Reference Dependence and Politicians' Credibility*

Edoardo Grillo†

This draft: March 2014.

Abstract

We consider a model of electoral competition in which two politicians compete to get elected. Each politician is characterized by a valence, which is unobservable to voters and can take one of two values: high or low. The electorate prefers politicians with high valence, but a random shock to candidates' popularity may lead it to elect low-valence ones. Candidates make statements concerning their valence. We show that if voters are standard expected utility maximizers, politicians' statements lack any credibility and no information transmission takes place. By introducing *reference-dependent preferences and loss aversion* a là Kőszegi and Rabin, we show that full revelation is possible. Indeed, if the electorate believes to candidates' announcements, such announcements will affect its reference point. As a result, if voters find out that a candidate lied pretending to be high valence when she is not, they may decide to elect the opponent in order to avoid the loss associated with electing a candidate worse than expected.

1 Introduction

"We must not promise what we ought not, lest we be called on to perform what we cannot."
Abraham Lincoln

Electoral announcements and candidates' promises concerning their ability to perform if elected are key aspects of electoral competitions: they polarize voters's attention, attract media's scrutiny and lead to heated debates about their feasibility and truthfulness.

Interestingly, conventional wisdom makes two, partially contradictory, statements about these announcements: on the one hand, it is claimed that they have no informational content and should

*A previous version of the paper circulated under the title: "Reference Dependence and Electoral Competition: Can Politicians'Announcements be Credible?". I am indebted to Stephen Morris, for his advice and encourgement throughout this project. I am also grateful to Roland Bénabou, Ennio Bilancini, Adam Meirowitz, Wolfgang Pesendorfer and Kristopher Ramsay for helpful comments and suggestions. I am also thankful to the participants to the GRASS VII annual conference, EEA-ESEM conference in Gothneburg, the IMT economic seminar in Lucca, Princeton Microeconomic Theory Lunch Seminar and the Princeton Political Economy Lunch Workshop at Princeton University. Obviously, I take full responsibility for all remaining errors. The author is grateful to Unicredit & Universities Knight of Labor Ugo Foscolo Foundation for its generous financial support through the Europe Foscolo research grant.

†Collegio Carlo Alberto, Via Real Collegio 30, Moncalieri (Turin), Italy, email: edoardo.grillo@carloalberto.org.

be ignored as politicians are ready to promise everything in order to be elected;[1] on the other hand, it is often suggested that excessive electoral promises may turn against the politician if the electorate realizes that the candidate is unable to deliver what she promised; the opening quote by Abraham Lincoln supports this latter view.

In this paper, we rationalize this second view by building a model of electoral competition in which the electorate is subject to *reference dependence and loss aversion a là Kőszegi and Rabin.* In our setting, two candidates (A and B) compete for a public office; each candidate is characterized by one of two possible valences, high (θ_H) or low (θ_L). Although valences are private information of candidates, politicians can make a public announcement about them.

Electoral competition is modelled in a probabilistic voting model; thus, although the electorate prefers high-valence candidates to low-valence ones, a random shock to candidates' popularity may lead a low-valence candidate to win elections.

If voters are standard expected utility maximizer, the only equilibrium entails no information transmission. Indeed, since both candidates are ready to claim to be high-valence in order to maximize their probability of winning, they lose any credibility and, in equilibrium, voters ignore their announcements.

We depart from the literature assuming that voters have reference-dependent preferences a là Kőszegi and Rabin, namely they evaluate outcomes with respect to a reference point determined through a rational expectation approach. Thus, whenever the utility they experience exceeds (respectively, falls short of) the reference utility determined according to equilibrium analysis, they incur a gain (respectively, a loss). We further assume that voters are loss averse, namely they dislike losses more than what they like equal-size gains.

Under these assumptions, we prove that a fully revealing equilibrium is possible and we fully characterize it. This enables us to highlight the forces behind its existence and to identify the circumstances under which truthtelling is more likely to arise. The mechanism behind this result can be described as follows. Candidates' announcements, if credible, modify voters' reference points concerning the utility the candidate will be able to deliver, if elected. Then, if the electorate were to find out that candidate j pretended to be high valence when she is not, it may decide to support candidate $i \neq j$ in order to avoid the loss associated with appointing a candidate worse than anticipated. Remarkably, this mechanism leads to the election of i even for realizations of the popularity shock that would have determined the victory of j, were she had been sincere from the beginning. Thus, the interaction of electoral announcements with reference dependence and loss aversion introduces a potential cost from lying and this may push candidates to fully reveal their valence in the first place.

Intuitively, if the electorate exhibits reference-dependence and loss aversion, a low-valence candidate who is deciding whether to announce her valence truthfully or to lie, faces a trade-off. If she lies and her lie goes undetected, her probability of winning increases with respect to the truthtelling strategy. If instead the lie is detected, such probability may decrease as voters may decide to sup-

[1] This view about politicians can be summarized with the following quote, attributed to Napoleon: "If you wish to be a success in the world, promise everything, deliver nothing."

port the other candidate in the attempt of avoiding losses. In particular, we can identify a range of realization of the popularity shock for which for which the electorate is willing to support a low-valence candidate against a high-valence one only if the former told the truth from the beginning. We refer to this as to the *switching range*.

Our main result is to prove that a fully revealing equilibrium exists if the switching range is sufficiently large and the probability of detecting a lie is high enough.

We want to stress that the existence of the switching range is associated with the joint effect of reference-dependence and loss aversion. Indeed, if the electorate were loss-neutral, gains and losses would receive the same weight and the incremental utility associated with electing a high-valence candidate as opposed to a low-valence one would be independent of the reference point. As a result, lies would not entail any cost and the only equilibrium of the game would still be uninformative.

Furthermore, the switching range varies non-monotonically with loss aversion. This is a consequence of the two-sided effect that this behavioral bias plays in our model. On the one hand, as we described above, it makes the electorate unwilling to accept *unexpected losses* leading to an increase in the switching range. On the other hand, it also makes voters unwilling to accept *expected losses* and pushes them to formulate equilibrium strategies which induce little volatility in payoffs. Thus, a high degree of loss aversion will result in equilibrium strategies that pick a high-valence candidate over a low-valence one for most realizations of the popularity shock and this will in turn reduce the switching range. Since the former effect dominates when loss aversion is low, while the latter prevails when loss aversion is high, starting from loss-neutrality, the switching range first increases and then decreases with the degree of loss aversion. As a result, truthtelling is most likely when the electorate is moderately loss averse.

Moreover, we also show that an increase in the uncertainty of the electoral outcomes (as measured by the support of the popularity shock that determines the electoral outcome) may increase the likelihood of a fully revealing equilibrium. Indeed, as the range of popularity shocks grows bigger, the electorate will be relatively more likely to support the low-valence candidate when the opponent is high-valence. This will, in turn, decrease the disadvantage associated with the announcement of being low-valence.

In addition to the full characterization of fully revealing equilibria, we show that other equilibria are possible. In particular, uninformative equilibria exist for every profile of parameters, while partially revealing equilibria arise when the ex-ante probability of high-valence candidates takes intermediate values. We further show that if there is sufficient uncertainty concerning the valence of candidates (for instance, if both types of valence are equally likely) and if the fully revealing equilibrium exists, this is also the only equilibrium that satisfies standard equilibrium refinements proposed for communication games.

The paper is organized as follows. In the remaining of the Introduction, we review the relevant literature. Section 2 describes the model and highlights the interaction between candidates' announcements and the formation of the reference point. In Section 3, we characterize the equilibria of the game. Section 4 discusses the assumptions of our model. Section 5 concludes. The Appendix collects all the proofs.

1.1 Related Literature

This paper focuses on the problem of strategic information transmission between candidates and voters. In this respect our paper is related to the literature on strategic information transmission pioneered by Crawford and Sobel, 1982 and Green and Stokey, 2007.[2] We depart from this literature assuming that the uninformed party (in our model, voters) exhibits reference dependence and loss aversion and we show how these assumptions can lead to credible information transmission.[3,4]

The political science literature has studied extensively how discrepancies between candidates' promises and actual performance can affect the electoral competition. A fruitful line of research started by Farejohn, 1986 addresses the conflict of interest between voters and politicians lacking any commitment power; in this context voters can discipline the incumbent politician by conditioning their electoral behavior on her performance while in office.[5] In this paper, we assume that candidates have an incentive to lie and lack any instrument to commit themselves to truthtelling; nevertheless credible information transmission can be attained thanks to the endogenous effect that announcements have on the reference point of the electorate.

In our model voters evaluate their actions based not only on the final outcome they induce, but also on the comparison between these outcomes and a reference point. This idea dates back at least to Kahneman and Tversky, 1979 and, since then, an extensive experimental evidence has confirmed the importance of reference points and loss aversion in determining agents' behavior.[6] In this paper, we follow Kőszegi and Rabin, 2006, 2007, 2009 and assume that the reference point is endogenously determined through a rational expectation approach;[7] however, we further embed the formation of the reference point into a communication game between informed and uninformed players (respectively, politicians and voters).[8] The relevance of players' behavior in the formation of reference points has been studied both theoretically and experimentally by Gill and Stone, 2010 and Gill and Prowse, 2012, who investigate tournament settings.

Our work is also related to Kőszegi, 2006 as it studies the role that communication and anticipatory utilities[9] can play in an agency problem; however, whereas Kőszegi, 2006 focuses on environments in which the interests of the two parties are perfectly aligned, our paper assume conflicting interests and tackle the issue of credible information transmission when two informed parties

[2] Farrell and Rabin, 1996 and Krishna and Morgan, 2008 provide a review of this literature.

[3] In a similar vein, Grillo, 2012 shows how reference dependence and loss aversion can yield to truthtelling through a change in the risk attitudes of players.

[4] In doing so, we assume that the the content of communication is verifiable with some probability. In this respect, our work is related to Dziuda, 2011, Ottaviani and Sorensen, 2006, Seidmann and Winter, 1997.

[5] See also Banks and Sundaram, 1998, Berganza, 2000, Duggan, 2000 and Schwabe, 2011.

[6] See, for instance Kahneman et al., 1990, Kahneman et al., 1991, van Dijk and van Knippenberg, 1996 and Fehr et al., 2011.

[7] Alternatively, the literature has also identified the reference point as the status quo. On this approach, see Kahneman and Tversky, 1991 and Sugden, 2003.,

[8] For a different approach, see Shalev, 2000.

[9] On anticipatory utilities see Loewenstein, 1987, Loewenstein and Prelec, 1992. For an axiomatic treatment see Caplin and Leahy, 2001, Epstein, 2008.

compete against each others.

Insofar we model a situation in which voters' beliefs concerning their own electoral behavior affect their preferences over final outcomes, our paper belongs to the literature on psychological games started by Geanakoplos et al., 1989 and extended to dynamic environments by Battigalli and Dufwenberg, 2009.[10] In particular, Battigalli et al., 2013 shows how guilt aversion can help attaining credible information transmission. The difference with our setting is not only semantic: our approach could be labelled as "independent of opponents' intentions", as voters' strategies do not depend on voters' intentions. On the contrary, guilt aversion requires modelling players' higher-order belief about opponents' intentions.

Political scientists have long recognized the role played by expectations management in electoral competitions. In particular, Kimball and Patterson, 1997 show that the gap between expectations and politicians' real performance play an important role in determining voters' attitude toward Congress.[11] Waterman et al., 1999 extend this analysis by showing that this expectation gap is important in explaining voters' electoral behavior.[12] On a similar note, a growing literature has documented the role played by expectations in the evaluation of public services.[13]

Nevertheless, to the best of our knowledge, there has been little theoretical work on the role played by reference points in determining electoral outcomes and affecting political equilibria. Some noticeable exceptions are Banks, 1990, Lindstadt and Staton, 2010 and Passarelli and Tabellini, 2013. Banks, 1990 builds a model in which candidates' valence is unknown and candidates incur a cost from delivering an outcome different from what announced; in this paper, we explicitly model the channel through which false announcements can generate such a cost. In Lindstadt and Staton, 2010 candidates are explicitly involved in expectations' manipulation and the paper shows how downward management of expectations can increase candidates' electoral prospects. By characterizing the actual channel through which expectations can affect electoral behavior, our model endogenizes the formation of the reference point and shows how upward management of expectations can be counterproductive. Finally, Passarelli and Tabellini, 2013 build a model in which losses with respect to the citizens' reference point may generate political unrest and use this channel to explain distortions in the level of public expenditure with respect to the Benthamite benchmark and excessive debt accumulation. Besides obvious differences in the research question, our model differs from Passarelli and Tabellini, 2013 also in the choice of the reference point. Whereas they assume that the reference point of a citizen is given by what a utilitarian social planner biased in favor of the citizen would choose, we assume that the reference point is determined in equilibrium by the strategic interaction between candidates and voters.

[10] On psycholigcal games see also Charness and Dufwenberg, 2006, 2010, 2011, Battigalli and Dufwenberg, 2007 and Rabin, 1994.

[11] See also G.R. Boynton and Patterson, 1969.

[12] See also Sigelman and Knight, 1983.

[13] See, for instance, James, 2009 and the references therein.

2 The Model

Two candidates, A and B, compete to get elected. Each candidate can be high or low valence. Formally, a candidate's valence is represented by her type $\theta \in \{\theta_L, \theta_H\}$: if type θ_k is elected, the electorate experiences a consumption utility equal to g_k, $k \in \{L, H\}$. We assume that $g_H > g_L$ and we refer to g_H (respectively, g_L) as to the level of consumption utility yielded by the high (respectively, low) valence candidate. To simplify notation, we define $G = g_H - g_L$. We can interpret candidates valences as the amount of public good that they can provide per unit of taxation.[14]

Candidates' types are determined independently according to the same distribution: each candidate has a probability q (respectively, $1-q$) to be high (respectively, low) valence. The type of a candidate is her own private information. At the beginning of the electoral competition, candidates can make simultaneous and public announcements concerning their types. We assume that communication is costless: candidates do not incur any direct cost from making these announcements. The candidates' utility from being elected and from losing the elections are equal to 1 and 0, respectively.

The electorate is represented by a single (median) voter whose electoral behavior depends not only on his beliefs about the valence of the candidates, but also on the realization of a random variable δ distributed uniformly in the interval $\left[-\frac{1}{2\psi}, \frac{1}{2\psi}\right]$. The assumptions of uniform distribution and of symmetricity about 0 are made for analytical convenience and the main results of the paper would generalize to other absolutely continuous cdf $F(\cdot)$.[15] δ can be thought as a popularity shock that hits voters' preferences after candidates made their announcements and could determine the electoral outcome.

Timing is as follows. In period 0, each candidate makes a statement concerning her own valence. In period 1, three random variables are independently realized: each candidate i generates a signal t^i, that may reveal her true valence to the electorate and the random variable δ is also realized. In period 2, elections take place and utilities are realized.

We want to stress that, whereas signals t^i ($i \in \{\text{A,B}\}$) may reveal something about the valence

[14] For instance, assume that voters have a constant income level y and that their utilty is given by:

$$(1-\tau) \cdot y + G(h), \text{ with } G'(\cdot) > 0 \text{ and } G''(\cdot) < 0,$$

where τ is a proportional tax rate and h is the level of public good provided by the politician in office. Assume futher that the pair (τ, h) is chosen by the elected politician in order to maximize voters' payoff subject to the budget constraint $\tau \cdot y = \frac{h}{\theta}$. Candidates' type are given by $\theta \in \{\theta_L, \theta_H\}$ with $\theta_L < \theta_H$; thus a high-valence candidate is able to provide higher levels of public good, h, for a given level of taxation, τ.

Then, we can define:

$$\begin{aligned} g_L &= (1-\tau^*(\theta_L))\, y + G(h^*(\theta_L)) \\ g_H &= (1-\tau^*(\theta_H))\, y + G(h^*(\theta_H)) \end{aligned}$$

where $(\tau^*(\theta), h^*(\theta))$ is the solution to the problem:

$$\arg\max_{(\tau,g)} (1-\tau)\, y + G(h) \text{ s.t. } \tau \cdot y = \frac{h}{\theta}.$$

It is immediate to verify that in this setting $g_L < g_H$.

[15] In particular, symmetricity about 0 can be easily relaxed. Instead, we need the pdf of $F(\cdot)$ to be uniformly bounded by some number $M(\eta, \lambda, q)$. This upper bound prevents small changes in the probability of being high-valence to result in large changes in the probability of winning.

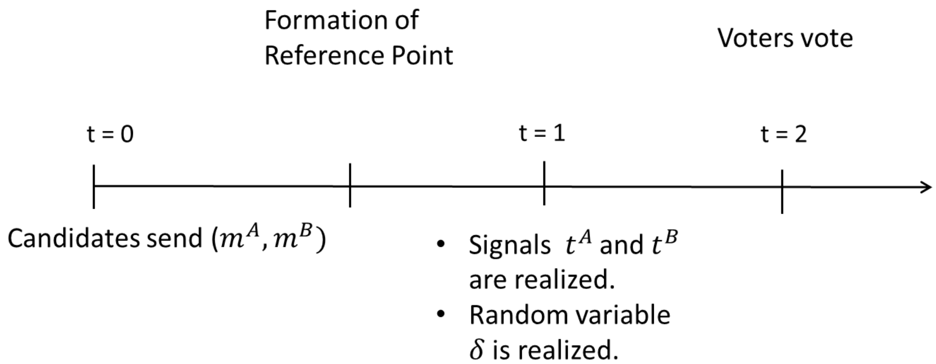

Figure 1: Timeline

of a candidate, δ is independent of candidates' actual types. For instance, δ could represent some personal trait of the candidate that is uncorrelated with her ability to provide the public good (e.g., her empathy or her ability to communicate effectively), some external event that makes the platform of one of the parties more appealing to the electorate, or some scandal that hits the party to which candidate i belongs without directly affecting the candidate. The timing of the model is summarized in Figure 1.

Signals t^i are generated according to the following technology. If a candidate has type g_k, she will send signal t_k with probability p and an uninformative signal, t_0, with probability $(1-p)$. Thus, the set of signals is given by $T = \{t_L, t_0, t_H\}$ and p captures the probability with which candidates' true type is revealed to the voter. In this setting, p can be interpreted as a measure of the degree of media's scrutiny (e.g., fact checking activity) and/or as some sort of blunder the candidate can make during the electoral campaign and reveal her incompetency.[16]

A pure strategy for candidate $i \in \{\text{A,B}\}$ is a function $s^i : \{\theta_L, \theta_H\} \to M$, where M is a finite set of messages. The set of pure strategies is denoted with S. The set of mixed strategies is given by $\Sigma = \Delta(S)$ and its generic element is denoted with σ^i. The voter's behavior can be described with a function $\gamma : M^2 \times T^2 \times \left[-\frac{1}{2\psi}, \frac{1}{2\psi}\right] \to [0,1]$, where $\gamma\left(m^A, m^B, t^A, t^B, d\right)$ is the probability that the electorate appoints candidate A when it listens to announcements $\left(m^A, m^B\right)$, receives signals $\left(t^A, t^B\right)$ and the realization of the popularity shock is d. Voting is costless and there is no abstention. The set of voters' strategies is denoted with Γ.

Departing from the previous literature, we assume that the voter has reference dependent preferences a là Kőszegi and Rabin. In particular, for any pair $(g, r) \in \{g_L, g_H\} \times \{g_L, g_H\}$ the utility function of the electorate is given by:

$$v(g \mid r) = g + \mu(g - r) \tag{1}$$

[16] As it will become clear, the choice of the actual signaling technology is irrelevant as long as, in equilibirum, there exists a positive probability, say p, of detecting the lies of low-valence candidates. Furthermore, the probability p can be endogenized assuming that candidates can distort p downward by exerting some effort e.

where:

$$\mu(x) = \eta \cdot \max\{0, x\} + \eta\lambda \min\{0, x\} \;\; \forall x \in \mathbb{R} \tag{2}$$

with $\eta \in [0,1)$, and $\lambda > 1$. Thus, the electorate's preferences are described by the function $v(\cdot \mid \cdot) : \{g_L, g_H\}^2 \to \mathbb{R}$, in which the first argument, g, represents the actual level of public good experienced by the electorate and the second argument, r, is its reference level. We refer to $v(\cdot \mid \cdot)$ as to the *total utility*. Electorate's total utility can be separated in two components: the *consumption utility*, represented by g, and the *gain/loss utility*, represented by $\mu(g - r)$. Intuitively, whenever the actual level of public good experienced by the electorate, g, exceeds (respectively, falls short of) the reference valence, r, the agent experiences a gain (respectively, a loss). In this setting η measures the relative weight of the gain/loss utility compared to the consumption utility, while $\lambda > 1$ captures *loss aversion*, namely the fact that voters dislike losses more than they like equal-size gains.

Following Kőszegi and Rabin, 2007, we extend the utility function to random outcomes and random reference points as follows. For every $(\tilde{g}, \tilde{r}) \in \Delta(\{g_L, g_H\}) \times \Delta(\{g_L, g_H\})$:[17]

$$V(\tilde{g} \mid \tilde{r}) = \sum_{g \in \{g_L, g_H\}} \sum_{r \in \{g_L, g_H\}} v(g \mid r) \cdot \tilde{g}[g] \cdot \tilde{r}[r] .^{18} \tag{3}$$

Obviously, if $\eta = 0$, the electorate behaves as a standard expected utility maximizer with linear vNM utility indexes. The assumption that deviations from the reference point are evaluated according to a piecewise linear function can be relaxed at the cost of an increase in analytical complexity.

Let $(\sigma^A, \sigma^B) \in \Sigma^2$ be the (independent) conjecture of the voter concerning the communication strategy followed by the candidates.[19] By Bayes rule, the probability the electorate assigns to candidate $i \in \{$A,B$\}$ being high-valence after announcement pair (m^A, m^B) is given by:[20]

$$\pi_1^i(m^i \mid \sigma^i) = \frac{q \cdot \sum_{s_i \in S_i} \sigma^i[s_i] \cdot s_i(\theta_H)[m^i]}{q \cdot \sum_{s_i \in S_i} \sigma^i[s_i] \cdot s_i(\theta_H)[m^i] + (1-q)\sum_{s_i \in S_i} \sigma^i[s_i] \cdot s_i(\theta_L)[m^i]} \tag{4}$$

if m^i has positive probability under σ^i and by any $\pi_1^i(m^i \mid \sigma^i) \in [0,1]$ if, instead, $q \cdot \sum_{s_i \in S_i} \sigma^i[s_i] \cdot s_i(\theta_H)[m^i] + (1-q)\sum_{s_i \in S_i} \sigma^i[s_i] \cdot s_i(\theta_L)[m^i] = 0$.

Similarly, $\pi_2^i(m^i, t^i \mid \sigma^i)$ is the probability that the electorate assigns to the candidate of party i being high valence after announcements m^i and signal t^i given conjecture σ^i. This can be written as:

$$\pi_2^i(m^i, t^i \mid \sigma^i) = \begin{cases} 0 & \text{if } t^i = t_L \\ \pi_1^i(m^i \mid \sigma^i) & \text{if } t^i = t_0 \\ 1 & \text{if } t^i = t_H \end{cases} . \tag{5}$$

[17] In what follows, we will sometimes abuse notation writing $V(g \mid \tilde{r})$ and $V(\tilde{g} \mid r)$ to denote the utility associated with a degenerate distribution over an actual outcome g or a reference outcome r.

[19] In our analysis, we take the shortcut of defining beliefs when players hold independent conjectures about their opponents' behavior. This approach is sufficient for the equilibrium analysis and simplifies the notation. However, it is straightforward to extend the notation to allow for correlation.

[20] Since the probability associated with candidate i depends neither on m^j, nor on σ^j, $j \neq i$, we write $\pi_1^i(m^i \mid \sigma^i)$ instead of $\pi_1^i(m^A, m^B \mid \sigma^A, \sigma^B)$.

In words, if the signal reveals the candidate's type, the electorate will update its belief accordingly; otherwise it will maintain the belief generated by the announcement.

Finally, given $\left\{\pi_1^i\left(m^i \mid \sigma^i\right)\right\}_{i\in A,B}$, let $\hat{\pi}\left(t^A, t^B \mid m^A, m^B, \sigma^A, \sigma^B\right)$ be the probability the electorate assigns to signals $\left(t^A, t^B\right)$ being generated when it holds conjectures $\left(\sigma^A, \sigma^B\right)$ and it received announcements $\left(m^A, m^B\right)$. $\hat{\pi}\left(t^A, t^B \mid m^A, m^B, \sigma^A, \sigma^B\right)$ is summarized in the following table:[21]

	t_L	t_0	t_H
t_L	$p^2\left(1-\pi_1^A\right)\left(1-\pi_1^B\right)$	$p\left(1-\pi_1^A\right)(1-p)$	$p^2\left(1-\pi_1^A\right)\pi_1^B$
t_0	$(1-p)\,p\left(1-\pi_1^B\right)$	$(1-p)^2$	$(1-p)\,p\pi_1^B$
t_H	$p^2\pi_1^A\left(1-\pi_1^B\right)$	$p\pi_1^A(1-p)$	$p^2\pi_1^A\pi_1^B$

The electorate's voting behavior depends on (i) the comparison between the (expected) total utility that candidates can provide, and (ii) the realization of the popularity shock, d. Formally, if the electorate has reference point $\tilde{r}$ and holds conjectures $\left(\sigma^A, \sigma^B\right)$ on the candidates' communication strategy, it will prefer voting for A (B) after history $\left(m^A, m^B, t^A, t^B, d\right)$ if:

$$\pi_2^A\left(m^A, t^A \mid \sigma^A\right) V\left(g_H \mid \tilde{r}\right) + \left(1-\pi_2^A\left(m^A, t^A \mid \sigma^A\right)\right) V\left(g_L \mid \tilde{r}\right) > \ (<) \\ (<) > \pi_2^B\left(m^B, t^B \mid \sigma^B\right) V\left(g_H \mid \tilde{r}\right) + \left(1-\pi_2^B\left(m^B, t^B \mid \sigma^B\right)\right) V\left(g_L \mid \tilde{r}\right) + d$$

and will be indifferent if the two sides are equal.

Although the electorate votes at the end of period 2 only, its reference point is determined after candidates make their initial announcements. This is a sensible assumption: if the announcements have some informational content, the electorate will update its beliefs according to these announcements and such a mental process will modify its reference point as well. In line with Kőszegi and Rabin, 2006, 2007, 2009, we endogenize the formation of the reference point assuming rational expectations: the reference point is determined by the electorate's belief concerning candidates' valence and by the strategy it plans to follow. This generates a loop between the formation of the reference point and the optimality of candidates' strategies: on the one hand, the optimality of a strategy is evaluated with respect to the reference point; on the other hand such a reference point is affected by the strategy chosen by the voter. This loop is closed imposing a natural consistency requirement: to be part of an equilibrium, a strategy must be optimal given that the reference point is the one induced by the strategy itself. Strategies satisfying this requirement are labelled as *reference-point consistent*.

To define reference-point consistency, we need some additional notation. Let $\left(\sigma^A, \sigma^B\right)$ be the electorate's (independent) conjecture about the candidates' communication strategy and $\gamma(\cdot)$ be the strategy it plans to follow. Then, after announcements $\left(m^A, m^B\right)$, it will assign probability $\hat{\pi}\left(t^A, t^B \mid m^A, m^B, \sigma^A, \sigma^B\right)$ to pair $\left(t^A, t^B\right)$ being generated. As a result, its reference point,

[21] To simplify notation we omit to specify the dependence of $\pi_1^i(\cdot)$ on m^i and σ^i.

$\tilde{r}\left(m^A, m^B \mid \gamma, \sigma^A, \sigma^B\right)$, will be given by a probability measure that assigns probability

$$\sum_{t^A, t^B} \hat{\pi}\left(t^A, t^B \mid m^A, m^B, \sigma^A, \sigma^B\right) \cdot \int_{-\frac{1}{2\psi}}^{\frac{1}{2\psi}} \Big[\pi_2^A\left(m^A, t^A \mid \sigma^A\right) \gamma\left(m^A, m^B, t^A, t^B, x\right) dF(x) + \pi_2^B\left(m^B, t^B \mid \sigma^B\right)\left(1 - \gamma\left(m^A, m^B, t^A, t^B, x\right)\right)\Big] dF(x)$$

to g_H and complementary probability to g_L. Intuitively, for every pair $\left(t^A, t^B\right)$ (which arise with probability $\hat{\pi}\left(t^A, t^B \mid \cdot\right)$) and every realization of d of the popularity shock, the electorate will get utility g_H either if it supports candidate A and this candidate is high-valence (which happens with probability $\gamma\left(m^A, m^B, t^A, t^B, d\right) \cdot \pi_2^A\left(m^A, t^A \mid \sigma^A\right)$) or if it supports candidate B and B is high-valence (which happens with probability $\left(1 - \gamma\left(m^A, m^B, t^A, t^B, d\right)\right) \cdot \pi_2^B\left(m^B, t^B \mid \sigma^B\right)$). Notice that the reference point is defined only with respect to the the consumption utility and not also with respect to the realization of the popularity shock. This is line with the idea that δ is a random shock independent of candidates' valence and on which voters have no information.

Furthermore, let $\tilde{g}\left(m^A, m^B, t^A, t^B, d \mid \gamma, \sigma^A, \sigma^B\right)$ be the distribution over $\{g_L, g_H\}$ induced by $\gamma(\cdot)$ after announcements $\left(m^A, m^B\right)$, signals $\left(t^A, t^B\right)$ and realization d, when the conjectures are $\left(\sigma^A, \sigma^B\right)$. Formally, $\tilde{g}\left(m^A, m^B, t^A, t^B, d \mid \gamma, \sigma^A, \sigma^B\right)$ assigns probability:

$$\pi_2^A\left(m^A, t^A \mid \sigma^A\right) \gamma\left(m^A, m^B, t^A, t^B, d\right) + \pi_2^B\left(m^B, t^B \mid \sigma^B\right)\left(1 - \gamma\left(m^A, m^B, t^A, t^B, d\right)\right)$$

to g_H and complementary probability to g_L.

Reference-point consistency requires that for every $\left(m^A, m^B, t^A, t^B, d\right)$, $\gamma\left(m^A, m^B, t^A, t^B, d\right)$ is optimal if the reference point is given by $\tilde{r}\left(m^A, m^B \mid \gamma, \sigma^A, \sigma^B\right)$.

Definition 1 *A strategy γ is reference-point consistent at $\left(m^A, m^B\right)$ given $\left(\sigma^A, \sigma^B\right)$ if for every $\left(t^A, t^B, d\right)$ and $\gamma' \in \Gamma$*

$$V\left(\tilde{g}\left(m^A, m^B, t^A, t^B, d \mid \gamma, \sigma^A, \sigma^B\right) \mid \tilde{r}\left(m^A, m^B \mid \gamma, \sigma^A, \sigma^B\right)\right) - d\gamma\left(m^A, m^B, t^A, t^B, d\right) \geq$$
$$\geq V\left(\tilde{g}\left(m^A, m^B, t^A, t^B, d \mid \gamma', \sigma^A, \sigma^B\right) \mid \tilde{r}\left(m^A, m^B \mid \gamma, \sigma^A, \sigma^B\right)\right) - d\gamma'\left(m^A, m^B, t^A, t^B, d\right)$$

A strategy is reference-point consistent given $\left(\sigma^A, \sigma^B\right)$ if it is reference-point consistent at $\left(m^A, m^B\right)$ given $\left(\sigma^A, \sigma^B\right)$ for every $\left(m^A, m^B\right)$.

We want to stress that the reference point of the electorate is determined through a *forward looking* approach: it is given by the distribution over outcomes induced by the electorate's beliefs and planned strategy after announcements $\left(m^A, m^B\right)$. Nevertheless, once established, the reference point does not change and, in particular, does not adjust to the additional information conveyed by signals $\left(t^A, t^B\right)$; in this respect, in period $t = 2$, the reference point is inherited from the previous periods. This is not a contradictory feature of our model. Indeed, our paper characterizes a channel through which communication may affect the behavior of the uninformed party and modify the equilibrium communication strategy of the informed one. Thus, although it is true that our mechanism work insofar *past* announcements have some persisting saliency in the mind of the electorate, such persistency stems from the fact that the electorate updates its *beliefs about the future* in response to such announcements. Moreover, if we allow for a partial revision of the reference point upon receiving signals $\left(t^A, t^B\right)$, the main qualitative findings of our model would go through.[22]

Candidates are standard expected utility maximizers. In particular, let $\bar{\pi}\left(t^A, t^B \mid \theta^A, \theta^B\right)$ be the probability that signal pair $\left(t^A, t^B\right)$ is generated conditional on candidates types being $\left(\theta^A, \theta^B\right)$. Thus, if agent A with type θ sends message m and believes the other players are following strategies $\left(\sigma^B, \gamma\right)$, her expected utility is given by:

$$\begin{aligned} & U^A\left(m, \sigma^B, \gamma \mid \theta\right) = \\ & \quad = \sum_{s\in S} \sigma^B[s] \Bigg[q \sum_{t^A,t^B} \bar{\pi}\left(t^A, t^B \mid \theta, \theta_H\right) \int_{-\frac{1}{2\psi}}^{\frac{1}{2\psi}} \gamma\left(m, s(\theta_H), t^A, t^B, x\right) dF(x) + \\ & \qquad + (1-q) \sum_{t^A,t^B} \bar{\pi}\left(t^A, t^B \mid \theta, \theta_L\right) \int_{-\frac{1}{2\psi}}^{\frac{1}{2\psi}} \gamma\left(m, s(\theta_L), t^A, t^B, x\right) dF(x) \Bigg]. \end{aligned}$$

Similarly, for candidate B:

$$\begin{aligned} & U^B\left(m, \sigma^A, \gamma \mid \theta\right) = \\ & \quad = \sum_{s\in S} \sigma^A[s] \Bigg[q \sum_{t^A,t^B} \bar{\pi}\left(t^A, t^B \mid \theta_H, \theta\right) \int_{-\frac{1}{2\psi}}^{\frac{1}{2\psi}} \left(1 - \gamma\left(s(\theta_H), m, t^A, t^B, x\right)\right) dF(x) + \\ & \qquad + (1-q) \sum_{t^A,t^B} \bar{\pi}\left(t^A, t^B \mid \theta_L, \theta\right) \int_{-\frac{1}{2\psi}}^{\frac{1}{2\psi}} \left(1 - \gamma\left(s(\theta_L), m, t^A, t^B, x\right)\right) dF(x) \Bigg]. \end{aligned}$$

We are now ready to define the solution concept we will be using throughout the paper.[23]

Definition 2 *A profile of strategies $\left(\gamma, \sigma^A, \sigma^B\right)$ is an equilibrium if:*
(i) for every $i \in \{A, B\}$, if $\sigma^i[s] > 0$ then

$$\forall\theta \in \{\theta_L, \theta_H\},\ s(\theta) \in \arg\max_{m\in M} U^i\left(m, \sigma^j, \gamma \mid \theta\right),\ i \neq j;$$

(ii) γ is a reference-point consistent strategy given $\left(\sigma^A, \sigma^B\right)$.

Notice that if we assume that the voter does not exhibit reference dependence ($\eta = 0$), the equilibrium definition collapses to the one of sequential equilibrium. In the paper, we will be

[22] For instance, we could assume that for each outcome $g \in \{g_L, g_H\}$, the electorate's reference point in period 2 is given by:

$$\alpha \cdot \tilde{r}\left(m^A, m^B \mid \gamma, \sigma^A, \sigma^B\right)[g] + (1-\alpha) \cdot \tilde{g}\left(m^A, m^B, t^A, t^B, d \mid \gamma, \sigma^A, \sigma^B\right)[g]$$

with $\alpha \in (0, 1)$.

[23] In the equilibrium definition, we omit to specify beliefs and to impose their consistency with Bayes rule as implied by (4) and (5).

particularly interested in two types of equilibria: *uninformative equilibria* and *fully revealing* ones. Their formal definition is given below..

Definition 3 *Let $(\gamma, \sigma^A, \sigma^B)$ be an equilibrium. Then:*
(i) the equilibrium is uninformative if for every $i \in \{A,B\}$ and every $m \in M$,

$$\sum_{s:s(\theta_L)=m} \sigma^i [s] = \sum_{s:s(\theta_H)=m} \sigma^i [s]$$

(ii) the equilibrium is fully informative if for every player i, every message m and every pair $\theta, \theta' \in \{\theta_L, \theta_H\}$ with $\theta \neq \theta'$, $\sigma^i [s] > 0$ and $s(\theta) = m$, then $s'(\theta') \neq m$, for every s' such that $\sigma^i [s'] > 0$.

In words, in an uninformative equilibrium, the electorate does not change its prior belief after any message m (namely, $\pi^i (m^i \mid \sigma^i) = q$ for every i and every message m). In this case, we can assume $M = \{\bar{m}\}$ and focus on uninformative communication strategies: $s^i_U(\theta) \equiv \bar{m}$.[24] On the contrary, in a *fully revealing* equilibrium, message pair (m^A, m^B) truthfully reveals candidates' type. In this case, we can assume $M = \{m_L, m_H\}$, where m_k should be interpreted as "my type is θ_k" and focus on fully revealing communication strategies: $s^i_R(\theta_L) = m_L$ and $s^i_R(\theta_H) = m_H$; obviously, $\pi^i (m^i_H \mid s^i_R) = 1$ and $\pi^i (m^i_L \mid s^i_R) = 0$.[25]

We conclude this section imposing an assumption that guarantees a sufficient degree of uncertainty in the electoral process: the support of the popularity shock must be sufficiently large. The actual role of this assumption will be discussed in more details in Section 3.

Assumption 1 $\frac{1}{2\psi} > G(1 + \eta\lambda)$.

Notice that Assumption 1 also puts an upper bound on the degree of loss aversion; indeed, Assumption 1 implies $\lambda < \frac{1}{\eta}\left(\frac{1}{2\psi G} - 1\right) = \bar{\lambda}$. Obviously, $\bar{\lambda} > 1$.

3 Equilibrium Analysis

In this section, we characterize the equilibria of the game. In particular, we first show that without reference dependence the unique equilibrium is uninformative (Proposition 1). Then, we introduce reference dependence and we show that fully revealing equilibria may arise (Proposition 3). Obviously, even under reference dependence, the uninformative equilibrium will still exists as the electorate is always free to ignore candidates' announcements (Proposition 4). Nonetheless, we show that fully revealing equilibria are better off both for high-valence candidates and for the electorate and, as a result, it satisfies most of the refinements proposed in the literature on strategic communication. Finally, we characterize symmetric partially revealing equilibria in which low-valence

[24] To this goal, we will assume that any out-of-equilibrium message $m \neq \bar{m}$ is interpreted by the electorate exactly as message $\bar{m}$, namely that $\pi^i_1 (m \mid s^i_U) = q$ for every $i \in \{\text{A,B}\}$.

[25] In this case, we can assume that any out-of-equilibrium message $m \notin \{m_L, m_H\}$ is interpreted by the electorate as message m_L, namely that $\pi^i_1 (m \mid s^i_R) = 0$ for every $i \in \{\text{A,B}\}$ and every message $m \neq m_L, m_H$.

candidates randomize between a message that reveals them as such and a message that is sent by high-valence candidates (Proposition 5).

Before proceeding with the equilibrium characterization, we prove some general properties of reference-point consistent strategies that will be useful in the analysis.

Lemma 1 *Let $\gamma(\cdot)$ be a reference-point consistent strategy given (σ^A, σ^B). Then, if assumption 1 holds, for every (m^A, m^B, t^A, t^B) there exists $d^*(m^A, m^B, t^A, t^B \mid \sigma^A, \sigma^B) \in \left[-\frac{1}{2\psi}, \frac{1}{2\psi}\right]$ such that:*[26]

$$\gamma\left(m^A, m^B, t^A, t^B, d \mid \sigma^A, \sigma^B\right) = \begin{cases} 1 & \text{if } d < d^*\left(m^A, m^B, t^A, t^B \mid \sigma^A, \sigma^B\right) \\ x \in [0,1] & \text{if } d = d^*\left(m^A, m^B, t^A, t^B \mid \sigma^A, \sigma^B\right) \\ 0 & \text{if } d > d^*\left(m^A, m^B, t^A, t^B \mid \sigma^A, \sigma^B\right) \end{cases} \tag{6}$$

Lemma 1 states that reference-point consistent strategies are cutoff strategies: the electorate will appoint A (respectively, B) if the realization of the popularity shock δ is below (respectively, above) the cutoff $d^*(\cdot \mid \cdot)$. As a result, reference-point consistent strategies can be represented by functions $(m^A, m^B, t^A, t^B) \longmapsto d^*(m^A, m^B, t^A, t^B) \in [-\bar{d}, \bar{d}]$.

A corollary of lemma 1 is that if $\gamma(\cdot)$ is a reference-point consistent strategy, the reference point, $\tilde{r}(m^A, m^B \mid \gamma, \sigma^A, \sigma^B)$, will assign probability

$$\sum_{t^A, t^B} \hat{\pi}\left(t^A, t^B \mid m^A, m^B, \sigma^A, \sigma^B\right) \cdot \Big[\pi_2^A\left(m^A, t^A \mid \sigma^A\right) \cdot F\left(d^*\left(m^A, m^B, t^A, t^B \mid \sigma^A, \sigma^B\right)\right) + \pi_2^B\left(m^B, t^B \mid \sigma^B\right) \cdot \left(1 - F\left(d^*\left(m^A, m^B, t^A, t^B \mid \sigma^A, \sigma^B\right)\right)\right)\Big]$$

to g_H and complementary probability to g_L.

3.1 Equilibrium without Reference Dependence

We begin considering the special case in which the electorate does not exhibit reference dependence ($\eta = 0$). Under this assumption, candidates' announcements have no long-lasting effect on the electorate's preference and, as a result, they lack any credibility. The intuition is straightforward: without reference dependence, claiming to be high-valence increases the probability of being elected if the lie is not detected. On the other hand, even if the electorate realizes that candidate i lied (by receiving a signal t^i that contradicts the initial announcement m^i), the candidate would not be worse off than if he had been sincere from the beginning. Consequently, if a message m^i could increase $\pi_1^i(\cdot \mid \cdot)$, both types of candidate i would send it and the announcement would not be credible. The next proposition formalizes this result.

Proposition 1 *Let $\eta = 0$. Then the unique equilibria of the game are uninformative. Thus, all*

[26] In the following expression, we make the simplyfying assumption that $x = 0$, if $d^*(m^A, m^B, t^A, t^B) = -\frac{1}{2\psi}$ and $x = 1$ if $d^*(m^A, m^B, t^A, t^B) = \frac{1}{2\psi}$.

equilibria are equivalent to $\left(\gamma, s_U^A, s_U^B,\right)$,[27] *where* γ *is characterized by the following thresholds:*[28]

$(\bar{m},\bar{m})$	t_L	t_0	t_H
t_L	0	$-qG$	$-G$
t_0	qG	0	$-(1-q)G$
t_H	G	$(1-q)G$	0

3.2 Full Revelation under Reference Dependence

Now assume that the electorate exhibits reference dependence, namely assume $\eta > 0$. Suppose further that it believes that candidates are following strategy $\left(s_R^A, s_R^B\right)$, namely that each candidate is announcing her valence truthfully. The following proposition characterizes the reference-point consistent strategy of the voter given $\left(s_R^A, s_R^B\right)$.

Proposition 2 *Suppose assumption 1 holds. Let*

$$\begin{aligned} d^+(\eta,\lambda,G) &= G\cdot(1+\eta\lambda) \\ d^-(\eta,G) &= G\cdot(1+\eta) \\ d_R(\eta,\lambda,G\mid\psi) &= \frac{G(2+\eta+\lambda\eta)}{2(1-G(\lambda-1)\psi\eta)} \in \left(d^-(\eta,G), d^+(\eta,\lambda,G)\right). \end{aligned}$$

Then, the reference point consistent strategy of the voter given $\left(s_R^A, s_R^B\right)$ *is characterized by the following thresholds:*

(m_H,m_H)	t_L	t_0	t_H
t_L	0	$-d^+$	$-d^+$
t_0	d^+	0	0
t_H	d^+	0	0

(m_H,m_L)	t_L	t_0	t_H
t_L	0	0	$-d_R$
t_0	d_R	d_R	0
t_H	d_R	d_R	0

,

(m_L,m_H)	t_L	t_0	t_H
t_L	0	$-d_R$	$-d_R$
t_0	0	$-d_R$	$-d_R$
t_H	d_R	0	0

(m_L,m_L)	t_L	t_0	t_H
t_L	0	0	$-d^-$
t_0	0	0	$-d^-$
t_H	d^-	d^-	0

The strategy described in proposition 2 have some interesting properties. First of all, it is easy to see that *initial announcements may have a long-lasting effect on electoral behavior*. Indeed, although $\pi_2^i(m_H,m_H,t_L,t_H) = \pi_2^i(m_L,m_H,t_L,t_H)$ for every $i\in\{A,B\}$, $d^*(m_H,m_H,t_L,t_H) \neq d^*(m_L,m_H,t_L,t_H)$ for every $\lambda > 1$. This happens because initial announcements modify not only the electorate's belief, but also its reference point and this latter change will play a persistent role on voters' behavior.

[27] Strategies s_U^i have been defined after Definition 3.

[28] The t_i-th row and t_j-th column in matrix $\left(m^A, m^B\right)$ represents $d^*\left(m^A, m^B, t_i, t_j\right)$. A similar notation holds for the other propositions as well.

Moreover, *telling a lie may hurt a candidate's electoral prospects.* More precisely, if a candidate lies, her probability of winning the election may fall below the one she could have guaranteed to herself by revealing her valence truthfully. To see this, suppose candidate A has low valence and believes B is following strategy $s_R^B(\cdot)$. Then, if she reveals her type truthfully, she wins either with probability $\frac{1}{2} - \psi \cdot d_R(\eta, \lambda, G \mid \psi)$ (if B has high valence) or with probability $\frac{1}{2}$ (if B has low valence). Instead, if she lies, her probability of winning depends both on her opponent's valence and on the signal she generates. In particular, if the lie is detected (that is, if she generates signal t_L), her probability of winning is either equal to $\frac{1}{2} - \psi \cdot d^+(\eta, \lambda, G)$ (if the opponent has high valence) or to $\frac{1}{2}$ (if the opponent is low valence). Since $d_R(\eta, \lambda, G \mid \psi) < d^+(\eta, \lambda, G)$, we can conclude that lies may lower the probability of winning the election.

Let $S(\eta, \lambda, G \mid \psi) \equiv \psi \cdot (d^+(\eta, \lambda, G) - d_R(\eta, \lambda, G \mid \psi))$; we will refer to $S(\eta, \lambda, G \mid \psi)$ as to the *switching range.*

Thus, we can summarize the previous discussion by saying that the switching range exists (formally, it has positive measure) only if voters exhibit reference dependence ($\eta > 0$) and are loss averse ($\lambda > 1$). Notice that this result holds only if reference dependence is paired with loss aversion. To understand why, observe that with loss-neutral voters ($\lambda = 1$), $S(\eta, 1, G \mid \psi) = 0$ and $d^+(\eta, 1, G) = d^-(\eta, G) = G(1 + \eta)$. Therefore, even if the lie were detected, the low-valence candidate would not be worse off than under a truthtelling strategy and, consequently, lies would have no downsides. Intuitively, the gain the electorate could get by supporting the high-valence candidate and the loss it would incur by supporting the low valence candidate have the same weight; thus, the net effect in favor of the high valence candidate will be the same independently of the reference point. Instead, if the agent is loss averse ($\lambda > 1$), the actual reference point matters: if the reference point assigns a high probability to g_H (because a low-valence candidate lied), the advantage in favor of high-valence candidates will be higher and, as a result, the detection of a lie could significantly decrease the electoral prospects of low valence candidates.

Furthermore, *the switching range is larger for intermediate values of loss aversion.* Recall that $\bar{\lambda}$ is the highest value of loss aversion compatible with assumption 1. Then, it is immediate to verify that $S(\eta, 1, G \mid \psi) = S\left(\eta, \bar{\lambda}, G \mid \psi\right) = 0$, $\left.\frac{\partial S(\eta,\lambda,G|\psi)}{\partial \lambda}\right|_{\lambda=1} > 0$ and $\left.\frac{\partial S(\eta,\lambda,G|\psi)}{\partial \lambda}\right|_{\lambda=\bar{\lambda}} < 0$. We conclude that the switching range is maximized for some value of $\lambda \in \left(1, \bar{\lambda}\right)$. In other words, the cost from lying (as measured by the measure of the switching range) is equal to 0 either if there is no loss aversion ($\lambda = 1$) or if the loss aversion is too high ($\lambda = \bar{\lambda}$) and it is maximal for some intermediate value.

The intuition behind this result relies on the double role played by loss aversion. On the one hand, a high level of loss aversion makes the voter less willing to accept *unexpected losses*; thus, if the electorate finds out that a candidate overstated her valence, it will be more willing (as measured by the cutoff $d^*(\cdot)$) to vote for her opponent as long as she can reduce potential losses. On the other hand, an increase in loss aversion makes the electorate less willing to formulate strategies that can yield *expected losses*; as a result, when loss aversion is maximal $\left(\lambda = \bar{\lambda}\right)$, the electorate will never support low-valence candidates when a high-valence candidate is available ($S\left(\eta, \bar{\lambda}, G \mid \psi\right) = 0$ and

$d^{+}(\eta, \lambda, G) = \frac{1}{2\psi}$).

Finally, it is useful to point out *that the electorate does not decrease its willingness to support a candidate just because she lied.* Indeed, if the candidate turns out to be better than what initially announced, such willingness could even increase. Instead, lying decreases the probability of winning if (i) the lie is detected, (ii) the lie would generate a loss for the voter, and (iii) this loss can be reduced or eliminated by supporting the other candidate. These features distinguish our setting from one in which voters exhibit preferences for honesty and enable us to highlight the circumstances under which truthtelling is more likely to arise.

Having characterized the optimal behavior of the electorate given $\left(s_R^A, s_R^B\right)$, we can focus on the candidates' communication strategies and characterize the conditions under which s_R^i is optimal.

Proposition 3 *Suppose assumption 1 holds. Then there exists $p^*(\eta, \lambda, G, q \mid \psi) < 1$ such that a fully revealing equilibrium exists if and only if $p \in (p^*(\eta, \lambda, G, q \mid \psi), 1)$. Furthermore $p^*(\eta, \cdot, G, q \mid \psi)$ is minimized at some $\lambda \in \left(1, \bar{\lambda}\right)$.*

Thus, under assumption 1, *a fully revealing equilibrium exists if and only if the probability of detecting a lie is sufficiently high.* The actual value of $p^*(\eta, \lambda, G, q \mid \psi)$ is given by:

$$\begin{aligned} p^*(\eta, \lambda, G, q \mid \psi) &= \frac{d_R(\eta, \lambda, G \mid \psi)}{d_R(\eta, \lambda, G \mid \psi) + q \cdot (d^{+}(\eta, \lambda, G) - d_R(\eta, \lambda, G \mid \psi))} \\ &= \frac{2 + \eta + \lambda\eta}{(2 + \eta + \lambda\eta) + q\eta(\lambda - 1)(1 - 2\psi G(1 + \eta\lambda))}. \end{aligned} \tag{7}$$

Equation (7) captures the key trade off faced by a low valence candidate. If she lies and the lie is not detected (which happens with probability $(1-p)$), her probability of winning increases by a positive amount equal to $\psi \cdot d_R(\eta, \lambda, G \mid \psi)$.[29] However, if the lie is detected (which happens with probability p), her probability of winning decreases by $q \cdot S(\eta, \lambda, G \mid \psi)$.[30]

Figure 2 captures the trade-off between truthtelling and lying. It depicts the realizations of δ for which candidate A would win if B follows a fully revealing communication strategy, the electorate plays the reference-point consistent strategy given $\left(\sigma_R^A, \sigma_R^B\right)$ and A is either high-valence (left) or low-valence (right). In this case the probability of winning depends on A's communication strategy (green lines correspond to truthtelling, while red lines correspond to lying), the valence of the other candidate and the possibility that the electorate detects a lie.

An implication of the previous result is that *the only relevant incentive compatibility constraint is the one associated to low-valence candidates.* To put it differently, a high valence candidate has no incentive to understate her actual valence in order to subsequently surprise the voter.[31] Intuitively, although claiming to have low valence and then positively surprising the electorate yields a winning

[29] In particular it goes from $\frac{1}{2} - \psi \cdot \hat{d}(\eta, \lambda, G)$ to $\frac{1}{2}$ if the opponent has high valence and from $\frac{1}{2}$ to $\psi \cdot \hat{d}(\eta, \lambda, G)$ if the opponent has low valence.

[30] In particular, the probability of winning stays constant at $\frac{1}{2}$ if the opponent has low valence and goes from $F\left(-\hat{d}(\eta, \lambda, G)\right)$ to $F\left(-d^{+}(\eta, \lambda, G)\right)$ if the opponent has high valence.

[31] You can see this by observing that $d^{-}(\eta, G) < d_R(\eta, \lambda, G \mid \psi)$.

probability equal to $\frac{1}{2}+\psi \cdot d^{-}(\eta, G)$ if the opponent has low valence and by $\frac{1}{2}+\psi \cdot d_R(\eta, \lambda, G \mid \psi)$ if the opponent has high-valence, announcing to have high valence from the beginning raises the electorate's expectation in favor of high-valence candidates and this, in turn, leads to an even larger winning probability ($\frac{1}{2}+\psi \cdot d_R(\eta, \lambda, G \mid \psi)$ if the opponent is low valence and $\frac{1}{2}+\psi \cdot d^{+}(\eta, \lambda, G)$ if the opponent is high valence).

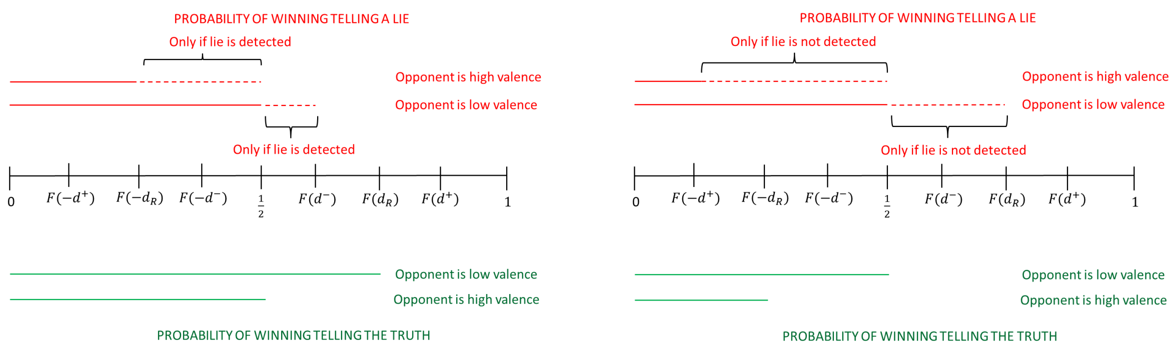

Figure 2: Candidate A's Trade-off between Truthtelling and Lying

Notice that $p^*(0, \lambda, G, q \mid \psi)=1$, $p^*(\eta, 1, G, q \mid \psi)=1$, $p^*\left(\eta, \bar{\lambda}, G, q \mid \psi\right)=1$. Thus, fully revealing equilibria do not exist in either cases. In particular, in the former case announcements do not have no long-lasting effect on the electorate's preferences, while in the latter one initial announcements do affect reference points, but overstating one's valence is not worse than being sincere from the beginning.

Instead, if $\lambda \in \left(1, \bar{\lambda}\right)$, $p^*(\eta, \lambda, G, q \mid \psi)<1$. Furthermore, $p^*(\eta, \lambda, G, q \mid \psi)$ is minimized at:

$$\lambda^*(\eta, \psi, G)=\frac{\left(\sqrt{G \psi(1+\eta)(2 \psi G(1+\eta)+1)}\right)}{G \psi \eta}-\frac{2+\eta}{\eta}$$

Taking the limit of as $\psi \rightarrow 0$, we get:

$$\lim _{\psi \rightarrow 0} p^*(\eta, \lambda, G, q \mid \psi)=\frac{2+\eta+\lambda \eta}{(2+\eta+\lambda \eta)+q \eta(\lambda-1)},$$

which is decreasing in λ. Letting loss aversion going to its limit, we also get:

$$\lim _{\lambda \rightarrow \infty}\left(\lim _{\psi \rightarrow 0} p^*(\eta, \lambda, G, q \mid \psi)\right)=\frac{1}{1+q}.$$

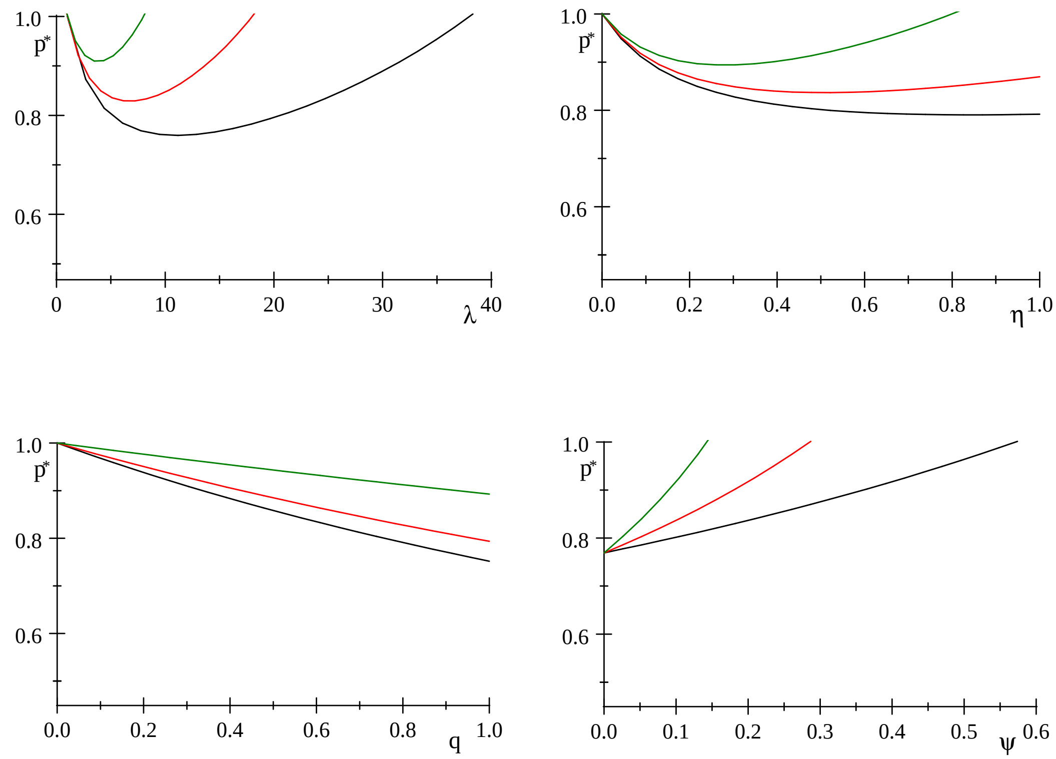

Figure 3: $p^*(\eta, \lambda, G, q \mid \psi)$. Not varying parameters are set equal to $\eta = \frac{1}{2}$, $\lambda = 5$, $q = \frac{3}{4}$, $\psi = \frac{1}{10}$, $G = \frac{1}{4}$ (black line), $G = \frac{1}{2}$ (red line), $G = 1$ (green line).

Thus, as we increase the uncertainty of the electoral outcome (i.e., we decrease ψ) and the degree of loss aversion (i.e., we increase λ), fully revealing equilibria can be supported for lower values of p. The lower bound on p is given by $\frac{1}{1+q} > \frac{1}{2}$. Notice that an increase in electoral uncertainty favors truthtelling equilibria. Indeed, as ψ increases extreme realizations of δ becomes relatively more likely; as a result, a low valence candidate is more likely to win against a high-valence one and this decreases the disadvantage a candidate would incur by announcing to be low-valence.

Figure 3 plots $p^*(\eta, \lambda, G, q \mid \psi)$ as a function of the parameters in our model. The non-monotonic pattern of $p^*(\eta, \lambda, G, q \mid \psi)$ in η and λ can be justified by the intuition we provided before. Furthermore, $p^*(\eta, \lambda, G, q \mid \psi)$ is increasing in G and decreasing in q. The former result follows from the fact that a bigger difference in the ability of the two types makes lies more attractive, whereas the latter one from the fact that an increase in the probability of facing a high-valence opponent increases the cost associated with lying.

3.3 Uninformative Equilibrium

Proposition 3 shows that reference dependence and loss aversion can yield truthful information transmission. Nevertheless, this equilibrium is not unique: an uninformative equilibrium also exists for any set of parameters. This is standard in communication games: if the electorate believes that candidates' announcements do not entail any relevant information and ignores them (that is, if it does not update its beliefs based on such announcements), uninformative communication strategies would be trivially optimal and this would, in turn, justify the electorate's initial conjectures. The characterization of the uninformative equilibrium is provided in the following proposition:

Proposition 4 *Let*

$$d_U = d_U(\eta, \lambda, G, q, p \mid \psi) = \frac{G(1 + \eta\lambda q + \eta(1-q))}{1 - 2G(\lambda - 1)(1-q)qp\psi\eta} \in \left[d^-(\eta, G), d^+(\eta, \lambda, G)\right]$$

Suppose assumption 1 holds. Then, there exists an uninformative equilibrium in which $M = \{\bar{m}\}$, $\left(s^A, s^B\right) = \left(s_U^A, s_U^B\right)$ *and the electorate's reference-point consistent strategy given* $\left(s_U^A, s_U^B\right)$ *is characterized by the following cutoffs:*

$(\bar{m}, \bar{m})$	t_L	t_0	t_H
t_L	0	$-qd_U$	$-d_U$
t_0	qd_U	0	$-(1-q)d_U$
t_H	d_U	$(1-q)d_U$	0

Obviously, in an uninformative equilibrium the behavior of the electorate depends only on the signals candidates generate; in particular, the strategy can be fully described using parameter q (which measures *ex-ante* uncertainty) and cutoff $d_U(\eta, \lambda, G, q, p \mid \psi) \geq G$. It is straightforward to check that $d_U(0, \lambda, G, q, p \mid \psi) = G$ and that $d_U(\cdot, \lambda, G, q, p \mid \psi)$ is increasing in η. Thus, the reference-point consistent strategy given $\left(s_U^A, s_U^B\right)$ is characterized by cutoffs which are larger in absolute value than the ones that would arise without reference dependence ($\eta = 0$). Indeed, in an uninformative equilibrium voters know that both types of candidates will be elected with some probability. This favours high-valence candidates for two reasons: on the one hand, low valence candidates generate losses vis-a-vis high valence candidate; on the other hand, high valence candidates generate gains vis-a-vis low valence candidate; as a result cutoff values will be increasing (in absolute value) with respect to the importance of reference dependence.

Furthermore, it is interesting to notice that $d_U(\eta, \lambda, G, q, p \mid \psi)$ can be higher or lower than $d_R(\eta, \lambda, G \mid \psi)$, the cutoff that arise in a fully revealing equilibrium when the message pair is (m_H, m_L). To understand why, recall that $d_R(\eta, \lambda, G \mid \psi) \in (d^-(\eta, G), d^+(\eta, \lambda, G))$ and observe that $d_U(\eta, \lambda, G, q, p \mid \psi)$ is increasing in q and that $d_U(\eta, \lambda, G, 0, p \mid \psi) = d^-(\eta, G)$ and $d_U(\eta, \lambda, G, 1, p \mid \psi) = d^+(\eta, \lambda, G)$. The intuition behind this result is as follows. Suppose that the candidate pair is (θ_H, θ_L) and that $p \cong 1$ so that types will be fully revealed before the electoral stage. In a fully revealing equilibrium, announcements pair (m_H, m_L) has two effects. On the one hand, it raises $\pi_2^A(\cdot)$ yielding to an increase in the cutoff $d^*(m_H, m_L, t_H, t_L)$. On the other hand, it makes the reference point adapt to the existence of a low valence candidate and mitigates the loss one would incur by voting for her. This second effect leads to a decrease in $d^*(m_H, m_L, t_H, t_L)$. On the contrary, in an uninformative equilibrium the reference point depends on the prior distribution only. Thus, if q is sufficiently high, the reference point given $\left(s_U^A, s_U^B\right)$ will assign a high probability to g_H and, for this reason, $d_U(\eta, \lambda, G, 1, p \mid \psi)$ will be high. As $q \to 1$, the latter channel we described before, will dominate and, for this reason, $d_U(\eta, \lambda, G, q, p \mid \psi)$ will eventually exceed $d_R(\eta, \lambda, G \mid \psi)$. Indeed, $d_R(\eta, \lambda, G \mid \psi) \geq d_U(\eta, \lambda, G, q, p \mid \psi)$ if and only if q is lower than $q^*(\eta, \lambda, G, p \mid \psi)$, where $q^* = q^*(\eta, \lambda, G, p \mid \psi)$ is implicitly defined by $d_R(\eta, \lambda, G \mid \psi) = d_U(\eta, \lambda, G, q^*(\eta, \lambda, G, p \mid \psi), p \mid \psi)$. It is easy to check that $q^*(\eta, \lambda, G, p \mid \psi) > \frac{1}{2}$.

3.4 Partially Revealing Equilibria

Although fully revealing and uninformative equilibria are useful benchmark, *partially revealing equilibria* may also exist. In these equilibria, unlike in uninformative ones, probabilities $\pi_1^i(\cdot)$ $i \in \{\text{A,B}\}$ depend on candidates' announcements, but, unlike in fully revealing equilibria, such probabilities do not jump to 0 or 1. In general, the class of partially revealing equilibria is large as each type may send many different messages.

To simplify the set of equilibria, we impose the following assumption which strengthen Assumption 1:

Assumption 2 $\frac{1}{2\psi} > 2 \cdot G(1 + \eta\lambda)$

Under this assumption, we can prove that the probability that the reference point assigns to g_H increases with the probability that each candidate i is high-valence

Lemma 2 *Let $i \in \{A,B\}$ and $\gamma(\cdot)$ be a reference point consistent strategy given $\left(\sigma^A, \sigma^B\right)$. Then, under assumption 1, for every $m^i, \left(m^i\right)' \in M$ and for every $m^j \in M$*

$$\pi_1^i\left(m^i \mid \sigma^i\right) > \pi_1^i\left(\left(m^i\right)' \mid \sigma^i\right) \Longrightarrow \tilde{r}\left(m^i, m^j \mid \gamma, \sigma^A, \sigma^B\right)[g_H] > \tilde{r}\left(\left(m^i\right)', m^j \mid \gamma, \sigma^A, \sigma^B\right)[g_H]$$

To understand the previous result, consider an increase in the probability that candidate A is high valence, $\pi_1^A\left(m^A \mid \sigma^A\right)$ (the reasoning for candidate B is equivalent and omitted). Obviously, this leads to an increase in the utility that the electorate thinks candidate i can provide. *Coeteris paribus*, this leads to an increase in the probability of winning of candidate i (namely, to an increase in $d^*(\cdot)$). However, if the electorate exhibits reference dependence, a raise in $\pi_1^i\left(m^i \mid \sigma^i\right)$ modifies voters' expectations and can result in harmful losses if the elected candidate is low-valence. As a result, loss aversion may push the electorate toward a decrease in $d^*(\cdot)$ in order to avoid subsequent losses (this will happen only if $\pi_1^A\left(m^A \mid \sigma^A\right)$ is sufficiently low and $\pi_1^B\left(m^A \mid \sigma^A\right)$ is sufficiently high). The net effect of these two forces is, in general, ambiguous. However, assumption 2 implies that an increase in $\pi_1^i\left(m^i \mid \sigma^i\right)$ will be unambiguously associated with a raise in $d^*(\cdot)$. Intuitively, since the electoral outcome is sufficiently uncertain, the electorate's reference point would assign probability to both types of politicians and changes in the probability of high-valence valence candidate will have a moderate effect on the formation of the reference point.

The next lemma shows, the equilibrium analysis can, without loss of generality, focus on profiles in which (i) for each candidate the set of messages, M, is equal to $\{m_*, m^*\}$, (ii) high-valence candidates always send m^*, (iii) low-valence candidates randomize between m^* and m_*.

Lemma 3 *Let* $(\gamma_P, \sigma_P^A, \sigma_P^B)$ *be a partially revealing equilibrium and suppose assumption 2 holds. Then, we can assume without loss of generality that: (i)* $M = \{m_*, m^*\}$, *(ii) if* $s^i(\theta_H) = m_*$, $\sigma_P^i[s^i] = 0$.

Lemma 3 confirms that in equilibrium the relevant incentive compatibility constraint is the one of low-valence candidates.[32] In particular, low-valence candidates face a trade-off similar to the one we have seen in the fully revealing equilibrium. By sending message m_*, the candidate reveals her type and the probability of being elected depends only on the information the electorate gathers on her opponent; if instead she sends message m^*, the probability of winning could raise or decrease depending on whether she generates signal t_0 or signal t_L.

A full characterization of partially revealing equilibria is beyond the scope of this paper. In what follows, we will focus on symmetric partially revealing equilibria, namely equilibria in which both candidates choose the same communication strategy, $\sigma_P^A = \sigma_P^B = \sigma_P$. By lemma 3, these equilibria can be indexed by the common probability $\pi_1(m^* \mid \sigma_P) = \pi$ and can be equivalently characterized by the probability with which the low-valence candidate sends message m^*; let this probability be z. By Bayes rule $z_\pi = \frac{q}{1-q} \cdot \frac{(1-\pi)}{\pi}$. We refer to these equilibria as to *π-symmetric partially revealing equilibria* and for every profile (m^A, m^B, t^A, t^B), we denote with $d_P^*(m^A, m^B, t^A, t^B \mid \pi)$ the cutoff that describes voters' reference point consistent strategy at information set (m^A, m^B, t^A, t^B).

Notice that, in a π-symmetric partially revealing equilibrium, the reference-point consistent strategy after message pair (m_*, m_*), is identical to the one arising in a fully revealing equilibrium after (m_L, m_L). Similarly, if the message pair is (m^*, m^*), the reference-point consistent strategy would be equal to the one of an uninformative equilibrium when the prior probability of high-valence candidates is π. Finally, the reference point consistent strategy following message pairs (m_*, m^*) and (m^*, m_*) is characterized by a threshold $d_P(\eta, \lambda, G, \pi, p \mid \psi) \in (d^-(\eta, G), d_U(\eta, \lambda, G, \pi, p \mid \psi))$. The next proposition provides a formal statement of these findings.

Proposition 5 *Let* (σ_P, σ_P) *be a symmetric profile of communication strategies and assume that* $\pi_1^i(m^* \mid \sigma_P) = \pi$ *for every* $i \in \{A,B\}$. *Define*

$$\begin{aligned} \hat{d} &= d_U(\eta, \lambda, G, \pi, p \mid \psi) = \frac{G(1 + \eta\lambda\pi + \eta(1-\pi))}{1 - 2G\eta(\lambda-1)(1-\pi)\pi p\psi} \\ d_P &= d_P(\eta, \lambda, G, \pi, p \mid \psi) = \frac{G \cdot (2 + \eta\lambda\pi + \eta(2-\pi))}{2 - 2G\eta(\lambda-1)(p + (1-p)\pi)\pi\psi} \end{aligned}$$

[32] The intuition is the same we provided after proposition 3.

Then, $\hat{d} \in (d_U(\eta,\lambda,G,q,p \mid \psi),\ d_R(\eta,\lambda,G \mid \psi))$ *and* $d_P \in \left(d^-(\eta,G),\hat{d}\right)$. *Furthermore, the reference-point consistent strategy given* (σ_P,σ_P) *is characterized by the following cutoffs:*

(m^*,m^*)	t_L	t_0	t_H
t_L	0	$-\pi\hat{d}$	$-\hat{d}$
t_0	$\pi\hat{d}$	0	$-(1-\pi)\hat{d}$
t_H	$\hat{d}$	$(1-\pi)\hat{d}$	0

(m_*,m_*)	t_L	t_0	t_H
t_L	0	0	$-d^-$
t_0	0	0	$-d^-$
t_H	d^-	d^-	0

,

(m_*,m^*)	t_L	t_0	t_H
t_L	0	$-\pi d_P$	$-d_P$
t_0	0	$-\pi d_P$	$-d_P$
t_H	d_P	$(1-\pi)d_P$	0

(m^*,m_*)	t_L	t_0	t_H
t_L	0	0	$-d_P$
t_0	πd_P	πd_P	$-(1-\pi)d_P$
t_H	d_P	d_P	0

In a π-symmetric partially revealing equilibrium, low-valence candidates randomize between two messages (m_* and m^*) and, consequently, they must be indifferent between them. If they send m_*, the politician reveals herself as a low-valence candidate and in this case her expected utility will be given by ($i,j \in \{\text{A,B}\}, i \neq j$):

$$U^i\left(m_*,\sigma_P^j,\gamma_P \mid \theta_L\right) = \frac{1}{2} - q\cdot\psi\cdot d_P(\eta,\lambda,G,\pi,p \mid \psi)$$

Instead, if she sends message m^*, the low-valence candidate pools with high-valence ones and her expected utility becomes ($i,j \in \{\text{A,B}\}, i \neq j$):

$$U^i\left(m^*,\sigma_P^j,\gamma_P \mid \theta_L\right) = \frac{1}{2} - q\cdot p\cdot\psi\cdot d_U(\eta,\lambda,G,\pi,p \mid \psi) + (1-p)(\pi-q)\cdot\psi\cdot d_P(\eta,\lambda,G,\pi,p \mid \psi)$$

Therefore, indifference requires:

$$(1-p)\pi\cdot d_P(\eta,\lambda,G,\pi,p \mid \psi) = qp\cdot\left(d_U(\eta,\lambda,G,\pi,p \mid \psi) - d_P(\eta,\lambda,G,\pi,p \mid \psi)\right) \tag{8}$$

From Proposition 5, we know that for every π and every profile of parameters (η,λ,G,p,ψ) satisfying assumption 2, $d_U(\eta,\lambda,G,\pi,p \mid \psi) > d_P(\eta,\lambda,G,\pi,p \mid \psi)$; thus if $p=1$, the right hand side of (8) is bigger than the left hand one and if $p=0$, the opposite is true. Moreover, keeping all the other parameters constant and exploiting Assumption 2, we can conclude that the left hand side of (8) is decreasing in p, while the right hand side is increasing in it.[33]. Therefore, for every π and every profile of parameters (η,λ,G,ψ), there exists a unique $p^{**}(\eta,\lambda,G,\pi,\psi) \in (0,1)$ such that (8) holds if and only if $p = p^{**}(\eta,\lambda,G,\pi,\psi)$.

[33] To see this last point, observe that:

$$\frac{\partial(qpd_U)}{\partial p} = Gq\cdot\frac{(1+\eta)+\pi(\lambda-1)\eta}{(1-2G(\lambda-1)(1-\pi)\pi p\psi\eta)^2}$$

whereas:

$$\frac{\partial(qpd_P)}{\partial p} = Gq\cdot\frac{\left(1-\pi^2\eta(\lambda-1)\psi G\right)}{(1-G\eta(\lambda-1)(p+(1-p)\pi)\pi)^2}\cdot\left((1+\eta)+\frac{1}{2}\pi(\lambda-1)\eta\right)$$

Now, consider the expected utility of a high-valence candidate in a π-symmetric partially revealing equilibrium. If she sends message m^*, her utility will be given by:

$$\frac{1}{2}+(1-q)\,\psi\cdot\left(z_\pi p\cdot d_U\left(\eta,\lambda,G,\pi,p\mid\psi\right)+\left(1-z_\pi\right)\cdot\left(\left(1-p\right)\pi+p\right)\cdot d_P\left(\eta,\lambda,G,\pi,p\mid\psi\right)\right).$$

Using equation (8), the previous expression can be rewritten as:[34]

$$\frac{1}{2}+(1-q)\,\psi\cdot d_P\left(\eta,\lambda,G,\pi,p\mid\psi\right)$$

On the contrary her utility if she sends message m_* her expected utility will be given by:

$$\frac{1}{2}-q\left(1-p\right)\pi\psi\cdot d_P\left(\eta,\lambda,G,\pi,p\mid\psi\right).$$

Since $d_P\left(\eta,\lambda,G,\pi,p\mid\psi\right)>0$, we conclude that the high valence candidate will be better off sending message m^*.

We summarize the previous discussion in the following proposition:

Proposition 6 *For every player $i\in\{A,B\}$, define two strategies s' and s'', such that $s'\left(\theta_H\right)=m^*$, $s'\left(\theta_L\right)=m_*$ and $s''\left(\theta_H\right)=s''\left(\theta_L\right)=m^*$. Suppose Assumption 2 holds. Then, for every profile of parameters $\left(\eta,\lambda,G,q,p,\psi\right)$, $\left(\gamma_P,\sigma_P,\sigma_P\right)$ is a π-symmetric partially revealing equilibrium if: (i) $\sigma_P\left[s''\right]=z_\pi=\frac{q\cdot(1-\pi)}{(1-q)\cdot\pi}$, (ii) $p=p^{**}\left(\eta,\lambda,G,\pi,\psi\right)$, and (iii) γ_P is characterized by the cutoffs given in Proposition 5. This is also the unique π-symmetric partially revealing equilibrium in the sense of Lemma 3.*

3.5 Equilibrium Comparison

The previous analysis highlights that in our model equilibrium multiplicity may arise. Indeed, whereas uninformative equilibria exist for every profile of parameters, fully revealing equilibria exist if and only if $p>p^*\left(\eta,\lambda,G,q\mid\psi\right)$. Finally, π-symmetric partially revealing equilibria exists when $p=p^{**}\left(\eta,\lambda,G,\pi,\psi\right)\in(0,1)$.

In this section, we will discuss how is it possible to select among different equilibria. To perform such analysis, we make an assumption that will simplify the welfare characterization of equilibria:

Assumption 3 *The* ex-ante *probability of high-types is not too high: $q\leq q^*\left(\eta,\lambda,G,p\mid\psi\right)$.*

Recall that Assumption 3 implies $d_R\left(\eta,\lambda,G\mid\psi\right)\geq d_U\left(\eta,\lambda,G,q,p\mid\psi\right)$.

To assess equilibria's welfare implications, we need to take a stance on the welfare criteria used to compare them. In particular, we will assume that expected utility of candidates is calculated at

The result follows from noticing that by Assumption 1 one can prove that

$$\frac{\left(1-\pi^2\eta\left(\lambda-1\right)\psi G\right)}{\left(1-G\eta\left(\lambda-1\right)\left(p+\left(1-p\right)\pi\right)\pi\right)^2}<\frac{1}{\left(1-2G\left(\lambda-1\right)\left(1-\pi\right)\pi p\psi\eta\right)^2}.$$

[34] See the proof of Proposition ?? for the details.

the *interim* stage, namely after that they learnt their type. Instead, as far as voters are concerned, we will assume that the welfare comparison is performed before voters listen to the actual messages sent by candidates. We regard both these assumptions as sensible. Indeed, equilibria can be distinguished on their degree of informativeness. Thus, on the one hand it makes sense to assume that candidates can choose how much information to reveal after learning their own type. On the other hand, the electorate is the receivers of such information and, consequently, it can decide whether to listen or ignore the announcements only before the actual messages are sent. Intuitively, if the electorate pays attention to what candidates say, this will by definition affect its beliefs and, through this channel, its reference point.[35]

Proposition 7 *Suppose Assumption 3 holds. Then, both the electorate and high-valence candidates are better off in a fully revealing equilibrium than in an uninformative one. is better off in the fully revealing equilibrium than in either the uninformative or the* π*-symmetric partially revealing equilibrium with* $\pi \in (q, 1)$*.*

Turning our attention to the comparison between fully revealing and uninformative equilibrium, we can prove the following proposition.

Proposition 8 *Suppose Assumption 2 holds. Then, both the electorate and high-valence candidates are better off in a fully revealing equilibrium than in a* π*-symmetric partially revealing equilibrium with* $\pi \in (q, 1)$*.*

Not surprisingly, one can easily check low-valence candidates rank equilibria in the opposite way.

Therefore, under Assumption 3 both the electorate and high-valence candidates are better off in the fully informative equilibrium whenever such equilibrium exists. As a result, standard equilibria refinements for communication games would select the fully revealing equilibrium whenever such an equilibrium exists. In particular, the fully revealing equilibrium would satisfy neologism-proofness (Farrell, 1993), announcement-proofness (Matthews et al., 1991) and NITS (namely, "No Incentive to Separate", Chen et al., 2008)., whereas uninformative equilibria would not.

Notice that propositions 7 holds under Assumption 3. If this is not the case, we can find profiles of parameters for which both candidates and the electorate is better off in the uninformative equilibrium than in the fully revealing equilibrium. In particular, if q is sufficiently high and $p \cong 1$, $d_U(\eta, \lambda, G, q, p \mid \psi) > d_R(\eta, \lambda, G \mid \psi)$.[36] As a result, high-valence candidates will prefer the uninformative equilibrium and the same may happen for voters. Whenever this happens, equilibria refinements will select the uninformative equilibrium.

[35] This is a by-product of the fact that in this model we abstract from the use of rethorical tools and from the choice of messages' clarity; thus, the informational content of a message is a property of the equilibrium construction and not of the actual announcement sent by candidates. For a model in which agents strategically choose the clarity of their messages, see Blume and Board, 2009

[36] See the discussion after Proposition 4 for the intuition behind this result.

4 Discussion of the Assumptions

4.1 Distribution over Types

In our model the valences of candidates are drawn independently from the same distribution. The model can be easily extended to deal with the case of different, but independent distribution and all the results would go through. In this case, one can easily show that the candidate with the lowest *ex-ante* probability of being low-valence is the one with the highest incentive to lie as it assigns a lower probability to her opponent being high-valence and, consequently, to her lie being harmful in terms of winning probability.

Our analysis would also go through if we assume that valences are independent conditional on some common shock, ζ. This would be a relevant assumption if, for instance, the cost of providing a certain public good depends both on the politician's idiosyncratic skills ($\theta \in \{\theta_L, \theta_H\}$) and by some macroeconomic shock capturing the status of the economy (ζ). In this case, our results could be easily generalized to the case in which ζ is observable to both candidates before they make their electoral announcements and candidates' skills remain sufficiently important in determining the total cost.[37]

Instead, our results are robust to the introduction of positive correlation among agents' types only if the degree of such correlation does not take extreme values.[38] To understand why, notice that in our model what prevents low-valence candidates from lying is the fear that such a lie could shift the preferences of the electorate in favor of the high-valence candidate. Thus, consider the extreme case in which valences are perfectly correlated (this could also be interpreted as a situation in which the amount of public good provided fully depends on the realization of the macroeconomic shock ζ and individual skills play role). In such a situation, a low-valence candidate is certain that her opponent is also low-valence and, consequently, lying does not entail any cost. As a result, only uninformative equilibria would be possible.

4.2 Heterogenous Voters

The model focuses on the interaction between a representative voter (labelled as "the electorate") and two politicians. This is done as we are more interested in the degree of information sharing between voters and politicians than in the redistributive conflicts within voters. However, voters' heterogeneity can be incorporated in our model.

For instance, consider a continuum of voters indexed by income and let the distribution of income levels be given by an absolutely continuous cdf $H(\cdot)$ with support in the interval $[0, \infty)$. Assume that the utility of a voter with income level y^i is given by $c^i + G(h)$ where c^i is individual consumption, $G(\cdot)$ is a continuous, strictly increasing and strictly concave function with $\lim_{x\to 0} G'(x) = \infty$ and h is a public good that must be financed through a proportional tax τ. Thus $c^i = y^i \cdot (1-\tau)$.

[37] The case in which ζ becomes known after candidates' announcements would require a better description of the message space as candidates may want formulate statements concerning their valence that depends on the actual realization of the macroeconomic shock.

[38] Negative correlation can be easily accomodated, but, obviously, it is not the most sensible assumption.

The government budget constraint depends on the politician in charge. In particular, it is given by: $\tau \cdot \int_0^\infty y dH(y) \geq \frac{h}{\theta}$. Notice that, a high-valence candidate ($\theta = \theta_H$) is more effective in transforming tax revenue into public good and, for this reason, all voters agree that, *coeteris paribus*, a high-valence candidate is better than a low-valence one. Income heterogeneity, however, yields to disagreement on the total amount of public good h that should be provided (in particular, high-income voters would prefer a lower provision of public good than low-income ones).

In this setting, we can interpret the announcements of candidates as a declaration of the total amount of public good they can generate for each level of taxation. Formally, we can model electoral competition assuming that each candidate announces the level of taxation τ with commitment and further tells the electorate how much public good she can provide with such a tax revenue. Once in office, he will then choose the actual pair (τ, h) in order to maximize political consensus or total welfare.[39] The analysis of Section 3 can be easily adapted to deal with this setting and would lead to the same conclusions concerning candidates' communication strategies.

4.3 Reference Dependence and Other Behavioral Biases

In our setting, false announcements modify the electorate's reference point and, through this channel, generate a cost for low-valence candidate. This cost, in turn, pushes candidates toward truthtelling. Other models explain truthtelling using different behavioral biases. For instance, Charness and Dufwenberg, 2006, 2010, 2011 and Battigalli et al., 2013 attain truthtelling through guilt aversion. In such a setting, candidates would not lie in order to avoid the guilt associated with letting the voters down. This approach would require to model players' higher order beliefs as voters expectations of candidates' intentions and candidates beliefs about these expectations would matter. Our modelling choice, instead, does not require the modelling of players' intentions: the change in the electorate's reference point depends only on the information content that voters assign to candidates' statements.

Alternatively, one could also assume that voters have preferences for honesty and punish candidates for not delivering what they promised.[40] Although this assumption would be sensible in many settings, we believe that our approach represents a step forward with respect to the previous literature. First of all, preferences for honesty would, strictly speaking, lead voters to punish candidates even when they positively surprise the electorate, whereas our approach is able to distinguish between gains and losses. Moreover, by modelling the formation of reference points and the mechanism through which it affects voters preferences, we provide a justification behind preferences for honesty and, consequently, we can make better predictions on the circumstances under which lies are most likely to hurt candidates electoral prospects. In particular, it is important to stress that in our model, although the *electorate may behave as if it were punishing candidates for their lies*, the reduction in the probability of winning associated with the detection of a lie does not stem from

[39] The actual choice of candidates' obejctive function is irrelevant as long as the value of the candidate's problem is increasing in her own type.

[40] See Banks, 1990.

the desire of punishing a candidate for her dishonesty. Instead, it follows from the joint effect of the change in reference point induced by the lie and of the desire of avoiding painful losses.

5 Conclusion

In this paper, we provide a model that could reconcile two apparently contradictory claims: (i) candidates' electoral speeches are not credible as politician would be ready to promise whatever it takes in order to be elected, and (ii) politicians are held accountable for their announcements. In particular, we built a simple probabilistic voting model in which two candidates compete to get elected. If voters care about consumption utility only, politicians' announcements would be uninformative: since politicians always have an incentive to pretend to be high valence, their statements will lack any credibility and voters will ignore them. The introduction of reference dependence and loss aversion overcome this problem by adding an additional channel through which politicians' announcements affect voters behavior, namely the formation of reference point. Indeed, if a candidate announce to be high valence, he induces his electorate to expect a high payoff; if voters subsequently find out that he cannot deliver this payoff (because his valence is lower than what initially claimed), they may vote for the opponent to avoid the disappointment associated with low-valence candidates. This effect may induce candidates to reveal their valence sincerely. Furthermore, the range of parameters for which full revelation arises is largest when voters are moderately loss averse. On the one hand, loss aversion must be sufficiently high to induce voters to change their electoral behavior after a lie. On the other hand, loss aversion must not me too high, as, otherwise, it would make voters unwilling to support a low-valence candidate destroying, in turn, her incentive to reveal the truth.

6 Appendix

6.1 Proof of Lemma 1

First, we show that $d \geq d'$ implies $\gamma\left(m^A, m^B, t^A, t^B, d \mid \sigma^A, \sigma^B\right) \geq \gamma\left(m^A, m^B, t^A, t^B, d' \mid \sigma^A, \sigma^B\right)$ for every $\left(m^A, m^B, t^A, t^B\right)$. Suppose not. Then, we can find $d, d' \in \left[-\bar{d}, \bar{d}\right]$ with $d \geq d'$ and a profile $\left(m^A, m^B, t^A, t^B\right)$ such that $\gamma\left(m^A, m^B, t^A, t^B, d \mid \sigma^A, \sigma^B\right) > \gamma\left(m^A, m^B, t^A, t^B, d' \mid \sigma^A, \sigma^B\right)$. To simplify notation, let $\tilde{r} = \tilde{r}\left(m^A, m^B \mid \gamma, \sigma^A, \sigma^B\right)$ and $\pi_2^i = \pi_2^i\left(m^i, t^i \mid \sigma^i\right)$.

Obviously, $\gamma\left(m^A, m^B, t^A, t^B, d \mid \sigma^A, \sigma^B\right) > 0$. Reference point consistency requires:

$$\sum_{r \in \{g_L, g_H\}} \tilde{r}\,[r] \cdot \left[\pi_2^A v\left(g_H \mid r\right) + \left(1 - \pi_2^A\right) v\left(g_L \mid r\right)\right] \geq$$

$$\geq \sum_{r \in \{g_L, g_H\}} \tilde{r}\,[r] \cdot \left[\pi_2^B v\left(g_H \mid r\right) + \left(1 - \pi_2^B\right) v\left(g_L \mid r\right)\right] + d$$

and

$$\sum_{r\in\{g_L,g_H\}} \tilde{r}[r]\cdot\left[\pi_2^A v(g_H \mid r) + \left(1-\pi_2^A\right) v(g_L \mid r)\right] \leq$$

$$\leq \sum_{r\in\{g_L,g_H\}} \tilde{r}[r]\cdot\left[\pi_2^B v(g_H \mid r) + \left(1-\pi_2^B\right) v(g_L \mid r)\right] + d'$$

Since $d > d'$, this is a contradiction.

Now, take any profile $\left(m^A, m^B, t^A, t^B\right)$ and simplify notation as before. If for some $d^* \in \left[-\bar{d}, \bar{d}\right]$,

$$\sum_{r\in\{g_L,g_H\}} \tilde{r}[r]\cdot\left[\pi_2^A v(g_H \mid r) + \left(1-\pi_2^A\right) v(g_L \mid r)\right] =$$

$$= \sum_{r\in\{g_L,g_H\}} \tilde{r}[r]\cdot\left[\pi_2^B v(g_H \mid r) + \left(1-\pi_2^B\right) v(g_L \mid r)\right] + d^*,$$

let $d^*\left(m^A, m^B, t^A, t^B\right) = d^*$.

If instead

$$\sum_{r\in\{g_L,g_H\}} \tilde{r}[r]\cdot\left[\pi_2^A v(g_H \mid r) + \left(1-\pi_2^A\right) v(g_L \mid r)\right] \geq$$

$$\geq \sum_{r\in\{g_L,g_H\}} \tilde{r}[r]\cdot\left[\pi_2^B v(g_H \mid r) + \left(1-\pi_2^B\right) v(g_L \mid r)\right] + \frac{1}{2\psi}$$

then, let $d^*\left(m^A, m^B, t^A, t^B\right) = \frac{1}{2\psi}$. On the contrary, if

$$\sum_{r\in\{g_L,g_H\}} \tilde{r}[r]\cdot\left[\pi_2^A v(g_H \mid r) + \left(1-\pi_2^A\right) v(g_L \mid r)\right] \leq$$

$$\leq \sum_{r\in\{g_L,g_H\}} \tilde{r}[r]\cdot\left[\pi_2^B v(g_H \mid r) + \left(1-\pi_2^B\right) v(g_L \mid r)\right] - \frac{1}{2\psi}$$

let $d^*\left(m^A, m^B, t^A, t^B\right) = -\frac{1}{2\psi}$. The statement of the lemma follows immediately. ∎

6.2 Proof of Proposition 1

Let $\left(\gamma, \sigma^A, \sigma^B\right)$ be an equilibrium. First notice that for every message m^i, $\pi_2^i\left(m^i, t_L \mid \sigma^i\right) = 0$, $\pi_2^i\left(m^i, t_0 \mid \sigma^i\right) = \pi_1^i\left(m^i \mid \sigma^i\right)$ and $\pi^i\left(m^i, t_H \mid \sigma^i\right) = 1$. Therefore, $\pi_2^i\left(m^i, t^i \mid \sigma^i\right)$ depends on m only if $t^i = t_0$.

Furthermore, at any information set $\left(m^A, m^B, t^A, t^B, d\right)$, the voter votes for A if:

$$\left(\pi_2^A\left(m^A, t^A \mid \sigma^A\right) - \pi_2^B\left(m^B, t^B \mid \sigma^B\right)\right) G > d$$

and for B if:

$$\left(\pi_2^A\left(m^A,t^A\mid\sigma^A\right)-\pi_2^B\left(m^B,t^B\mid\sigma^B\right)\right)G\leq d.$$[41]

Thus, following Lemma 1, for every profile $\left(m^A,m^B,t^A,t^B\right)$, we can define cutoff:

$$d^*\left(m^A,m^B,t^A,t^B\right)=\left(\pi_2^A\left(m^A,t^A\mid\sigma^A\right)-\pi_2^B\left(m^B,t^B\mid\sigma^B\right)\right)\cdot G$$

Now consider candidate A (the reasoning for B is analogous and omitted). Her expected utility when her type is θ and she sends message $m\in M$ is given by:

$$V_A\left(m,\sigma^B,\gamma\mid\theta\right)=\sum_{s\in S}\sigma^B\left[s\right]q\left[\sum_{t^A,t^B}\bar{\pi}\left(t^A,t^B\mid\theta,\theta_H\right)F\left(d^*\left(m^A,s\left(\theta_H\right),t^A,t^B\right)\right)+\right.$$
$$\left.+\left(1-q\right)\sum_{t^A,t^B}\bar{\pi}\left(t^A,t^B\mid\theta,\theta_L\right)F\left(d^*\left(m^A,s\left(\theta_L\right),t^A,t^B\right)\right)\right].$$

Notice that this expression depends on m^A only through $d^*\left(\cdot\right)$, which is increasing in $\pi_2^A\left(\cdot\right)$. Furthermore, it is strictly increasing whenever

$$\left|\left(\pi_2^A\left(m^A,t^A\mid\sigma^A\right)-\pi_2^B\left(m^B,t^B\mid\sigma^B\right)\right)G\right|<\frac{1}{2\psi},$$

which holds by Assumption 1.

Suppose there exists a message m_H sent with positive probability such that $\pi_1^A\left(m_H\mid\sigma^A\right)>q$. Then, there must exists another message m_L, sent with positive probability, such that $\pi_1^A\left(m_L\mid\sigma^A\right)<q$. Therefore, message m_L must be sent with positive probability by a candidate with low valence. Then, the low valence candidate could modify her strategy and send message m_H every time she was supposed to send message m_L. Obviously, $\pi_1^A\left(m_H\mid\sigma^A\right)>\pi_1^A\left(m_L\mid\sigma^A\right)$ and, consequently, $\pi_2^A\left(m_H,t_0\mid\sigma^A\right)>\pi_2^A\left(m_L,t_0\mid\sigma^A\right)$. As a result, this deviation would increase the candidates' expected utility and contradict the definition of equilibrium.

We conclude that for every message m, $\pi_1^A\left(m\mid\sigma^A\right)=\pi_1^B\left(m\mid\sigma^B\right)=q$ and that

$$\pi_2^A\left(m^i,t^i\mid\sigma^i\right)=\begin{cases}1 & t^i=1\\ q & t^i=t_0\\ 0 & t^i=t_L\end{cases}$$

The statement of the theorem follows from the cutoffs we defined before. ∎

6.3 Proof of Proposition 2

Suppose the electorate holds conjecture $\left(s_R^A,s_R^B\right)$. Then, for every $i\in\{\text{A,B}\}$, $\pi_1^i\left(m_H,\mid s_R^i\right)=1$, $\pi_1^i\left(m_L\mid s_R^i\right)=0$. We will analyze each messages pair separately (we can assume that any message $m\notin\{m_L,m_H\}$ is interpreted as coming from a low valence candidate).

First, consider message pair is $\left(m_H,m_H\right)$. Notice that $\hat{\pi}\left(t^A,t^B\mid m_H,m_H,s_R^A,s_R^B\right)>0$ if and only if $\left(t^A, t^B\right) \in \left\{\left(t_H, t_H\right), \left(t_H, t_0\right), \left(t_0, t_H\right), \left(t_0, t_0\right)\right\}$ and that, for all these signals pairs, $\pi_2^i\left(m_H, t^i \mid s_R^i\right) = 1$ for every $i \in \{\text{A,B}\}$. Thus, the reference point of the voter will be a degenerate probability measure that assigns probability 1 to g_H. Therefore, one can easily check that $d^*\left(m_H, m_H, t^A, t^B\right) = 0$ for all signal pairs $\left(t^A, t^B\right) \in \left\{\left(t_H, t_H\right), \left(t_H, t_0\right), \left(t_0, t_H\right), \left(t_0, t_0\right)\right\}$. Suppose instead that A generates signal t_L, while B generates either t_0 or t_H. Then $\pi^A\left(m_H, t_L \mid s_T^A\right) = 0$, while $\pi^B\left(m_H, t_H \mid s_T^B\right) = \left(m_H, t_0 \mid s_T^B\right) = 1$. Thus, the cutoff will be defined by $g_L + \eta\lambda\left(g_L - g_H\right) = g_H + d$. Thus,

$$d^*\left(m_H, m_H, t_L, t_0\right) = d^*\left(m_H, m_H, t_L, t_H\right) = -G \cdot (1 + \eta\lambda) = -d^+(\eta, \lambda, G),$$

which is greater than $-\frac{1}{2\psi}$ by assumption 1. A symmetric reasoning yields $d^*\left(m_H, m_H, t_0, t_L\right) = d^*\left(m_H, m_H, t_H, t_L\right) = G \cdot (1 + \eta\lambda) = d^+(\eta, \lambda, G)$. Finally if both candidates generate signal t_L, then $d^*\left(m_H, m_H, t_L, t_L\right) = 0$ as the threshold is defined by:

$$g_L + \eta\lambda\left(g_L - g_H\right) = g_L + \eta\lambda\left(g_L - g_H\right) + d.$$

Now, consider message pair $\left(m_L, m_L\right)$. In this case, $\hat{\pi}\left(t^A, t^B \mid m_L, m_L, s_R^A, s_R^B\right) > 0$ if and only if $\left(t^A, t^B\right) \in \left\{\left(t_L, t_L\right), \left(t_L, t_0\right), \left(t_0, t_L\right), \left(t_0, t_0\right)\right\}$ and after these signals pairs $\pi^i\left(m_L, t^i \mid s_R^i\right) = 0$ for every candidate i. Then, the reference point at $\left(m_L, m_L\right)$ will be a degenerate measure that assigns probability 1 to g_L. As a result, for every signal pair $\left(t^A, t^B\right) \in \left\{\left(t_L, t_L\right), \left(t_L, t_0\right), \left(t_0, t_L\right), \left(t_0, t_0\right)\right\}$, $d^*\left(m_L, m_L, t^A, t^B\right) = 0$. Furthermore, if candidate A generates signal t_H, while candidate B generates signal t_L or t_0, the cutoff will be defined by equation: $g_H + \eta\left(g_H - g_L\right) = g_L + d$. Thus, $d^*\left(m_L, m_L, t_H, t_L\right) = d^*\left(m_L, m_L, t_H, t_0\right) = G \cdot (1 + \eta) = d^-(\eta, G)$. Symmetrically, $d^*\left(m_L, m_L, t_L, t_H\right) = d^*\left(m_L, m_L, t_0, t_H\right) = -G \cdot (1 + \eta) = -d^-(\eta, G)$. Finally if both candidates generates signal t_H, $d^*\left(m_L, m_L, t_H, t_H\right)$ is the solution of the following equation:

$$g_H + \eta\left(g_H - g_L\right) = g_H + \eta\left(g_H - g_L\right) + d,$$

so that $d^*\left(m_L, m_L, t_H, t_H\right) = 0$.

Finally, let the messages pair be $\left(m_H, m_L\right)$ (the case $\left(m_L, m_H\right)$ is symmetric and omitted). Then $\hat{\pi}\left(t^A, t^B \mid m_H, m_L, s_R^A, s_R^B\right) > 0$ if and only if $\left(t^A, t^B\right) \in \left\{\left(t_0, t_0\right), \left(t_H, t_L\right), \left(t_H, t_0\right), \left(t_0, t_L\right)\right\}$ and after all these signal pairs we have $\pi_2^A\left(m_H, t^A \mid s_R^A\right) = 1$ and $\pi_2^B\left(m_L, t^B \mid s_R^B\right) = 1$. Thus, the reference point, $\tilde{r}\left(m_H, m_L \mid \gamma, s_R^A, s_R^B\right)$ will be given by a probability measure that assigns probability $\sum_{t^A, t^B}\left[\hat{\pi}\left(t^A, t^B \mid m^A, m^B, \sigma_R^A, \sigma_R^B\right) \cdot F\left(d^*\left(m^A, m^B, t^A, t^B \mid \sigma_R^A, \sigma_R^B\right)\right)\right]$ to g_H and complementary probability to g_L. Reference point consistency requires that after signal pairs $\left(t^A, t^B\right) \in$

$\{(t_0, t_0), (t_H, t_L), (t_H, t_0), (t_0, t_L)\}$,

$$G + \eta \left(1 - \sum_{t^A, t^B} \hat{\pi} \left(t^A, t^B \mid m^A, m^B, \sigma^A, \sigma^B \right) \cdot F \left(d^* \left(m^A, m^B, t^A, t^B \mid \sigma^A, \sigma^B \right) \right) \right) G +$$

$$+ \eta\lambda \left(\sum_{t^A, t^B} \hat{\pi} \left(t^A, t^B \mid m^A, m^B, \sigma^A, \sigma^B \right) \cdot F \left(d^* \left(m^A, m^B, t^A, t^B \mid \sigma^A, \sigma^B \right) \right) \right) G =$$

$$= d^* \left(m_H, m_L, t^A, t^B \right)$$

We conclude that for all $\left(t^A, t^B\right) \in \{(t_0, t_0), (t_H, t_L), (t_H, t_0), (t_0, t_L)\}$, $d^* \left(m_H, m_L, t^A, t^B\right)$ must be the same. Let this common value be equal to $d_R = d_R (\eta, \lambda, G \mid \psi)$. Then $\tilde{r} \left(m_H, m_L \mid \gamma, s_R^A, s_R^B\right)$ will assign probability $F(d_R)$ to g_H and $(1 - F(d_R))$ to g_L. Thus for all signal pairs $\left(t^A, t^B\right) \in \{(t_0, t_0), (t_H, t_L), (t_H, t_0), (t_0, t_L)\}$, reference point consistency requires $d^* \left(m_H, m_L, t^A, t^B\right) = d_R$ where d_R is given by the value that solves:

$$g_H + \eta (1 - F(d)) (g_H - g_L) = g_L + \eta\lambda F(d) (g_L - g_H) + d$$

We conclude that:

$$d_R (\eta, \lambda, G \mid \psi) = \frac{G (2 + \eta + \eta\lambda)}{2 (1 - \eta (\lambda - 1) \psi G)}$$

By Assumption 1 $(1 - \eta (\lambda - 1) \psi G) > 0$ and $d_R (\eta, \lambda, G \mid \psi) \in [d^- (\eta, G), d^+ (\eta, \lambda, G)]$. Furthermore, $d_R (\eta, 1, G \mid \psi) = d^- (\eta, G)$ and $d_R \left(\eta, \bar{\lambda}, G \mid \psi\right) = d^+ \left(\eta, \bar{\lambda}, G\right) = \frac{1}{2\psi}$. Now suppose A generates signal t_H or t_0, while candidate B generates signal t_H. Then $\pi_2^A \left(m_H, t_H \mid s_R^A\right) = \pi_2^A \left(m_H, t_0 \mid s_R^A\right) = 1$ and $\pi_2^B \left(m_L, t_H \mid s_R^B\right) = 1$ and the cutoff on δ will be defined by:

$$g_H + \eta \cdot (1 - F(d)) \cdot G = g_H + \eta \cdot (1 - F(d)) \cdot G + d$$

so that $d^* (m_H, m_L, t_H, t_H) = d^* (m_H, m_L, t_0, t_H) = 0$. Following a similar reasoning, we can conclude that $d^* (m_H, m_L, t_L, t_L) = d^* (m_H, m_L, t_L, t_0) = 0$. ∎

6.4 Proof of Proposition 3

Consider candidate A (the analysis for candidate B is similar and omitted) and suppose she conjectures that B is following communication strategy s_R^B and that the voter is playing the reference-point consistent strategy given $\left(s_R^A, s_R^B\right)$. By proposition 2, the difference in expected utility between truthtelling and lying is given by:

$$U^A \left(m_H, s_R^B, \gamma \mid \theta_H\right) - U^A \left(m_L, s_R^B, \gamma \mid \theta_H\right) = \frac{q}{2} + (1 - q) F (d_R (\eta, \lambda, G \mid \psi)) -$$

$$- q \left(\frac{p}{2} + (1 - p) F (-d_R (\eta, \lambda, G \mid \psi))\right) - (1 - q) \left(pF \left(d^- (\eta, G)\right) + \frac{(1 - p)}{2}\right)$$

if the candidate has high valence and by

$$U^A\left(m_L, s_R^B, \gamma \mid \theta_L\right) - U^A\left(m_H, s_R^B, \gamma \mid \theta_L\right) = qF\left(-d_R\left(\eta, \lambda, G \mid \psi\right)\right) + \frac{(1-q)}{2} -$$
$$- q\left(pF\left(-d^+\left(\eta, \lambda, G\right)\right) + \frac{(1-p)}{2}\right) - (1-q)\left(\frac{p}{2} + (1-p)F\left(d_R\left(\eta, \lambda, G \mid \psi\right)\right)\right)$$

if the candidate has low valence.

Since $d_R(\eta, \lambda, G \mid \psi) > d^-(\eta, G) > 0$ and F is symmetric about 0, we can easily conclude that $V^A\left(m_H, s_R^B, \gamma \mid g_H\right) > V^A\left(m_L, s_R^B, \gamma \mid g_H\right)$. Therefore truthtelling is optimal for high-valence candidates.

Define $h(p) = U^A\left(m_L, s_R^B, \gamma \mid \theta_L\right) - U^A\left(m_H, s_R^B, \gamma \mid \theta_L\right)$. Thus,

$$h(p) = q\left(F\left(-d_R\left(\eta, \lambda, G \mid \psi\right)\right) - pF\left(-d^+\left(\eta, \lambda, G\right)\right) - \frac{(1-p)}{2}\right) -$$
$$- (1-q)(1-p)\left(F\left(d_R\left(\eta, \lambda, G \mid \psi\right)\right) - \frac{1}{2}\right)$$

Obviously, $h(\cdot)$ is continuous in p. Furthermore since $d^+(\eta, \lambda, G) > \hat{d}(\eta, \lambda, G) > 0$ and F is symmetric about 0,

$$\begin{aligned} h(0) &= q\left(F\left(-d_R\left(\eta, \lambda, G \mid \psi\right)\right) - \frac{1}{2}\right) - (1-q)\left(F\left(d_R\left(\eta, \lambda, G \mid \psi\right)\right) - \frac{1}{2}\right) < 0 \\ h(1) &= q\left(F\left(-d_R\left(\eta, \lambda, G \mid \psi\right)\right) - F\left(-d^+\left(\eta, \lambda, G\right)\right)\right) > 0 \end{aligned}$$

Furthermore:

$$h'(p) = q\left(\frac{1}{2} - F\left(-d^+\left(\eta, \lambda, G\right)\right)\right) + (1-q)\left(F\left(d_R\left(\eta, \lambda, G \mid \psi\right)\right) - \frac{1}{2}\right) > 0$$

Thus, there exists a unique $p^*(\eta, \lambda, G, q \mid \psi) < 1$, such that $U^A\left(m_L, s_R^B, \gamma \mid \theta_L\right) = U^A\left(m_H, s_R^B, \gamma \mid \theta_L\right)$. We conclude that if $p > (<)\ p^*(\eta, \lambda, G, q \mid \psi)$, then $U^A\left(m_L, s_R^B, \gamma \mid \theta_L\right) > (<)\ U^A\left(m_H, s_R^B, \gamma \mid \theta_L\right)$. As a result, if $p \in [p^*(\eta, \lambda, G, q \mid \psi), 1]$, then $U^A\left(m_L, s_R^B, \gamma \mid \theta_L\right) \geq U^A\left(m_H, s_R^B, \gamma \mid \theta_L\right)$ and $U^A\left(m_H, s_R^B, \gamma \mid \theta_H\right) \geq U^A\left(m_L, s_R^B, \gamma \mid \theta_H\right)$ so that a fully revealing equilibrium exists.[42]

On the contrary, if a fully revealing equilibrium exists, then we need $U^A\left(m_L, s_R^B, \gamma \mid \theta_L\right) \geq U^A\left(m_H, s_R^B, \gamma \mid \theta_L\right)$ and $U^A\left(m_H, s_R^B, \gamma \mid \theta_H\right) \geq U^A\left(m_L, s_R^B, \gamma \mid \theta_H\right)$. By the previous reasoning, we can conclude that p must belong to the interval $[p^*(\eta, \lambda, G, q \mid \psi), 1]$.

Also notice that $\lim_{\lambda \to 1} p^*(\eta, \lambda, G, q \mid \psi) = 1$ and similarly, $\lim_{\lambda \to \bar{\lambda}} p^*(\eta, \lambda, G, q \mid \psi) = 1$ (this last result follows from the fact that $\lim_{\lambda \to \bar{\lambda}} d_R(\eta, \lambda, G \mid \psi) = \frac{1}{2\psi}$). Furthermore, by the implicit

[42] In the knife-edge case in which $p = p^*(\eta, \lambda, G, q \mid F)$, we assume that type θ_L sends message m_L. Obviously, none of our results hinges on this tie-breaking rule.

function theorem

$$\left.\frac{\partial p^{*}\left(\eta,\lambda,G,q\mid\psi\right)}{\lambda}\right|_{\lambda=1}<0,\quad \left.\frac{\partial p^{*}\left(\eta,\lambda,G,q\mid\psi\right)}{\lambda}\right|_{\lambda=\lambda}>0,$$

so that $p^{*}\left(\eta,\lambda,G,q\mid\psi\right)$ is minimized for some value of loss aversion $\lambda\in\left(1,\bar{\lambda}\right)$. ∎

6.5 Proof of Proposition 4

To prove the existence of the uninformative equilibrium, it is sufficient to show that the electorate's reference point consistent strategy given $\left(s_U^A,s_U^B\right)$ is the one described in the proposition. Obviously, $\pi_1^i\left(\bar{m}\mid s_U^i\right)=q$, $\pi_2^i\left(\bar{m},t_L\mid s_U^i\right)=0$, $\pi_2^i\left(\bar{m},t_0\mid s_U^i\right)=q$, $\pi_2^i\left(\bar{m},t_H\mid s_U^i\right)=1$ for every $i\in$ {A,B}. Furthermore, $\hat{\pi}\left(t^A,t^B\mid m^A,m^B,s_U^A,s_U^B\right)$ is given by:

$\left(m^A,m^B\right)$	t_L	t_0	t_H
t_L	$p^2\left(1-q\right)^2$	$p\left(1-q\right)\left(1-p\right)$	$p^2\left(1-q\right)q$
t_0	$\left(1-p\right)p\left(1-q\right)$	$\left(1-p\right)^2$	$\left(1-p\right)pq$
t_H	$p^2q\left(1-q\right)$	$pq\left(1-p\right)$	p^2q^2

Pick any reference point and notice that for any signal t, $\pi_2^A\left(\bar{m},t\mid s_U^A\right)=\pi_2^B\left(\bar{m},t\mid s_U^B\right)$. Thus, voting for candidate A leads the same consumption utility and gain/loss utility then voting for candidate B. As a result,

$$d^{*}\left(\bar{m},\bar{m},t_L,t_L\right)=d^{*}\left(\bar{m},\bar{m},t_0,t_0\right)=d^{*}\left(\bar{m},\bar{m},t_H,t_H\right)=0$$

Now suppose that the voter receives signals $\left(t_H,t_L\right)$. Then, $\pi_2^A\left(\bar{m},t_H\mid s_U^A\right)=1$, $\pi_2^A\left(\bar{m},t_L\mid s_U^A\right)=0$ and the relevant cutoff, $d^{*}\left(\bar{m},\bar{m},t_H,t_L\right)$, is defined by:

$$\left(1+\eta\cdot\tilde{r}\left(\bar{m},\bar{m}\mid\gamma,s_U^A,s_U^B\right)\left[g_L\right]+\eta\lambda\cdot r\left(\bar{m},\bar{m}\mid\gamma,s_U^A,s_U^B\right)\left[g_H\right]\right)\cdot G=d^{*}\left(\bar{m},\bar{m},t_H,t_L\right),$$

where $r\left(\bar{m},\bar{m}\mid\gamma,s_U^A,s_U^B\right)$ has to be determined. Rearranging terms, we get:

$$d^{*}\left(\bar{m},\bar{m},t_H,t_L\right)=\left(1+\eta+\eta\left(\lambda-1\right)\cdot r\left(\bar{m},\bar{m}\mid\gamma,s_U^A,s_U^B\right)\left[g_H\right]\right)\cdot G \tag{9}$$

The symmetric nature of our model yields $d^{*}\left(\bar{m},\bar{m},t_L,t_H\right)=-d^{*}\left(\bar{m},\bar{m},t_H,t_L\right)$.

Furthermore, following similar steps, we can easily conclude that:

$$d^{*}\left(\bar{m},\bar{m},t_H,t_0\right)=\left(1-q\right)\cdot d^{*}\left(\bar{m},\bar{m},t_H,t_L\right)=-d^{*}\left(\bar{m},\bar{m},t_0,t_H\right) \tag{10}$$

and

$$d^{*}\left(\bar{m},\bar{m},t_0,t_L\right)=q\cdot d^{*}\left(\bar{m},\bar{m},t_H,t_L\right)=-d^{*}\left(\bar{m},\bar{m},t_L,t_0\right). \tag{11}$$

Equations (9)-(11) define the reference point consistent strategy as a function of the reference point. Substituting $\hat{\pi}\left(t^A, t^B \mid \bar{m}, \bar{m}, s_U^A, s_U^B\right)$ in the definition of the reference point, we get:

$$\begin{aligned}
\tilde{r}\left(\bar{m}, \bar{m} \mid \gamma, s_U^A, s_U^B\right)[g_H] &= p^2 q (1-q) F\left(d^*\left(\bar{m}, \bar{m}, t_H, t_L\right)\right) + \\
&+ p^2 (1-q) q \left(1 - F\left(d^*\left(\bar{m}, \bar{m}, t_L, t_0\right)\right)\right) + \\
&+ (1-p) p (1-q) q F\left(d^*\left(\bar{m}, \bar{m}, t_0, t_L\right)\right) + \\
&+ (1-p) pq \left(q F\left(d^*\left(\bar{m}, \bar{m}, t_0, t_H\right)\right) + 1 - F\left(d^*\left(\bar{m}, \bar{m}, t_0, t_H\right)\right)\right) \\
&+ p (1-p) q \left(F\left(d^*\left(\bar{m}, \bar{m}, t_H, t_0\right)\right) + q\left(1 - F\left(d^*\left(\bar{m}, \bar{m}, t_H, t_0\right)\right)\right)\right) \\
&+ p (1-q)(1-p) q \left(1 - F\left(d^*\left(\bar{m}, \bar{m}, t_L, t_0\right)\right)\right) + (1-p)^2 q + p^2 q^2
\end{aligned}$$

Using the symmetric nature of our model, we can simplify the previous expression into:

$$\begin{aligned}
&\tilde{r}\left(m^A, m^B \mid \gamma, s_U^A, s_U^B\right)[g_H] = \\
&\quad = q\left(1 - (1-q) p (2-p)\right) + 2(1-q) q p^2 F\left(d^*\left(\bar{m}, \bar{m}, t_H, t_L\right)\right) + \\
&\quad\quad + 2(1-q) qp (1-p)\left(F\left(d^*\left(\bar{m}, \bar{m}, t_0, t_L\right)\right) + F\left(d^*\left(\bar{m}, \bar{m}, t_H, t_0\right)\right)\right)
\end{aligned}$$

and using equations (9)-(11), we can conclude that:

$$\begin{aligned}
&\tilde{r}\left(m^A, m^B \mid \gamma, s_U^A, s_U^B\right)[g_H] = \\
&\quad = q\left(1 - (1-q) p (2-p)\right) + 2(1-q) q p^2 F\left(d^*\left(\bar{m}, \bar{m}, t_H, t_L\right)\right) + \\
&\quad\quad + 2(1-q) qp (1-p)\left(F\left(q d^*\left(\bar{m}, \bar{m}, t_H, t_L\right)\right) + F\left((1-q) d^*\left(\bar{m}, \bar{m}, t_H, t_L\right)\right)\right)
\end{aligned}$$

Thus, $d^*\left(\bar{m}, \bar{m}, t_H, t_L\right)$ and $\tilde{r}\left(\bar{m}, \bar{m} \mid \gamma, s_U^A, s_U^B\right)[g_H]$ are defined by the following two equations in two unknowns:

$$d^*\left(\bar{m}, \bar{m}, t_H, t_L\right) = G\left(1 + \eta + \eta(\lambda - 1)\tilde{r}\left(\bar{m}, \bar{m} \mid \gamma, s_U^A, s_U^B\right)[g_H]\right)$$

$$\begin{aligned}
&\tilde{r}\left(m^A, m^B \mid \gamma, s_U^A, s_U^B\right)[g_H] = \\
&\quad = q\left(1 - (1-q) p (2-p)\right) + 2(1-q) q p^2 F\left(d^*\left(\bar{m}, \bar{m}, t_H, t_L\right)\right) + \\
&\quad\quad + 2(1-q) qp (1-p)\left(F\left(q d^*\left(\bar{m}, \bar{m}, t_H, t_L\right)\right) + F\left((1-q) d^*\left(\bar{m}, \bar{m}, t_H, t_L\right)\right)\right)
\end{aligned}$$

In particular, $d^*\left(\bar{m}, \bar{m}, t_H, t_L\right)$ solves:

$$\begin{aligned}
d^*\left(\bar{m}, \bar{m}, t_H, t_L\right) = G\Big(&1 + \eta + \eta(\lambda-1) q\left(1 - (1-q) p (2-p)\right) + \\
&+ 2\eta(\lambda-1) q (1-q) p\Big(p F\left(d^*\left(\bar{m}, \bar{m}, t_H, t_L\right)\right) + \\
&+ (1-p)\left(F\left((1-q) d^*\left(\bar{m}, \bar{m}, t_H, t_L\right)\right) + F\left(q d^*\left(\bar{m}, \bar{m}, t_H, t_L\right)\right)\right)\Big)\Big)
\end{aligned}$$

so that $d^*\left(\bar{m}, \bar{m}, t_H, t_L\right) = d_U\left(\eta, \lambda, G, q, p \mid \psi\right) = \frac{G(1+\eta\lambda q+\eta(1-q))}{1-2G(\lambda-1)(1-q)qp\psi\eta}$.

By assumption 1, $1-2G(\lambda-1)(1-q)qp\psi\eta > 0$ and consequently $d_U(\eta,\lambda,G,q,p \mid \psi) > d^-(\eta,G)$ if and only if $\eta(\lambda-1)q(1+2G(1-q)p\psi(1+\eta)) > 0$, which always holds. Furthermore, $d_U(\eta,\lambda,G,q,p \mid \psi)$ is increasing in q and equals $d^+(\eta,\lambda,G)$ when $q=1$. We conclude that $d_U(\eta,\lambda,G,q,p \mid \psi) \in [d^-(\eta,G), d^+(\eta,\lambda,G)]$. ∎

6.6 Proof of Lemma 2

Consider candidate A (the proof for candidate B is identical and omitted). Let γ be a reference-point consistent strategy given (σ^A,σ^B). Then, by Lemma 1 we know that for any (m^A,m^B):

$$\tilde{r}\left(m^A,m^B \mid \gamma,\sigma^A,\sigma^B\right)[g_H] = \sum_{t^A,t^B} \hat{\pi}\left(t^A,t^B \mid m^A,m^B,\sigma^A,\sigma^B\right) \cdot \Big[\pi_2^A\left(m^A,t^A \mid \sigma^A\right) F\left(d^*\left(m^A,m^B,t^A,t^B \mid \sigma^A,\sigma^B\right)\right) + \\ + \pi_2^B\left(m^B,t^B \mid \sigma^B\right)\left(1-F\left(d^*\left(m^A,m^B,t^A,t^B \mid \sigma^A,\sigma^B\right)\right)\right)\Big]$$

and that for any (m^A,m^B,t^A,t^B):

$$\begin{aligned} d^*\left(m^A,m^B,t^A,t^B \mid \sigma^A,\sigma^B\right) &= \left(\pi_2^A\left(m^A,t^A \mid \sigma^A\right)-\pi_2^B\left(m^B,t^B \mid \sigma^B\right)\right) \cdot \\ &\quad \cdot G \cdot \left(1+\eta+\eta(\lambda-1)\cdot \tilde{r}\left(m^A,m^B \mid \gamma,\sigma^A,\sigma^B\right)[g_H]\right) \end{aligned}$$

Substituting $d^*\left(m^A,m^B,t^A,t^B \mid \sigma^A,\sigma^B\right)$ in $\tilde{r}\left(m^A,m^B \mid \gamma_P,\sigma_P^A,\sigma_P^B\right)[g_H]$, using the definition of $\hat{\pi}\left(t^A,t^B \mid m^A,m^B,\sigma^A,\sigma^B\right)$ and exploiting the symmetricity of $F(\cdot)$, we conclude that:[43]

$$\begin{aligned} \tilde{r}[g_H] = &(1-p)p\left(1-\pi_1^A\right)\pi_1^B F\left(\pi_1^B \cdot \Lambda\right) + \\ &+ p\pi_1^A(1-p)\left(\pi_1^B+\left(1-\pi_1^B\right)F\left(\left(1-\pi_1^B\right)\cdot\Lambda\right)\right) + \\ &+ (1-p)p\pi_1^B\left(\pi_1^A+\left(1-\pi_1^A\right)F\left(\left(1-\pi_1^A\right)\cdot\Lambda\right)\right) + \\ &+ p^2\left(\pi_1^A+\pi_1^B-\pi_1^A\pi_1^B\right)F(\Lambda) + (1-p)p\left(1-\pi_1^B\right)\pi_1^A F\left(\pi_1^A\cdot\Lambda\right) + \\ &+ p^2\pi_1^A\pi_1^B(1-F(\Lambda)) + (1-p)^2\left(\pi_1^B+\left(\pi_1^A-\pi_1^B\right)F\left(\left(\pi_1^A-\pi_1^B\right)\cdot\Lambda\right)\right), \end{aligned}$$

where $\Lambda = G(1+\eta+\eta(\lambda-1)\cdot\tilde{r}[g_H])$.

Applying the implicit function theorem and using Assumption 2 and the fact that $F(\cdot)$ is uniform in $\left[-\frac{1}{2\psi},\frac{1}{2\psi}\right]$, we can conclude that the derivative of $\tilde{r}[g_H]$ with respect to π_1^A is positive.

6.7 Proof of Lemma 3

Let $(\gamma,\sigma^A,\sigma^B)$ be an equilibrium. Pick any player i and define $\pi^{i,*} = \max_{m\in M}\pi_1^i\left(\cdot \mid \sigma_P^i\right)$. The maximum exists since the message space is finite. Obviously $\pi^{i*} > 0$.

First, we show that high-valence candidates will only send messages that yield probability $\pi^{i,*}$.

[43] To simplify notation, we omit to specify the dependency on messages (m^A,m^B), signals (t^A,t^B) and players' behavior $(\gamma,\sigma^A,\sigma^B)$.

Formally, we show that for every i, $\sigma^i\left[s^i\right] > 0$ if and only if $\pi_1^i\left(s^i\left(g_H\right) \mid \sigma^i\right) = \pi^{i*}$. Then, we will be able to focus without loss of generality on equilibria in which high-valence candidates send a message m^* such that $\pi_1^i\left(m^* \mid \sigma_P^i\right) = \pi^{i,*}$

Suppose not. Then we can find a message m such that: (i) $s^i\left(g_H\right) = m$ for some s^i such that $\sigma^i\left[s^i\right] > 0$, and (ii) $\pi_1^i\left(m \mid \sigma^i\right) < \pi^{i*}$. Thus, high valence candidate must be indifferent between inducing belief π^{i*} and inducing belief $\pi_1^i\left(m \mid \sigma^i\right)$. Pick any message m^* such that $\pi_1^i\left(m^* \mid \sigma^i\right) = \pi^{i*}$. Since $\pi_1^i\left(m \mid \sigma^i\right) < \pi^{i*}$, lemma 2 implies that for any message m^j sent by the candidate j $(j \neq i)$, $\tilde{r}\left(m^*, m^j \mid \gamma, \sigma^A, \sigma^B\right)\left[g_H\right] > \tilde{r}\left(m, m^j \mid \gamma, \sigma^A, \sigma^B\right)\left[g_H\right]$. Furthermore, for every profile $\left(m^A, m^B, t^A, t^B\right)$, the cutoff $d^*\left(m^A, m^B, t^A, t^B \mid \sigma^A, \sigma^B\right)$ is equal to:

$$\begin{aligned} d^*\left(m^A, m^B, t^A, t^B \mid \sigma^A, \sigma^B\right) &= \left(\pi_2^A\left(m^A, t^A \mid \sigma^A\right) - \pi_2^B\left(m^B, t^B \mid \sigma^B\right)\right) \cdot \\ &\quad \cdot G \cdot \left(1 + \eta + \eta\left(\lambda - 1\right) \cdot \tilde{r}\left(m^A, m^B \mid \gamma, \sigma^A, \sigma^B\right)\left[g_H\right]\right). \end{aligned}$$

Therefore, for any profile $\left(m^B, t^A, t^B\right)$, $d^*\left(m^*, m^B, t^A, t^B \mid \sigma^A, \sigma^B\right) \geq d^*\left(m, m^B, t^A, t^B \mid \sigma^A, \sigma^B\right)$ with strict inequality whenever $\pi_2^A\left(m^A, t^A \mid \sigma^A\right) \neq \pi_2^B\left(m^B, t^B \mid \sigma^B\right)$, which happens with positive probability. Thus, the high valence candidate will always prefer sending message m^* instead of message m, contradicting our initial hypothesis. As a result, $\sigma^i\left[s^i\right] > 0$ if and only if $\pi_1^i\left(s^i\left(g_H\right) \mid \sigma^i\right) = \pi^{i*}$.

Now, consider low-valence candidate. Three cases are possible. If the low-valence candidate of party i never (respectively, always) sends message m^*, then the equilibrium is a fully informative, (respectively, uninformative) one. Then consider the case in which the low-valence candidate sends both m^* and some other messages. In the former case, we can assume that she is playing only one message $m_* \neq m^*$ as our previous result implies that $\pi_1^i\left(m \mid \sigma^i\right) = 0$ for every $m \neq m^*$. Thus, we can assume that the low-valence candidate sends only two messages: m^* and an additional message m_*, such that $\pi_1^i\left(m_* \mid \sigma^i\right) = 0$. ∎

6.8 Proof of Proposition 5

Let $(\gamma_P, \sigma_P, \sigma_P)$ be a π-symmetric partially revealing equilibrium. We will consider the four possible message pairs independently.

If the message pair is (m_*, m_*), the cutoffs are identical to the one generated in a fully revealing equilibrium after message pair (m_L, m_L). If instead the message pair is (m^*, m^*), the analysis is identical to the one of an uninformative equilibrium in which $q = \pi$. Thus, cutoffs can be described in the following table:

(m^*, m^*)	t_L	t_0	t_H
t_L	0	$-\pi\hat{d}$	$-\hat{d}$
t_0	$\pi\hat{d}$	0	$-\left(1-\pi\right)\hat{d}$
t_H	$\hat{d}$	$\left(1-\pi\right)\hat{d}$	0

where $\hat{d} = d_U\left(\eta, \lambda, G, \pi, p \mid \psi\right) \in \left[d_U\left(\eta, \lambda, G, q, p \mid \psi\right), d^+\left(\eta, \lambda, G\right)\right]$. In particular, the inclusions follow from $\pi > q$ and the fact that $d_U\left(\eta, \lambda, G, q, p \mid \psi\right)$ is increasing in q and reaches $d^+\left(\eta, \lambda, G\right)$

when $q = 1$.

Consider, message pair (m^*, m_*) (the case (m_*, m^*) is symmetric and omitted). In this case $\hat{\pi}\left(t^A, t^B \mid m^*, m_*, \sigma_P, \sigma_P\right)$ is given by

(m^*, m_*)	t_L	t_0	t_H
t_L	$p^2(1-\pi)$	$p(1-p)(1-\pi)$	0
t_0	$(1-p)p$	$(1-p)^2$	0
t_H	$p^2\pi$	$p(1-p)\pi$	0

In this case, the cutoffs associated with various signal pairs are given by:

$$d^*(m^*, m_*, t_H, t_0) = d^*(m^*, m_*, t_H, t_L) = \\ = G \cdot (1 + \eta \cdot \tilde{r}(m^*, m_* \mid \sigma_P, \sigma_P, \gamma_P)[g_L] + \eta\lambda \cdot \tilde{r}(m^*, m_* \mid \sigma_P, \sigma_P, \gamma_P)[g_H])$$

$$d_P^*(m^*, m_*, t_0, t_0) = d_P^*(m^*, m_*, t_0, t_L) = \pi \cdot d_P^*(m^*, m_*, t_H, t_L)$$
$$d_P^*(m^*, m_*, t_L, t_0) = d_P^*(m^*, m_*, t_L, t_L) = 0$$

Furthermore, one can also verify that the reference point $\tilde{r}(m_*, m^* \mid \sigma_P, \sigma_P, \gamma_P)$ assigns probability

$$p \cdot \pi \cdot F(d_P^*(m^*, m_*, t_H, t_L)) + (1-p) \cdot \pi \cdot F(\pi \cdot d_P^*(m^*, m_*, t_H, t_L)) = \\ = \frac{\pi}{2} + \psi \cdot \pi \cdot d_P^*(m^*, m_*, t_H, t_L) \cdot (p + (1-p)\pi)$$

to g_H and complementary probability to g_L.

Solving for $\tilde{r}(m^*, m_* \mid \sigma_P, \sigma_P, \gamma_P)$ and $d^*(m^*, m_*, t_H, t_L)$, we get:

$$d^*(m^*, m_*, t_H, t_L) = \frac{G \cdot \left(1 + \eta + \eta \cdot (\lambda - 1)\frac{\pi}{2}\right)}{(1 - G\eta \cdot (\lambda - 1) \cdot (p + (1-p)\pi)\pi\psi)}$$

$$\tilde{r}(m^*, m_* \mid \sigma_P, \sigma_P, \gamma_P)[g_H] = \pi \cdot \left(\frac{1}{2} + \frac{(p + (1-p)\pi) \cdot \pi\psi \cdot G \cdot \left(1 + \eta + \eta \cdot (\lambda - 1)\frac{\pi}{2}\right)}{(1 - G\eta \cdot (\lambda - 1) \cdot (p + (1-p)\pi)\pi\psi)}\right)$$

Given these results, one can immediately conclude that:

$$d^*(m^*, m_*, t_H, t_H) = 0$$
$$d^*(m^*, m_*, t_0, t_H) = -(1-\pi) \cdot d^*(m^*, m_*, t_H, t_L)$$
$$d^*(m^*, m_*, t_L, t_H) = -d^*(m^*, m_*, t_H, t_L)$$

Define $d_P(\eta, \lambda, G, \pi, p \mid \psi) = d^*(m^*, m_*, t_H, t_L)$ and notice that this cutoff is increasing in π and reaches $d_R(\eta, \lambda, G \mid \psi)$ when $\pi = 1$. Thus, $d_P(\eta, \lambda, G, \pi, p \mid \psi) < d_R(\eta, \lambda, G \mid \psi)$ for every π and p.

Furthermore $d_P(\eta,\lambda,G,\pi,p\mid\psi) < d_U(\eta,\lambda,G,\pi,p\mid\psi)$ if and only if:

$$\frac{\left(1+\eta\cdot\left(1-\frac{\pi}{2}\right)+\eta\lambda\cdot\frac{\pi}{2}\right)}{\left(1-G\eta\cdot(\lambda-1)\cdot(p+(1-p)\,\pi)\cdot\pi\psi\right)} < \frac{\left(1+\eta\lambda\cdot\pi+\eta\cdot(1-\pi)\right)}{1-2G\eta\cdot(\lambda-1)\cdot(1-\pi)\,\pi p\psi}$$

The previous inequality is satisfied if and only if:

$$2G\psi\left(\pi_1(1+\eta)-p(1-\pi_1)(1+\eta)+(\lambda-1)\,\eta\pi_1^2\right)<1$$

which is guaranteed by assumption 1. ∎

6.9 Proof of Proposition 7

The utility of the electorate in the fully revealing equilibrium is equal to:

$$\begin{aligned} &q^2 g_H+(1-q)^2 g_L+2\cdot q\cdot(1-q)\cdot\left(g_L+F\left(d_R(\eta,\lambda,G\mid\psi)\right)\cdot G\right)\\ &\qquad -2\cdot q\cdot(1-q)\cdot\eta(\lambda-1)\cdot F\left(d_R(\eta,\lambda,G\mid\psi)\right)\cdot\left(1-F\left(d_R(\eta,\lambda,G\mid\psi)\right)\right)G, \end{aligned}$$

Instead, in an uninformative equilibrium it is given by:

$$\begin{aligned} &q^2 g_H+(1-q)^2 g_L+2\cdot q\cdot(1-q)\cdot(g_L+K\cdot G)+\\ &\quad+\left(q^2\eta\cdot\tilde{r}\left(\bar{m},\bar{m}\mid\gamma_U,s_U^A,s_U^B\right)[g_L]-(1-q)^2\eta\lambda\cdot\tilde{r}\left(\bar{m},\bar{m}\mid\gamma_U,s_U^A,s_U^B\right)[g_H]\right)\cdot G+\\ &\quad+2q(1-q)\cdot\left(\eta K\cdot\tilde{r}\left(\bar{m},\bar{m}\mid\gamma_U,s_U^A,s_U^B\right)[g_L]-\eta\lambda(1-K)\cdot\tilde{r}\left(\bar{m},\bar{m}\mid\gamma_U,s_U^A,s_U^B\right)[g_H]\right)G \end{aligned}$$

where K is the probability with which the high-valence candidate is chosen when one candidate is high-valence and the other one is low valence. Formally:

$$\begin{aligned} K &= p^2\cdot F\left(d_U(\eta,\lambda,G,q,p\mid\psi)\right)+\frac{(1-p)^2}{2}+\\ &\quad+p(1-p)\cdot F\left((1-q)\cdot d_U(\eta,\lambda,G,q,p\mid\psi)\right)+(1-p)\,p\cdot F\left(q\cdot d_U(\eta,\lambda,G,q,p\mid\psi)\right) \end{aligned}$$

and, consequently,

$$K=\frac{1}{2}+p\cdot\psi\cdot d_U(\eta,\lambda,G,q,p\mid\psi)$$

Furthermore, $\tilde{r}\left(\bar{m},\bar{m}\mid\gamma_U,s_U^A,s_U^B\right)[g_H]=q^2+2q(1-q)\,K$.

Using the definition of reference points and simplifying, we can conclude that the total utility of the electorate in the fully revealing equilibrium is greater than the one in the uninformative equilibrium if and only if:

$$\begin{aligned} 2\cdot(F(d_R)-K) > \eta(\lambda-1)\cdot\Big(&2\cdot F(d_R)\cdot(1-F(d_R))\\ &-q(1-q)-2q^2(1-K)-2(1-q)^2K-4q(1-q)K(1-K)\Big) \end{aligned}$$

Recall that $F(d_R), K > \frac{1}{2}$.

Since Assumption 3 holds, $d_R(\eta, \lambda, G \mid \psi) \geq d_U(\eta, \lambda, G, q, p \mid \psi)$ and, consequently, $F(d_R) > K$. Thus, the left hand side of the previous inequality is positive whereas the right-hand side is lower or equal than:

$$\eta(\lambda - 1) \cdot \Big(2 \cdot K \cdot (1 - K) - q(1 - q) - 2q^2(1 - K) - 2(1 - q)^2 K - 4q(1 - q) K(1 - K)\Big)$$

which is always negative (indeed, the previous expression is maximized when $K = \max\left\{\frac{1}{2}, \frac{q^2}{1-2q(1-q)}\right\}$ and, in both these cases, it is negative).

Now, consider high-valence politicians. In a fully revealing equilibrium, such candidates would get an expected utility equal to:

$$U^A\left(m_H, s_R^B, \gamma_R \mid \theta_H\right) = \frac{1}{2} + (1 - q) \cdot \psi \cdot d_R(\eta, \lambda, G \mid \psi).$$

On the other hand, their expected utility in an uninformative equilibrium would be given by:

$$U^A\left(\bar{m}, s_U^B, \gamma_U \mid \theta_H\right) = \frac{1}{2} + (1 - q) \cdot p \cdot \psi \cdot d_U(\eta, \lambda, G, q, p \mid \psi)$$

By Assumption 3, it is immediate to verify that $U^A\left(m_H, s_R^B, \gamma_R \mid \theta_H\right) > U^A\left(\bar{m}, s_U^B, \gamma_U \mid \theta_H\right)$. ∎

6.10 Proof of Proposition 8

Fix a π-symmetric partially revealing equilibrium. To simplify notation, we will denote the threshold of Proposition with d_P^π and d_U^π instead of $d_P(\eta, \lambda, G, \pi, p \mid \psi)$ and $d_U(\eta, \lambda, G, \pi, p \mid \psi)$. Moreover, let $d_R = d_R(\eta, \lambda, G \mid \psi)$.

Notice that in a π-partially revealing equilibrium, the probability with which a high-valence candidate gets elected is given by:

$$\begin{aligned} &q^2 + 2q(1 - q)(1 - z) \cdot (pF(d_P^\pi) + (1 - p) F(d_P^\pi)) + \\ &\quad + 2q \cdot (1 - q) \cdot z \cdot \Big(p^2 F(d_U^\pi) + (1 - p)^2 F(0) + p(1 - p) F((1 - \pi) d_U^\pi) + (1 - p) pF(\pi d_U^\pi)\Big) \end{aligned}$$

where $z = \frac{(1-\pi) \cdot q}{\pi \cdot (1-q)}$. The previous expression can be simplified to:

$$q^2 + 2q(1 - q) z \left(\frac{1}{2} + p\psi d_U^\pi\right) + 2q(1 - q)(1 - z)\left(\frac{1}{2} + p\psi d_P^\pi + (1 - p)\pi\psi d_P^\pi\right)$$

Similarly, the probability of electing a low-valence candidate is given by:

$$(1 - q)^2 + 2q(1 - q) z \cdot \left(\frac{1}{2} - p\psi d_U^\pi\right) + 2q(1 - q)(1 - z)\left(\frac{1}{2} - p\psi d_P^\pi - (1 - p)\psi\pi d_P^\pi\right)$$

As a result, the consumption utility that the electorate gets in a π-symmetric partially revealing equilibrium is given by:

$$g_L + qG + 2q(1-q)\left((1-z)\cdot(p+(1-p)\pi)\cdot d_P^{\pi} + z\cdot p\cdot d_U^{\pi}\right)\cdot\psi G$$

while the gain/loss utility is given by:

$$-\eta(\lambda-1)\left(q^2 + 2q(1-q)z\left(\frac{1}{2}+p\psi d_U^{\pi}\right) + 2q(1-q)(1-z)\left(\frac{1}{2}+p\psi d_P^{\pi} + (1-p)\pi\psi d_P^{\pi}\right)\right)\cdot$$
$$\cdot\left((1-q)^2 + 2q(1-q)z\cdot\left(\frac{1}{2}-p\psi d_U^{\pi}\right) + 2q(1-q)(1-z)\left(\frac{1}{2}-p\psi d_P^{\pi} - (1-p)\psi\pi d_P^{\pi}\right)\right)G$$

Observe that the consumption utility in a fully revealing equilibrium is higher than the one in a π-symmetric partially revealing equilibrium if and only if:

$$d_R \geq (1-z)\cdot(p+(1-p)\pi)\cdot d_P^{\pi} + z\cdot p\cdot d_U^{\pi} \tag{12}$$

Now, notice that the right-hand side of 12 can be written as

$$(p+(1-p)\pi)\cdot d_P^{\pi} + z\cdot p\cdot(d_U^{\pi}-d_P^{\pi}) - z\cdot(1-p)\pi\cdot d_P^{\pi} =$$
$$= (p+(1-p)\pi)\cdot d_P^{\pi} + z\cdot\frac{(1-p)\pi\cdot d_P^{\pi}}{q} - z\cdot(1-p)\pi\cdot d_P^{\pi} = d_P^{\pi}$$

where the first inequality follows from equation (8). Thus, (12) is always satisfied as $d_P^{\pi} < d_R$.

Moreover the gain/loss utility in a fully revealing equilibrium will be higher (namely, lower in absolute value) than in a π-symmetric partially revealing one if and only if:

$$2\left(\frac{1}{2}+\psi d_R\right)\cdot\left(\frac{1}{2}-\psi d_R\right) < \left(1+2\psi(1-q)\left(zpd_U^{\pi} + (1-z)\left(pd_P^{\pi} + (1-p)\pi d_P^{\pi}\right)\right)\right)\cdot$$
$$\cdot\left(1-2\psi q\left(zpd_U^{\pi} + (1-z)\left(pd_P^{\pi} + (1-p)\pi d_P^{\pi}\right)\right)\right)$$

Using equation (8), the right-hand side simplifies to $\left(1+2\psi(1-q)d_P^{\pi}\right)\cdot\left(1-2\psi q d_P^{\pi}\right)$. Thus, the previous inequality can be written as:

$$(1+2\psi d_R)\cdot(1-2\psi d_R) < 2\left(1+2\psi(1-q)d_P^{\pi}\right)\cdot\left(1-2\psi q d_P^{\pi}\right)$$

Since the right-hand side is decreasing in q, a sufficient condition for the previous inequality is that:

$$2\cdot\left(1-\psi\frac{G(2+\eta\lambda\pi+\eta(2-\pi))}{1-G(\lambda-1)(p+(1-p)\pi)\pi\psi\eta}\right) >$$
$$> \left(1+\psi\frac{G(2+\eta+\lambda\eta)}{(1-G(\lambda-1)\psi\eta)}\right)\cdot\left(1-\psi\frac{G(2+\eta+\lambda\eta)}{(1-G(\lambda-1)\psi\eta)}\right)$$

The statement of the proposition follows noticing that $2 > \left(1+\psi\frac{G(2+\eta+\lambda\eta)}{(1-G(\lambda-1)\psi\eta)}\right)$ and that

$\frac{G(2+\eta\lambda\pi+\eta(2-\pi))}{1-G(\lambda-1)(p+(1-p)\pi)\pi\psi\eta}$ is increasing in π and is equal to $\frac{G(2+\eta+\lambda\eta)}{(1-G(\lambda-1)\psi\eta)}$ when $\pi = 1$.

As far as high-valence candidates are concerned, one can easily verify that the expected utility of a high-valence candidate in a π-symmetric partially revealing equilibrium is given by:

$$\begin{aligned} U^A\left(m^*, \sigma_P^B, \gamma_P \mid \theta_H\right) &= \frac{1}{2} + (1-q)\cdot\psi\cdot z\cdot p\cdot d_U\left(\eta,\lambda,G,\pi,p\mid\psi\right) + \\ &\quad + (1-q)\cdot\psi\cdot(1-z)\cdot\left((1-p)\,\pi + p\right)\cdot d_P\left(\eta,\lambda,G,\pi,p\mid\psi\right) \end{aligned}$$

which, using equation (8), can be rewritten as:[44]

$$U^A\left(m^*, \sigma_P^B, \gamma_P \mid \theta_H\right) = \frac{1}{2} + (1-q)\cdot\psi\cdot d_P\left(\eta,\lambda,G,\pi,p\mid\psi\right).$$

Since $d_P\left(\eta,\lambda,G,\pi,p\mid\psi\right)$ is increasing in π and equals $d_R\left(\eta,\lambda,G\mid\psi\right)$ when $\pi = 1$, we can further conclude that $U^A\left(m_H, s_R^B, \gamma_R \mid \theta_H\right) > U^A\left(m^*, \sigma_P^B, \gamma_P \mid \theta_H\right)$.

References

Banks, J. S. (1990). A model of electoral competition with incomplete information. *Journal of Economic Theory 50*(2), 309–325.

Banks, J. S. and R. K. Sundaram (1998). Optimal retention in agency problems. *Journal of Economic Theory 82*(2), 293–323.

Battigalli, P., G. Charness, and M. Dufwenberg (2013). Deception: The role of guilt. *Journal of Economic Behavior & Organization, Forthcoming 93*, 227–232.

Battigalli, P. and M. Dufwenberg (2007). Guilt in games. *American Economic Review 97*(2), 170–176.

Battigalli, P. and M. Dufwenberg (2009). Dynamic psychological games. *Journal of Economic Theory 144*(1), 1–35.

Berganza, J. C. (2000). Two roles for elections: Disciplining the incumbent and selecting a competent candidate. *Public Choice 105*, 165–194.

Blume, A. and O. Board (2009). Intentional vagueness. *Erkenntnis*, 1–45.

Caplin, A. and J. Leahy (2001). Psychological expected utility theory and anticipatory feelings. *The Quarterly Journal of Economics 116*(1), 55–79.

Charness, G. and M. Dufwenberg (2006). Promises and partnership. *Econometrica 74*(6), 1579–1601.

[44] See the proof of Proposition 7 for the details.

Charness, G. and M. Dufwenberg (2010). Bare promises: An experiment. *Economics Letters 107*(2), 281–283.

Charness, G. and M. Dufwenberg (2011). Participation. *American Economic Review 101*(4), 1211–37.

Chen, Y., N. Kartik, and J. Sobel (2008). Selecting cheap-talk equilibria. *Econometrica 76*(1), 117–136.

Crawford, V. P. and J. Sobel (1982). Strategic information transmission. *Econometrica 50*(6), 1431–51.

Duggan, J. (2000). Repeated elections with asymmetric information. *Economics and Politics 12*(2), 109–135.

Dziuda, W. (2011). Communication with detectable deceit. Mimeo.

Epstein, L. G. (2008). Living with risk. *Review of Economic Studies 75*(4), 1121–1141.

Farejohn, J. (1986). Incumbent performance and electoral control. *Public Choice 50*(1-3), 5–25.

Farrell, J. (1993). Meaning and credibility in cheap-talk games. *Games and Economic Behavior 5*(4), 514–531.

Farrell, J. and M. Rabin (1996). Cheap talk. *Journal of Economic Perspectives 10*(3), 103–18.

Fehr, E., O. Hart, and C. Zehnder (2011). Contracts as reference points–experimental evidence. *American Economic Review 101*(2), 493–525.

Geanakoplos, J., D. Pearce, and E. Stacchetti (1989). Psychological games and sequential rationality. *Games and Economic Behavior 1*(1), 60 – 79.

Gill, D. and V. Prowse (2012). A structural analysis of disappointment aversion in a real effort competition. *American Economic Review 102*(1), 469–503.

Gill, D. and R. Stone (2010). Fairness and desert in tournaments. *Games and Economic Behavior 69*(2), 346–364.

G.R. Boynton, R. D. H. and S. C. Patterson (1969). Perceptions and expectations of the legislature and support for it. *American Journal of Sociology 75*(1), 62–76.

Green, J. R. and N. L. Stokey (2007). A two-person game of information transmission. *Journal of Economic Theory 135*(1), 90–104.

Grillo, E. (2012). Reference dependence and politicians' credibility. Mimeo.

James, O. (2009). Evaluating the expectations disconfirmation and expectations anchoring approaches to citizen satisfaction with local public services. *Journal of Public Administration Research and Theory 19*(1), 107–123.

Kahneman, D., J. L. Knetsch, and R. H. Thaler (1990). Experimental tests of the endowment effect and the coase theorem. *Journal of Political Economy 98*(6), 1325–48.

Kahneman, D., J. L. Knetsch, and R. H. Thaler (1991). The endowment effect, loss aversion, and status quo bias: Anomalies. *Journal of Economic Perspectives 5*(1), 193–206.

Kahneman, D. and A. Tversky (1979). Prospect theory: An analysis of decision under risk. *Econometrica 47*(2), 263–91.

Kahneman, D. and A. Tversky (1991). Loss aversion in riskless choice: A reference-dependent model. *The Quarterly Journal of Economics 106*(4), 1039–61.

Kőszegi, B. (2006). Emotional agency. *The Quarterly Journal of Economics 121*(1), 121–156.

Kőszegi, B. and M. Rabin (2006). A model of reference-dependent preferences. *The Quarterly Journal of Economics 121*(4), 1133–1165.

Kőszegi, B. and M. Rabin (2007). Reference-dependent risk attitudes. *American Economic Review 97*(4), 1047–1073.

Kőszegi, B. and M. Rabin (2009). Reference-dependent consumption plans. *American Economic Review 99*(3), 909–36.

Kimball, D. C. and S. C. Patterson (1997). Living up to expectations: Public attitudes toward congress. *Journal of Politics 59*(3), 701–728.

Krishna, V. and J. Morgan (2008). Cheap talk. In S. N. Durlauf and L. E. Blume (Eds.), *The New Palgrave Dictionary of Economics*. Basingstoke: Palgrave Macmillan.

Lindstadt, R. and J. K. Staton (2010). Managing expectations: When does it work and what are the consequences? Mimeo.

Loewenstein, G. (1987). Anticipation and the valuation of delayed consumption. *Economic Journal 97*(387), 666–84.

Loewenstein, G. and D. Prelec (1992). Anomalies in intertemporal choice: Evidence and an interpretation. *The Quarterly Journal of Economics 107*(2), pp. 573–597.

Matthews, S. A., M. Okuno-Fujiwara, and A. Postlewaite (1991). Refining cheap-talk equilibria. *Journal of Economic Theory 55*(2), 247–273.

Ottaviani, M. and P. N. Sorensen (2006). Reputational cheap talk. *The RAND Journal of Economics 37*(1), 155–175.

Passarelli, F. and G. Tabellini (2013). Emotions and political unrest. IGIER Working Paper n. 474.

Rabin, M. (1994). A model of pre-game communication. *Journal of Economic Theory 63*(2), 370–391.

Schwabe, R. (2011). Reputation and accountability in repeated elections. Mimeo.

Seidmann, D. J. and E. Winter (1997). Strategic information transmission with verifiable messages. *Econometrica 65*(1), pp. 163–169.

Shalev, J. (2000). Loss aversion equilibrium. *International Journal of Game Theory 29*(2), 269–287.

Sigelman, L. and K. Knight (1983). Why does presidential popularity decline? a test of the expectation/disillusion theory. *Public Opinion Quarterly 47*(3), 310–324.

Sugden, R. (2003). Reference-dependent subjective expected utility. *Journal of Economic Theory 111*(2), 172–191.

van Dijk, E. and D. van Knippenberg (1996). Buying and selling exchange goods: Loss aversion and the endowment effect. *Journal of Economic Psychology 17*(4), 517 – 524.

Waterman, R., H. C. Jenkins-Smith, and C. L. Silva (1999). The expectations gap thesis: Public attitudes toward an incumbent president. *Journal of Politics 61*(4), 944–966.