Número 290

Does Charisma Discourage Experience and Encourage Extremism?

The Electoral Strategies of a Populist Candidate

GILLES SERRA

Octubre 2017

CENTRO DE INVESTIGACIÓN Y DOCENCIA ECONONÓMICAS

CIDE

Advertencia

Los Documentos de Trabajo del CIDE son una herramienta para fomentar la discusión entre las comunidades académicas. A partir de la difusión, en este formato, de los avances de investigación se busca que los autores puedan recibir comentarios y retroalimentación de sus pares nacionales e internacionales en un estado aún temprano de la investigación.

De acuerdo con esta práctica internacional congruente con el trabajo académico contemporáneo, muchos de estos documentos buscan convertirse posteriormente en una publicación formal, como libro, capítulo de libro o artículo en revista especializada.

Acknowledgements

Gilles Serra is assistant professor at the Division of Political Studies at CIDE. He would like to thank his research assistant Gerardo Manrique De Lara Ruiz for great help.
Email: gilles.serra@cide.edu
Personal page: http://www.investigadores.cide.edu/gilles.serra.

Abstract

I model an election between a populist candidate with little government experience and high charisma, and a mainstream candidate with much government experience and low charisma. Taking a step back in time, I also model the career choices of this populist candidate: he must consider how much governing experience to acquire before running for high office, and then he must decide how extremist his campaign platform should be. The model finds two major trade-offs that are unfortunate for the median voter: candidates that are attractive in terms of their high charisma will be unattractive in terms of their low experience and high extremism. The model also finds that popular discontent, coming from an economic or political crisis, makes an inexperienced outsider more likely to win the election with an extremist agenda, which helps explain the recent "rise of populism" identified by several authors around the world. Another contribution is explaining, within a unified theory, numerous empirical findings: I connect the model to the literature from different academic approaches (behavioral, comparative and institutional) and different geographical regions (the United States, Latin America and Europe). Special reference is made to four prominent outsiders: Donald Trump, Hugo Chávez, Alberto Fujimori and Jean-Marie Le Pen.

Keywords: populism, charisma, experience, elections, democracy

Resumen

Aquí modelo una elección entre un candidato populista con poca experiencia de gobierno y alto carisma, y un candidato "mainstream" con mucha experiencia de gobierno y bajo carisma. Tomando un paso atrás en el tiempo, también modelo las decisiones de carrera de este candidato populista: debe considerar cuánta experiencia de gobierno adquirir antes de competir por un alto cargo, y entonces debe decidir cuan extremista debe ser su plataforma de campaña. El modelo encuentra dos grandes compromisos que son desafortunados para el votante mediano: los candidatos que son atractivos en cuestión de su carisma serán decepcionantes en cuestión de su baja experiencia y su alto extremismo. El modelo también encuentra que el descontento popular, proveniente de alguna crisis económica o política, incremente la posibilidad e que un "outsider" gane la elección con una agenda extremista, lo cual ayuda a explicar el reciente "ola populista" identificada por varios autores alrededor del mundo. Otra contribución es explicar, dentro de una teoría unificada, numerosos hallazgos empíricos: conecto el modelo a la literatura proveniente de distintos enfoques académicos (conductual, comparativo e institucional) y distintas áreas geográficas (Estados Unidos, Latinoamérica y Europa). Se hace especial referencia a cuatro prominentes "outsiders": Donald Trump, Hugo Chávez, Alberto Fujimori y Jean-Marie Le Pen.

Palabras clave: populismo, carisma, experiencia, elecciones, democracia

The Electoral Strategies of a Populist Candidate: Does Charisma Discourage Experience and Encourage Extremism?

Gilles Serra*

Abstract

I model an election between a populist candidate with little government experience and high charisma, and a mainstream candidate with much government experience and low charisma. Taking a step back in time, I also model the career choices of this populist candidate: he must consider how much governing experience to acquire before running for high office, and then he must decide how extremist his campaign platform should be. The model finds two major trade-offs that are unfortunate for the median voter: candidates that are attractive in terms of their high charisma will be unattractive in terms of their low experience and high extremism. The model also finds that popular discontent, coming from an economic or political crisis, makes an inexperienced outsider more likely to win the election with an extremist agenda, which helps explain the recent "rise of populism" identified by several authors around the world. Another contribution is explaining, within a unified theory, numerous empirical findings: I connect the model to the literature from different academic approaches (behavioral, comparative and institutional) and different geographical regions (the United States, Latin America and Europe). Special reference is made to four prominent outsiders: Donald Trump, Hugo Chávez, Alberto Fujimori and Jean-Marie Le Pen.

*Department of Political Science, Centre for Economics Research and Teaching (CIDE), Mexico City, Mexico, gilles.serra@cide.edu

1 The downside of charismatic populism

What kind of leader will govern a country following a democratic election? Voters value a number of features in their authorities, such as their competence, their communication skills, their capacity to unify the country, their respect for the rule of law, their commitment to democracy, and their responsiveness to regular citizens. However, those features might not be compatible with each other, and it might not be realistic to expect all of them from a single candidate. This essay will argue that democratic competition itself creates major trade-offs in the types of candidates available to voters, such that any election winner should be expected to lack some important qualities. For instance, voters may agree on the benefits of electing a chief executive who is a competent technocrat with experience in government who can design effective policies to resolve the nation's problems. They may also agree on the prudence of electing a moderate politician whose mainstream views can unify the country by making the centrist majority of the population feel represented. But will voters rally around an experienced moderate if she is boring, if she seems distant, or if she comes across as elitist? Regular citizens might be swayed by someone closer to them, someone they can connect with, and someone they believe will defend the common folk. They might flock around an inspiring orator who promises to liberate the masses from the self-serving elite that has captured the government through corruption and cronyism. Some important questions then are: if this populist outsider manages to win the election by beating the establishment candidate, will he lack the required experience to govern effectively? And will he implement extremist policies well beyond the median voter's preferences? According to empirical observation suggesting a rise in populism, these dilemmas are increasingly frequent around the world.

The theory in this paper explores the ways in which a new politician may shape his profile with the goal of reaching high office at a future election. The main goal of the model is to make predictions about three features that characterize a candidate: his policies, his experience and his charisma, which I define more precisely below. Concretely, I wish to explore the relationship between these characteristics to uncover potential conflicts between them. Indeed, the results reveal two trade-offs that are unfortunate for voters, whereby candidates with high levels of charisma will tend to have low levels of experience and high levels of extremism. A further goal of the model is to predict the conditions leading to the election of a populist outsider instead of a mainstream politician, and the levels of expertise that we can expect from each. In accordance to intuition, my theoretical model finds that popular disaffection, coming for example from a severe economic or political crisis, makes the successful election of a populist outsider more likely. More surprisingly, such crises are predicted to decrease the relevant experience acquired by the populist; increase his extremism; and increase the extremism of the government.

Another contribution of this paper is to discuss and explain several important regularities from the empirical literature. I demonstrate how my theoretical results are consistent with a large number of empirical observations from different academic approaches (e.g. behavioral, comparative and institutional) and different geographical regions (e.g. the United States, Latin America and Europe). I make special reference to four prominent outsider politicians whose biographies are consistent with my model in several regards: Donald Trump, Hugo Chávez, Alberto Fujimori and Jean-Marie Le Pen. The literature review will illustrate the utility of a unified formal theory to connect and organize observations from a diversity of empirical fields.

With these goals in mind, I develop a theory about the career decisions of a new politician who wishes to compete eventually for high office. At the outset, only two parameters characterize the politician. On one hand, he has intense and well-defined policy preferences in the left-right political spectrum; to be concrete, he will have a quadratic utility function with an ideal point to the right of the median voter. (Of course, all the results can be easily inverted by assuming an ideal point to the left of the median voter.) On the other hand, he may enjoy a certain amount of talent to articulate an anti-elite rhetoric that will resonate with voters for electoral support; to be concrete, he has a *valence* parameter due to his personal qualities. Donald Stokes (1963) coined the expression "valence" in reference to issues that all voters agree to value positively, in contrast to "positional" issues where voters may disagree depending on their ideologies. In this model, the outsider candidate may have some valence corresponding to a specific type of *charisma*.

Ordinary use of the word "charisma" in contemporary language differs somewhat from the way I wish to use it in this essay. While all the results in the model work well by interpreting the valence parameter as charm, celebrity or physical beauty, I rather have in mind an interpretation of charisma that is closer to its usage in political theory. In particular, I am referring to a more Weberian type of charisma, understood as an intimate and direct communion between the leader and his followers. Max Weber saw charisma as a very rare power endowing its holder with the capacity to elicit passionate popular support (Weber 1978). The charismatic leader is able to inspire true faith in the mission that he claims to embody.[1] My model assumes that an outsider candidate who possesses this type of charisma will engage in *populist* rhetoric. Populism is characterized by a Manichean discourse painting society as divided in two antagonistic groups: the pure and defenseless people against the corrupt and privileged elite, the former being a victim of the latter.[2] In recent history, the

[1] Eatwell (2017a) summarized this concept the following way: "In the pioneering approach established by Max Weber during the early twentieth century, charisma was seen as a quasi-religious phenomenon in which confident, prophetic leaders inspired an affective mass at times of crisis and against a background of secular modernisation." (p. 4)

[2] According to Inglehart and Norris (2016), "*populism* is understood as a philosophy that emphasizes faith in the wisdom and virtue of ordinary people (the silent majority) over the 'corrupt' establishment. Populism

successful populists have usually been charismatic. For example, Juan Perón in Argentina, who is considered the quintessential Latin American populist, was alleged to enjoy almost blind support from the masses because of his charisma. Micozzi and Saiegh (2016) interpret the emotional dimension of Peronism as valence, given its overlapping cultural, political, and economic breadth in the population. To be as precise as possible about the type of valence that I refer to in my model, I will call it *populist charisma,* by which I will understand the direct and non-mediated appeal that an outsider candidate may enjoy among voters by virtue of his credible anti-establishment credentials. As an important feature of the model, the effect of populist charisma will depend on the level of discontent with the government among voters.

The model has three stages. In a first stage, the new politician needs to decide how much government experience to acquire before seeking high office, knowing that such experience could be valued by voters. In effect he needs to make a career decision regarding the effort he is willing to put into preparing for the big election. To increase his electoral appeal, he may choose to spend time in relevant positions, such as taking a cabinet ministry in the administration or running for lower office such as mayor or legislator. A rational candidate will choose his amount of prior experience making a cost-benefit analysis, which will turn out to depend on his amount of populist charisma and the level of popular discontent. In a second stage, this new candidate finally starts campaigning for high office, facing a mainstream rival with a high level of office experience but no populist charisma. Hence the election exhibits a left-wing candidate from the establishment against a right-wing newcomer. In the third stage, voters elect one of the two candidates based on their three characteristics: their policy platforms, their experience in government and their populist charisma. The results of the election, such as the candidate platforms and the policy implemented by the winner, will all depend on the primitives of the model such as the existing amount of popular discontent.

The paper proceeds as follows. Section 2 will place my model within the formal literature, especially the previous models of valence, leadership and populism. Section 3 will model an election with a policy dimension and a valence dimension which is based on the more general policy-valence model of Serra (2010). Section 4 extends the model to study the interaction between two types of valence: an endogenous one understood as experience, and an exogenous one understood as charisma. Section 5 presents a realistic modification of the model whereby acquiring government experience reduces populist charisma. Given that this modification yields similar results to the main model, it serves as a robustness check. Section 6 shows how the model connects to the existing empirical evidence, by showing how its assumptions

reflects deep cynicism and resentment of existing authorities, whether big business, big banks, multinational corporations, media pundits, elected politicians and government officials, intellectual elites and scientific experts, and the arrogant and privileged rich. Ordinary people are regarded as homogeneous and inherently 'good' or 'decent', in counterpart to dishonest elites."

and results find support in a large body of literature from several subfields and different regions. Section 7 discusses the normative implications of the theoretical results, namely the trade-offs that voters can expect in democratic elections. An appendix to the paper that is posted online provides two more extensions that also serve a robustness checks, and it includes the proofs of all the theorems and corollaries.[3]

2 Previous theoretical literature

At its core, this model belongs to the formal literature on valence, especially the class of models studying the relationship between valence dimensions and policy dimensions. Some of this literature has sought to address a fundamental debate about whether valence leads to extremism or moderation. Within such debate, some models such as Groseclose (2001) find that a valence-advantaged candidate will be more moderate than a valence-disadvantaged one, while other models such as Serra (2010) find the opposite result. Like most of this literature, my paper includes an ideology dimension following the standard Downsian framework. Unlike the existing literature, my model adds, not one, but two valence parameters corresponding to two different dimensions that are valued by voters, namely experience and charisma. This aspect brings my paper close to Adams, Merrill, Simas and Stone (2011). These authors also consider two types of valence that voters may value in addition to policy. Their two concepts of valence have in fact some resemblance to mine. On one hand, they distinguish characteristics that are merely useful to winning elections such as name recognition, fund-raising ability and campaigning skills, which they called *strategic valence*. On the other hand they distinguish characteristics that are actually valuable for elected officials to govern, such as integrity, competence, and dedication, which they called *character valence*. I see my concept of *populist charisma* as being close to their concept of *strategic valence*, and my concept of *government experience* as being close to their concept of *character valence*.[4]

Another part of this literature has endeavored to endogenize the valence dimension by allowing political agents to modify the valence parameters through their actions. The most frequent interpretations for an endogenous valence relate to raising money, running advertisements, or exerting other kinds of effort during campaigns. My paper departs from the established cannon by giving a novel interpretation. I will interpret the endogenous valence

[3] The online appendix is available at http://www.investigadores.cide.edu/gilles.serra/.

[4] Other effects of valence have also been studied in the formal literature. For example, Calvo and Murillo (2017) allow the valence dimension to be correlated to the policy dimension. In other words, they make the assumption that voters that already like a party for its ideology will be more sensitive to this party's valence. Carter and Patty (2015) make the observation that some candidates may wish to skip campaigning altogether, depending on the amount of valence they count on. Surprisingly, they demonstrate that candidates may choose to remain on the ballot (and perhaps win) without exerting any effort at campaigning. Other formal models studying valence and ideology include Calvo and Hellwig (2011); Aragonès and Xefteris (2013); Crisp, Patty, Penn and Schibber (2014); Hitt, Volden and Wiseman (2017).

as the experience in government that a candidate may acquire throughout his career. This way I aim to shed light on the career decisions of a candidate through a longer timespan than other formal models of endogenous valence.[5]

By proposing the concept of *populist charisma*, my model focuses on studying citizens who have leadership skills. As such, it is related to a budding formal literature on leadership. Torun Dewan and David Myatt have modeled the conditions for "effective leadership" to arise. I beleive the authors' views are largely compatible with the Weberian approach to charisma that I espouse in this paper. On one hand, a number of special qualities are needed from the individual seeking to become a leader. They claim that a leader's influence increases with his judgment (i.e. his sense of direction) and his ability to convey ideas (i.e. his clarity of communication). On the other hand, a number of conditions need to exist among the potential followers: they must be seeking direction and guidance, they must be facing a coordination problem, and they must be listening to potential leaders to learn about their environment (Dewan and Myatt 2007; Dewan and Myatt 2008; Dewan and Myatt 2012).

Other formalizations of *populism* also relate to my model, such as Acemoglu, Egorov and Sonin (2013) and Fox and Stephenson (2015). As in my paper, these authors remark that antipathy toward a powerful elite can give rise to populist politicians making extremist proposals. For example, when mainstream politicians are perceived by the population to be corrupt and beholden to the wealthy, an independent candidate can thrive by proposing measures that are extreme left, as has often been in the case in Latin America. My paper is also motivated by this observation. However, my explanation differs from the one in those papers. In their models, a populist politician chooses an extremist platform to send an informative signal about his independence from the rich. In my model, he does so because of his preferences and his valence: the populist politician adopts a platform close to his extremist ideal point knowing that he can still win the election based on his charisma.

3 An election over ideology, experience and charisma

[5] Other interesting interpretations of endogenous valence can be found in the literature. Schofield, Claassen, Ozdemir and Zakharov (2011) study the endogenous spending by candidates on valence, which is interpreted as campaign advertising. Meirowitz (2008) studies the amount of money that will be spent by an incumbent and a challenger to increase their respective valence when they both have different marginal costs. The candidates in Carrillo and Castanheira (2008) need to select a policy platform, which is observable, and make an investment in quality, which is unobservable. In Callander (2008), valence is observed when the election is over: after getting elected, the candidate chooses a level of effort that is valued by voters. Ashworth and Bueno de Mesquita (2009) study how the endogenous adoption of platforms affects the endogenous adoption of valence. Penn (2009) offers a reinterpretation of the valence parameter: she postulates that individuals derive utility not only from their own welfare (akin to the policy loss function), but also from the welfare of the group they psychologically identify with (akin to the valence dimension).

This section models a democratic election for high office, perhaps for president or prime minister. At this stage, the candidates have already made their career choices in years past, so their respective amounts of experience in government are taken as fixed. In short, the candidates count on certain amounts of experience, and perhaps some charisma, which are exogenously given. Upon starting their campaigns, their only strategic choice is the ideological platforms they decide to adopt for competing in the election. The situation is game-theoretic in the sense that each candidate wishes to anticipate what the other candidate will choose – and their predicted choices are the Nash equilibrium of a simultaneous game. Next section will take a step back in time to analyze the career decision that one of these candidates will make, in terms of deciding how much work experience to acquire in government-related offices before the big election that he wishes to participate in.

3.1 The election

There is a contest for office between two candidates who propose different ideological platforms and have different amounts of experience in government. There might have been more contenders initially, but we assume that in the last stretch of the campaign only two serious contenders remain with realistic possibilities of winning. A candidate's experience for governing can be understood as *valence*. In formal political theory, valence is often treated as a dimension that is valued positively by all voters, meaning they all prefer higher values in this dimension. Accordingly, I will assume that a candidate enjoys higher support from the general electoral by virtue of having held previous office or cabinet positions. I will denote by e the level of experience of a candidate, where e is a number between zero (no experience whatsoever) and one (the most experience that can be expected).

In contrast, the candidates' policy proposals in the left-right political spectrum are valued differently by different voters because each voter has a different ideal point in this dimension. Each candidate needs to design an ideological platform to compete in the election, which does not need to be identical to his own ideological preferences. In fact it is feasible, as will occur in this model, for a candidate to adopt a platform that differs from the one he would ideally prefer. So in this model, a candidate with extremist preferences can choose whether to announce a moderate or an extremist platform depending on his optimal electoral strategy. I assume, like most spatial models in this tradition, that a platform becomes binding once it is announced, meaning that a candidate will be forced to implement the platform he promised. I will denote an ideological platform by x, where x can be any negative or positive number.

In addition to his office experience and his policy platform, an outsider candidate may have *populist charisma*. I assume that this is equally valuable for all voters across the ideological spectrum, but not every new candidate is equally charismatic. Populist charisma will be treated as a variable c that can take any value between zero (no charisma at all)

and one (the highest level of charisma). In short, c can be understood as a second type of valence.

The impact that such charisma may have on voters depends on the context. In particular, I will assume that the electorate's appetite for a charismatic populist will depend on the economic, political and social conditions before the election. I will summarize those macro-conditions with a variable indicating the level of *popular discontent* in the country, by which I mean the degree to which people in general have developed resentment against all branches of government and the political elite. The level of people's discontent with the political system will be called δ, which can take any value between zero (perfect contentment) and one (utmost disappointment).

3.2 The voters

Voters care about three separate dimensions – ideology, experience and charisma – where the importance of charisma relative to the other dimensions depends on a fourth dimension – popular discontent. Hence their utility functions will depend on four parameters. First, the policy implemented after the election, which is labeled x with $x \in \mathbb{R}$. Second, the amount of government experience of the elected candidate, which is labeled e with $e \in [0,1]$. Third, the amount of populist charisma of the elected candidate, which is labeled c with $c \in [0,1]$. And fourth, the amount of popular discontent before the election, which is labeled δ with $\delta \in [0,1]$. The effect of charisma is mediated by popular discontent, such that voters perceive this type of valence to be δc. Adding experience gives the candidate's total amount of valence, which is $e + \delta c$.

The electorate has a known median voter, which we call M, whose preferences are decisive in the election. Regarding ideology, voters have linear and single-peaked utility functions around their ideal point. We normalize the ideal point of the median voter to zero, meaning that her disutility from policy distance is $|x|$. The utility function of M is thus given by

$$U_M(x, e, c, \delta) = -|x| + e + \delta c \tag{1}$$

3.3 The candidates

There are two candidates competing in this election, labeled R and L for the right-wing candidate and the left-wing candidate, respectively. I assume that candidates are *policy-motivated*, meaning that they care about the policy implemented after the election. To be concrete, both candidates have clearly different preferences on opposite sides of the median voter, with L having a negative ideal point and R having a positive one. Hence, irrespective of the platforms they promise to voters, we know that one candidate has genuinely leftist

preferences and the other one has genuinely rightist preferences. It should be noted that other formal models frequently assume that candidates only care about winning elections regardless of their policy platforms, meaning they are *office-motivated.* In my model, assuming instead that candidates care about the policy implemented by the government makes most sense to analyze the desires of extremist candidates who wish to use their charisma to influence the election.[6]

Given that I wish to focus all attention on other variables, I will simplify the specification of candidates' preferences by normalizing the ideal point of R to 1 and the ideal point of L to -1. The assumption implies that candidates have equally extremist preferences, given that both ideal points are equidistant from the center. This is convenient, as we can thus be sure that any asymmetries that we will find in candidates' choices come from parameters other than their true ideological preferences. In particular, if one candidate chooses a moderate platform while the other chooses an extremist one, we know it will be due entirely to their different levels of experience and charisma as well as the people's discontent – not to their ideal points. Both candidates have single-peaked utility functions over policy. In contrast to citizens, whose utility functions are linear, I will assume that the utility functions of candidates are quadratic. This allows candidates to be highly sensitive to different parameters that affect the policy outcome. This difference can be justified by thinking of candidates as having very intense preferences, making them more sensitive to policy changes than the average non-politicized citizen.[7] In sum, their utilities are given by:

$$U_R(x) = -(1-x)^2 \tag{2}$$

$$U_L(x) = -(-1-x)^2 \tag{3}$$

Before the election, candidates R and L formulate policy platforms x_R and x_L, which might be different from their ideal points, with $x_R, x_L \in \mathbb{R}$. Any promise a candidate makes to voters in terms of policy will need to be implemented if he is elected to office; in other words, platform announcements are binding.

In this election, one of the candidates is an outsider while the other one is an insider, meaning that the former does not entirely belong to the established political elite while the

[6] In any event, I have studied this same model when candidates have mixed desires, being simultaneously policy-motivated and office-motivated. As shown by this extension in the online appendix, all the results in the model remain intact. The reason is that both motivations provide incentives to candidates in the same direction. Hence such extension serves as a robustness check. I thank an anonymous reviewer for suggesting this extension.

[7] In any case, I have also derived all the results of the model when candidates have linear utility functions like voters. As indicated in the corresponding extension in the online appendix, some of the results remain, such as the existence of unique equilibria for all parameter values; and the effect of charisma on extremism. But the effect of charisma on experience disappears because candidates stop being sensitive to initial conditions.

latter is very clearly identified with the establishment. Without loss of generality, R will be the outsider while L will be the insider. As an outsider candidate, R is able to run an anti-establishment campaign, trying to connect directly with voters as common folk. His success, however, depends on his level of populist charisma, which we label c_R with $c_R \in [0,1]$. On the other hand, given that L is a mainstream candidate we assume that she cannot credibly run an antiestablishment campaign, so her level of populist charisma is zero.

Candidates are also characterized by a parameter e denoting each candidate's prior experience in government. We call e_R the experience level of R, with $e_R \in [0,1]$. On the other hand, we will assume that L has the highest possible level of experience, which is one. Hence, according to each candidate's levels of experience and charisma, R's total amount of valence is given by $e_R + \delta c_R$ while L's total amount of valence is 1.

I start by assuming that charisma and experience are compatible with each other, meaning they can coexist in a candidate without contradiction. In technical terms, I am treating c and e as perfect substitutes. This is convenient, as we can be sure that any trade-off between them comes from the candidate's career constraints, rather than some inherent incompatibility in the voters' minds. However, in a later section at the end of this article I analyze the full model again with the assumption that experience in government is to some degree incompatible with populist charisma.[8]

The following variable will be useful for future calculations: we define A_R as the non-policy advantage of candidate R compared to L due to his experience and charisma; it is the extra utility that he brings to voters in dimensions other than policy. So A_R is the valence advantage of R over L defined as $A_R \equiv e_R + \delta c_R - 1$. Note of course that this number could be negative, in which case A_R would represent a valence *dis*advantage for R compared to L. Given the range of values that all relevant variables can take, it can easily be proved that $-1 \leq A_R \leq 1$.

3.4 Timing, information and solution concept

The timing of this election is the following:

1. All exogenous variables are observed: the people's discontent (δ), R's charisma (c_R), and the experience of each candidate (e_R for R, and 1 for L).

2. Candidates simultaneously choose their platforms (x_R and x_L).

3. The median voter elects a candidate (R or L)

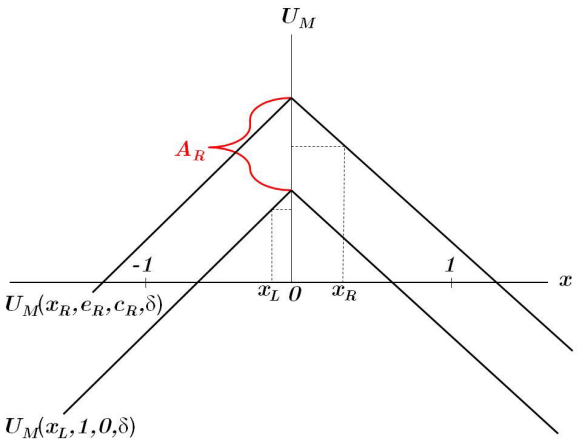
[8] To be concrete, it is possible that a populist candidate may lose some of her initial charisma as she spends time acquiring government experience. In Section 5, I prove that adding this feature does not change the essence of the results. In particular, the effects of charisma on all the variables remain the same.

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

3.5 The last stage: voters' choice

Before providing equilibrium results, it is worth looking more closely at how the median voter makes her decision in this kind of model.[9] At Stage 3 of this election, M will vote for the candidate maximizing her utility. I will make the following indifference assumptions. If M is indifferent between the two candidates, she will vote for the one with highest valence – this is exactly the situation that will occur in equilibrium. If both candidates are not only indifferent but also have the same valence, M will randomize equally between the two.[10]

Figure 1: The effect of an advantage for R over L in charisma and experience, A_R



As can be seen in Figure 1, M's appreciation for a candidate decreases with the distance between her ideal point and that candidate's platform, and increases with the candidate's experience and charisma. In essence, the parameters e and c "shift up" the utility function for this candidate, acting as valence. The figure depicts an example of how M evaluates R and L, where it is assumed that $A_R > 0$ and $|x_L| < |x_R|$. In this case, candidate R is strictly preferred to candidate L in spite of having a more extremist platform. Candidate R would win the election because his higher scores in the valence dimensions more than compensate his extremism in the policy dimension. As proved in the next section, the situation depicted

[9] A more general discussion of the model in this section can be found in Serra (2010).

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

in this graph would not be an equilibrium, however, because candidate R would benefit from choosing an even more extremist platform closer to his ideal point.

3.6 The effect of experience and charisma on the voters' decision

We now turn our attention to the behavior of candidates when they must formulate their policies at Stage 2 of the election. The exogenous parameters in this election are e_R (the right-wing candidate's level of government experience), c_R (the right-wing candidate's level of populist charisma), and δ (the people's disenchantment with political institutions). Given that all these parameters are fixed at this stage of the game, the equilibrium platforms and equilibrium outcomes are contingent on their values. In other words, there is a different subgame for each combination of values of e_R, c_R and δ.

Anticipating each other's decision, what platforms will candidates formulate? Our solution concept, subgame-perfect equilibrium (SPE), imposes that R and L must play a Nash equilibrium (NE) in every subgame. We call x_R^* and x_L^* this equilibrium and x^* the winning platform. As it turns out, a unique equilibrium exists for all parameter values.[11] In the following theorem, remember that A_R is the advantage in terms of valence that candidate R has over candidate L, defined as $A_R = e_R + \delta c_R - 1$, and whose values are in the range $A_R \in [-1, 1]$.

Theorem 1 *The equilibrium platforms of candidates and the policy outcomes of this election, as a function of A_R, are given in Table 1.*

Table 1: Equilibrium outcomes of the election

Value of A_R	**Equilibrium platforms** x_R^* and x_L^*	**Winning platform** x^*	**Winning candidate**
$0 < A_R \leq 1$	$x_R^* = A_R$ $x_L^* = 0$	A_R	R
$A_R = 0$	$x_R^* = 0$ $x_L^* = 0$	0	R or L with equal probability
$-1 \leq A_R < 0$	$x_R^* = 0$ $x_L^* = A_R$	A_R	L

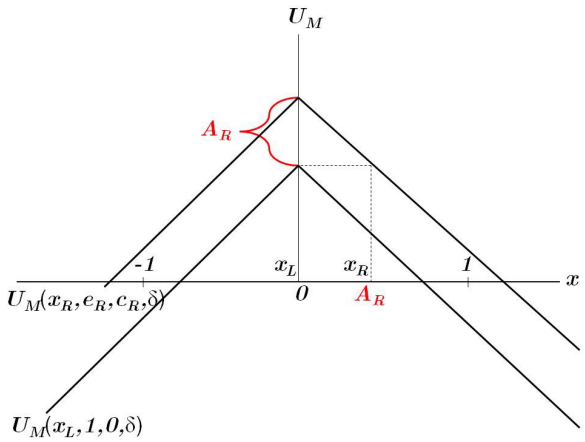
There are several remarks to make about the results in this table. First note the results when $A_R = 0$, that is, when there is no valence difference between the candidates. This

[11] The proofs of all the results in this paper come in the appendix posted online at http://www.investigadores.cide.edu/gilles.serra/.

corresponds to a standard election between two candidates who are policy motivated and compete only in the policy dimension: the centripetal forces in the election drive both candidates to converge completely to the median voter's ideal point (Calvert 1985).

Whenever $A_R \neq 0$, however, the results depart from the standard outcome in notable ways. Most importantly, the candidate with highest valence is able to diverge away from the median voter toward his ideal point, and increasingly so as his valence advantage increases. For example when candidate R has the higher valence, meaning that $0 < A_R$, he is able to diverge from the center and still win the election based on his superior valence. In the meantime, the candidate with lowest valence, say L, will converge to the center of the spectrum. The reason why L adopts the median voter's ideal point is to force R, who will win the election anyway, to converge as much as possible. By adopting $x_L = 0$ she constrains R to diverge no further than $x_R = A_R$. This equilibrium is depicted in Figure 2.

Figure 2: Equilibrium platforms x_L^* and x_R^* when there is a valence advantage for R over L



From this result, a corollary about the policy implemented in this election can be immediately stated.

Corollary 1 *The policy implemented after this election will be* $x^* = A_R$.

As a final remark, my model predicts that valence leads to extremism in the following sense. Imagine that A_R increases from zero to positive values. Then candidate L will locate at the center but will lose the election; meanwhile candidate R will adopt an increasingly right-wing platform that he will implement upon winning the election. Note that such extremism of R compared to L is independent of their true preferences, given that we assumed both candidates to have equally extremist ideal points at 1 and -1 respectively.

4 Acquiring experience in government

The results above could be used to take a step back in time. Now we can analyze the career decisions by an inexperienced candidate who is aiming for a high-level position in government such as president or prime minister. At this stage, he must decide how to prepare for a future run. One of the major decisions of this neophyte candidate is whether to engage in public service now to increase his likelihood of getting elected later. For example, in preparation for the specific race that really interests him, he could run for lower office first. Or he could seek a cabinet position for the sitting administration. As I document later, there exists ample empirical evidence that such prior experience helps candidates obtain electoral support. There is also evidence that candidates know this empirical fact, which motivates many of them to engage in progressive careers to move up the political echelons.

So in this section I allow the outsider candidate, R, to choose his amount of office experience before running in the election that he is actually interested in. In other words, in contrast with the previous section, here the parameter e_R is endogenous. The potential candidate R, who at this stage is only a hopeful from outside the political establishment, needs to choose the amount of e_R that maximizes his future payoffs. Being a rational actor, R will forecast the consequences of his choice in the future election that he will participate in. The results in the previous section should therefore be taken here as the "average" or the "typical" election that R can expect to face once he reaches that stage.[12]

How does this potential candidate calculate the costs and benefits of his career choices? I continue to assume that R is motivated by the policy that will be implemented by the government, as given in the utility function in Equation 2. So any benefit from acquiring experience would come from a more favorable policy implemented after the election.[13] As the results below will show, higher amounts of e_R would allow the outsider candidate R to pull policy closer to his ideal point; so he has an incentive to acquire as much experience as possible before the election. However, the exact payoff from his effort will depend on the context, namely his level of charisma c_R, and the level of people's discontent, δ. In addition, there is a cost in acquiring experience, since doing so requires effort and sacrifice. Hence his optimal effort will come out of a cost-benefit analysis that I analyze below.

4.1 Timing, information and solution concept

Suppose that an election for high office will be held at a specific date, and candidate R must decide how much to prepare for it. To be precise, the outsider candidate R needs to choose

[12] To avoid the issue of risk aversion entering R's calculations, we could assume that he is sure of the parameters he will face at this future election.

[13] An extension in the online appendix proves that all the results of the model still hold when we add a payoff from winning the election to the candidates' preferences.

his amount of government experience before running in the election that he is interested in. The timing of this game is the following:

1. All exogenous variables are observed: the people's discontent (δ), candidate R's charisma (c_R) and the expected experience of candidate L (which is 1).
2. Candidate R chooses how much government experience (e_R) to acquire.
3. R's experience (e_R) is observed.
4. Candidates simultaneously choose their platforms (x_R and x_L).
5. The median voter elects a candidate (R or L)

The game must be solved by backward induction, but stages 3, 4 and 5 are identical to the game where experience is exogenous, which was studied in the previous section. So we can take the results from that game as given (namely Theorem 1), and directly study the *reduced game* at Stage 2 of the timing above.

4.2 The benefit and cost of acquiring experience

Candidate R starts off with no office-related experience at all; that is, without any effort on his part, he would enter the election with $e_R = 0$. What would his benefit of increasing e_R be? His main goal is to influence the policy implemented, x^*, after the election. We assume that he can foresee how the election would play out for each one of his possible choices, that is, he understands the election outcomes for any given value of e_R as given in Table 1. To be concrete, assume that R has formulated some beliefs about the parameters that he expects to face in the future contest. He expects the mainstream candidate L to have an ideal point of -1; to have a level of experience of 1; and to have no populist charisma at all. At the same time, he expects popular discontent to be δ and he knows his ideal point to be 1. Last but not least, he knows his charisma to be c_R. He expects those values with certainty.

Given those fixed parameters, Theorem 1 can be restated to give all the election outcomes as a function of R's decision variable e_R. From Equation 2 we know that R's payoff after the election will be $U_R(x^*) = -(1 - x^*)^2$. From Corollary 1, we know that $x^* = A_R$ with $A_R \equiv e_R + \delta c_R - 1$. It can be thus calculated that:

$$U_R(x^*) = -(2 - e_R - \delta c_R)^2$$

From this result we can see that R's payoff from the policy implemented is increasing with prior experience e_R. This creates incentives to increase his office experience throughout his career. It will not come for free, however. Acquiring office experience will have a cost in

terms of effort and resources.[14] We will assume that the cost that R incurs in acquiring any additional experience is given by a quadratic function. To be concrete, the cost of acquiring e_R will be e_R^2.

4.3 The effects of charisma and discontent

We have thus specified in full the costs and benefits to R of acquiring government experience, and we are able to analyze his maximization problem. We will denote by $W(e_R)$ the function of benefits minus costs of choosing a certain level of e_R; it corresponds to the total payoffs to R from pursuing the office that he is interested in – and this is what R will aim to maximize. We have that

$$W(e_R) = -(2 - e_R - \delta c_R)^2 - e_R^2$$

We will call e_R^* the optimal amount of government experience for R. Its value comes from maximizing $W(e_R)$ as given in the previous expression. The following theorem provides this optimal choice as a function of the exogenous parameters.

Theorem 2 *The outsider candidate has a unique optimal choice of government experience for each of the possible values of c_R and δ. The optimal choice is*

$$e_R^* = 1 - \frac{\delta}{2} c_R$$

which is a straight line with respect to c_R, with negative slope and positive values in the relevant interval.

This expression relates the amount of government experience to the level of populist charisma that an outsider candidate is known to have. It implies a remarkable result that has never been stated in the formal literature to my knowledge: there is an inverse relationship between a candidate's charisma and the government experience he will acquire. As can be seen in Theorem 2, higher levels of c_R induce lower levels of e_R^*. In different words, charisma has a *crowding out effect* on experience, whereby charismatic candidates will exert less effort in becoming experienced before an election. This represents an unfortunate trade-off for voters, who can expect their candidates to be experienced or charismatic but not both.

This result allows in turn to calculate the outcome from the high-level election once the outsider candidate is ready to participate. Upon choosing a level of e_R^*, the outsider candidate has in essence determined all the subsequent election outcomes. These are simply given by Table 1 by replacing the value of e_R^*. One of the outcomes that is particularly interesting to

[14] In addition to this cost of effort, at a later section I study another cost of acquiring government experience corresponding to the candidate's lost reputation as an anti-system outsider.

predict is the level of extremism of the outsider candidate R. We will call x_R^{**} the platform that he will choose following his experience in government. This corresponds to the value of x_R^* once e_R^* is chosen. How far from the center will his platform be? As the following theorem shows, this will also depend on his level of populist charisma, c_R.

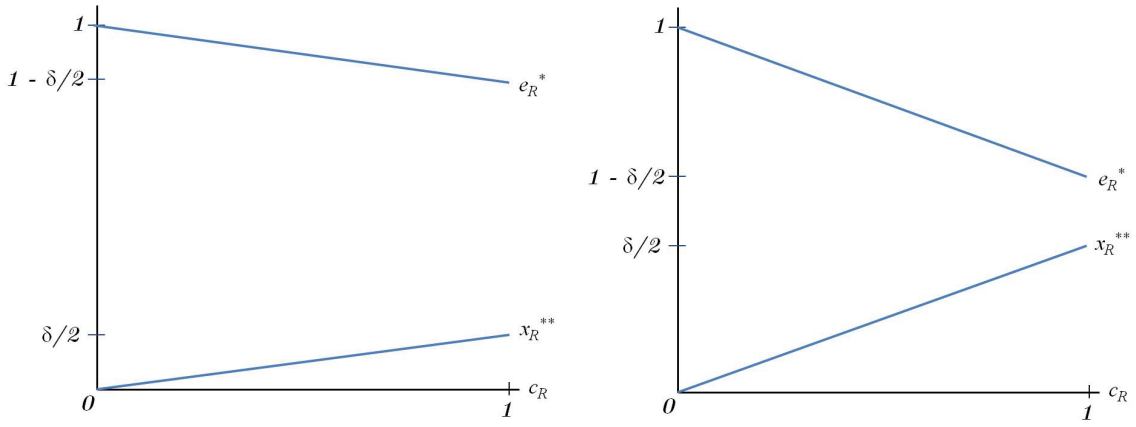
Theorem 3 *Following his optimal amount of government experience, e_R^*, the outsider candidate has a unique optimal choice of policy platform x_R^{**} for each of the possible values of c_R and δ. The optimal is*

$$x_R^{**} = \frac{\delta}{2} c_R$$

which is a straight line with respect to c_R, with positive slope and positive values in the relevant interval.

This expression relates the extremism of the policy offered to voters to the level of populist charisma that an outsider candidate is known to have. The theorem above implies another remarkable result that is worth having in the theoretical literature: there is a positive relationship between a candidate's charisma and the divergence of his platform. From Theorem 3 we can see that higher levels of c_R induce higher levels of x_R^{**}. This occurs because more charismatic candidates can afford more extremist platforms and still win the election. This represents an unfortunate trade-off for centrist voters, who can expect their candidates to be moderate or charismatic but not both.

Figure 3: Charisma discourages experience and encourages extremism



Low popular discontent, δ High popular discontent, δ

Both effects can be seen in Figure 3. Candidates that are more charismatic, i.e. having

higher c_R, will acquire less experience e_R^* (according to Theorem 2) and will adopt a more extremist platform x_R^{**} (according to Theorem 3). These two effects are magnified by people's discontent, δ: a higher discontent increases charisma's negative effect on experience and positive effect on extremism.

As a way to summarize succinctly the total effects of the two main primitive variables, populist charisma and popular discontent, let me state the following corollary which comes directly from the previous results in this paper.

Corollary 2 *All things equal, increasing the amount of populist charisma, c_R, or increasing the amount of popular discontent, δ, has the following effects:*

1. *The optimal amount of experience that the outsider candidate acquires, e_R^*, decreases.*
2. *The extremism of the equilibrium platform that the outsider candidate adopts, x_R^{**}, increases.*

5 A robustness check: government experience is incompatible with populist charisma

The previous analysis assumed that charisma and experience are compatible with each other in the voters' minds, meaning that voters see them as perfect substitutes of each other. To be concrete, voters calculate the valence of a given candidate simply by adding up his level of experience and his level of charisma, without seeing any contradiction between these two traits – this can be seen in Equation 1. Such assumption was convenient to isolate the effects stemming from the candidate's career restrictions, rather than the voters' psychology. An important result of the model was that a candidate's charisma discourages him from acquiring experience. This was due to the restrictions in the candidate's career development: acquiring experience in lower office is costly in terms of time, effort and opportunity cost for the outsider candidate. This cost was enough to yield the result that a candidate will happily trade-off some of his valence, in the form of high charisma, for a lower effort, in the form of low experience. In other words, charisma will crowd-out experience in the candidate's career choices.

Yet, the voters' psychology could introduce an additional restriction in the acquisition of experience. If a populist's appeal to voters is based on his anti-system credentials, will that appeal decrease if the populist gets involved in government activities? It is possible that voters would see the two traits as partially incompatible: if an outsider starts acquiring government experience, voters might on one hand appreciate his increased competence, while on the other hand they might deplore his increased elitism. In particular, the candidate's

critiques of the governing elite might lose credibility when voters notice the amount of years that he has spent as part of this elite. Thus his anti-system rhetoric might lose some legitimacy to the degree he is seen as part of this system. In terms of my model, this alternative viewpoint would postulate that acquiring e_R could somehow decrease c_R.

In real life it is difficult to know how voters compare those two traits in their minds. As I mention in the empirical review later, there are very few statistical studies of charisma and experience together, and I am not aware of any that correlates the two. So it is worth having a theoretical exploration of this issue. In this section I redo all the analysis with the new assumption that experience in government reduces populist charisma. Any similarities in the results can serve as robustness checks of the original model above.

5.1 The new preferences of voters: charisma decreases with experience

In this section I assume that experience is to some degree incompatible with charisma in the following sense: each increase in government experience will decrease the populist charisma of the candidate. Concretely, instead of δc as before, I will assume that populist charisma is now given by $\delta (c - e)$. This represents a very direct way in which experience will be discouraged. In fact this assumption will operate as a second cost to increasing e, in addition to the cost of effort that was postulated in the original model.

Accordingly, instead of Equation 1, the preferences of the median voter are now given by

$$U_M (x, e, c, \delta) = - |x| + e + \delta (c - e) \tag{4}$$

5.2 New equilibrium results: charisma still discourages experience and encourages extremism

We proceed to check how this new utility function for the median voter affects the analysis. How will this change in the voters' view of charisma and experience affect the strategic behavior of the outsider candidate? I will show that the main results are essentially preserved. In particular, the effect of charisma is in all cases the same as before; and the effect of popular discontent is in many cases the same as before. While the calculations carried out by all players are more complex, we still find a unique equilibrium for each set of parameter values.[15]

[15] To present the results more cleanly, in this section I will assume that the expected experience of the mainstream left candidate L is minimal, namely zero. Given that this candidate has no charisma either, this is equivalent to assuming that her valence is zero. This assumption allows showing the effect of R's charisma in its full range. The assumption could be relaxed, but some equilibria would then fail to show the effect of charisma because experience would remain at zero for large intervals, given how costly it is to acquire.

I start with the result that voters and candidates follow basically the same strategies in the election as before. The main difference is how A_R, i.e. the valence advantage of R over L, is calculated. Now that L's valence is zero, and that R's charisma decreases with his experience, we have that $A_R \equiv e_R + \delta (c_R - e_R)$. With this definition, it can easily be proved that we still have $-1 \leq A_R \leq 1$. This is enough to prove that the candidates' choice of platforms in equilibrium follows exactly the same rules as before, albeit with a different definition for A_R, as stated in the following result.

Theorem 4 *If the median voter now has a utility function as in Equation 4, the equilibrium platforms of candidates and the policy outcomes of this election are still given in Table 1 from Theorem 1, but now with* $A_R \equiv e_R + \delta (c_R - e_R)$. *In consequence, we still have* $x^* = A_R$ *as in Corollary 1.*

This allows studying the career choices of the populist candidate at Stage 2 of the game when he needs to choose a level of government experience. The main difference is that acquiring experience is now more costly than it was in the original model. In addition to the cost of effort, now the candidate is wary that acquiring too much experience in government might undermine the anti-elite charisma he initially enjoyed. This will be reflected in a lower investment in experience. Apart from this, the main result of the paper still holds, namely that charisma will discourage experience. As indicated by the following theorem, increasing c_R will decrease the choice of e_R, just as in the original model. The reason is the same as before: high charisma will crowd out effort.

Theorem 5 *If the median voter now has a utility function as in Equation 4, the outsider candidate again has a unique optimal choice of government experience for each of the possible values of* c_R *and* δ. *The optimal choice is*

$$e_R^* = \frac{1-\delta}{\delta^2 - 2\delta + 2} - \frac{\delta (1-\delta)}{\delta^2 - 2\delta + 2} c_R$$

which is still a straight line with respect to c_R *with negative slope and positive values in the relevant interval.*

This in turn allows calculating the extremism that can be expected from the populist candidate as a function of the parameters. The main difference with the original model will be the effect of popular discontent, which will now be ambiguous. In this section, a higher δ will have two countervailing effects on the valence of the populist candidate (and hence on the extremism of his platform): on one hand it will make populist charisma more effective, but on the other hand it will make experience more detrimental to his charisma, resulting in the ambiguous total effect. Other than this, the important result that charisma encourages

extremism still holds. As indicated by the following result, increasing c_R will increase the choice of x_R, just as in the original model. The reason is the same as before: all things equal, higher charisma allows the outsider to adopt a more extremist platform and still win the election.

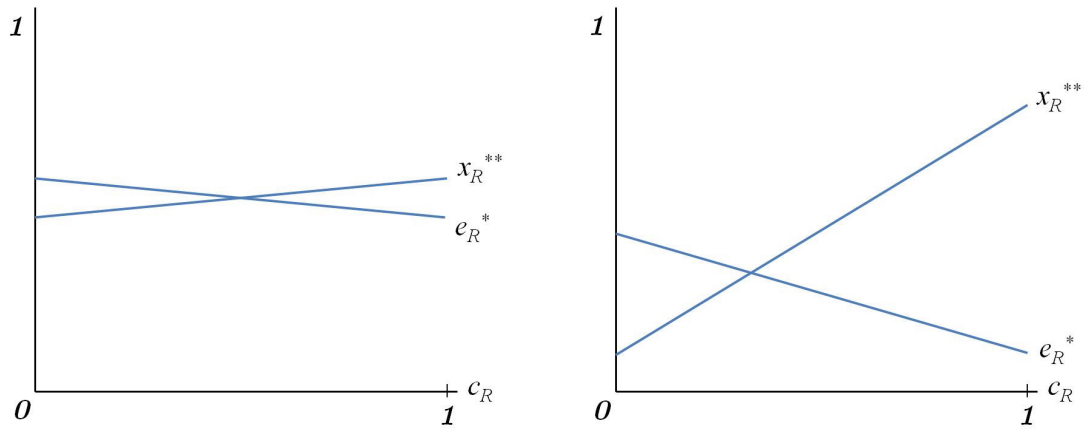
Theorem 6 *If the median voter now has a utility function as in Equation 4, the outsider candidate has a unique optimal choice of policy platform x_R^{**} for each of the possible values of c_R and δ. The optimal is*

$$x_R^{**} = \frac{(1-\delta)^2}{\delta^2 - 2\delta + 2} + \frac{\delta}{\delta^2 - 2\delta + 2} c_R$$

which is still a straight line with respect to c_R with positive slope and positive values in the relevant interval.

The effects of c_R and δ in this new setting can be visualized in the graphs below. As depicted in the graphs, while the effects of popular discontent are more ambiguous than before, the effects of populist charisma remain intact: all things equal, candidates having higher c_R will still acquire less experience e_R^* (according to Theorem 5) and will still adopt a more extremist platform x_R^{**} (according to Theorem 6).

Figure 4: Charisma discourages experience and encourages extremism even when charisma is incompatible with experience



Low popular discontent High popular discontent

The following two corollaries serve to summarize the total effects of each of the main

primitive variables in this section.

Corollary 3 *All things equal, increasing the amount of populist charisma, c_R, has the following effects:*

1. *The optimal amount of experience that the outsider candidate acquires, e_R^*, decreases.*
2. *The extremism of the equilibrium platform that the outsider candidate adopts, x_R^{**}, increases.*

Corollary 4 *All things equal, increasing the amount of popular discontent, δ, has the following effects:*

1. *The optimal amount of experience that the outsider candidate acquires, e_R^*, decreases.*
2. *The extremism of the equilibrium platform that the outsider candidate adopts, x_R^{**}, decreases for $\delta \in \left(0, \frac{1-\sqrt{-2c_R+2c_R^2+1}}{c_R}\right)$, and increases for $\delta \in \left(\frac{1-\sqrt{-2c_R+2c_R^2+1}}{c_R}, 1\right)$.*

In sum, the following effects are preserved in this section compared to the original model:

- The candidate's charisma discourages his investment in experience.
- The candidate's charisma encourages him to adopt a more extremist platform.
- Popular discontent discourages the candidate's investment in experience.
- For high values in the allowed interval, popular discontent encourages the candidate to adopt a more extremist platform.

The only divergent result in this section is that for low values in the allowed interval, popular discontent encourages the candidate to adopt a more centrist platform. Overall, the essence of the model proved to be robust to this new specification.

6 Contrasting the theory to the empirical literature

Interestingly, many of the assumptions and results in this theory are supported by the empirical literature on these topics. Indeed, they find validation in a large number of observations from different subfields, using different methodologies, and coming from different geographical areas. While I do not attempt a comprehensive literature review, in this section I show how several representative publications are consistent with a number of theoretical aspects of my model. Furthermore, this review will illustrate the usefulness of formal theory to organize and explain a number of seemingly disconnected empirical patterns within a single framework. The empirical regularities that find echo in my model are the following.

6.1 Outsider newcomers often adopt an anti-establishment rhetoric

The goal of this model was to study the career decisions made by political outsiders who are interested in competing for high office. I focused on those special citizens who are not identified with the political elite and do not initially have any significant experience in government, and yet have enough resources to run a credible campaign. I presumed that such outsiders often adopt an anti-establishment rhetoric – in effect claiming that my model studied the behavior of *populists*. To sustain this assumption, I review some compelling observations by diverse academics illustrating how political newcomers, more often than not, choose to campaign with an anti-elite discourse.

Several academics remarked that Donald Trump, as a candidate, displayed some of the classic features of populism (Eatwell 2017b). He claimed, for example, to embody the struggle of the neglected common man against the corrupt political elite ("the election is about who runs this country, the special interests or the people.") Inglehart and Norris (2016) explain that in Trump's view, he led an insurgency movement on behalf of ordinary Americans upset with a supposedly corrupt and dishonest establishment (such as crook'd Hillary or lyin' Ted). Another textbook example of anti-system candidate is Hugo Chávez. He had attempted to overthrow the Venezuelan government by leading a military coup barely six years before joining the electoral competition for president. Roberts (2012) explains that *Chavismo* was a moralistic ideology constructed around an alleged dualism between the "virtuous people" and the "rancid and venal oligarchy". As another example, the populist nature of Alberto Fujimori as a candidate is well explained in Levitsky and Loxton (2012). Indeed, his presidential campaign grew increasingly Manichean, for example by promising to sweep away the elite on behalf of the real Peru ("we are the real people"). Finally, fringe parties in Western Europe are also very critical of the cultural and political elite. According to De Lange and Art (2011), radical European parties are populist in their unscrupulous manipulation of public sentiments of anxiety among ordinary men and women who have an allegedly superior common sense. All these examples have served to inspire my model, leading me to assume that the outsider citizen R may have some valence that is valued by voters based on his anti-elite reputation.

6.2 Voters value government experience

One of the assumptions in this paper is that voters value the past governing experience of a candidate. In other words, the electoral support of a candidate increases when he has previously held other elected office or a relevant cabinet position. There is actually a large literature from U.S. politics studying the impact of past political experience on the electoral support of candidates. The experience variable often falls under the rubric of "quality", given

that authors use it as a proxy for the governing quality of different candidates. The recurrent finding is that, all things equal, candidates with more office-holding experience receive more votes.[16] Evidence of this effect has also begun to surface in European politics. Hobolt and Høyland (2011) analyze a data set on the political experience of party representatives in national elections to the European Parliament to evaluate the extent to which voters prefer candidates with more political experience. They find that, yes, parties that choose experienced candidates are rewarded by voters. This justifies including a parameter for experience, e, that is valued positively by voters in Equation 1.

6.3 Experience can be acquired to win future elections

My model is original in treating a candidate's office experience as endogenous, given that I study the decision to acquire its optimal amount to make progress into higher office. This theoretical choice should be reminiscent of a large empirical literature falling under the rubric of *progressive ambition* studying the gradual career decisions that many politicians make to move up the system echelons. This has been observed in the U.S. for a long time,[17] and in other regions such as Latin America more recently. In pioneering work about Brazil, Samuels (2003) shows that aspiring to a seat in the federal legislature is not usually motivated by *static ambition*, but rather by the desire to attain higher office subsequently, such as mayor, senator or governor. In essence, many Brazilian politicians do not seek to occupy a congressional seat for a long time as an end goal *per se*; rather they view it as a potential means for seeking more powerful office. Similar dynamics were found in Uruguay by Chasquetti and Micozzi (2014) and in Argentina by Jones, Saiegh, Spiller and Tommasi (2002). Therefore, assuming that our inexperienced candidate R can choose to acquire some experience e_R in lower office to increase his chances at a higher office reflects very real dynamics around the world.

6.4 Voters value charisma

Another assumption in this paper is that voters are swayed by a candidate's charisma. To be concrete, I assumed that charisma increases electoral support irrespective of other parameters such as the ideological position of the candidate. As I mentioned in the introduction, there are different definitions of charisma and hence different approaches to measuring it. In this paper I preferred following a rather "Weberian" interpretation of charisma, whereby voters are swayed by the populist appeals of a leader who might have a particular talent for connecting to them in ways that create excitement. Evidence of this phenomenon can come from four famous outsiders whose success was in large part based on this type of appeal.

[16] Early contributions were Jacobson and Kernell (1983) and Stone, Maisel and Maestas (2004).
[17] Schlesinger (1966); Abramson, Aldrich and Rohde (1987).

In his presidential campaign of 2016 (and still today as president), Donald Trump was an effective user of old media and new media, strategically using provocation and his celebrity status from the television program 'The Apprentice' (Eatwell 2017b). In spite of being a billionaire, he was remarkably successful at identifying with the average citizen in many localities. To build a connection with his audience at a campaign rally in Pennsylvania, he said "I love blue collar workers, and I consider myself in a certain way to be a blue collar worker. I treat them with dignity, they are great people." Defying most of the polls, it seems that such charismatic connection created sufficient excitement among millions of voters to grant him victory in this state.

In a survey study of the electoral support for Hugo Chávez in 1998, Weyland (2003) shows that voters' optimism about the candidate was inflated by his charisma. The author believes that "Chávez's charismatic personality probably helped to instill hope in his followers." His connection to people depended in part on personal characteristics, such as his crude diction and belligerent rhetoric. According to Weyland's pre-electoral survey, Chávez had a broad base of supporters that was heterogeneous, multi-class and ideologically diverse (which I believe is similar to assuming that he had valence).

Alberto Fujimori was also successful at presenting himself as an outsider and a "man of the people" in his presidential campaign. According to Levitsky and Loxton (2012), Fujimori was at ease in this role – riding a bicycle, wearing a poncho, and speaking his folksy ungrammatical Spanish. As a non-white child of working-class immigrants, he could credibly introduce himself as a Peruvian everyman. Surveys showed that Fujimori's "newness" and lack of partisan ties were his greatest electoral assets.

The success of the French party Front National is often attributed to the personal characteristics of its founder Jean-Marie Le Pen, who was also able to connect with many voters directly. One of Le Pen's most-cited aphorisms claimed that he only said out loud what ordinary French people thought in private (Eatwell 2017a). Experts have argued that "the originality, creativity and charisma of Le Pen had been the predominant factors in Front National's success" (Pedahzur and Brichta 2002). Perhaps for this reason, Jean-Marie Le Pen consistently polled higher in public opinion surveys than his party did.

More broadly, the personal appeal of party leaders has been key to the survival of certain extreme parties in Europe. As argued by De Lange and Art (2011):

> "It has generally been acknowledged that most radical right parties have charismatic leaders who have strong rhetorical skills, are media savvy, and know how to appeal to the ordinary man in the street. These external qualities are important to attract voters." (p. 1233)

All these observations serve to validate placing the variable c in the voters' utility function.

6.5 Charismatic populists tend to be inexperienced

One of the main results in this paper is the negative relationship between charisma and experience (Theorem 2). My model predicts that a charismatic candidate will not bother working in government much. He prefers running for high office directly instead of running for low office first. In essence, charisma has a *crowding out effect* on experience. In contrast, an uncharismatic candidate or party leader can be expected to work harder at the lower levels of government to gain the voters' respect in a high-level election. It is hard to find statistical studies looking at this question. There exist many publications about charisma and many others about experience, but very few with both variables together, and even fewer explicitly correlating them. The qualitative evidence about specific candidates who are deemed charismatic is more informative.

Donald Trump did not have a track record of public service when he launched his presidential campaign in 2015. Academics have described him as inexperienced (Carreras 2016) and a neophyte (Inglehart and Norris 2016). His absence of office experience was not due to a lack of opportunities. In October 2013, GOP assemblymen from New York circulated a memo suggesting enthusiastically that Trump should run for governor of the state.[18] Trump was not consulted before the Republicans circulated their public invitation, but, in any case, he responded that being governor at that stage "is not something that is of great interest to me." So he declined the invitation to run. His lack of appetite for public service was even more salient compared to the extensive record of his main rival in 2016, Hillary Clinton. Eatwell (2017a) claims that Clinton campaigned heavily on experience rather than personality. This was clearly conveyed at one of her debates when she famously said "do you know what else I prepared for? I also prepared to be president." My model suggests a specific causal explanation for this observation about Hillary Clinton: my results predict that an uncharismatic politician who wishes to become president will acquire a lot of office experience.

Hugo Chávez served in the military but did not serve in the civilian government before running for president. He achieved notable popularity by organizing a failed coup d'état against a reviled government. So upon being released from jail, Chávez entered the first presidential election that was held. Kurt Weyland observed that people's expectations of Chávez were strikingly divorced from a dispassionate assessment of his past performance.

> "The very strength of optimism among Chávez supporters is difficult to explain in those terms, given the grave, longstanding problems facing Venezuela and the questionable track record of this candidate, who lacked administrative experience and failed to present a clear socioeconomic program. It seems that

[18] *USA Today*, "N.Y. Republicans want Donald Trump to run for governor", October 14, 2013.

> many citizens simply felt compelled to believe in Chávez's charisma, independent of realistic assessments of his likely success." (Weyland 2003: 825)

Fujimori was described as "the man from nowhere". Before running for president, he was a mathematics professor and an agricultural engineer at a public rural university (Carreras 2016). He had no political experience and no political connections, save for becoming rector of his university and then president of the National Assembly of Public Rectors. It is noteworthy that he contemplated running for the Senate in 1990, but he decided to run for the Presidency instead (Levitsky and Loxton 2012).

As mentioned before, an observation about extreme-right parties in Europe is that they are often headed by charismatic leaders. Another observation is that such leaders are often inexperienced. Pim Fortuyn, a populist rightist in the Netherlands, was eloquent and physically attractive. At televised debates, he always looked directly into the camera to explain his vision in plain terms (De Lange and Art 2011). But he was a political novice and most of his advisors and friends had no political experience whatsoever. This did not stop him from deciding to form his own national party while he had been on the political stage for less than a year. On the other hand, when extreme-right parties lack charismatic leadership, such as the British National Party or the Greek Golden Dawn, they often focus on local campaigns to gain experience before moving on to the national stage (Eatwell 2017a). The contrast between the assertiveness of parties with and without charismatic leaders seems consistent with my theoretical predictions.

6.6 Charismatic populists tend to offer extremist platforms

Another important result of this model is that charisma encourages extremism, in the sense that a charismatic candidate is predicted to adopt an ideological platform further from the ideological center (Theorem 3). I am not aware of statistical studies relating candidates' charisma with their policy positions, but there is significant qualitative evidence suggesting there might be a positive relationship.

Donald Trump's positions on issues have been fluid and sometimes inconsistent – part of his rhetoric in 2016 was even categorized as "leftist". But regarding immigration, which was, and continues to be, one of his main topics, he is most often described as radical right. Compared to most of his mainstream rivals in 2016, even those from his own party, Trump's campaign proposals were more extreme. Inglehart and Norris (2016) described his rhetoric as stirring up a potent mix of racial resentment, intolerance of multiculturalism, nationalistic isolationism, mistrust of outsiders, and anti-Muslim animus. Then they ask: How could such a polarizing figure become the standard-bearer for the GOP – much less have any chance of entering the White House? Inglehart and Norris proposed an answer

that is actually consistent with my theory: it is possible that among many moderate voters, Trump's charisma may have more-than-compensated his extreme views.

Hugo Chávez campaigned on a clear left-wing platform in opposition to market liberalization. He accused global multinationals of being rapacious exploiters at the service of a North American empire, and he advocated a statist economy where large strategic industries should remain in public hands. As summarized in Roberts (2012), "*Chavismo* combined heavy doses of nationalism, socialism, and a charismatic style of political mobilization." His main rival, Henrique Salas, was viewed by voters as being more moderate than Chávez according to pre-electoral polls – but he still lost the election. These are exactly the positions predicted of the winner and the loser in Theorem 1.

While he was leader of his party, Jean-Marie Le Pen displayed all the credentials of a right-wing extremist leader, as he participated in all the struggles of the extreme right including the most violent ones (Pedahzur and Brichta 2002). Since the early 1980s, halting immigration and multiculturalism were his signature themes. Under his charismatic leadership, the Front National steadily improved at the polls until the climactic election of 2002 where he even reached the second round of the presidential elections. In 2011, Jean-Marie was succeeded at the helm of the party by his daughter Marine, a lawyer by profession who has a managerial style that is very different from that of her provocative father (Eatwell 2017a). There is a consensus that Marine Le Pen has moved the party somewhat toward the center. My model offers a possible explanation for Marine Le Pen's moderation compared to her father: my results predict that a less charismatic leader is forced to become more moderate in order to remain competitive.

More broadly, the success of extreme right movements, and their ethnically exclusive ideologies, is often attributed to the appealing personalities of their leaders. For example, Nigel Farage is in part credited for the remarkable rise since 2014 of the UK Independence Party (Goodwin and Milazzo 2015). Indeed, according to Van der Brug and Mughan (2007), the new radical right parties and their fascist predecessors in Europe share in common the prevalence of charismatic leadership, which is held to sway voters. For this reason, the significance of leaders' attributes has been central to debates about the rise since the 1980s of populist parties in Europe: charisma is often seen as an important factor in explaining their success (Eatwell 2017a).

6.7 Populist outsiders are more likely to compete successfully in elections when there is popular discontent against the state

My model studied the effects of popular disenchantment on the career decisions by a new candidate seeking high office. A parameter for popular discontent, δ, was meant to cap-

ture the degree to which the electorate was receptive to anti-establishment appeals against the political system as a whole. The main result was that high levels of popular discontent, stemming for example from bad economic, political or social conditions, lead to a higher likelihood of such citizens winning the race while choosing to acquire very little prior experience in office (Theorem 4 and Corollary 2). There exist numerous empirical studies suggesting that candidates considered to be non-mainstream can thrive when there is popular discontent. A number of politicians who have been labeled newcomers, outsiders, populists, anti-establishment, etc., have succeeded in becoming the chief executive during times of economic downturn, political crisis or social unrest.

Academics believe that Donald Trump benefitted from a protest vote that helped him to victory. The premise is that worsening economic and social conditions led important sectors of the population to lose faith in the capacity of mainstream politicians to respond to their concerns. In America, there is still resentment for the 2008 global financial crisis, which compounded a long-term transformation in the workforce stemming from technological automation, the collapse of manufacturing industry, the inflow of migrant labor, and historically high levels of income inequality (Inglehart and Norris 2016). Economic concerns might have made large strata of society more susceptible to the anti-establishment appeals of a charismatic populist. In addition, Eatwell (2017a) believes that recent fears in the USA about immigration raised existential concerns about the social order; consequently, many voters who felt threatened were "willing to take a risk on an inexperienced politician like Trump."

Venezuela had endured two decades of chronic economic hardship when Hugo Chávez was elected president. Excessive spending during the mid 1970s quickly degenerated into a debt crisis when oil prices dropped. The 1980s saw further economic decline due to ill-conceived liberalization attempts; and inflation reached 8% per month before the election in 1998 (Weyland 2003). People attributed their economic difficulties to the two oligarchic parties, AD and COPEI, which had a track record of incompetence and malfeasance, having colluded to monopolize the political arena (Roberts 2012). Their poor performance was a prelude to a popular backlash against the party system. Enter Chávez. A statistical analysis of a large pre-electoral survey concluded that:

> "Popular dissatisfaction with the actual state of democracy had a highly significant impact on vote intentions for Hugo Chávez. (...) Supporters of this radical populist clearly tended to reject the way in which the political class had been running the country and to agree with the mounting criticism of the established political system. Thus political discontent played an important role in inducing Venezuelans to vote for a risky outsider." (Weyland 2003: 833)

Alberto Fujimori came to power in the midst of the most challenging economic, social and

political times of Peru's contemporary history. According to Levitsky and Loxton (2012), this political newcomer's rise from obscurity to the presidency was rooted in a double crisis. First, Peru was facing economic collapse. The heterodox policies attempted during the late 1980s had catastrophic consequences as they resulted in hyperinflation, a major increase in foreign debt, and a significant drop in GDP before the presidential campaign of 1990. Second, Peru suffered the rise of a powerful guerrilla movement, the Shining Path, which was one of the most violent insurgent groups in Latin America. By the end of the 1980s, this Maoist group had killed more than 25,000 people and was controlling a quarter of Peru's municipalities. This brought the Peruvian state to the brink of collapse, raising the specter of a Shining Path victory. In this context of double crisis, public disaffection with the established political parties soared. Carreras (2016) explains that Fujimori exploited this popular disaffection with the political class to pave his way toward victory.

Much work has studied the reasons why extreme right parties have been successful in Western Europe. Some of it argues that globalization, through a variety of interrelated processes, has created a fertile breeding ground for extremism (Mudde 2007). Indeed, many academics hold that charismatic leaders are most likely to emerge at times of major social change, especially when economic crisis coincides with political crisis (Eatwell 2017a). The literature has argued that voting for right-wing populist parties is partly a protest vote. In this view, populist party supporters abandon their traditional parties to send a message of protest against inefficiency, incompetence and incumbents in general (Van der Brug and Mughan 2007). Economic problems in many countries since 2008 have further increased support for policies advocated by these parties, such as restricting immigration. For example, this has been true in France, where the radical right has gathered momentum following the global financial crisis. Eatwell (2017a) attributes the sudden take-off of the Front National to growing concerns about immigration among French voters. All this is consistent with Figure 3 showing that all effects in my model are steeper with a higher δ.

7 The trade-offs between charismatic populism, government experience and policy moderation

This paper studied some likely causes and consequences of charismatic populism. I presumed in my theory that outsider politicians will tend to adopt an anti-establishment rhetoric, making this assumption in my model and citing empirical evidence. But when will these populist appeals by outsider politicians be effective? The model posited a variable which was meant to capture the personal characteristics of the outsider candidate that would determine the effectiveness of his populist appeals to voters; I called it *populist charisma* and labeled it c. In the model, this type of charisma is "populist" in the following ways: it is valued

positively by a broad class of voters across ideological lines as a valence issue; only the outsider newcomer may possess a certain amount of it; the mainstream candidates do not have any populist charisma; and the impact of this parameter is increasing in the amount of popular discontent at the beginning of the game, called δ. Having defined these terms, the focus of my paper was in analyzing the effects of populist charisma on the democratic process. To be concrete, I focused on studying its effect on the career decisions of an outsider politician, with the ensuing effects on the behavior of voters and other candidates in an eventual election.

Among the most relevant results in the model are two unfortunate trade-offs that voters are expected to face. On one hand, the electorate may find appeal in charismatic populism. Part of the appeal may be emotional: as mentioned before, the most successful populists have been able to inspire passionate support thanks to their charismatic connection to ordinary people. But voters can also conceive of instrumental and tangible benefits of electing an anti-establishment outsider. These politicians often try to deliver on their promise of weakening the state capture by vested interests representing a corrupt political class. They promote fresh faces in government, enabling some elite rotation. And they frequently attempt to establish more channels of direct democracy, such as plebiscites and bottom-up institutions to make the government more responsive.

But my model suggested two costs that have not been frequently pointed out in the existing literature, and had not been derived together in a unified theory. A first cost is electing an inexperienced politician with little previous exposure to governing. According to my results, if an outsider politician has large amounts of charisma, he will optimally choose to seek high office immediately without much prior preparation; in particular he will choose to skip the effort of acquiring office experience by taking a cabinet position or running for lower office. In other words, a candidate's initial charisma will crowd out his effort at acquiring experience. A second cost is forming an extremist government that will implement policies far from the median voter's ideology. According to another result from the model, if an outsider politician has high levels of charisma he will be able to adopt an extremist platform and still win the election. He can do so because his valence as a populist more than compensates his far-out ideological platform. Upon getting elected, he will then carry out his program of extreme (left or right) policies that could alienate a mass of moderate voters.

These new results suggest that voters will be systematically disappointed in some of the major issues they care about in elections. In particular, it is unrealistic to expect a winning candidate to possess all of the main qualities desired from a head of government, such as experience, moderation and charisma. The results here uncover some fundamental contradictions among these three qualities stemming from the democratic process itself. This should give pause to centrist voters before electing a charismatic populist, as they risk

placing an inexperienced extremist at the helm. If they wish to be governed by a competent technocrat with prudently moderate policies, they might have to accept a boring workhorse with ties to the established elite.

References

[1] Abramson, Paul R., John H. Aldrich, and David W. Rohde. "Progressive ambition among United States senators: 1972-1988." The Journal of Politics 49, no. 1 (1987): 3-35.

[2] Acemoglu, Daron, Georgy Egorov, and Konstantin Sonin. "A political theory of populism." The Quarterly Journal of Economics 128, no. 2 (2013): 771-805.

[3] Adams, James, Samuel Merrill III, Elizabeth N. Simas, and Walter J. Stone. "When candidates value good character: A spatial model with applications to congressional elections." The Journal of Politics 73, no. 1 (2011): 17-30.

[4] Aragonès, Enriqueta, and Dimitrios Xefteris. "Imperfectly informed voters and strategic extremism." (2013), manuscript.

[5] Ashworth, Scott, and Ethan Bueno De Mesquita. "Elections with platform and valence competition." Games and Economic Behavior 67, no. 1 (2009): 191-216.

[6] Buttice, Matthew K., and Caitlin Milazzo. "Candidate positioning in Britain." Electoral Studies 30, no. 4 (2011): 848-857.

[7] Callander, Steven. 2008. "Political Motivations." Review of Economics Studies 75: 671-697.

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

[9] Calvo, Ernesto, and Maria Victoria Murillo (2017), *Non-Policy Politics: Richer voter, poorer voter, and the diversification of electoral strategies*, book manuscript.

[10] Calvo, Ernesto, and Timothy Hellwig. "Centripetal and centrifugal incentives under different electoral systems." American Journal of Political Science 55, no. 1 (2011): 27-41.

[11] Carreras, Miguel., 2016. Institutions, governmental performance and the rise of political newcomers. European Journal of Political Research.

[12] Carter, Jennifer, and John W. Patty. "Valence and campaigns." American Journal of Political Science 59, no. 4 (2015): 825-840.

[13] Carrillo, Juan and Micael Castanheira. 2008. "Information and Strategic Political Polarization." Economic Journal 118 July: 845-874.

[14] Chasquetti, Daniel, and Juan Pablo Micozzi. "The subnational connection in unitary regimes: Progressive ambition and legislative behavior in Uruguay." Legislative Studies Quarterly 39, no. 1 (2014): 87-112.

[15] Crisp, Brian F., John W. Patty, Elizabeth Maggie Penn and Constanza Schibber (2014), "Valence, Partisan Competition, & Legislative Institutions," paper presented at the annual meeting of the Midwest Political Science Association.

[16] De Lange, Sarah L., and David Art (2011) "Fortuyn versus Wilders: An agency-based approach to radical right party building." West European Politics 34, no. 6 : 1229-1249.

[17] Dewan, Torun, and David P. Myatt. "Leading the party: Coordination, direction, and communication." American Political Science Review 101, no. 04 (2007): 827-845.

[18] Dewan, Torun, and David P. Myatt. "The qualities of leadership: Direction, communication, and obfuscation." American Political science review 102, no. 03 (2008): 351-368.

[19] Dewan, Torun, and David P. Myatt. "On the rhetorical strategies of leaders: Speaking clearly, standing back, and stepping down." Journal of Theoretical Politics 24, no. 4 (2012): 431-460.

[20] Eatwell, Roger (2017a). "Charisma and the Radical Right". In The Oxford Handbook of the Radical Right, Editor Jens Rydgren, Oxford University Press, Oxford. Forthcoming.

[21] Eatwell, Roger (2017b). "Populism and Fascism". In The Oxford Handbook of Populism, Editors Rovira Kaltwasser Taggart and Ochoa Espejo, Oxford University Press, Oxford. Forthcoming.

[22] Fox, Justin, and Matthew C. Stephenson. "The welfare effects of minority-protective judicial review." Journal of Theoretical Politics 27, no. 4 (2015): 499-521.

[23] Goodwin, Matthew, and Caitlin Milazzo. 2015. *UKIP: Inside the campaign to redraw the map of British politics*. Oxford University Press,

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

[25] Hitt, Matthew P., Craig Volden, and Alan E. Wiseman. "Spatial models of legislative effectiveness." American Journal of Political Science (2017).

[26] Hobolt, Sara B., and Bjørn Høyland. "Selection and sanctioning in European Parliamentary elections." British Journal of Political Science 41, no. 03 (2011): 477-498.

[27] Inglehart, Ronald, and Pippa Norris. (2016), "Trump, Brexit, and the rise of Populism: Economic have-nots and cultural backlash." Harvard Kennedy School, Faculty Research Working Paper RWP16-026.

[28] Jacobson, G.C., Kernell, S., 1983. Strategy and Choice in Congressional Elections. Yale University Press, New Haven.

[29] Jones, Mark P., Sebastián Saiegh, Pablo T. Spiller, and Mariano Tommasi. "Amateur legislators–professional politicians: The consequences of party-centered electoral rules in a federal system." American Journal of Political Science (2002): 656-669.

[30] Levitsky, Steven, and James Loxton (2012). "Populism and Competitive Authoritarianism: The Case of Fujimori's Peru'." Populism in Europe and the Americas: Threat or Corrective for Democracy: 160-181.

[31] Meirowitz, Adam. 2008. "Electoral Contests, Incumbency Advantages, and Campaign Finance." *Journal of Politics* 70 (03) July: 681-699.

[32] Micozzi, Juan Pablo, and Sebastián M. Saiegh. "An empirical stochastic model of Argentina's Impossible Game (1955–1966)." Journal of Theoretical Politics 28, no. 2 (2016): 266-287.

[33] Mudde, Cas (2007). *Populist radical right parties in Europe*. Cambridge University Press. Cambridge.

[34] Pedahzur, Ami, and Avraham Brichta (2002). "The institutionalization of extreme right-wing charismatic parties: a paradox?." Party Politics 8, no. 1: 31-49.

[35] Penn, Elizabeth Maggie (2009), "From Many, One: State Representation and the Construction of an American Identity," *Journal of Theoretical Politics,* 21(3): 343–364.

[36] Roberts, Kenneth, 2012. "Populism and democracy in Venezuela under Hugo Chávez." In Mudde, Cas, and Cristóbal Rovira Kaltwasser. Populism in Europe and the Americas: Threat or corrective for democracy?. Cambridge University Press: 136-59.

[37] Samuels, David. (2003), Ambition, Federalism, and Legislative Politics in Brazil. New York: Cambridge University Press.

[38] Schlesinger, Joseph. 1966. Ambition and Politics: Political Careers in the United States. Chicago: Rand McNally.

[39] Schofield, Norman, Christopher Claassen, Ugur Ozdemir, and Alexei Zakharov. "Estimating the effects of activists in two-party and multi-party systems: comparing the United States and Israel." Social choice and welfare 36, no. 3-4 (2011): 483-518.

[40] Serra, Gilles. 2010. "Polarization of What? A Model of Elections with Endogenous Valence," The Journal of Politics, Vol. 72, Issue 2, April, pp. 426-437.

[41] Stokes, Donald. 1963. "Spatial Models of candidate Competition." American Political Science Review 57: 368-377.

[42] Stone, Walter J., L. Sandy Maisel, and Cherie D. Maestas. "Quality counts: Extending the strategic politician model of incumbent deterrence." American Journal of Political Science 48, no. 3 (2004): 479-495.

[43] Van der Brug, Wouter, and Anthony Mughan.(2007) "Charisma, leader effects and support for right-wing populist parties." Party Politics 13, no. 1 : 29-51.

[44] Weber, Max. 1978. "Charisma and Its Transformation." Pp. 1111-57 in *Economy and Society: An Outline of Interpretive Sociology*, vol. 2, edited by G. Roth and C. Wittich. Berkeley: University of California Press.

[45] Weyland, Kurt. "Economic voting reconsidered: Crisis and charisma in the election of Hugo Chávez." *Comparative Political Studies* 36, no. 7 (2003): 822-848.

A Appendix

A.1 Extension: Mixed office and policy motivations

The original specification assumed that candidates had only one motivation: influencing the platform implemented by the government in their favorite direction. In my model, candidates have preferences over policies, just like citizens do. Akin to voters who have ideal points in the left-right political spectrum, each candidate also has an ideal point reflecting his ideology. When candidates have this type of preferences, they are often said to be "policy motivated" (as in Calvert 1985). This assumption makes sense to analyze the desires of outsider candidates who wish to use their charisma to influence the election. However it is worth considering a more complex scenario that includes the candidates' desire to win the election per se. A majority of formal models of spatial elections assume that candidates receive a direct payoff from being victorious, which can be interpreted in many ways, such enjoying power and the perquisites of office. Generically, the literature refers to those benefits as "ego-rents", and candidates who have this type of preferences are often said to be "office motivated" (Calvert 1985). In this section I study the more complete scenario were candidates have both types of preferences, being simultaneously policy-motivated and office-motivated. I will show that adding this realistic mixture of preferences does not change the results at all; in fact the theorems and corollaries remain intact because the two types of motivations do not conflict with each other in equilibrium. In all relevant situations, a candidate either wishes to move from losing to winning in order to obtain a better policy outcome; or he wishes to move from winning with a less preferred platform to winning with a more preferred platform. So adding office motivation to the model does not conflict with policy motivation, which allows all the results to be preserved. Therefore this section serves as a robustness check.

A.1.1 The new preferences of candidates: office and policy motivations

Except for the preferences of candidates, in this section I follow all the specifications of the original model. In particular, the median voter's utility function is still given by Equation 1, the mainstream left candidate is still expected to have a high experience of 1 and a low charisma of 0; charisma and experience are compatible as perfect substitutes of each other; and acquiring the amount of experience e_R carries a unique quadratic cost of effort e_R^2. The only difference is that candidates now have the following preferences, which replace Equations 2 and 3.

$$
\begin{aligned}
U_R(x) &= -(1-x)^2 + G \qquad (5)\\
U_L(x) &= -(-1-x)^2 + G \qquad (6)\\
\text{where } G &= \begin{cases} g \text{ if the candidate wins, with } g > 0 \\ 0 \text{ if the candidate loses} \end{cases}
\end{aligned}
$$

Hence the value G represent the gains from victory per se, irrespective of the policy implemented by the government after the election: if the candidate wins, he gets a strictly positive amount g of ego-rents; otherwise he gets no ego-rents. This is in addition to the utility experienced from policy.

A.1.2 All equilibrium results still hold

Now I state the equilibrium results in this new scenario, showing that all the results in the model still hold intact (the proofs come at the end of this appendix). We start with the last stages of the game, i.e. the strategic choices of platforms by candidates, and the choice of an election winner by voters.

Theorem 7 *If candidates now value winning office per se in addition to the policy implemented, as in Equations 5 and 6, the equilibrium platforms of candidates and the policy outcomes of this election are still given in Table 1 from Theorem 1.*

This allows deriving the optimal strategies by candidate R in earlier stages of the game. The optimal level of experience is the following.

Theorem 8 *If candidates now value winning office per se in addition to the policy implemented, as in Equations 5 and 6, the optimal choice of e_R is the same as in Theorem 2, namely*

$$e_R^* = 1 - \frac{\delta}{2} c_R$$

Therefore adding a gain g from winning the election does not change this result.

And the level of extremism adopted by R in equilibrium is the following.

Theorem 9 *If candidates now value winning office per se in addition to the policy implemented, as in Equations 5 and 6, the optimal choice of x_R^{**} is the same as in Theorem 3, namely*

$$x_R^{**} = \frac{1}{2} \delta c_R$$

Therefore adding a gain g from winning the election does not change this result.

Hence charisma discourages experience and encourages extremism in the same way as in the original model. Accordingly, all the results regarding the effect of δ remain the same as well.

A.2 Extension: Candidates with linear utility functions over policy

The model in this paper postulated that candidates have quadratic loss functions over policy, as specified in Equations 2 and 3. In contrast, the readers may note that I postulated that voters have linear loss functions over policy, as specified in 1. I justified this difference explaining that a quadratic specification allows candidates to be highly sensitive to different parameters that affect the policy outcome. For instance, we could think of candidates as having very intense preferences that make them more sensitive to policy changes than the average non-politicized citizen. However, for consistency, it is worth exploring how the model would change if the candidates had the same utility functions over policy as voters, namely, linear loss functions. In this section I determine which results remain and which ones disappear with such change.

A.2.1 The new preferences of candidates: linear loss functions over policy

Except for the preferences of candidates, in this section I follow all the specifications of the original model. In particular, the median voter's utility function is still given by Equation 1, the mainstream left candidate is still expected to have a high experience of 1 and a low charisma of 0; charisma and experience are compatible as perfect substitutes of each other; and acquiring the amount of experience e_R carries a unique quadratic cost of effort e_R^2. The only difference is that candidates now have the following preferences, which replace Equations 2 and 3.

$$U_R(x) = -|1-x| \tag{7}$$

$$U_L(x) = -|-1-x| \tag{8}$$

A.2.2 New equilibrium results: charisma encourages extremism but has no effect on experience

I start by showing that the last stage of the game, i.e. the election, is played the same way as in the original model.

Theorem 10 *If candidates now have linear loss functions over policy, as in Equations 7*

and 8, the equilibrium platforms of candidates and the policy outcomes of this election are still given in Table 1 from Theorem 1.

This allows deriving the optimal strategies by candidate R in earlier stages of the game. As indicated in the following result, the optimal level of experience does not depend on charisma or popular discontent anymore.

Theorem 11 *If candidates now have linear loss functions over policy, as in Equations 7 and 8, the optimal choice of e_R is always*

$$e_R^* = \frac{1}{2}$$

which is a constant value that does not change with c_R or δ.

Hence the optimal amount of experience does not depend on the primitive parameters in the model, and this occurs because the marginal benefit of experience does not depend on those parameters anymore. To be precise, the benefit of experience in terms of obtaining a better policy is now insensitive to whether this policy is far or close to the candidate's ideal point. In contrast, in the original model with quadratic loss functions, the candidate was more desperate for experience when he had low charisma than when he had high charisma, because policy would be farther from its ideal point in the former case than the latter case.

In consequence we lose the effect that charisma discourages experience in this setup. However, the effect that charisma encourages extremism is partially preserved, as indicated in the following result.

Theorem 12 *If candidates now have linear loss functions over policy, as in Equations 7 and 8, the optimal choice of x_R^{**} for each possible value of c_R and δ is*

$$x_R^{**} = \begin{cases} 0 \text{ if } c_R < \frac{1}{2\delta} \\ \delta c_R - \frac{1}{2} \text{ if } c_R \geq \frac{1}{2\delta} \end{cases}$$

which is flat line for low values of c_R up until $c_R = \frac{1}{2\delta}$, after which it is an upward sloping line with positive values.

Hence, in this scenario, very high levels charisma will still encourage extremism.

A.3 Proofs of all the results

A.3.1 Theorem 1

Proof. First we need to derive how the median voter M chooses who to vote for. If $U_M(x_L, 1, 0, \delta) < U_M(x_R, e_R, c_R, \delta)$ then M will vote for R. If $U_M(x_R, e_R, c_R, \delta) < U_M(x_L, 1, 0, \delta)$

then M will vote for L. If $U_M(x_R, e_R, c_R, \delta) = U_M(x_L, 1, 0, \delta)$ then M will vote according to the indifference assumptions stated in the text: if candidates have different valences then M will vote for the highest-valence one; if candidates have the same valences then M will randomize equally between the two.

To be more precise, M will vote according to the following three rules which depend on the value of $A_R \equiv e_R + \delta c_R - 1$.

If $A_R > 0$, then M will vote for R if and only if $U_M(x_L, 1, 0, \delta) \leq U_M(x_R, e_R, c_R, \delta)$ $\Leftrightarrow -|x_L| + 1 \leq -|x_R| + e_R + \delta c_R \Leftrightarrow |x_R| - |x_L| \leq e_R + \delta c_R - 1 \Leftrightarrow |x_R| - |x_L| \leq A_R$. Otherwise she will vote for L.

If $A_R < 0$, then M will vote for L if and only if $U_M(x_R, e_R, c_R, \delta) \leq U_M(x_L, 1, 0, \delta)$ $\Leftrightarrow A_R \leq |x_R| - |x_L|$. Otherwise she will vote for R.

If $A_R = 0$, then M will vote for R if $|x_R| - |x_L| < A_R$; she will vote for L if $|x_R| - |x_L| > A_R$; or she will randomize equally between R and L if $|x_R| - |x_L| = A_R$.

This allows calculating how far the valence-advantaged candidate can pull his platform away from the median voter while still winning the election. From the equations above we can see that if R has higher valence, he can pull his platform x_R toward the right all the way to $A_R + |x_L|$ and still win the election. If L has higher valence, he can pull his platform x_L toward the left all the way to $A_R - |x_R|$. If both candidates have the same valence, neither can pull his platform farther than the other, otherwise he will lose for sure. All this leads to different cases that must be analyzed separately.

We can now derive how the candidates will choose their platforms in equilibrium. To do so we need to divide A_R in the following intervals.

- Case $A_R = 0$:

In this case neither of the candidates has a valence advantage over the other. M will therefore vote for the candidate whose platform is closest to zero, or will randomize equally between the two if both platforms are equidistant from zero (as established by the indifference assumptions). It is well known in this setting that the unique Nash equilibrium is for both candidates to converge to the median voter's ideal point (see for example Calvert 1985).

- Case $0 < A_R \leq 1$:

Let us study all the possible locations of x_L to see which ones can be sustained in a NE. If x_L is too extreme, namely $|x_L| \geq 1 - A_R$, then R can propose $x_R = 1$, which is his ideal point, and still win the election. That is because $|x_L| \geq 1 - A_R \Rightarrow |x_L| \geq 1 - (e_R + \delta c_R - 1)$ $\Rightarrow -|x_L| + 1 \leq -1 + e_R + \delta c_R \Rightarrow U_M(x_L, 1, 0, \delta) \leq U_M(1, e_R, c_R, \delta)$. But then L could deviate unilaterally to a more moderate platform and win the election, and therefore this cannot be an equilibrium.

On the other hand, if x_L is moderate enough without reaching zero, namely $0 < |x_L| < 1 - A_R$, then R's best response is to propose the rightmost platform that allows it to win the election, which is $x_R = |x_L| + A_R$. By doing so, $U_M(x_L, 1, 0, \delta) = U_M(|x_L| + A_R, e_R, c_R, \delta)$ and R wins the election (due to our indifference assumption when $A_R > 0$). But then L could adopt a more centrist platform and win the election, so this cannot be an equilibrium.

Only if $x_L = 0$ and $x_R = A_R$ we have R best-responding to x_L and L best-responding to x_R, and this is because L has become as centrist as possible in order to constrain R's extremism. This is therefore the only NE.

- Case $-1 \leq A_R < 0$:

This is the mirror image of the case $0 < A_R < 1$. Hence the unique NE is for candidate R to converge to the median voter's ideal point, $x_R^* = 0$, and for candidate L to adopt the leftmost platform that can win the election, i.e. $x_L^* = A_R$. ■

A.3.2 Corollary 1

Proof. The result comes directly from looking at the third column of Table 1. Note that this is only true because $A_R \in [-1, 1]$, which comes from the allowed intervals for c, e and δ. A more general result including values of A_R outside of this interval can be found in Serra (2010). ■

A.3.3 Theorem 2

Proof. We start by recalling from Corollary 1 that in equilibrium we have $x^* = A_R$ with $A_R \equiv e_R + \delta c_R - 1$. This allows us to calculate the utility that R derives from policy in equilibrium, which is $U_R(x^*) = -(1 - e_R - \delta c_R + 1)^2$. By subtracting the cost of acquiring e_R from this expression, we obtain $W(e_R)$, which is $W(e_R) = -(2 - e_R - \delta c_R)^2 - e_R^2$.

Given this result, what is the optimal e_R that R could adopt? R needs to solve the problem

$$\max_{e_R} W(e_R) \text{ subject to } 0 \leq e_R \leq 1$$

The first-order conditions give us the critical point $e_R = \frac{2-\delta c_R}{2}$. The proof is the following: $\frac{\partial W(e_R)}{\partial e_R} = 0 \Rightarrow -2(2 - e_R - \delta c_R)(-1) - 2e_R = 0 \Rightarrow 4 - 2e_R - 2\delta c_R - 2e_R = 0 \Rightarrow 4 - 4e_R - 2\delta c_R = 0 \Rightarrow \frac{2-\delta c_R}{2} = e_R$. A quick look at the second-order conditions proves that this critical point is a maximum. The proof is the following: $\frac{\partial^2 W(e_R)}{\partial e_R^2} = -4 < 0$.

Is this critical point in the required interval?

Given our assumption that $c_R \leq 1$ and $\delta \leq 1$, we have that $\frac{2-\delta c_R}{2} \geq 0$. The proof is the following: $\delta c_R \leq 1 \Rightarrow \delta c_R \leq 2 \Rightarrow 0 \leq 2 - \delta c_R \Rightarrow 0 \leq \frac{2-\delta c_R}{2}$.

On the other hand, given our assumptions that $c_R \geq 0$ and $\delta \geq 0$, we have that $\frac{2-\delta c_R}{2} \leq 1$. The proof is the following: $\delta c_R \geq 0 \Rightarrow -\delta c_R \leq 0 \Rightarrow 2 - \delta c_R \leq 2 \Rightarrow \frac{2-\delta c_R}{2} \leq 1$.

Therefore $\frac{2-\delta c_R}{2}$ satisfies the constraints and is an interior maximum. We can prove it is positive the following way: $\frac{2-\delta c_R}{2} > 0 \Leftrightarrow 2 - \delta c_R > 0 \Leftrightarrow 2 > \delta c_R$ which is always true in the ranges $c_R \in [0,1]$ and $\delta \in [0,1]$.

Rewriting this expression we get that $e_R = 1 - \frac{\delta}{2} c_R$ which is a straight line with slope $-\frac{\delta}{2}$, which is strictly negative for all possible values of δ except $\delta = 0$. ■

A.3.4 Theorem 3

Proof. Now that experience is endogenous, candidate R will start the election with an optimally chosen level of experience, e_R^*. From Table 1, we know that $x_R^* = A_R$ if $A_R \geq 0$, and $x_R^* = 0$ if $A_R \leq 0$. So we should start by calculating the equilibrium value of A_R, which we will call A_R^*. We know from Theorem 2 that $e_R^* = 1 - \frac{\delta}{2} c_R$. Plugging this value into A_R we obtain $A_R^* = e_R^* + \delta c_R - 1 = 1 - \frac{\delta}{2} c_R + \delta c_R - 1 = \frac{\delta}{2} c_R$. Given that $c_R \geq 0$ and $\delta \geq 0$, we have that $A_R^* \geq 0$ and therefore $x_R^{**} = A_R^* = \frac{\delta}{2} c_R$. We can see that this value is positive in the ranges $c_R \in [0,1]$ and $\delta \in [0,1]$. This expression is clearly a straight line with slope $\frac{\delta}{2}$, which is strictly positive for all possible values of δ except $\delta = 0$. ■

A.3.5 Corollary 2

Proof.

1. From Theorem 2, we know that $e_R^* = 1 - \frac{\delta}{2} c_R$. This expression is linearly decreasing in c_R and δ.

2. From Theorem 3, we know that $x_R^{**} = \frac{\delta}{2} c_R$. This expression is linearly increasing in c_R and δ.

■

A.3.6 Theorem 4

Proof. Throughout Section 5, we must recall the assumption I am making about the total valence of L, namely that it is zero. This implies that in this section we have $A_R \equiv e_R + \delta c_R (c_R - e_R) = (1 - \delta) e_R + \delta c_R$.

For Table 1 to apply in this new setup, we need to prove that indeed we have $-1 \leq A_R \leq 1$. So let me prove both equalities.

On one hand,

$A_R \leq 1$

$\Leftrightarrow (1 - \delta) e_R + \delta c_R \leq 1$ for any values of e_R and c_R such that $e_R \in [0,1]$ and $c_R \in [0,1]$

$\Leftrightarrow (1-\delta)(1)+\delta(1)\leq 1$

$\Leftrightarrow 1\leq 1$ which is always true.

On the other hand,

$-1\leq A_R$

$\Leftrightarrow -1\leq (1-\delta)e_R+\delta c_R$ for any values of e_R and c_R such that $e_R\in[0,1]$ and $c_R\in[0,1]$

$\Leftrightarrow -1\leq (1-\delta)(0)+\delta(0)$

$\Leftrightarrow -1\leq 0$ which is always true.

Therefore we have that $A_R\in[-1,1]$.

By looking at the third column of Table 1 it is still true that $x^*=A_R$. ■

A.3.7 Theorem 5

Proof. We start by recalling that in equilibrium we have $x^*=A_R$, which comes from simply looking at the third column of Table 1. We also recall that $A_R=e_R+\delta(c_R-e_R)$. This allows us to calculate the utility that R derives from policy in equilibrium, which is $U_R(x^*)=-(1-e_R+\delta e_R-\delta c_R)^2$. By subtracting the cost of acquiring e_R from this expression, we obtain $W(e_R)$, which is $W(e_R)=-(1-e_R-\delta(c_R-e_R))^2-e_R^2$.

Given this result, what is the optimal e_R that R could adopt? R needs to solve the following problem:

$$\max_{e_R} W(e_R) \text{ subject to } 0\leq e_R\leq 1$$

The first-order conditions give us the critical point $e_R=\frac{(1-\delta)(1-\delta c_R)}{\delta^2-2\delta+2}$. This comes from the following calculations:

$\frac{\partial W(e_R)}{\partial e_R}=0\Leftrightarrow 4\delta e_R-4e_R-2\delta-2\delta c_R-2\delta^2e_R+2\delta^2c_R+2=0\Leftrightarrow e_R=\frac{(1-\delta)(1-\delta c_R)}{\delta^2-2\delta+2}$
$\Leftrightarrow e_R=\frac{1-\delta}{\delta^2-2\delta+2}-\frac{\delta(1-\delta)}{\delta^2-2\delta+2}c_R$

The second order conditions show this critical point is a maximum. This comes from the following calculations:

$\frac{\partial^2 W(e_R)}{\partial e_R^2}<0\Leftrightarrow 4\delta-2\delta^2-4<0$ which is always true.

Is this critical point in the required interval? We must prove that $0\leq e_R^*\Leftrightarrow\frac{(1-\delta)(1-\delta c_R)}{\delta^2-2\delta+2}$
$\Leftrightarrow 0\leq(1-\delta)(1-\delta c_R)\Leftrightarrow 0\leq 1-\delta c_R\Leftrightarrow\delta c_R\leq 1$ which is always true. We must also prove that $e_R^*\leq 1\Leftrightarrow\frac{(1-\delta)(1-\delta c_R)}{\delta^2-2\delta+2}\leq 1\Leftrightarrow(1-\delta)(1-\delta c_R)\leq\delta^2-2\delta+2\Leftrightarrow-\delta^2+\delta-1\leq\delta(1-\delta)c_R$
$\Leftrightarrow\frac{-\delta^2+\delta-1}{\delta(1-\delta)}\leq c_R$ which is always true bc $\delta(1-\delta)\geq 0$ but $-\delta^2+\delta-1<0$. Therefore $\frac{(1-\delta)(1-\delta c_R)}{\delta^2-2\delta+2}$ satisfies the constraints and is an interior maximum.

To prove that e_R^* is a straight line with respect to c_R we simply expand the formula above to obtain $\frac{(1-\delta)(1-\delta c_R)}{\delta^2-2\delta+2}=\frac{1-\delta}{\delta^2-2\delta+2}-\frac{\delta(1-\delta)}{\delta^2-2\delta+2}c_R$. I already proved that all the values of e_R^* are positive. To prove that its slope is negative we note that $\delta\geq 0$, $(1-\delta)\geq 0$, and the polynomial $\delta^2-2\delta+2$ does not have roots and is always positive. ■

A.3.8 Theorem 6

Proof. From Table 1, we know that $x_R^* = A_R$ if $A_R \geq 0$. So we should start by calculating the equilibrium value A_R^* remembering that $A_R = e_R + \delta\left(c_R - e_R\right)$. Now that experience is endogenous, candidate R will start the election with an optimally chosen level of experience, e_R^*. We know from Theorem 5 that $e_R^* = \frac{1-\delta}{\delta^2-2\delta+2} - \frac{\delta(1-\delta)}{\delta^2-2\delta+2}c_R = \frac{(1-\delta)(1-\delta c_R)}{\delta^2-2\delta+2}$.

Plugging this value into A_R, and with some algebra, we obtain:

$A_R^* = \frac{(1-\delta)(1-\delta c_R)}{\delta^2-2\delta+2} + \delta\left(c_R - \frac{(1-\delta)(1-\delta c_R)}{\delta^2-2\delta+2}\right) = \frac{\delta^2+\delta c_R-2\delta+1}{\delta^2-2\delta+2}$

Now we prove that A_R^* will always be positive.

$A_R^* \geq 0 \Leftrightarrow \frac{\delta^2+\delta c_R-2\delta+1}{\delta^2-2\delta+2} \geq 0$

$\Leftrightarrow \delta^2 + \delta c_R - 2\delta + 1 \geq 0$ because the polynomial $\delta^2 - 2\delta + 2$ is always strictly positive

$\Leftrightarrow \left(\delta^2\right) + \left(\delta c_R - 2\delta\right) + (1) \geq 0$ which is always true because this is a positive polynomial with no roots.

So, given that A_R^* is positive, we have that $x_R^{**} = A_R^*$.

To prove that x_R^{**} is a straight line with respect to c_R we simply expand the formula above to obtain $x_R^{**} = \frac{\delta^2+\delta c_R-2\delta+1}{\delta^2-2\delta+2} = \frac{(1-\delta)^2}{\delta^2-2\delta+2} + \frac{\delta}{\delta^2-2\delta+2}c_R$. I already proved that all the values of A_R^* are positive. To prove that its slope is positive we note that $\delta \geq 0$, and the polynomial $\delta^2 - 2\delta + 2$ does not have roots and is always positive. ■

A.3.9 Corollary 3

Proof.

1. From Theorem 5, we know that e_R^* is a downward sloping line with respect to c_R, so the optimal experience decreases with charisma.

2. From Theorem 6, we know that x_R^{**} is an upward sloping line with respect to c_R, so the extremism chosen by the candidate increases with his charisma.

■

A.3.10 Corollary 4

Proof.

1. We know from Theorem 5 that, $e_R^* = \frac{1-\delta}{\delta^2-2\delta+2} - \frac{\delta(1-\delta)}{\delta^2-2\delta+2}c_R$. Differentiating this expression with respect to δ we obtain $\frac{\partial e_R^*}{\partial \delta} = \frac{4\delta c_R-2c_R-\delta^2 c_R+\delta^2-2\delta}{\left(\delta^2-2\delta+2\right)^2}$. We need to prove that this derivative is negative. Note that the denominator is strictly positive because $\left(\delta^2-2\delta+2\right)^2$ is a positive parabola with no roots. So we need to prove that $4\delta c_R - 2c_R - \delta^2 c_R + \delta^2 - 2\delta < 0$ for the allowed intervals of δ and c_R. We have that:

   $4\delta c_R - 2c_R - \delta^2 c_R + \delta^2 - 2\delta < 0$

$\Leftrightarrow \left(-\delta^2 c_R + \delta^2\right) + (4\delta c_R - 2\delta) + (-2c_R) < 0$

$\Leftrightarrow (1 - c_R)\,\delta^2 + (4c_R - 2)\,\delta + (-2c_R) < 0$

As long as $c_R < 1$, this is an upward parabola. It will be negative only between its roots, which are:

Low root: $\frac{-(4c_R-2)-2\sqrt{2c_R^2-2c_R+1}}{2(1-c_R)} = \frac{2c_R+\sqrt{-2c_R+2c_R^2+1}-1}{c_R-1} = \frac{1-2c_R-\sqrt{-2c_R+2c_R^2+1}}{1-c_R}$

High root: $\frac{-(4c_R-2)+2\sqrt{2c_R^2-2c_R+1}}{2(1-c_R)} = \frac{2c_R-\sqrt{-2c_R+2c_R^2+1}-1}{c_R-1} = \frac{1-2c_R+\sqrt{-2c_R+2c_R^2+1}}{1-c_R}$

We can prove that δ always falls within these roots, given that $\delta \in (0,1)$. For this, we first prove that the low root of the previous parabola is smaller than zero.

$\frac{1-2c_R-\sqrt{-2c_R+2c_R^2+1}}{1-c_R} < 0$

$\Leftrightarrow 1 - 2c_R - \sqrt{-2c_R + 2c_R^2 + 1} < 0$ because $1 - c_R > 0$

$\Leftrightarrow (1 - 2c_R)^2 < \left(\sqrt{-2c_R + 2c_R^2 + 1}\right)^2 \Leftrightarrow 4c_R^2 - 4c_R + 1 < 2c_R^2 - 2c_R + 1$

$\Leftrightarrow 0 < 2c_R - 2c_R^2 \Leftrightarrow 0 < 2c_R\,(1 - c_R)$ as long as $c_R > 0$ and $c_R < 1$

$\Leftrightarrow 0 < 2$ which is always true.

Then we prove that the high root of the previous parabola is larger than one.

$\frac{1-2c_R+\sqrt{-2c_R+2c_R^2+1}}{1-c_R} > 1 \Leftrightarrow 1-2c_R+\sqrt{-2c_R + 2c_R^2 + 1} > 1-c_R \Leftrightarrow \sqrt{-2c_R + 2c_R^2 + 1} > 1 - c_R - (1 - 2c_R)$

$\Leftrightarrow \sqrt{-2c_R + 2c_R^2 + 1} > c_R \Leftrightarrow -2c_R + 2c_R^2 + 1 > c_R^2 \Leftrightarrow -c_R^2 - 2c_R + 2c_R^2 + 1 > 0 \Leftrightarrow -2c_R + c_R^2 + 1 > 0$

$\Leftrightarrow (1 - c_R)^2 > 0$ which is always true for $c_R \in (0,1)$.

In sum, for all values of $\delta \in (0,1)$ we have that $\frac{\partial e_R^*}{\partial \delta} < 0$. Hence, as δ increases from zero to one, the optimal experience e_R^* decreases monotonically.

2. We know from Theorem 6 that, $x_R^{**} = \frac{(1-\delta)^2}{\delta^2-2\delta+2} + \frac{\delta}{\delta^2-2\delta+2}c_R$. Differentiating this expression with respect to δ we obtain $\frac{\partial x_R^{**}}{\partial \delta} = \frac{2c_R+2\delta-\delta^2 c_R-2}{\left(\delta^2-2\delta+2\right)^2}$. We need to determine when this derivative is positive. Note that the denominator is strictly positive because $\left(\delta^2 - 2\delta + 2\right)^2$ is a positive parabola with no roots. So we need to determine when the numerator is positive. We have that:

$2c_R+2\delta-\delta^2 c_R-2 > 0 \Leftrightarrow \left(-\delta^2 c_R\right)+(2\delta)+(2c_R - 2) > 0 \Leftrightarrow (-c_R)\,\delta^2+(2)\,\delta+(2c_R - 2) > 0$. As long as $c_R > 0$, this is an downward parabola. It will be positive only inside its roots, which are:

Low root: $\frac{-(2)+2\sqrt{2c_R^2-2c_R+1}}{2(-c_R)} = -\frac{\sqrt{-2c_R+2c_R^2+1}-1}{c_R} = \frac{1-\sqrt{-2c_R+2c_R^2+1}}{c_R}$

High root: $\frac{-(2)-2\sqrt{2c_R^2-2c_R+1}}{2(-c_R)} = \frac{\sqrt{-2c_R+2c_R^2+1}+1}{c_R} = \frac{1+\sqrt{-2c_R+2c_R^2+1}}{c_R}$

We now prove that the low root of the previous parabola is larger than zero.

$\frac{1-\sqrt{-2c_R+2c_R^2+1}}{c_R} > 0$

$\Leftrightarrow 1 - \sqrt{-2c_R + 2c_R^2 + 1} > 0$ as long as $c_R > 0$

$\Leftrightarrow 1 > \sqrt{-2c_R + 2c_R^2 + 1} \Leftrightarrow 1 > -2c_R + 2c_R^2 + 1 \Leftrightarrow 0 > 2c_R\,(c_R - 1) \Leftrightarrow 0 > (c_R - 1)$

$\Leftrightarrow 1 > c_R$ which is true by assumption.

Next we prove that the low root smaller than one.

$\frac{1-\sqrt{-2c_R+2c_R^2+1}}{c_R} < 1 \Leftrightarrow 1 - \sqrt{-2c_R + 2c_R^2 + 1} < c_R$ as long as $c_R > 0$

$\Leftrightarrow 1 - c_R < \sqrt{-2c_R + 2c_R^2 + 1} \Leftrightarrow c_R^2 - 2c_R + 1 < 2c_R^2 - 2c_R + 1 \Leftrightarrow 0 < c_R^2$ which is always true.

Now we prove that the high root larger than one.

$\frac{1+\sqrt{-2c_R+2c_R^2+1}}{c_R} > 1$

$\Leftrightarrow 1 + \sqrt{-2c_R + 2c_R^2 + 1} > c_R$ as long as $c_R > 0$

$\Leftrightarrow \sqrt{-2c_R + 2c_R^2 + 1} > c_R - 1$ which is always true bc the left-hand side is positive while the right-hand side is negative.

This proves that $\frac{\partial x_R^{**}}{\partial \delta} < 0$ if δ is between zero and the low root calculated above; and $\frac{\partial x_R^{**}}{\partial \delta} > 0$ if δ is between the low root calculated above and one. Hence, as δ increases from zero to one, we have that x_R^{**} decreases for $\delta \in \left(0, \frac{1-\sqrt{-2c_R+2c_R^2+1}}{c_R}\right)$, and increases for $\delta \in \left(\frac{1-\sqrt{-2c_R+2c_R^2+1}}{c_R}, 1\right)$.

■

A.3.11 Theorem 7

Proof. We need to reassess all the configurations of x_R and x_L to show they are Nash equilibria if and only if they were Nash equilibria in the proof of Theorem 1. There I studied all the possible configurations of x_R and x_L for each possible value of A_R. In each case, I analyzed the possible unilateral deviations by each player, R and L, to determine whether they were profitable and hence discarded such configurations. Only the configurations without profitable unilateral deviations are Nash equilibria. All the profitable unilateral deviations were of one the following kinds: (1) The candidate prefers to move his platform in order to win rather than lose the election. (2) The candidate prefers to move his platform in order to win with a more, rather than less, favorable platform. Both types of deviations

remain profitable after we add an ego rent G to the candidate's utility functions. So all the configurations that were discarded for Theorem 1 are also discarded in this specification.

Moreover, all the configurations that survived as equilibria from Theorem 1 also survive as equilibria in this specification. That is because any deviation from such configurations involves either losing the election, or winning with a less favorable platform, or both. In those configurations, there does not exist a deviation that will improve policy while worsening ego rents, or improve ego rents while worsening policy. Therefore they remain equilibria in this scenario. ■

A.3.12 Theorem 8

Proof. The non-policy advantage of the right-wing candidate, A_R, can be calculated to be the same as before, namely $A_R \equiv e_R + \delta c_R - e_L - \delta c_L = e_R + \delta c_R - 1 - 0 = e_R + \delta c_R - 1$. Given that this value is always such that $A_R \in [-1, 1]$, we have that $x^* = A_R$.

An important difference in this section is the utility derived by candidate R after the election, which is now given by $U_R(x^*) = -(1 - x^*)^2 + G = -(1 - A_R)^2 + G = -(1 - e_R - \delta c_R + 1)^2 + G = -(2 - e_R - \delta c_R)^2 + G$.

The cost of acquiring e_R is still e_R^2. So objective function is $W(e_R) = -(2 - e_R - \delta c_R)^2 + G - e_R^2$.

In sum, candidate R wants to $\max_{e_R} W(e_R) = -(2 - e_R - \delta c_R)^2 + G - e_R^2$. Expanding this expression we get $W(e_R) = (-2) e_R^2 + (-2y\delta + 4) e_R + \left(-(c_R\delta - 2)^2 + G\right)$. We need to divide this problem in cases, depending on whether R will win the election, tie, or lose the election, which correspond to the cases $A_R > 0$; $A_R = 0$; and $A_R < 0$. These in turn correspond to the cases $e_R > 1 - \delta c_R$; $e_R = 1 - \delta c_R$; and $e_R < 1 - \delta c_R$.

- Case R wins for sure ($A_R > 0$), which implies $G = g$.

In this case $W(e_R) = (-2) e_R^2 + (-2c_R\delta + 4) e_R + \left(-(c_R\delta - 2)^2 + g\right)$. To solve this maximization problem we can differentiate the expression, which gives us $W'(e_R) = -4x - 2\delta c_R + 4$. The first order conditions require that $-4e_R - 2\delta c_R + 4 = 0$ which occurs at $e_R = 1 - \frac{1}{2} c_R\delta$. Is this solution within the interval $e_R > 1 - \delta c_R$? We have that $1 - \frac{1}{2} c_R\delta > 1 - \delta c_R \Leftrightarrow -\frac{1}{2} c_R\delta > -\delta c_R \Leftrightarrow \frac{1}{2} < 1$ which is always true. The second order conditions tell us that $W''(e_R) = -4$, which is negative, so this critical point is a maximum. At this value we have $W(e_R) = (-2)\left(1 - \frac{1}{2} c_R\delta\right)^2 + (-2c_R\delta + 4)\left(1 - \frac{1}{2} c_R\delta\right) + \left(-(c_R\delta - 2)^2 + g\right)$.

- Case R ties with L ($A_R = 0$), which implies $G = \frac{g}{2}$.

In this case $W(e_R) = (-2) e_R^2 + (-2c_R\delta + 4) e_R + \left(-(c_R\delta - 2)^2 + \frac{g}{2}\right)$ with $e_R = 1 - \delta c_R$. Therefore

$W(e_R) = (-2)(1 - \delta c_R)^2 + (-2c_R\delta + 4)(1 - \delta c_R) + \left(-(c_R\delta - 2)^2 + \frac{g}{2}\right)$

- Case R loses for sure ($A_R < 0$), which implies $G = 0$.

In this case $W(e_R) = (-2)\, e_R^2 + (-2c_R\delta + 4)\, e_R - (c_R\delta - 2)^2$. To solve this maximization problem we can differentiate the expression, which gives us $W'(e_R) = -4x - 2\delta c_R + 4$. The first order conditions require that $-4e_R - 2\delta c_R + 4 = 0$ which is impossible. In fact, in this interval we always have that $W'(e_R) > 0$. So R prefers higher values of e_R. Given that this interval is open, a maximum does not exist, but the maximal value tends to $\lim_{e_R \to 1-\delta c_R} (-2)\, e_R^2 + (-2\delta c_R + 4)\, e_R - (\delta c_R - 2)^2$, which is $(-2)\,(1 - \delta c_R)^2 + (-2\delta c_R + 4)\,(1 - \delta c_R) - (\delta c_R - 2)^2$.

With some algebra, it can be proved that the $W(e_R)$ is highest in the first interval, then the second one and then the third one. Hence the global maximum is the one calculated for the first interval, namely, $e_R^* = 1 - \frac{\delta}{2} c_R$. It should be noted that this is the same result as in Theorem 2. ■

A.3.13 Theorem 9

Proof. The previous result allows calculating the valence advantage to R in equilibrium, A_R^*, which is:

$A_R^* = e_R^* + \delta c_R - 1 = 1 - \frac{1}{2}\delta c_R + \delta c_R - 1 = \frac{1}{2}\delta c_R$

Since $A_R^* \geq 0$, we have from Table 1 that $x_R^{**} = A_R^* = \frac{1}{2}\delta c_R$. ■

A.3.14 Theorem 10

Proof. For Table 1 to apply in this new setup, we need to prove that indeed we have $-1 \leq A_R \leq 1$. But this happens exactly for the same reasons as in the original model, because A_R is calculated exactly the same way as $A_R \equiv e_R + \delta c_R - e_L - \delta c_L = e_R + \delta c_R - 1$. The proof of Theorem 1 relied only on candidates have single-peaked preferences around their ideal points; it did no rely at any moment on those preferences being quadratic. So the proof remains valid in this setup. ■

A.3.15 Theorem 11

Proof. We start by recalling that in equilibrium we have $x^* = A_R$, which comes from simply looking at the third column of Table 1. We also recall that $A_R = e_R + \delta c_R - 1$. This allows us to calculate the utility that R derives from policy in equilibrium, which is $U_R(x^*) = -|1 - (e_R + \delta c_R - 1)| = e_R + \delta c_R - 2$. By subtracting the cost of acquiring e_R from this expression, we obtain $W(e_R)$, which is $W(e_R) = -|1 - (e_R + \delta c_R - 1)| - e_R^2 = e_R + \delta c_R - 2 - e_R^2$.

Given this result, what is the optimal e_R that R could adopt? R needs to maximize the function $W(e_R)$. The optimum is found by solving the problem

$$\max_{e_R} W(e_R) \text{ subject to } 0 \leq e_R \leq 1$$

The first-order conditions give us the critical point $e_R = \frac{1}{2}$. This comes from the following calculations:

$\frac{\partial W(e_R)}{\partial e_R} = 0 \Leftrightarrow -2e_R + 1 = 0 \Leftrightarrow 1 = 2e_R \Leftrightarrow e_R = \frac{1}{2}$. Note that this value falls within the required interval $e_R \in [0, 1]$.

The second order conditions show this critical point is a maximum. This comes from the following calculations:

$\frac{\partial^2 W(e_R)}{\partial e_R^2} < 0 \Leftrightarrow -2 < 0$ which is always true. ■

A.3.16 Theorem 12

Proof. From Table 1, we know that $x_R^{**} = \begin{cases} 0 \text{ if } A_R^* < 0 \\ A_R^* \text{ if } A_R^* \geq 0 \end{cases}$. So we should start by calculating the equilibrium value A_R^* remembering that $A_R = e_R + \delta c_R - 1$. From Theorem 11 we know that that $e_R^* = \frac{1}{2}$. Plugging this value into A_R, and with some algebra, we obtain:

$A_R^* = e_R + \delta c_R - 1 = \frac{1}{2} + \delta c_R - 1 = \delta c_R - \frac{1}{2}$.

We now determine when A_R^* is positive:

$A_R^* \geq 0 \Leftrightarrow \delta c_R - \frac{1}{2} \geq 0 \Leftrightarrow c_R \geq \frac{1}{2\delta}$.

What this is implies for x_R^{**} is that:

$x_R^{**} = \begin{cases} 0 \text{ if } \delta c_R - \frac{1}{2} < 0 \\ \delta c_R - \frac{1}{2} \text{ if } \delta c_R - \frac{1}{2} \geq 0 \end{cases} \Leftrightarrow x_R^{**} = \begin{cases} 0 \text{ if } c_R < \frac{1}{2\delta} \\ \delta c_R - \frac{1}{2} \text{ if } c_R \geq \frac{1}{2\delta} \end{cases}$ which is what the theorem claims. ■