

The electoral strategies of a populist candidate: Does charisma discourage experience and encourage extremism?

Journal of Theoretical Politics
2018, Vol. 30(1) 45–73
©The Author(s) 2017
Reprints and permissions:
sagepub.com/uk/journalsPermissions.nav

DOI: 10.1177/0951629817737860
journals.sagepub.com/home/jtp



Gilles Serra

Department of Political Science, Centre for Economics Research and Teaching (CIDE), Mexico

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 who 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 an election with an extremist agenda; this helps explain the recent ‘rise of populism’ identified by several authors around the world. This theory is also able to explain 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

Charisma; democracy; elections; experience; populism

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, these features

Corresponding author:

Gilles Serra, Department of Political Science, Centre for Economics Research and Teaching (CIDE), Mexico City, Mexico.

Email: gilles.serra@cide.edu

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 candidate 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 throughout the paper. 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 with 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 the 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 in articulating an anti-elite rhetoric that will resonate with voters for electoral support; to be concrete, he has a *valence* parameter, owing to his personal qualities. Stokes (1963) coined the expression ‘valence’ to refer to issues that all voters agree to value positively, in contrast with ‘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. Weber (1978) saw charisma as a very rare power, endowing its holder with the capacity to elicit passionate popular support. 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 into 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 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 a 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 policies 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 context of the formal literature, especially previous models of valence, leadership, and populism. Section 3 will model an election with a policy dimension and a valence dimension 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 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 online appendix to this paper provides two more extensions that also serve a robustness checks, and 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 work has sought to address a fundamental debate about whether valence leads to extremism or moderation. Within such a debate, some models, such as that of Groseclose (2001), find that a valence-advantaged candidate will be more moderate than a valence-disadvantaged one, while other models, such as that of Serra (2010), find the opposite result. Like most of that 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 that of Adams et al. (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 the one hand, they distinguish characteristics that are merely useful in winning elections, such as name recognition, fund-raising ability, and campaigning skills, which they call *strategic valence*. On the other hand, they distinguish characteristics that are actually valuable for elected officials in governing, 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 canon by giving a novel interpretation. I will interpret the endogenous valence as the experience in government that a candidate may acquire throughout his career. In this way, I aim to shed light on the career decisions of a candidate over a longer timespan than other formal models of endogenous valence.⁵

By proposing the concept of *populist charisma*, my model focuses on citizens who have leadership skills. As such, it is related to a budding formal literature on leadership. Dewan and Myatt (2007, 2008, 2012) have modeled the conditions for ‘effective leadership’ to arise. I believe their views are largely compatible with the Weberian approach to charisma that I espouse in this paper. On the one hand, a number of special qualities are needed from the individual seeking to become a leader. These two authors 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 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, 2008, 2012).

Other formalizations of *populism* also relate to my model, such as those of Acemoglu et al. (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 the case in Latin America. My paper is also motivated by this observation. However, my explanation differs from the kind of explanation offered by Acemoglu et al. (2013) and Fox and Stephenson (2015). 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

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. On 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. The 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 in which he wishes to participate.

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 that 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 that 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

$$U_M(x, e, c, \delta) = -|x| + e + \delta c \quad (1)$$

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 left-wing preferences and the other one has genuinely right-wing preferences. It should be noted that other formal models frequently assume that candidates only care about winning elections regardless of their policy platforms, meaning that 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 . This 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 with 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

$$U_R(x) = -(1 - x)^2 \quad (2)$$

$$U_L(x) = -(-1 - x)^2 \quad (3)$$

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 is an insider, meaning that the former does not entirely belong to the established political elite, while the 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 $c_R \in [0, 1]$. Conversely, given that L is a mainstream candidate, we assume that she cannot credibly run an anti-establishment 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 $e_R \in [0, 1]$. Conversely, 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 that 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 Section 5, 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 with 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 *disadvantage* for R compared with 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 as follows:

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).

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, which requires that strategies form a Nash equilibrium 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

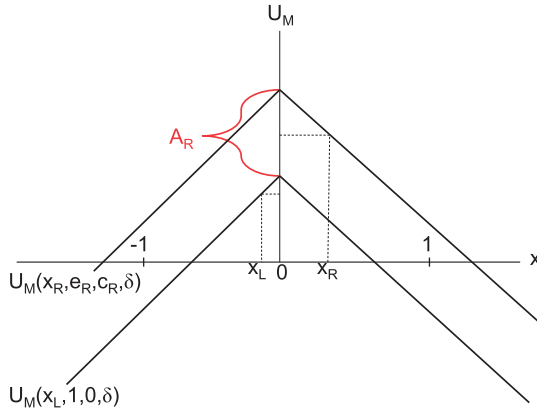


Figure 1. Effect of an advantage for R over L in charisma and experience, A_R .

not only indifferent but also have the same valence, M will randomize equally between the two.¹⁰

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 over 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 for his extremism in the policy dimension. As proved in the next section, the situation depicted 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. 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, imposes that R and L must play a Nash equilibrium in every subgame. We call this equilibrium x_R^* and x_L^* and the winning platform x^* . 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 $A_R = e_R + \delta c_R - 1$, the values of which 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 |

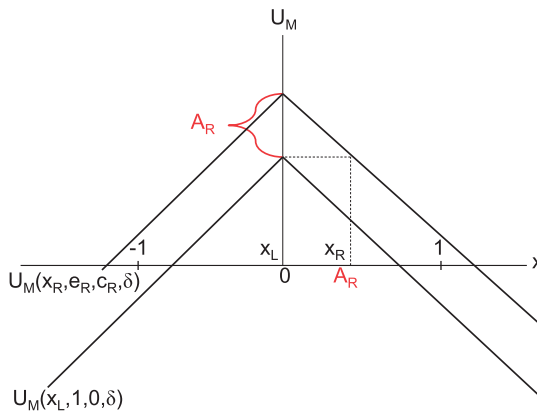


Figure 2. Equilibrium platforms x_L^* and x_R^* when there is a valence advantage for R over L .

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 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.

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 on winning the election. Note that such extremism of R compared with 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

These results could be used to take a step back in time. Now we can analyze the career decisions made 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 following results will show, a larger value of e_R would allow the outsider candidate R to pull the winning 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 next.

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 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 this timing.

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 thus be 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 ; this 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 other 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 us in turn to calculate the outcome from the high-level election once the outsider candidate is ready to participate. On 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 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.

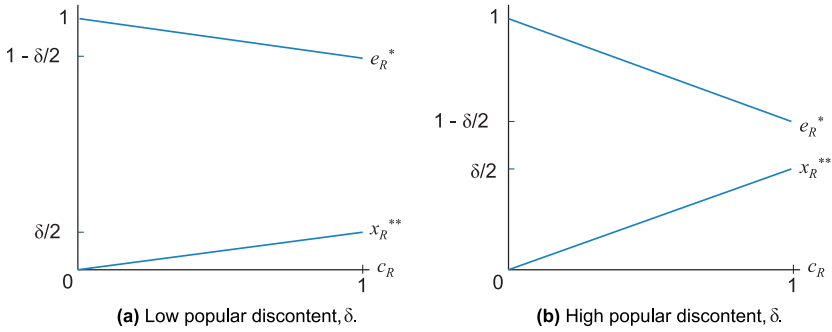


Figure 3. Charisma discourages experience and encourages extremism.

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. Theorem 3 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.

Both effects can be seen in Figure 3. Candidates who are more charismatic, i.e. have 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 being 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). This original assumption was convenient in isolating 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 becomes 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 the 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 that 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 study that correlates the two. So it is worth having a theoretical exploration of this issue. In this section, I repeat 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.

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 as

$$U_M(x, e, c, \delta) = -|x| + e + \delta(c - e) \quad (4)$$

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 made by all players are more complex, we still find a unique equilibrium for each set of parameter values.¹⁵

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 us to study 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, the candidate is now wary that acquiring too much experience in government might undermine the anti-elite charisma he initially enjoyed. This will be reflected in a smaller 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 us to calculate the extremism that can be expected from the populist candidate as a function of the parameters. The main difference from the original model will be the effect of popular discontent, which will now be ambiguous. In this section, a higher value of δ will have two countervailing effects on the valence of the populist candidate (and hence on the extremism of his platform): on the 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 being equal, higher charisma allows the outsider to adopt a more extremist platform and still win the election.

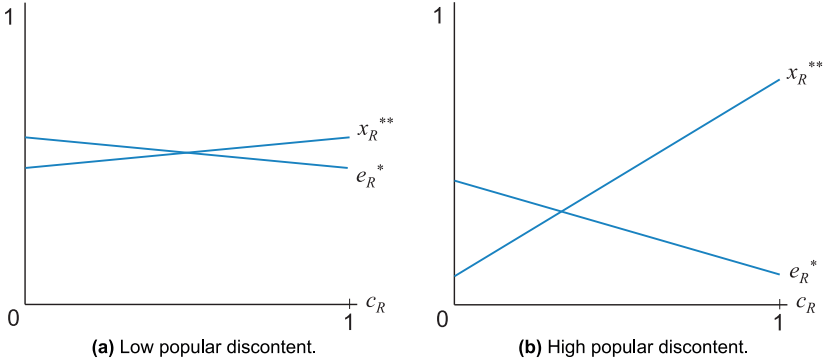


Figure 4. Charisma discourages experience and encourages extremism even when charisma is incompatible with experience.

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 are illustrated in Figure 4. 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 being 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).

The following two corollaries serve to summarize the total effects of each of the main primitive variables in this section.

Corollary 3. *All things being 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 being 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 model, compared with 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 with 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 in organizing and explaining a number of seemingly disconnected empirical patterns within a single framework. The empirical regularities that find echoes 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 interested in competing for high office. I focused on those particular 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. He claimed, for example, to embody the struggle of the neglected common man against the corrupt political elite, insisting that the election was about 'who runs this country, the special interests or the people' (Eatwell, 2017b). Inglehart and Norris (2016 pp. 5–7) 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

an 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 by 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 offices or a relevant cabinet position. There is actually a large literature from US 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 being 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 dataset of 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 body of 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 USA 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 et al. (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 to follow 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.

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' (Edelman, 2016). 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. This 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 the Front National's success (Pedahzur and Brichta, 2002, p. 41). 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, p. 1233):

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.

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 that study 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 lack of office experience was not due to a lack of opportunities. In October 2013, Republican assemblymen from New York circulated a memo suggesting enthusiastically that Trump should run for governor of the state (Spector, 2013). 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 with 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' (Hafner, 2016). 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. Immediately upon being released from jail, Chávez entered the first presidential election that was held. Weyland (2003, p. 825) 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 socio-economic program. It seems that many citizens simply felt compelled to believe in Chávez's charisma, independent of realistic assessments of his likely success.

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. In 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 when he had been on the political stage for less than a year. Conversely, 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 to suggest that 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 with 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 (2016) 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 for 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 Le Pen 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 with 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 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 capture 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 chief executive during times of economic downturn, political crisis, or social unrest.

Academics believe that Donald Trump benefited 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 the 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, owing 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, Acción Democrática (AD) and Comité de Organización Política Electoral Independiente (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 (Weyland, 2003, p. 833) 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.

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 reduction in gross domestic product 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 been done to study 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, which shows that all effects in my model are steeper for a higher value of δ .

7. 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 that was intended to capture the personal characteristics of the outsider candidate, which 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 the 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. Moreover 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 on the other hand, my model suggested two costs that have not been frequently mentioned in the existing literature, and had not been derived together in a unified theory. A first cost lies in 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 comes from 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. On getting elected, he will then carry out his program of extreme (left or right) policies, which 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 they consider 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.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Notes

1. Eatwell (2017a) summarized this concept in 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.’
2. According to Inglehart and Norris (2016, pp. 6–7), ‘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 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.’
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, 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 in campaigning. Other formal models studying valence and ideology include those of Aragonès and Xefteris (2013); Calvo and Hellwig (2011); Crisp et al. (2014); and Hitt et al. (2017).
5. Other interesting interpretations of endogenous valence can be found in the literature. Schofield et al. (2011) study endogenous spending by candidates on valence, which is interpreted as campaign advertising. Meiorowitz (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 posited by 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).

6. In any event, I have studied this same model for the case 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. This is because both motivations provide incentives to candidates in the same direction. Hence, such an 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 for the case when candidates have linear utility functions, as for 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.
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.
9. A more general discussion of the model presented 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.
11. The proofs of all the results in this paper are given in the online appendix at <http://www.investigadores.cide.edu/gilles.serra/>.
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.
14. In addition to this cost of effort, in Section 5, I study another cost of acquiring government experience, corresponding to the candidate's lost reputation as an anti-system outsider.
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 the effect of R 's charisma to be shown 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.
16. Early contributions were those of Jacobson and Kernell (1983) and Stone et al. (2004).
17. Abramson et al. (1987); Schlesinger (1966).

References

- Abramson PR, Aldrich JH and Rohde DW (1987) Progressive ambition among United States senators: 1972–1988. *The Journal of Politics* 49(1): 3–35.
- Acemoglu D, Egorov G and Sonin K (2013) A political theory of populism. *The Quarterly Journal of Economics* 128(2): 771–805.
- Adams J, Merrill S III, Simas EN, et al. (2011) When candidates value good character: A spatial model with applications to congressional elections. *The Journal of Politics* 73(1): 17–30.

- Aragonès E and Xefteris D (2017) Imperfectly informed voters and strategic extremism. *International Economic Review* 58(2): 439–471.
- Ashworth S and Bueno De Mesquita E (2009) Elections with platform and valence competition. *Games and Economic Behavior* 67(1): 191–216.
- Callander S (2008) Political motivations. *Review of Economics Studies* 75: 671–697.
- Calvert RL (1985) Robustness of the multidimensional voting model: Candidate motivations, uncertainty, and convergence. *American Journal of Political Science* 29: 69–95.
- Calvo E and Hellwig T (2011) Centripetal and centrifugal incentives under different electoral systems. *American Journal of Political Science* 55(1): 27–41.
- Carreras M (2016) Institutions, governmental performance and the rise of political newcomers. *European Journal of Political Research* 56(2): 364–380.
- Carrillo J and Castanheira M (2008) Information and strategic political polarization. *Economic Journal* 118: 845–874.
- Carter J and Patty JW (2015) Valence and campaigns. *American Journal of Political Science* 59(4): 825–840.
- Chasquetti D and Micozzi JP (2014) The subnational connection in unitary regimes: Progressive ambition and legislative behavior in Uruguay. *Legislative Studies Quarterly* 39(1): 87–112.
- Crisp BF, Patty JW, Penn EM, et al. (2014) Valence, partisan competition, & legislative institutions. In: *72nd annual conference of the Midwest Political Science Association*, Chicago, IL, 3–6 April 2014. Bloomington, IN: MPSA.
- De Lange SL and Art D (2011) Fortuyn versus Wilders: An agency-based approach to radical right party building. *West European Politics* 34(6): 1229–1249.
- Dewan T and Myatt DP (2007) Leading the party: Coordination, direction, and communication. *American Political Science Review* 101(04): 827–845.
- Dewan T and Myatt DP (2008) The qualities of leadership: Direction, communication, and obfuscation. *American Political Science Review* 102(03): 351–368.
- Dewan T and Myatt DP (2012) On the rhetorical strategies of leaders: Speaking clearly, standing back, and stepping down. *Journal of Theoretical Politics* 24(4): 431–460.
- Eatwell R (2017a) Charisma and the Radical Right. In: Rydgren J (ed.) *The Oxford Handbook of the Radical Right*. Oxford, UK: Oxford University Press (in press).
- Eatwell R (2017b) Populism and fascism. In: Rovira Kaltwasser C, Taggart P, Ochoa Espejo P, et al. (eds) *The Oxford Handbook of Populism*. Oxford, UK: Oxford University Press (in press).
- Edelman A (2016) Billionaire Donald Trump: I am a blue collar worker. *New York Daily News*, 10 October, 2016.
- Fox J and Stephenson MC (2015) The welfare effects of minority-protective judicial review. *Journal of Theoretical Politics* 27(4): 499–521.
- Goodwin M and Milazzo C (2015) *UKIP: Inside the Campaign to Redraw the Map of British Politics*. Oxford, UK: Oxford University Press.
- Groseclose T (2001) A model of candidate location when one candidate has a valence advantage. *American Journal of Political Science* 45(4): 862–886.
- Hafner J (2016) ‘I prepared to be president,’ Clinton tells Trump in debate zinger. *USA Today*, 27 September, 2016.
- Hitt MP, Volden C and Wiseman AE (2017) Spatial models of legislative effectiveness. *American Journal of Political Science* 61(3): 575–590.
- Hobolt SB and Høyland B (2011) Selection and sanctioning in European Parliamentary elections. *British Journal of Political Science* 41(03): 477–498.
- Inglehart R and Norris P (2016) Trump, Brexit, and the rise of Populism: Economic have-nots and cultural backlash. Harvard Kennedy School, Cambridge, MA, USA, Faculty Research Working Paper no. RWP16-026.
- Jacobson GC and Kernell S (1983) *Strategy and Choice in Congressional Elections*. New Haven, CT: Yale University Press.

- Jones MP, Saiegh S, Spiller PT, et al. (2002) Amateur legislators—professional politicians: The consequences of party-centered electoral rules in a federal system. *American Journal of Political Science* 46(3): 656–669.
- Levitsky S and Loxton J (2012) Populism and competitive authoritarianism: The case of Fujimori's Peru. In: Mudde C and Rovira Kaltwasser C (eds.) *Populism in Europe and the Americas: Threat or Corrective for Democracy?* Cambridge, UK: Cambridge University Press, pp. 160–181.
- Meirowitz A (2008) Electoral contests, incumbency advantages, and campaign finance. *Journal of Politics* 70(03): 681–699.
- Micozzi JP and Saiegh SM (2016) An empirical stochastic model of Argentina's Impossible Game (1955–1966). *Journal of Theoretical Politics* 28(2): 266–287.
- Mudde C (2007) *Populist Radical Right Parties in Europe*. Cambridge, UK: Cambridge University Press.
- Pedahzur A and Brichta A (2002) The institutionalization of extreme right-wing charismatic parties: A paradox? *Party Politics* 8(1): 31–49.
- Penn EM (2009) From many, one: State representation and the construction of an American identity. *Journal of Theoretical Politics* 21(3): 343–364.
- Roberts K (2012) Populism and democracy in Venezuela under Hugo Chávez. In: Mudde C and Rovira Kaltwasser C (eds.) *Populism in Europe and the Americas: Threat or Corrective for Democracy?* Cambridge, UK: Cambridge University Press, pp. 136–159.
- Samuels D (2003) *Ambition, Federalism, and Legislative Politics in Brazil*. New York, NY: Cambridge University Press.
- Schlesinger J (1966) *Ambition and Politics: Political Careers in the United States*. Chicago, IL: Rand McNally.
- Schofield N, Claassen C, Ozdemir U, et al. (2011) Estimating the effects of activists in two-party and multi-party systems: Comparing the United States and Israel. *Social Choice and Welfare* 36(3–4): 483–518.
- Serra G (2010) Polarization of what? A model of elections with endogenous valence. *The Journal of Politics* 72(2): 426–437.
- Spector J (2013) N.Y. Republicans want Donald Trump to run for governor. *USA Today*, 14 October, 2013.
- Stokes D (1963) Spatial models of candidate competition. *American Political Science Review* 57: 368–377.
- Stone WJ, Maisel LS and Maestas CD (2004) Quality counts: Extending the strategic politician model of incumbent deterrence. *American Journal of Political Science* 48(3): 479–495.
- Van der Brug W and Mughan A (2007) Charisma, leader effects and support for right-wing populist parties. *Party Politics* 13(1): 29–51.
- Weber M (1978) Charisma and its transformation. In: Roth G and Wittich C (eds.) *Economy and Society: An Outline of Interpretive Sociology*, vol. 2. Berkeley, CA: University of California Press, pp. 1111–1157.
- Weyland K (2003) Economic voting reconsidered: Crisis and charisma in the election of Hugo Chávez. *Comparative Political Studies* 36(7): 822–848.