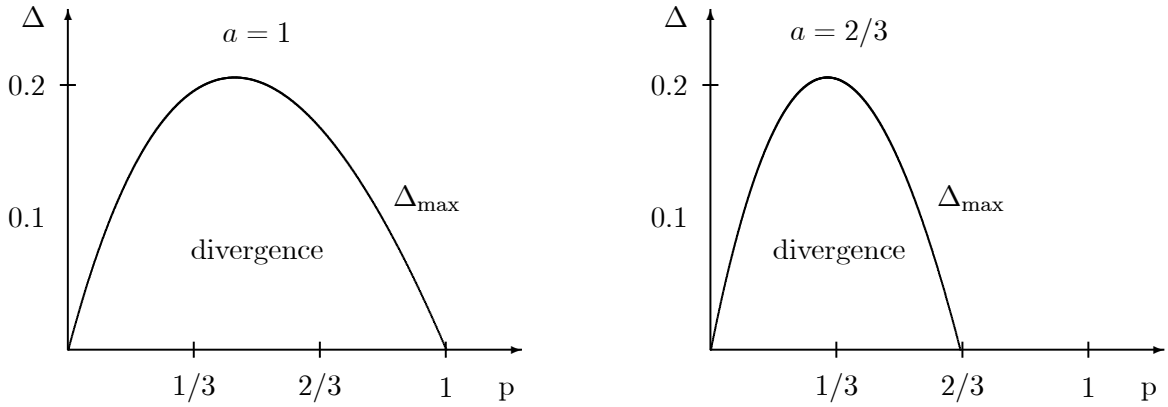

Figure 2. Maximum distance Δ compatible with divergence

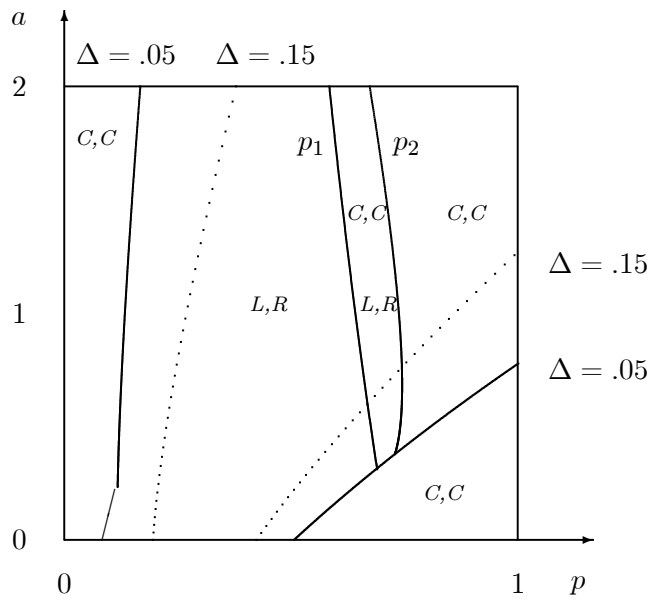

Figure 3. Positioning equilibria as a function of (p, a)

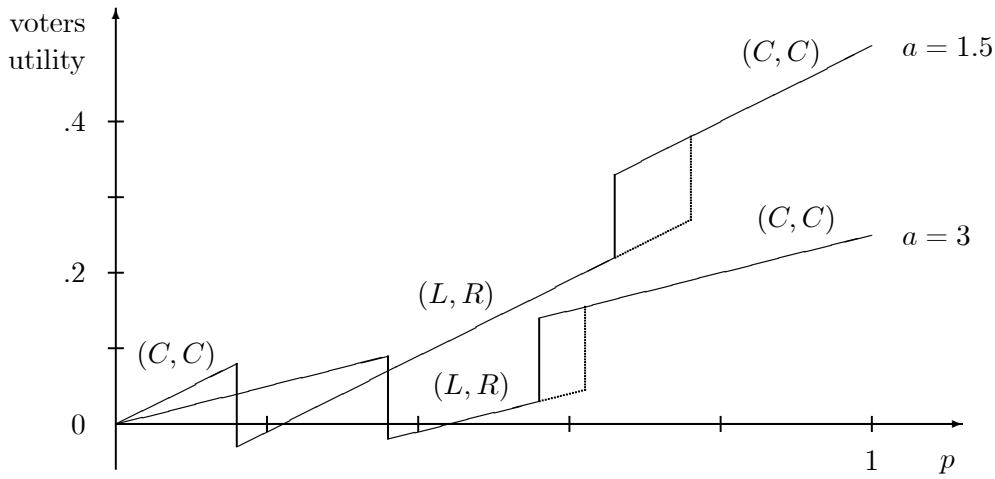

Figure 4. Voters' expected welfare ($\Delta = 0.1$)