# A model of Punitive Voting *

Manuel Lleonart-Anguix†

October 1, 2024

**Abstract**

I analyze a two-period model of political competition where voters care about candidates' integrity. Candidates must trade off implementing their preferred policy against maintaining their electoral promises. Voters punish candidates that deviate from electoral promises by voting for their opponent. I find that punitive voting can exert political discipline only if candidates face low levels uncertainty about voters preferences. In this case candidates' electoral promises are a compromise between their preferred policy and voters' preferences, and when elected they implement their promise. Finally, I show that when one candidate's ideal policy is closer to the median voter, an equilibrium exists where one candidate is disciplined and the other is not.

**Keywords:** Reputation, elections, political promises.

*I thank Enriqueta Aragonés and Pau Milán for their guidance during the redaction of this paper. I also thank Esther Hauk, Malachy James, Gilat Levy, Raquel Lorenzo, Antonin Macé, Jordi Massó, Amedeo Piolatto, Jesús Sánchez, and Dimitrios Xefteris for their comments, and the participants at the BSE Jamboree, UAB MicroLab, and Red Española de Elección Social (REES) meeting.

†Universitat Autònoma de Barcelona. Contact: manuel.lleonart@bse.eu

# 1 Introduction

> *Our country is being run by incompetent people. And I won't be angry when we fix it, but until we fix it, I'm very, very angry.* [1]
> **Donald J. Trump**

Voters often cast their ballots out of frustration and disillusionment. Wuthnow (2018) documented that the political discontent of rural voters in America significantly fueled populist movements, including the support for Donald Trump in the 2016 election. He argues that many rural voters turned to Trump not necessarily because they endorsed his policies, but because he resonated with their anger and resentment towards a political establishment they felt had abandoned them. Born, van Eck, and Johannesson (2018) analyzed the impact of electoral promises on voter behavior in a controlled setting. They show that when politicians break their promises, voters retaliate by using their votes to punish them. Similarly Galeotti and Zizzo (2015) experimentally measures voters' reactions when candidates face a trade-off between competence and honesty, observing that voters tend to prioritize honesty, even when it leads to lower payoffs. These experiments suggests a strong voter bias towards integrity, reflecting deeper emotional and ethical concerns in electoral behavior. This pattern aligns with broader trends. As (Tambe, 2018) shows, in countries where voters lack trust in elections or party responses, voter turnout and participation tend to be low.

Voters are never pivotal in determining election outcomes, yet they are often motivated by underlying behavioral factors such as civic duty and psychological satisfaction, among others (Blais, 2000). However, one key incentive that has been largely overlooked is anger. While ubiquitous in media portrayals, social scientists have often ignored the impact of voter disenchantment and how perceived dishonesty can "flip" voters' decisions.

In this paper, I propose a new theory of electoral competition where voters' concern about dishonesty plays a central role, leading them to punish candidates who cannot be trusted. The model features two candidates, $L$ and $R$, each with distinct policy preferences, and a voter positioned between them. Candidates are unsure about the exact position of the voter, but they share a common distribution over it. In the first period, candidates make campaign promises, and the voter chooses her preferred option. Once elected, a candidate may implement a policy that deviates from her promise, leading the voter to downgrade their perceived integrity. Although the voter has single-peaked preferences over policies, her preference for a candidate is influenced by the integrity associated with that candidate's promises. Consequently, a dishonest candidate who pledges to enact the voter's ideal policy may be less appealing than an honest candidate who promises a less ideal policy. The

[1] As seen in (Boylan, 2016).

first challenge candidates face is deciding on a campaign promise that strikes a balance between aligning closely with the voter's preferred policy—necessary to win elections—and staying true to their preferred policy, allowing them to act with integrity if elected. Once in office, incumbents must face the tension between the short-term temptation to pursue self-interest, which could jeopardize their future electoral success, and the long-term benefits of maintaining honesty to maximize their prospects in future elections. In the second period, another election takes place, during which candidates once again choose their promises and policies, and the voter updates his evaluation of each candidate's integrity. Payoffs are determined by the policies enacted. In an extension of the model, I assume that the voter's expected ideal policy is not equidistant between the two candidates' bliss points, but instead closer to one of them. This proximity influences the more distant candidate's likelihood of reneging on their promise once elected.

This paper demonstrates that concerns about punitive voting play a crucial role in shaping electoral dynamics and policy implementation. The main finding reveals a non-monotonic relationship between electoral promises and uncertainty regarding voter preferences in equilibrium. When there is an intermediate level of uncertainty about voter preferences and the voter is sufficiently sensitive to integrity, the threat of punitive voting serves as a disciplining mechanism. In this scenario, candidates are incentivized to craft promises that balance the voter's ideal policy with their own preferences. Given their uncertainty about how voters will respond in equilibrium, candidates may be willing to risk a small chance of losing in exchange for making promises closer to their preferred policy. If these promises are close enough to their preferences—and assuming that candidates value future elections and voters prioritize integrity—the winning candidate in the first period will fulfill their promise. This leads to a second period where both candidates are perceived as having similar levels of integrity.

However, sustaining political integrity is challenging and only likely to emerge under the specific conditions outlined above. When these conditions are not met, candidates often fail to implement their campaign promises, with their decisions being heavily influenced by the level of uncertainty regarding voter preferences. When uncertainty is low, candidates understand the high risk of deviating from the voter's ideal policy, which significantly increases their chances of losing the election, leading them to converge on similar promises. Yet, these promises are typically far from the candidates' true preferences, prompting the eventual winner to act dishonestly after being elected, sacrificing future payoffs. Conversely, when uncertainty is high, dishonesty is encouraged in a different way. Due to the lack of precise information about the voter's ideal point, the perceived importance of integrity in the second stage is overshadowed by the uncertainty surrounding the preferred policy, reducing the emphasis candidates place on honesty. As a result, acting dishonestly in the first stage does not significantly harm a candidate's prospects in the final stage. Despite this high uncertainty, candidates are still motivated to make promises that moderately align with both the voter's and their ideal policies to minimize the cost of dishonesty in the second period.

**Literature Review** This paper contributes to the extensive literature studying electoral competition via promises. The first economic model of the political parties' behavior is presented in Downs (1957), where office-motivated candidates compete as rational agents to win election. In this paper, parties have an intrinsic value of gaining office and are not motivated by policies *per se*. On it, governments try to maximize the number of votes to stay in power as in Hotelling (1929) with consumers. This analysis misses the ideological motivation of parties. Barro (1973) presents a model of repeated election to study the behavior of elected public office-holders when they care both about reelection and policies. He states that candidates act selfishly choosing their preferred policy if there is no political control by citizens and apply the promised policies when the public control of candidates is present. As in my paper, in Barro (1973) voters are backwards-looking[2] and lying is a factor that *discards* political representatives from being reelected when they act in their own interest. The two previous models predict a fight for the median voter that might fail to explain some of the electoral promise location that is observed in reality. Wittman (1983) and Calvert (1985) introduce candidates that care both about the probability of winning and the policy that is applied. A similar result appears in this paper, where the joint role of integrity and uncertainty is enough to create polarization in the electoral promises. In this line, Alesina (1988) develops a model of forward-looking voters and ideologically motivated candidates without commitment. The absence of commitment creates an equilibrium of the one-shot game where each candidate promises her ideal policy due to a problem of credibility. Convergence only appears when the game is infinitely repeated. My setup allows for both types of equilibria depending on the importance of the reputation and the level of uncertainty within the model. Banks (1990) develops a model where candidates have ideal points, announce political positions and then, the winner needs to apply a policy. His model considers a continuous cost of deviating from the promises as I do in this paper. The main difference is that, in my model, that cost is introduced by the probability of reelection. Aragonés, Palfrey, and Postlewaite (2007) study how reputation is used by voters to form beliefs about the credibility of candidates' campaign promises. They show that there is an equilibrium where candidates announce policies different from their ideal point and keep their promises. The voter I present in this paper is, instead backward-looking. In my model, candidates incur reputation costs when they apply a policy that differs from their promises. This cost harms the predisposition of citizens to vote for the candidates that they consider reliable as a sort of punishment.[3]

[2] See also Austen-Smith and Banks (1989).

[3] For a literature review on retrospective voting see Healy and Malhotra (2013).

Weelden (2013) analyzes the trade-off between applying rent-seeking policies or increasing the probabilities of reelection with an infinite number of candidates. He extends the analysis of credibility of campaign promises when there is no full commitment. In the same line, Bischoff and Siemers (2013) explain how retrospective voting and biased beliefs affect the policy outcomes and the probabilities of reelection. List and Sturm (2006) analyze the use of policy promises in campaign in a two-dimension policy space. They find that there exist strong effects to approach to the voter in secondary policy issues. The relation between the promises and the policy intentions of candidates is also studied in Schnakenberg (2016). He proves that cheap talk can also be informative even if there is no commitment. In this line, Kartik and Van Weelden (2018) prove that cheap talk during the elections can affect candidates' behavior when they are in power. My model, is different from the *cheap-talk* model, as there is an endogenous *cost of lying* when candidates promise things that are different from the policies that they apply. Kartik and McAfee (2007) (through character) and Callander and Wilkie (2007) (through the cost of lying) develop models where candidates differ in their willingness to lie in elections and voters interpret promises in a signaling game to disentangle the different types.

Andreottola (2020) studies the incentives of a politician to hide information in order to not be considered as a *flip-flopper*. In his model, the optimal policy might change given new private information. Differently from my paper, he focuses on the information updating of an incumbent politician and how a strong reputation concern affects the application of the optimal policies, while I focus on the campaign polarization and the capacity of the repeated election to force honest behaviors. Rivas (2015) develops a model where a politician has to take several decisions during his term office. In his model, citizens are backwards looking when evaluating the quality of the politician. Under this assumptions, he sees incentives to take selfish decisions first and act honestly later. This differs with my papers because he considers multiple decisions in one electoral period, so there is no threaten of no-reelection.

Differently from previous papers on re-election, I model reputation as an additive term that decreases the utility of the voter to choose certain candidate. This way of modeling is close to the literature of electoral competition and *valence*.[4] Ashworth and Bueno de Mesquita (2009) develop a model where candidates can invest in costly valences and show that valences are more important the less polarized is the society. In their model, a shock in one candidate's valence leads to complete platform convergence, while in mine, reputation can keep polarization. Gouret, Hollard, and Rossignol (2011) analyze in a survey prior to the French presidential election of 2007 several spatial voting models with valence. Taking into account the empirical evidence, they develop a model of intensity valence that diverges

[4]See Stokes (1963), Ansolabehere and Snyder (2000) Groseclose (2001), Aragones and Palfrey (2002), Hummel (2010), Aragonès and Xefteris (2017). Xefteris (2013), Denter (2021), and Buisseret and Van Weelden (2022) among others.

from the additive valence models.

The rest of the paper is organized as follows: Section 2 presents the model, including the players and the timing. Section 3 develops the equilibrium of the model and, discusses the results and their implications. All the proofs are relegated to the Appendix.

# 2 Model

In this section, I present the main features of the model: players, their payoffs, and the dynamics of information in the game.

Two *candidates* compete in an election to decide who will apply a policy. I define a policy as a real number, $x \in \mathbb{R}$. The timing of the election is the following: first, each candidate simultaneously announces a promise $p_i \in \mathbb{R}$ about the policy she will implement. The voter observes these promises and votes for his preferred candidate; the candidate who wins the vote is elected and applies a policy, $\pi \in \mathbb{R}$. Then, the voter observes the policy implemented and computes the cost of lying to the winning candidate. Once this process finishes, it repeats itself. I analyze two periods: $t = 0$ and $t = 1$.

There are three players: two candidates, $L$ and $R$ (she), and a voter (he). I use $i$ to refer to a generic candidate and $v$ to the voter.

**The voter** Assuming the existence of just one voter is a simplification without loss of generality that allows me to focus on the behaviour of candidates. It is equivalent to take any distribution of voters and consider that the winner is elected by majority voting. In this case, the analysis focuses on the actions of the median voter.

The voter will choose the candidate that he thinks is closer to him. For simplicity, I will not consider the possibility of abstention. Therefore, he must vote for one and only one candidate. He has an innate preference over the policy, that I normalize to zero, and it is common knowledge to all players. The voter decides his vote without thinking in future election; he will punish those candidates who lie. I assume this as my main goal is to study how the promises are shaped in the presence of reelection. I want to capture the punishment effect that deviations from the promises made in the campaign provoke; this is why I don't allow the voter to make any inference about the preferred policy of candidates.

**Cost of lying** The *cost of lying* is the key factor in this model. I will refer to it as cost of lying, integrity cost, or reputation cost. It measures the difference between a candidate's promises and what she effectively does. The voter calculates the cost of lying of each candidate at every period, which will affect the chances of election. By assumption, this term at $t = 0$ will be zero for both candidates.

Equation (1) describes the cost of lying of candidate $i$ at the beginning of period 1.

$$\mathfrak{C}_i^1 = \delta |p_i^0 - \pi_i^0| \tag{1}$$

where $\delta \in \mathbb{R}^+$. Equation (1) states that the reputation cost of a candidate consists of the positive difference between the promise made during election and the policy applied after winning. The interpretation of the term $\mathfrak{C}_i^1$ is the following: it increases when a candidate deviates from her promise. $\delta$ captures the effect of the difference between promises and policies on voter's decisions. The term $\mathfrak{C}$ distinguishes this model from a cheap-talk model, as it is and endogenous cost of lying that relates promises and policies through the probability of winning.

Candidates do not know exactly the preferred candidate for the voter given $\mathbf{p}$ (the full vector of promises) and $\boldsymbol{\pi}$ (the full vector of policies); I assume that the voter has a private preference for one of the candidates that affects his willingness to vote for her but does not depend on $\mathbf{p}_i$ or $\boldsymbol{\pi}_i$. This preference is a pair of random variables $A^0$, $A^1$ *iid* that follow a uniform distribution in the interval $[-a, a]$, where $a \geq 0$ and affects the utility for the voter of choosing candidate $R$. The utility of the voter at the period $t$ is given by the next equation.

$$\mathcal{U}_v^t(p_i^t, \mathfrak{C}_i^t) = \begin{cases} -|p_L^t| - \mathfrak{C}_L^t & \text{if he votes for candidate } L \\ -|p_R^t| - \mathfrak{C}_R^t - A^t & \text{if he votes for candidate } R \end{cases} \tag{2}$$

where $p_i^t$ are the promises of candidate $i$ at period $t$ (see below). As (2) shows, the voter receives utility from each candidate considering two factors. The first factor is the distance from the promises of each candidate to her preferred policy, 0. The second is this candidate's cost of lying. An increase in $\mathfrak{C}_i^t$ decreases the utility of choosing candidate $i$ for the voter. Hence, given a pair of promises, it will diminish the candidate's chances of being elected. As there are two candidates, there is always at least one maximizing the utility of voter $v$. The probability of candidate $R$ winning is equivalent to $P\left(A^t \leq -|p_R^t| + |p_L^t| + \mathfrak{C}^t\right)$. Using the fact that $A^t$ is uniformly distributed, the probability of candidate $R$ winning reads as,

$$P(\text{Win R}) = \begin{cases} 0 & \text{if } -a \geq -|p_R^t| + |p_L^t| + \mathfrak{C}^t \\ \frac{1}{2a}\left(-|p_R^t| + |p_L^t| + \mathfrak{C}^t + a\right) & \text{if } -a \leq -|p_R^t| + |p_L^t| + \mathfrak{C}^t \leq a \\ 1 & \text{if } a \leq -|p_R^t| + |p_L^t| + \mathfrak{C}^t \end{cases}$$

where $\mathfrak{C}^t = \mathfrak{C}_L^t - \mathfrak{C}_R^t$. The cost of lying is what makes promises *costly*, as they can endogenously reduce the probability of election of a candidate. Candidates are punished by the voter if they move away from their preferred policy, however, a candidate can win with positive probability even if she does not promise what the voter wants. As a final remark, note that the random variable that affects the utility of the voter is realized twice (once per period). These two variables are *iid*. This allows me to focus the study on the role of uncertainty in the decision-making of politicians and ignore the information updating that would arise assuming some correlation between both realizations of the random variable.

**Candidates** Each candidate has an exogenously given preferred policy,[5] $x_R$ and $x_L \in \mathbb{R}$ for candidates $R$ and $L$, respectively. These points are assumed to be symmetric around zero. Without loss of generality I assume $x_R = 1$ and $x_L = -1$. At the beginning of the period, each candidate publicly announces $p_i^t$, which is a promise about the policy that the candidate will implement if elected. The winner of the election decides $\pi_i^t$. For notation simplicity, I set $\pi_i^t = p_i^t$ for the losing candidate, as she does not face a cost of lying. I say that a candidate is dishonest whenever she chooses $p_i^t \neq \pi_i^t$.

To simplify notation, I will use trough the paper $\mathfrak{C}^t$ to refer to the integrity cost for both candidates, $\mathbf{p^t}$ for the promises, and $\boldsymbol{\pi^t}$ for the policies at time $t$. I will remove the superscript $t$ when it refers to the complete history.

**Flow utility** Candidates act as rational agents, maximizing their utility. The utility function of candidate $i$ for period $t$ is as follows:

$$u_i^t(\pi_i^t; \pi_{-i}^t, \mathbf{p^t}, \mathfrak{R}^\mathbf{t}) = \begin{cases} -|x_i - \pi_i^t| & \text{if } i \text{ wins} \\ -|x_i - \pi_{-i}^t| & \text{if } i \text{ loses} \end{cases} \tag{3}$$

Where $\pi_{-i}^t$ is the policy applied by the other candidate when she wins. The utility that candidates perceive from winning election comes from the opportunity to implement

[5]The ideal point of each candidate can be also interpreted not as a preferred policy, but as the policy that they think is ideal to everyone.

a policy. For this reason, despite the candidates are not office motivated, they seek to win. With this, candidates maximize their total utility for the game.

$$U_i(\boldsymbol{p}, \boldsymbol{\pi}) = \sum_{t=0}^{1} \gamma^t u_i^t(\pi_i^t; x_i, \pi_{-i}^t) \tag{4}$$

for $0 < \gamma < 1$ a time discount factor. The utility of the game for candidates is the discounted sum of utilities for each period.

To sum up, candidates must choose every period a promise, $p_i^t$, and a policy to implement, $\pi_i^t$. Hence, the strategy of the candidate $i$ at time $t$ is a pair $\{p_i^t, \pi_i^t\}$. Consequently, the candidate's strategy for the game will consist of a sequence of promises and actions $\{p_i^t, \pi_i^t\}_{t=0}^1$. The structure of one period is as follows:

1. First, on *Campaign*, candidates publicly announce their policies, interpreted as intentions about policy implementations, $p_i$.
2. Second, in the *Voting stage*, the voter decides which candidate gives him a higher utility and votes.
3. Finally, in the *Office stage,* the elected candidate chooses a policy $\pi_i$. The voter observes it and update candidate's cost of lying.

# 3 Analysis

In this section, I develop all the necessary steps to solve the model and present the equilibrium to the game. The equilibrium is a sub-game perfect Nash equilibrium for the candidates' strategies. It is a pair of promises and policies for each individual and each period:

$$\{p_R^0, p_R^1, \pi_R^0, \pi_R^1\}, \ \{p_L^0, p_L^1, \pi_L^0, \pi_L^1\}$$

I find the equilibria of the game by backward induction. The next set of results characterizes the equilibrium strategies in the last period, $t = 1$. Later, I will solve the first period. All proofs can be found in Appendix A.

## 3.1 Last period

In the last period, each candidate chooses a promise and a policy. This decision determines candidates' integrity cost, $\mathfrak{C}_i^1$. For this reason, the decision on $p_i^1$ and $\pi_i^1$ are not motivated by their effect on future lying costs as will happen in the first period. Instead, when candidates decide on a promise and a policy to implement, they only care about the impact on that period's utility. Proposition 3.1 states the policy that candidates apply in equilibrium using this reasoning. As there is no mechanism that the voter can use to punish her, the candidate that wins election in the last period maximizes the flow utility of $t = 1$.

**Proposition 3.1** (Optimal policies for the last period)**.** *In the last period, the candidate that wins the election best replies with her preferred policy in equilibrium:*

$$\pi_i^1 = x_i$$

The best reply of the previous proposition holds for every promise, $p_i^1$, and integrity cost, $\mathfrak{C}_i^1$, given that she has won the election. As candidate $i$ is the winner, there is no possible punishment from the voter. Therefore, she chooses in equilibrium the policy that maximizes his utility. Once a candidate is elected to choose the policy, the probability of winning election does not play a role in her decisions. As the policy in the last stage does not depend on the promise that made the candidate win, she will apply her preferred policy. This is the only result independent on the variable $A^t$.

Using the previous result for the optimal policies for the last period, I can state the equilibrium payoffs for each candidate conditional on winning and losing. If candidate $i$ wins the final period's election, she will apply her preferred policy and get a utility given by:

$$u_i(\pi_i^1; p_i^1, p_{-i}^1, \pi_{-i}^1) = -|x_i - \pi_i^1| = -|x_i - x_i| = 0 \tag{5}$$

If instead, candidate $i$ loses in the last period, the other candidate will apply her preferred policy, and then the utility for losing is

$$u_i(\pi_i^1; p_i^1, p_{-i}^1, \pi_{-i}^1) = -|x_i - \pi_{-i}^1| = -|x_i - x_{-i}| = -2 \tag{6}$$

This result allows me to find the promises in equilibrium, as the optimal policies determine the flow utilities of both losing and winning. If candidates arrive in period one with the same $\mathfrak{C}_i$, they must choose the voter's preferred policy in order to win, as the voter can only differentiate between them by their promises. This is not the case if they have different costs of lying. If, for example, candidate $L$ arrives at the last period with big lying costs, her opponent has a *relative advantage* that she can exploit to win that period's election. In this case, candidate $L$ can not do anything to win, as candidate $R$ will always win in equilibrium promising the voter's ideal policy. As only one candidate wins the election at each period and both begin the game with zero integrity cost, in equilibrium, there is only one case in which both candidates have the same reputation in the last period. This is if $\mathfrak{C}_L^1 = \mathfrak{C}_R^1 = 0$. Proposition 3.2 states the optimal policy promises for both candidates in the last period of the game depending on the integrity costs.

**Proposition 3.2** (Optimal promise in the last period)**.** *If $\mathfrak{C}_L^1 \geq a$ there exist multiple equilibria of the form*

$$p_R^1 \in [a - \mathfrak{C}_L^1, a + \mathfrak{C}_L^1]$$

*If $a > \mathfrak{C}_L^1 \geq 0$, $\mathfrak{C}_R^1 = 0$, the unique equilibrium is*

$$p_R = p_L = 0$$

The promises made in equilibrium in the last period depend on the difference in lying costs between the two candidates. If this difference is big enough, $\mathfrak{C}_L^1 \geq a$, the candidate with lower costs will have a range of optimal promises that makes her win with certainty. If not, both candidates fight for the median voter. With a positive, but small enough, integrity cost both candidates have a positive probability of winning in equilibrium and they still promise the ideal point of the voter. This enlarges the strategies that might be optimal in the first period with respect a case where there is full information, as incurring in integrity costs is not a disqualifier. Next section characterizes the results for the first period of the game and states the equilibria of the game.

## 3.2 First period

Given the results of the last period of the game, I can calculate the equilibrium of the game. The structure of this section is similar to the previous one. First, I will characterize the optimal policies for any possible promise. Second, I will use this result to find the promises of both candidates in equilibrium. One particularity of the first period is that, by assumption, both candidates are symmetric for the voter's ideal point; they have a zero integrity cost, and their preferred policy is symmetric around the voter.

The candidate that wins the election must choose which policy to apply. This decision can make her keep her initial integrity or incur in a cost by lying. This decision is one of the keys to the model. The policy that the candidate applies determines her level of integrity for the next period and consequently her chances to win. At the same time, the policies that the candidate applies in the first period depend on the promise that made her win. If this promise is close enough to the ideal point of the candidate, she will keep her promise and she will choose her ideal point otherwise. However, there is an extra condition for both candidates to keep their promise as proposition 3.3 shows; the level of uncertainty about voter's preferences, $a$, must be smaller than the future discounted salience of the integrity cost, $\gamma\delta$.[6]

**Proposition 3.3** (Optimal policies in the first period)**.** *If candidate $R$ wins at $t = 0$ with $p^0_R$, she chooses, in equilibrium,*

$$\pi^0_R = \begin{cases} p^0_R & \textit{if } a < \gamma\delta \textit{ and } 1+\gamma \geq p^0_R \geq 1-\gamma \\ 1 & \textit{Otherwise} \end{cases}$$

This proposition splits the model into two opposite cases depending on whether the level of uncertainty is smaller than the future salience of integrity, $a < \gamma\delta$, or not. The optimal policy for candidate $L$ is symmetric to proposition 3.3. The optimal policies in the first period are only the ideal point or keeping the promise. As the utility function is linear, there is no possibility of finding an optimal promise in the middle between these two options. Assuming a quadratic utility for the candidates, for the voter, or for both can make different equilibria appear where $\pi^0_R \in (p^0_R, 1)$.

Before stating the proposition that characterizes the optimal promises in equilibrium for the first period, it is important to note that, at the beginning of the game, the expected utility of candidates can have three different shapes. These shapes depend on the relation between the parameters of the model and the promise $p^0_R$. If the candidate is honest in the first period, $\pi^0_R = p^0_R$, the expected utility of candidate $R$ reads as

$$\mathbb{E}\left[U_R(\mathbf{p}, \boldsymbol{\pi})\right] = -\frac{1}{2a}\left(-|p^0_R| + |p^0_L| + a\right)\left[|1-p^0_R| + \gamma\right] - \frac{1}{2a}\left(a + |p^0_R| - |p^0_L|\right)\left[|1-p^0_L| + \gamma\right]$$

where $\mathbf{p}$ is the vector of promises and $\boldsymbol{\pi}$ the vector of policies. The first term corresponds

[6] As both $A^t$ follow a uniform distribution, $a$ is a measure of the variance of $A^t$.

to the probability of winning in the first period multiplied by the utility of winning and applying the promise $-|1-p_R^0|$ plus the expected utility of the last period, that is $-2$ with probability one half. As both candidates have the same level of integrity in the last stage, both have the same chances of winning in equilibrium. The second term is the probability of losing times the utility of candidate $L$ applying her policy.

If the candidate applies a different policy to her promise, the utility function for candidate $R$ is

$$\mathbb{E}\left[U_R(\mathbf{p},\boldsymbol{\pi})\right]=\begin{cases}-\frac{1}{2a}\left(-|p_R^0|+|p_L^0|+a\right)\left[\frac{\gamma}{a}\left(\delta|1-p_R^0|+a\right)\right] & \\ -\frac{1}{2a}\left(a+|p_R^0|-|p_L^0|\right)\left[2+\frac{\gamma}{a}\left(-\delta|1-p_L^0|+a\right)\right] & \text{if } \delta|1-p_R^0|<a \\ -\frac{1}{2a}\left(-|p_R^0|+|p_L^0|+a\right)2\gamma & \\ -\frac{1}{2a}\left(a+|p_R^0|-|p_L^0|\right)2 & \text{if } \delta|1-p_R^0|\geq a\end{cases}$$

Notice two things. First, the fact that the separation between both expressions depend on the relation between $\delta|1-p_R^0|$ and $a$ is because of the optimal promises in the last stage. Second, this is only true in the case of symmetry.

The first term of both expressions correspond with the probability of winning multiplied by the utility of winning. The first period utility is zero, as candidate applies her ideal point. The only difference is the probability in equilibrium of winning in the second stage. If $\delta|1-p_R^0|<a$ this probability is positive. In the opposite case, this probability is zero and the candidate loses with certainty. Notice that the expected utility is a continuous function in the three cases. If $\delta|1-p_R^0|=a$, the two previous expressions are equal. The same holds for $p_R^0=\pi_R^0$. The next proposition characterizes the optimal promises in the first period as a function of the relation between $a$ and $\gamma\delta$.

**Proposition 3.4** (Optimal promises in the first period). *Assume $a<\gamma\delta$.*

- *If $a\geq 2(1-\gamma)$ candidates choose in equilibrium*

$$p_R^0=\min\left\{1,\frac{a}{2}\right\},\quad p_L^0=\max\left\{-1,-\frac{a}{2}\right\}$$

- *If $a<2(1-\gamma)$ candidates choose in equilibrium*

$$p_R^0=p_L^0=0$$

*Assume $a>\gamma\delta$.*

*If* $a \leq \frac{2\gamma\delta}{2-\gamma\delta}$ *and* $2(1-\gamma) < \gamma\delta$*,*

$$p_R^0 = 1 + \frac{a}{2} - \frac{a}{\gamma\delta},\ p_L^0 = -1 - \frac{a}{2} + \frac{a}{\gamma\delta}$$

*Otherwise,*

$$p_R^0 = p_L^0 = 0$$

In total, three different equilibria exist for the sub-game of the promises. The first arises if the level of uncertainty, $a$ is between $\gamma\delta$ and $2(1-\gamma)$. In this equilibrium candidates' promise is a function of the level of uncertainty and it goes further from the ideal policy of the voter if the uncertainty increases. Notice that, when the importance of the future enlarges, the interval $[2(1-\gamma), \gamma\delta]$ also does. The appearance of this equilibrium implies that $\gamma$ is relatively high, at least higher than $\frac{2}{3}$, as if not the interval is empty for any $\delta$. In the case where $\delta \geq \frac{a}{\gamma}$, candidates will promise their bliss point. This happens because of two forces: the cost of lying is high, so candidates want to be honest, but also the level of uncertainty about the preferences of the voter is high enough with respect to the future discount factor. So candidates put a relatively high value on both elections and they have little information about the preferences of the voter.

The second equilibrium for the promises require a high level of uncertainty, as it only appears when $a > \gamma\delta > \max\left\{\frac{2a}{2+a}, 2(1-\gamma)\right\}$. Candidates' promises are weekly between the ideal policy of the voter and their preferred policy. In this case, the level of uncertainty has the inverse effect than in the previous equilibrium. Here, when the uncertainty increases, candidates converge to the ideal policy of the voter. If the future discounted salience of the level of integrity and the level of uncertainty are close, this equilibrium converges to the previous one.

The third appears in any other case. This equilibria were both candidates promise the voter's ideal policy might happen for two reasons. First, because the future is important ($\gamma$ high) and the level of uncertainty about the preferences of the voter, $a$, is small. Second, because the future salience of the integrity is small relative to the rest of the parameters.

The next theorem presents all three different types of equilibria that exist in this model. Notice that equilibria in the second stage depend on the relation between $\mathfrak{C}_i^1$ and $a$, as stated at the beginning of this section. If the cost of lying of one candidate exceeds the value of the uncertainty, the other candidate has a margin of movements on her promises around the ideal policy of the median voter that allows him to win with certainty in equilibrium.

In the opposite case, both candidates will promise the voter's ideal policy.

**Theorem 3.1** (Equilibria of the game)**.** *This game has three different equilibria depending on the parameters.*

1. *Honest equilibrium. If* $2(1-\gamma) < a < \gamma\delta$*, in the first stage both candidates promise* $p_R^0 = p_R^0 = \min\left\{1, \frac{a}{2}\right\}$*, and* $p_L^0 = p_L^0 = \max\left\{-1, -\frac{a}{2}\right\}$*. The winner applies her promise,* $\pi_i^0 = p_i^0$ *and keeps her integrity. In the second stage, both candidates promise the voter's ideal policy and, if they win, apply her preferred policy.*
2. *High uncertainty equilibrium. If* $\frac{2\gamma\delta}{2-\gamma\delta} > a > \gamma\delta > 2(1-\gamma)$*, in the first stage both candidates promise* $p_R^0 = 1 + \frac{a}{2} - \frac{a}{\gamma\delta}$ *and* $p_L^0 = -1 - \frac{a}{2} + \frac{a}{\gamma\delta}$*. The winner implements her ideal policy. On the second stage, both candidates promise the median's voter ideal policy and the winner applies her ideal policy.*
3. *Complete information equilibrium. In any other case, in the first stage, both candidate promise voter's ideal policy,* $p_R^0 = p_L^0 = 0$*, and the winner implements her ideal policy. If* $\delta > a$*, the candidate with zero lying cost makes a promise in the interval* $[a-\delta, \delta-a]$*, wins with probability one and applies her ideal policy. If, instead,* $\delta \leq a$*, both candidates choose the ideal policy of the median voter and the winner applies her ideal policy.*

See proof on page 29.

To understand this theorem, I will consider the different cases based on the values of $\gamma$ and $\delta$. When $2(1-\gamma) > \gamma\delta$, the optimal promises for both candidates in the first stage are the voter's ideal policy independently on $a$, the level of uncertainty. Notice that, if $\gamma < \frac{2}{3}$, this will always be the case. However, candidates implement their ideal policy when they win the election. The integrity effects appear also in the second stage. After the first election, either one candidate has zero probabilities of winning (when $\delta > a$) or both candidates still have positive probability of winning (when $\delta \leq a$).

If the candidates still have a positive probability of winning in the second stage, then their optimal strategies are to promise the voter's ideal policy. In this case, the level of uncertainty about the decision of the median voter is high relative to the importance of integrity. In other words, the exogenous stochastic part of the voter's utility is high relative to the endogenous one. Any deviation from the voter's ideal policy is small relative to the unknown preferences of the median voter. Therefore, both candidates still want to fight for the voter in equilibrium.

If the lying cost of one candidate is high enough and her probability of winning for one candidate is zero in the second stage, the equilibrium is equivalent to the case where candidates have full information on voter's preferences. I call it the *complete information equilibrium*. The candidate that loses in the first election can choose a range of promises around the voter's ideal policy to win the second stage election with certainty. I can assume an infinitesimally small *legacy effect* that pushes candidates towards being honest even when the game is over. This effect would make the right candidate choose the promise $p_R^1 = \delta - a$ and the left candidate $p_L^1 = a - \delta$. Then, incurring in integrity costs in the first stage makes the advantaged candidate polarize in the second stage. Given his advantage, she can win with certainty, choosing a more polarized position. In this case, promises are a mechanism that reveals the policy that the candidate will apply in the second stage.

Let's analyze now the two types of equilibria that arise when $2(1-\gamma) < \gamma\delta$. If the level of uncertainty is in between those two terms, candidates are willing to polarize in their promises and move from the voter's ideal policy. As the future weight of the integrity is high concerning the uncertainty, they prefer to risk some of their probability of winning in exchange for a promise that will implement if elected. The effect of integrity also appears in the policies applied by the winner in equilibrium. Only in this equilibrium do candidates implement their promise when winning. This is due to two reasons: if the uncertainty about the utility of the median voter is low, candidates know that incurring high lying costs makes them lose in the second stage with certainty. Therefore, under low levels of variance, candidates prefer to apply not their ideal policy, but the promise that made them win. This effect is insufficient, it is also necessary that the promises are not too far from the ideal policy of the candidate, as otherwise, candidates prefer to sacrifice future utility in exchange for the actual one. Theorem 3.1 ensures that assuming an infinitesimally low level of uncertainty can make candidates go honest when both periods are equally valuable for them, $\gamma = 1$. As they are truthful in their promises, in the last stage, both candidates compete with zero integrity costs.

If the level of uncertainty is high enough, $a > \gamma\delta > 2(1-\gamma)$, but smaller than $\frac{2\gamma\delta}{2-\gamma\delta}$ a different equilibrium appears; the so called *high uncertainty equilibrium*. Promises in the first stage are a decreasing function of the level of uncertainty. Two effects act together on the promises of this equilibrium. On the one hand, candidates do not apply their promises if they win, which creates an integrity cost for the next stage that will decrease the probability of winning. On the other hand, the high level of uncertainty makes candidates unsure about the result of the election and, consequently, about the effect of integrity on the probability of winning. These two effects make candidates move away from the voter's preferred policy; they are willing to risk the probability of winning in the first stage in exchange for incurring fewer lying costs when they deceive and having a higher probability of winning in the second. However, when the level of uncertainty becomes higher than $\frac{2\gamma\delta}{2-\gamma\delta}$, the ignorance about the voter's behavior dominates the effect of the integrity on the second stage, which makes candidates promise the voter's preferred policy.

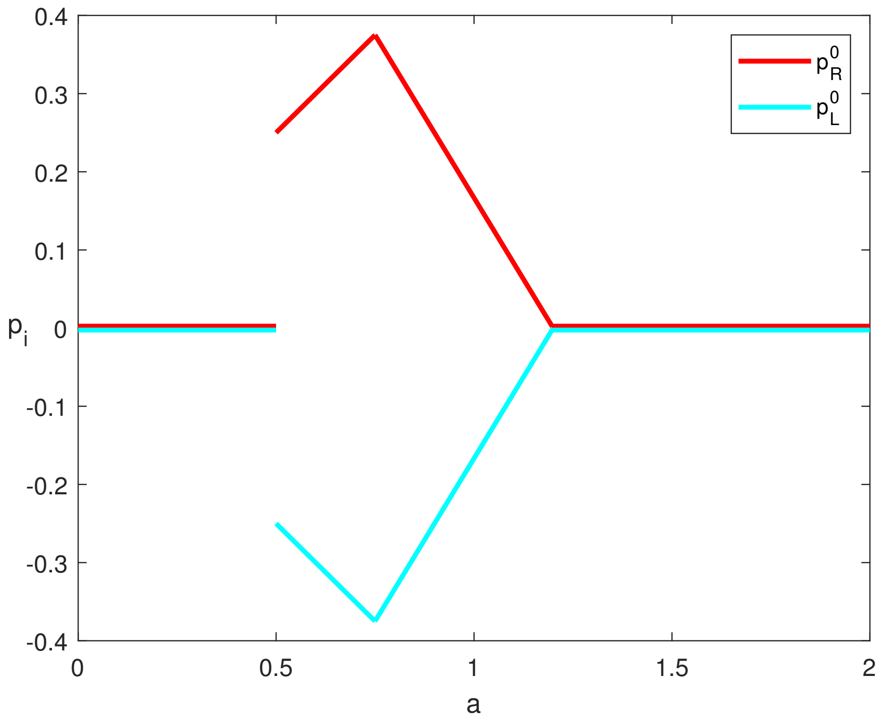


Figure 1: The promises in the first stage for candidate $R$ as a function of $a$. Parameters: $\gamma = 0.75, \delta = 1$

## 3.3 Descriptive statics

Figure 1 shows the evolution of the promises in the first stage as a function of the level of uncertainty. First of all, note that promises are continuous when $\gamma\delta = a$, as

$$\lim_{a \to \gamma\delta} 1 + \frac{a}{2} - \frac{a}{\gamma\delta} = \frac{a}{2}$$

and $\gamma\delta > \frac{2\gamma\delta}{2+\gamma\delta}$. Therefore, for small levels of uncertainty, $a < 2(1-\gamma)$ candidates choose still the ideal policy of the voter as their promise corresponding with the *complete information equilibrium*. However, when the level of uncertainty increases, they move towards more polarized positions. If the level of uncertainty about voter's preferences continues increasing $a > \gamma\delta$, the effect is the opposite, as candidates chose more moderate promises. This effect continues until the level of uncertainty is high enough that candidates do not want to risk probabilities of winning in exchange of less integrity costs and go back to promise the voter's ideal policy. In consequence, being closer to the voter in the first stage (via promises) makes candidates go further in the second stage when they apply their ideal policy (via integrity).

Hence, integrity plays a double role. In the second stage, if candidates have kept a relatively small lying cost, they can still win with some positive probability. Knowing this, candidates do forwards looking reasoning and introduce this effect on their decision-making. Therefore, in the first stage, the integrity effect can polarize candidates in their promises and not choose only the ideal policy for the voter. From the point of view of *honesty*, how promises and policies relate, this polarization can be positive, as it produces a lower deviation between them.

### 3.4 Non-symmetric bliss points

For this section, assume that the ideal policies of both candidates differ. I will study the case where $x_R > 1$ and $x_L = -1$. This assumption means that the right candidate is further from the voter than her opponent. I will prove that there exists an equilibrium where one candidate sticks to her promise while the other prefers to lie. As the strategies for the last period are not affected by the bliss points, all the results of that period will stay the same.[7] However, under this new assumption, the policies stated in proposition 3.3 change in the following way:

$$\pi_R^0 = \begin{cases} p_R^0 & \text{if } a < \gamma\delta \text{ and } p_R^0 \geq x_R - \gamma \\ x_R & \text{Otherwise} \end{cases}$$

$$\pi_L^0 = \begin{cases} p_L^0 & \text{if } a < \gamma\delta \text{ and } p_L^0 \leq -1 + \gamma \\ -1 & \text{Otherwise} \end{cases}$$

On the one hand, when candidate $R$ decides to be dishonest, she chooses $x_R$ instead of one. On the other hand, notice that there is a difference in the boundaries for the parameters that make both candidates honest. This interval is not symmetric anymore, as when $x_R > 1$, it takes promises farther from zero to make candidate $R$ apply her promise in equilibrium. Notice that the analysis is the opposite of $x_R < 1$.

As the arguments of symmetry do not hold anymore, there exists the possibility that, in equilibrium, one candidate decides to be honest and apply her policy while the other shirks. Following the proof of proposition 3.4 I can state that if $a < \gamma\delta$ and $a \geq 2(x_R - \gamma)$ both candidates are honest in equilibrium. However, a sufficient condition for candidate $R$ shirking is $a < 2(x_R - \gamma)$. My goal in this section is to show that for a range of parameters $x_R, \gamma, \delta$, and $a$, there exists an equilibrium where candidate $L$ wants to keep her promise.

[7] With the exception that $\pi_R^1 = x_R$.

In the case where candidate $L$ is honest, and candidate $R$ is dishonest, the expected utility for each candidate reads as:

$$\mathbb{E}\left[U_L(\mathbf{p},\boldsymbol{\pi})\right] = \begin{cases} -\frac{1}{2a}(a-p_R^0-p_L^0)(1+x_R) - \frac{1}{2a}(a+p_R^0+p_L^0)(1+p_L^0+\gamma) & \delta|x_R-p_R|>a \\ -\frac{1}{2a}(a-p_R^0-p_L^0)(1+x_R+\frac{\gamma}{a}(a-\delta(x_R-p_R^0))) & \\ -\frac{1}{2a}(a+p_R^0+p_L^0)(1+p_L^0+\gamma) & \delta|x_R-p_R^0|<a \end{cases}$$

$$\mathbb{E}\left[U_R(\mathbf{p},\boldsymbol{\pi})\right] = \begin{cases} -\frac{1}{2a}(a-p_R^0-p_L^0)2\gamma - \frac{1}{2a}(a+p_R^0+p_L^0)(x_R-p_L^0+\gamma) & \delta|x_R-p_R^0|>a \\ -\frac{1}{2a}(a-p_R^0-p_L^0)(\frac{\gamma}{a}(a+\delta(x_R-p_R^0))) & \\ -\frac{1}{2a}(a+p_R^0+p_L^0)(x_R-p_L^0+\gamma) & \delta|x_R-p_R^0|<a \end{cases}$$

Notice that, the utility for candidate $R$ when $\delta|x_R-p_R^0|>a$ is maximized at $p_R^0=0$. Hence, when $\delta x_R>a$, there is a unique sub-equilibrium for the promises where $p_R^0=0$ and candidate $L$ answers with the $p_L^0$ that maximizes her expected utility. Solving the unconstrained maximization problem,

$$p_L^0 = \frac{x_R-\gamma-p_R^0-a}{2}$$

Intersecting the two best replies I get $p_R^0=0$ and $p_L^0=\frac{x_R-\gamma-a}{2}$. It is necessary to check whether two more conditions hold: first, $p_L^0\leq 0$ and $p_L^0<-1+\gamma$. Where the second is the condition for the honesty of candidate $L$. Notice that, as $-1+\gamma$ is always negative, it is enough to check the second condition.

$$\frac{x_r-\gamma-a}{2}<-1+\gamma \Rightarrow 2+x_R<a+3\gamma$$

Summing up, when $2(x_R-\gamma)>a>2+x_R-3\gamma$ and $a<\gamma\delta$, candidate $L$ chooses a promise weekly between $-1$ and $0$ and applies it whenever she wins while candidate $R$ chooses the preferred point of the voter and is dishonest in equilibrium.

# 4 Conclusion

In this paper, I develop a theoretical model to examine the polarization of politicians as a function of uncertainty regarding voter preferences. The model offers predictions about candidate behavior, suggesting that integrity influences campaign promises when its salience is sufficiently high and uncertainty about voter preferences is moderate. The relationship between uncertainty and campaign promises is non-monotonic. When uncertainty about voter preferences is moderate, candidates tend to diverge from the median voter towards their own preferred policies. However, when uncertainty is too high, the expected cost in integrity diminishes, leading candidates to revert to more median positions.

The findings of this paper complement existing literature on electoral competition, particularly in the areas of electoral rhetoric and backward-looking voter behavior. From a theoretical standpoint, the next logical step would be to explore this mechanism in an infinitely repeated game, examining how the role of integrity evolves in the absence of last-period effects. Additionally, extending the model to a multi-policy framework could address intriguing questions about how politicians might strategically lie on secondary issues while maintaining their promises on key issues, thereby preserving much of their integrity.

The theoretical framework of this model can be extended with two empirical approaches to measure the role of integrity in policy selection. A natural next step would be a laboratory experiment in which participants are assigned roles as politicians and voters. Politicians would make campaign promises to secure election, then select policies, replicating the mechanism outlined in the model. In a repeated-game setting, this experiment could explore how integrity influences the selection of promises and how the cost of lying impacts the chances of reelection, thereby testing the model's predictions.

Another valuable empirical extension would involve using real-world data to measure politicians' deviations from their campaign promises,[8] combined with survey data on voter behavior. This would provide insights into how the precision of information available to politicians shapes their campaign promises and influences their likelihood of flip-flopping in a real-world context.

[8] Websites like politifact.com track this information.

# References


Alesina, A. (1988). Credibility and policy convergence in a two-party system with rational voters. *The American Economic Review*, *78*(4), 796–805. Retrieved 2023-04-11, from http://www.jstor.org/stable/1811177

Andreottola, G. (2020, March). Flip-flopping and Electoral Concerns [CSEF Working Papers]. (558). Retrieved from https://ideas.repec.org/p/sef/csefwp/558.html

Ansolabehere, S., & Snyder, J. M. (2000). Valence politics and equilibrium in spatial election models. *American Journal of Political Science*, *45*. Retrieved 2023-04-11, from https://doi.org/10.1023/A:1005020606153

Aragones, E., & Palfrey, T. (2002). Mixed equilibrium in a downsian model with a favored candidate. *Journal of Economic Theory*, *103*(1), 131-161. Retrieved from https://EconPapers.repec.org/RePEc:eee:jetheo:v:103:y:2002:i:1:p:131-161

Aragonés, E., Palfrey, T., & Postlewaite, A. (2007). Political reputations and campaign promises. *Journal of the European Economic Association*, *5*, 846-884.

Aragonès, E., & Xefteris, D. (2017). Imperfectly informed voters and strategic extremism. *International Economic Review*, *58*(2), 439–471. Retrieved 2023-12-01, from http://www.jstor.org/stable/44280181

Ashworth, S., & Bueno de Mesquita, E. (2009). Elections with platform and valence competition. *Games and Economic Behavior*, *67*(1), 191-216. Retrieved from https://EconPapers.repec.org/RePEc:eee:gamebe:v:67:y:2009:i:1:p:191-216

Austen-Smith, D., & Banks, J. (1989). Electoral accountability and incumbency. In P. Ordeshook (Ed.), *Models of strategic choice in politics.* University of Michigan Press.

Banks, J. S. (1990). A model of electoral competition with incomplete information. *Journal of Economic Theory*, *50*(2), 309-325. Retrieved from https://www.sciencedirect.com/science/article/pii/0022053190900055

Barro, R. J. (1973). The control of politicians: An economic model. *Public Choice*, *14*, 19–42. Retrieved from http://www.jstor.org/stable/30022701

Bischoff, I., & Siemers, L.-H. (2013). Biased beliefs and retrospective voting: why democracies choose mediocre policies. *Public Choice*, *156*(1/2), 163–180. Retrieved 2023-04-11, from http://www.jstor.org/stable/42003153

Blais, A. (2000). *To vote or not to vote: The merits and limits of rational choice theory.* University of Pittsburgh Press. Retrieved 2024-08-19, from http://www.jstor.org/stable/j.ctt5hjrrf

Born, A., van Eck, P., & Johannesson, M. (2018). An experimental investigation of election promises. *Political Psychology*, *39*(3), 685–705. Retrieved 2023-04-11, from http://www.jstor.org/stable/45095198

Boylan, J. F. (2016). *The year of the angry voter.* Retrieved 2016, from https://www.nytimes.com/2016/02/11/opinion/the-year-of-the-angry-voter.html1/

3https://www.nytimes.com/2016/02/11/opinion/the-year-of-the-angry-voter.html

Buisseret, P., & Van Weelden, R. (2022). Polarization, valence, and policy competition. *American Economic Review: Insights*, *4*(3), 341-52. Retrieved from https://www.aeaweb.org/articles?id=10.1257/aeri.20210169

Callander, S., & Wilkie, S. (2007). Lies, damned lies, and political campaigns. *Games and Economic Behavior*, *60*(2), 262-286. Retrieved from https://www.sciencedirect.com/science/article/pii/S089982560600176X doi: https://doi.org/10.1016/j.geb.2006.12.003

Calvert, R. L. (1985). Robustness of the multidimensional voting model: Candidate motivations, uncertainty, and convergence. *American Journal of Political Science*, *29*(1), 69–95. Retrieved from http://www.jstor.org/stable/2111212

Denter, P. (2021). Valence, complementarities, and political polarization. *Games and Economic Behavior*, *128*, 39-57. Retrieved from https://www.sciencedirect.com/science/article/pii/S0899825621000427

Downs, A. (1957). An economic theory of political action in a democracy. *Journal of Political Economy*, *65*. Retrieved from https://EconPapers.repec.org/RePEc:ucp:jpolec:v:65:y:1957:p:135

Galeotti, F., & Zizzo, D. J. (2015). Competence versus honesty: What do voters care about? Retrieved from https://halshs.archives-ouvertes.fr/halshs-01180812/document

Gouret, F., Hollard, G., & Rossignol, S. (2011). An empirical analysis of valence in electoral competition. Retrieved from https://EconPapers.repec.org/RePEc:hal:journl:halshs-00867711

Groseclose, T. (2001). A model of candidate location when one candidate has a valence advantage. *American Journal of Political Science*, *45*(4), 862–886. Retrieved 2022-04-11, from http://www.jstor.org/stable/2669329

Healy, A., & Malhotra, N. (2013). Retrospective voting reconsidered. *Annual Review of Political Science*, *16*, 285-306. Retrieved from https://doi.org/10.1146/annurev-polisci-032211-212920

Hotelling, H. (1929). Stability in competition. *The Economic Journal*, *39*(153), 41–57. Retrieved 2023-04-11, from http://www.jstor.org/stable/2224214

Hummel, P. (2010). On the nature of equilibria in a downsian model with candidate valence. *Games and Economic Behavior*, *70*(2), 425-445. Retrieved from https://www.sciencedirect.com/science/article/pii/S089982561000028X

Kartik, N., & McAfee, R. P. (2007, June). Signaling character in electoral competition. *American Economic Review*, *97*(3), 852-870. Retrieved from https://www.aeaweb.org/articles?id=10.1257/aer.97.3.852 doi: 10.1257/aer.97.3.852

Kartik, N., & Van Weelden, R. (2018, 01). Informative Cheap Talk in Elections. *The Review of Economic Studies*, *86*(2), 755-784. Retrieved from https://doi.org/10.1093/

restud/rdy009

List, J. A., & Sturm, D. M. (2006, 11). How Elections Matter: Theory and Evidence from Environmental Policy*. *The Quarterly Journal of Economics*, *121*(4), 1249-1281. Retrieved from https://doi.org/10.1093/qje/121.4.1249

Rivas, J. (2015, 12). Private agenda and re-election incentives. *Social Choice and Welfare*(46), 899-915.

Schnakenberg, K. E. (2016). Directional cheap talk in electoral campaigns. *The Journal of Politics*, *78*(2), 527–541. Retrieved 2023-04-11, from https://www.jstor.org/stable/26550729

Stokes, D. E. (1963). Spatial models of party competition. *American Political Science Review*, *57*(2), 368–377.

Tambe, E. B. (2018). Who does not vote and why? implication for new democracies. In R. M. Yonk (Ed.), *Elections* (chap. 1). Rijeka: IntechOpen. Retrieved from https://doi.org/10.5772/intechopen.80988 doi: 10.5772/intechopen.80988

Weelden, R. V. (2013). Candidates, credibility, and re-election incentives. *The Review of Economic Studies*, *80*(4 (285)), 1622–1651. Retrieved 2023-04-11, from http://www.jstor.org/stable/43551569

Wittman, D. (1983). Candidate motivation: A synthesis of alternative theories. *The American Political Science Review*, *77*(1), 142–157. Retrieved 2023-04-11, from http://www.jstor.org/stable/1956016

Wuthnow, R. (2018). *The left behind: Decline and rage in rural america.* Princeton University Press.

Xefteris, D. (2013). Equilibrium in a discrete downsian model given a non-minimal valence advantage and linear loss functions. *Mathematical Social Sciences*, *65*(2), 150-153. Retrieved from https://EconPapers.repec.org/RePEc:eee:matsoc:v:65:y:2013:i:2:p:150-153

# A Mathematical appendix

*Proof of Proposition 3.1.* The last period flow utility of candidate $i$ if she wins at $t = 1$ is

$$u_i(\pi_i^1; p_i^1, p_{-i}^1, \pi_{-i}^1) = -(|x_i - \pi_i^1|)$$

Notice that this utility function only depends on $\pi_i^1$, a variable that the candidate chooses, and $x_i$ which is exogenous. Then, this expression is maximized at $\pi_i^1 = x_i$ for every pair $p_i^1$, $p_{-i}^1$. $\square$

*Proof of Proposition 3.2.* The expected utility in equilibrium for candidate $R$ given $p_L^1$ at period 1 is:

$$\mathbb{E}(u_R^1(p_R^1; p_L^1, \pi_L^1, \pi_R^1)) = \begin{cases} -\frac{1}{a}\left(a + |p_R^1| - |p_L^1| - \mathfrak{C}_L^1\right) & \text{if } a + |p_R^1| - |p_L^1| - \mathfrak{C}_L^1 \in [-a, a] \\ 0 & \text{if } a + |p_R^1| - |p_L^1| - \mathfrak{C}_L^1 > a \\ -2 & \text{if } a + |p_R^1| - |p_L^1| - \mathfrak{C}_L^1 < -a \end{cases}$$

If $\mathfrak{C}_L \geq a$, for every $p_L^1$, there is a $p_R^1$ close enough to zero such that candidate $R$ wins with probability one and gets zero utility at $t = 1$. In particular, given $p_L^1$, candidate $R$'s best reply must meet the following condition

$$|p_R^1| \leq |p_L^1| + \mathfrak{C}_L^1 - a$$

If candidate L plays a promise different from zero in equilibrium, candidate R can best reply with a promise bigger, in absolute value, than $\mathfrak{C}_L^1 - a$. Notice that this can not happen in equilibrium, as candidate L will be willing to deviate and choose a $p_L^1$ closer to zero such that the previous condition fails. Hence, for $\mathfrak{C}_L^1 > a$ there are multiple equilibria where candidate R plays $p_R^1 \in [a - \mathfrak{C}_L^1, \mathfrak{C}_L^1 - a]$. As candidate $L$ can not choose any promise to win, she is indifferent to any promise that she can choose. In equilibrium, $p_L^1 \in \mathbb{R}$.

If $a > \mathfrak{C}_L^1 \geq 0$, candidate $L$ can guarantee herself a positive probability of winning choosing $p_L^1 = 0$, as $\mathfrak{C}_L^1 - a < 0$ and $|p_R^1| \leq 0$. Both candidates, choosing a promise close to the voter's ideal policy, can have a strictly positive probability of winning. With this, candidate's R utility maximization is equivalent to maximize

$$\frac{-1}{a}(a + |p_R^1| - |p_L^1| - \mathfrak{C}_L^1)$$

which is decreasing for $|p_R^1|$. Reasoning equivalently for candidate L, I can state that both candidates best reply to each other with $p_R^1 = p_L^1 = 0$. $\square$

*Proof of Proposition 3.3.* The expected utility of candidate $R$ at $t = 0$ when she wins with $p_R^0$ reads as

$$\mathbb{E}\left[U_R(\pi_R^0; p_R^0)\right] = \begin{cases} -|1 - \pi_R^0| - \frac{\gamma}{a}(a + \delta|\pi_R^0 - p_R^0|) & \text{if } \delta|\pi_R^0 - p_R^0| < a \\ -|1 - \pi_R^0| - 2\gamma & \text{if } \delta|\pi_R^0 - p_R^0| > a \end{cases} \tag{7}$$

The two parts of the function depend on the integrity cost. Equivalently, I can say that the difference between the policy and the promise determines the boundaries of the utility function. For both parts, the first term corresponds to the utility derived from applying the policy $\pi_R^0$. The second term is the expected utility in the last period of the game. It's important to note two issues with this function. First, the function is continuous for every $\pi_R^0$. In particular, when $\delta|\pi_R^0 - p_R^0| = a$, as

$$\lim_{\mathfrak{C}_R^0 \longrightarrow a^+} \mathbb{E}\left[U_R(\pi_R^0; p_R^0)\right] = \lim_{\mathfrak{C}_R^0 \longrightarrow a^-} \mathbb{E}\left[U_R(\pi_R^0; p_R^0)\right]$$

Second, the function has a unique maximum for $\mathfrak{C}_R^0 > a$, and it is $\pi_R^0 = 1$. Consequently, to find the maximum, it is enough to study the $\pi_R^0$ maximizing the function when $\mathfrak{C}_R^0 \leq a$ and comparing the values of the expected utility.

The expected utility when $\mathfrak{C}_R^0 \leq a$ is increasing on $\pi_R^0 \in [p_R^0, 1]$ if $1 > \frac{\gamma\delta}{a}$. In this case, $\pi_R^0 = 1$ is the maximum of this function.

If $1 < \frac{\gamma\delta}{a}$, the function is decreasing in the same interval for $\pi_R^0$. Hence, the maximum of the function will be $p_R^0$ if the first part of the function is the biggest, and $\pi_R^0$ otherwise:

$$\pi_R^0 = \begin{cases} p_R^0 & \text{if } -|1 - p_R^0| - \gamma \geq -2\gamma \\ 1 & \text{if } -|1 - p_R^0| - \gamma < -2\gamma \end{cases}$$

So $p_R^0$ is the maximum if

$$|1 - p_R^0| \leq \gamma$$

Rearranging terms, we have

$$1 + \gamma \geq p_R^0 \geq 1 - \gamma$$

Therefore,

$$\pi_R^0 = \begin{cases} p_R^0 & \text{if } a < \gamma\delta \text{ and } 1 + \gamma \geq p_R^0 \geq 1 - \gamma \\ 1 & \text{Otherwise} \end{cases}$$

□

*Proof of Proposition 3.4.* If $a < \gamma\delta$ and $p_R^0 \geq (1 - \gamma)$, Proposition 3.3 states that the candidates will apply their promise if they win. The expected utility for candidate $R$ is then,

$$\begin{aligned} \mathbb{E}\left[U_R(\mathbf{p}, \boldsymbol{\pi})\right] = &- \frac{1}{2a}\left(-|p_R^0| + |p_L^0| + a\right)\left[|1 - p_R^0| + \gamma\right] \\ &- \frac{1}{2a}\left(a + |p_R^0| - |p_L^0|\right)\left[|1 - p_L^0| + \gamma\right] \end{aligned}$$

Notice that $|p_R^0| > 1$ can not happen in equilibrium, as $p_R^0 = 1$ gives higher utility when winning and increases or keeps the probability of winning. In equilibrium, it must be that $p_R^0 \geq 0$, as for any promise smaller than zero, promising the policy of the voter increases the probability and the utility of winning. Reasoning similarly for candidate $L$, the expected utility can be rewritten as

$$\begin{aligned} \mathbb{E}\left[U_R(\mathbf{p}, \boldsymbol{\pi})\right] = &- \frac{1}{2a}\left(-p_R^0 - p_L^0 + a\right)\left[1 - p_R^0 + \gamma\right] \\ &- \frac{1}{2a}\left(a + p_R^0 + p_L^0\right)\left[1 - p_L^0 + \gamma\right] \end{aligned}$$

Solving the utility maximization problem for $p_R^0$ and doing the same for candidate $L$, I get the best replies of each candidate.

$$p_R^0 = \begin{cases} \frac{a}{2} & \text{if } a < 2 \\ 1 & \text{if } a \geq 2 \end{cases}, \quad p_L^0 = \begin{cases} -\frac{a}{2} & \text{if } a < 2 \\ -1 & \text{if } a \geq 2 \end{cases}$$

But notice that if $a \geq 2$,

$$a \geq \frac{a}{2} \geq 1 \geq \gamma\delta$$

This statement is false by assumption whenever $\delta \leq 1$. However, it can be true when $\delta > 1$. Therefore, if $\gamma\delta > a$ and $a \geq 2$, $p_R^0 = 1 = -p_L^0$ and, in any other case, $p_R^0 = \frac{a}{2}$. Therefore, if $a < \gamma\delta$ and $a \geq 2(1-\gamma)$

$$p_R^0 = \frac{a}{2},\ p_L^0 = -\frac{a}{2}$$

Assume now that $a < \gamma\delta$ and $a < 2(1-\gamma)$. Candidate $R$ will only apply her promise if $p_R^0 > 1-\gamma$. If not, the expected utility of candidate $R$ as a function of $p_R^0$ is

$$\begin{aligned} \mathbb{E}\left[U_R(\mathbf{p}, \boldsymbol{\pi})\right] = &- \frac{1}{2a}\left(-|p_R^0| + |p_L^0| + a\right)\left[\frac{\gamma}{a}\left(\delta|1-p_R^0| + a\right)\right] \\ &- \frac{1}{2a}\left(a + |p_R^0| - |p_L^0|\right)\left[2 + \frac{\gamma}{a}\left(-\delta|-1-p_L^0| + a\right)\right] \end{aligned} \tag{8}$$

when $\delta|1-p_R^0| < a$, and

$$\mathbb{E}\left[U_R(\mathbf{p}, \boldsymbol{\pi})\right] = -\frac{1}{a}\left(-|p_R^0| + |p_L^0| + a\right)\gamma - \frac{1}{a}\left(a + |p_R^0| - |p_L^0|\right) \tag{9}$$

if $\delta|1-p_R^0| > a$.

Notice that the expected utility in (9) reaches a maximum at $p_R^0 = 0$. On the other hand, using the same reasoning of the previous part of the proof, $p_R^0 \in [0,1]$ and $p_L^0 \in [-1,0]$. And then, I can write equation (8) as

$$\begin{aligned} \mathbb{E}\left[U_R(\mathbf{p}, \boldsymbol{\pi})\right] = &- \frac{1}{2a}\left(-p_R^0 - p_L^0 + a\right)\left[\frac{\gamma}{a}\left(\delta(1-p_R^0) + a\right)\right] \\ &- \frac{1}{2a}\left(a + p_R^0 + p_L^0\right)\left[2 + \frac{\gamma}{a}\left(-\delta(1+p_L^0) + a\right)\right] \end{aligned}$$

The unconstrained maximization problem of the previous equation gives as a maximum $p_R^* = 1 + \frac{a}{2} - \frac{a}{\gamma\delta}$. By symmetry, we can see that $p_L^0 = -1 - \frac{a}{2} + \frac{a}{\gamma\delta}$.

If $\delta(1 - p_R^0) \geq a$, candidate $R$ chooses in equilibrium $p_R^0 = 0$. This is because the integrity cost in which candidate $R$ incurs given his optimal promise is too high, and therefore she can choose a better promise. The same holds for candidate $L$. If instead, $\delta(1 - p_R^0) < a$, we have that developing the inequality,

$$1 > \frac{2 - \gamma\delta}{2\gamma} \tag{10}$$

This means that $p_R^* < 0$, as

$$1 \leq \frac{a}{\delta}\left(\frac{2 - \gamma\delta}{2\gamma}\right)$$

where $\frac{a}{\delta} < \gamma \leq 1$ and the right term is smaller than one. Notice that, condition (10) can be rewritten as $\gamma\delta > 2(1 - \gamma)$. Therefore, if $\gamma\delta > 2(1 - \gamma) > a$ both candidates choose $p_i^0 = 0$ in equilibrium.

If $a > \gamma\delta$, candidate $R$ will apply the policy $\pi_R^0 = 1$ if she wins. Hence, the expected utility of candidate $R$ reads as

$$\begin{aligned}\mathbb{E}\left[U_R(\mathbf{p}, \boldsymbol{\pi})\right] = &-\frac{1}{2a}\left(-|p_R^0| + |p_L^0| + a\right)\left[\frac{\gamma}{a}\left(\delta|1 - p_R^0| + a\right)\right] \\ &-\frac{1}{2a}\left(a + |p_R^0| - |p_L^0|\right)\left[2 + \frac{\gamma}{a}\left(-\delta| - 1 - p_L^0| + a\right)\right]\end{aligned} \tag{11}$$

when $\delta|1 - p_R^0| < a$. Using the same reasoning than previously,

$$\begin{aligned}\mathbb{E}\left[U_R(\mathbf{p}, \boldsymbol{\pi})\right] = &-\frac{1}{2a}\left(-p_R^0 - p_L^0 + a\right)\left[\gamma\left(\delta(1 - p_R^0) + a\right)\right] \\ &-\frac{1}{2a}\left(a + p_R^0 + p_L^0\right)\left[2 + \frac{\gamma}{a}\left(-\delta(1 + p_L^0) + a\right)\right]\end{aligned}$$

If $\delta|1 - p_R^0| > a$, the expected utility is

$$\mathbb{E}\left[U_R(\mathbf{p}, \boldsymbol{\pi})\right] = -\frac{1}{a}\left(-|p_R^0| + |p_L^0| + a\right)\gamma - \frac{1}{a}\left(a + |p_R^0| - |p_L^0|\right)$$

This function, when $\delta|1 - p_R^0| > a$, has as maximum at $p_R^0 = 0$. The unconstrained maximization problem of the expected utility when $\delta|1 - p_R^0| < a$ gives as a maximum $p_R^* = 1 + \frac{a}{2} - \frac{a}{\gamma\delta}$. By symmetry, we can see that $p_L^0 = -1 - \frac{a}{2} + \frac{a}{\gamma\delta}$.

Notice that, as the expected utility function is continuous when $\delta|1 - p_R^0| = a$, I only need to compare the utility of $p_R^0 = 0$ against $p_R^*$ to know which is the maximum. Two conditions must hold for $p_R^*$ being a best reply in equilibrium:

1. $p_R^* \geq 0$. This implies, using the same procedure of the previous proposition,

$$\gamma\delta \geq \frac{2a}{2+a}$$

2. $\delta(1 - p_R^*) < a$. This implies

$$1 > \frac{2 - \gamma\delta}{2\gamma}$$

If any of these two condition fails, candidate $R$ prefers to choose in equilibrium $p_R^0 = 0$. This can be either because $\delta < a$ and the candidate in equilibrium has a positive probability of winning in the second stage or either because even having zero probability of winning, the candidate values more increasing the probability of winning than the integrity cost.

□

*Proof of Theorem 3.1.* The proof is done in three different parts, each one corresponding to the three different types of equilibria.

1. If $2(1 - \gamma) < \gamma\delta$ the optimal promises in equilibrium are determined by proposition 3.4, and the policies by proposition 3.3. As candidates apply their promise, they arrive to the second stage with same level of integrity, and hence, they will play the most moderate policy and chose their ideal policy if they win.

2. In this case, promises are those stated at the beginning of this section. Proposition 3.3 ensure that the winner will apply her ideal policy. The lying costs of the candidate that wins are

$$\mathfrak{C}_i^t = \delta\left|1 + \frac{a}{2} - \frac{a}{\gamma\delta} - 1\right| = \delta a\left|\frac{2 - \gamma\delta}{2\gamma\delta}\right|$$

   At the beginning of this section I characterize the promises in equilibrium depending on the relation between integrity and uncertainty. By assumption,

$$\delta a \left| \frac{2-\gamma\delta}{2\gamma\delta} \right| < a$$

or, equivalently,

$$\left| \frac{2-\gamma\delta}{2\gamma} \right| < 1 \Rightarrow 2(1-\gamma) < \gamma\delta$$

Therefore, both candidates promise the ideal policy of the median voter and apply their ideal policy if they win.

3. For the third type of equilibrium, either $\gamma\delta < 2(1-\gamma)$ or $\gamma\delta < \frac{2a}{2+a}$. In the both cases, the optimal promises are determined by proposition 3.4. Proposition 3.3 guarantees that the winner applies her ideal policy. She incurs on an integrity cost equal to $\delta$. From the characterization of promises, if $\delta < a$ there is a unique equilibrium in which both candidates play $p_i^1 = 0$ and the winner applies her preferred policy. If, instead, $\delta > a$, the advantaged candidate can play anything in the interval $[a-\delta, \delta-a]$ to win and apply her ideal policy with probability one and the other candidate can not play any promise to win.

□

*Proof.* (Promises in the first stage) As shown previously, payoffs in equilibrium do not depend on promises. The expected utility in equilibrium for candidate $R$ given $p_L^1$ at period 1 is

$$\mathbb{E}(u_R^1(p_R^1; p_L^1, \pi_L^1, \pi_R^1)) = P(Win_L)(-2) = \begin{cases} \frac{-1}{a}(a + |p_R^1| - |p_L^1| - \mathfrak{C}_L^t) & \text{if } -|p_R^1| + |p_L^1| + \mathfrak{C}_L^t \in [-a, a] \\ 0 & \text{if } -|p_R^1| + |p_L^1| + \mathfrak{C}_L^t > a \\ -2 & \text{if } -|p_R^1| + |p_L^1| + \mathfrak{C}_L^t < -a \end{cases}$$

If $\mathfrak{C}_L \geq a$, for every $p_L^1$ there is a $p_R^1$ close enough to zero such that candidate $R$ wins with probability one and gets zero utility at $t = 1$. In particular, given $p_L^1$, candidate $R$ best reply must meet the following condition

$$|p_R^1| \leq |p_L^1| + \mathfrak{C}_L^1 - a$$

If candidate $L$ plays a promise different from zero in equilibrium, candidate $R$ can best reply with a promise bigger, in absolute value, than $\mathfrak{C}_L^1 - a$. Notice that this can not happen in equilibrium, as candidate $L$ will be willing to deviate and choose a $p_L^1$ closer to zero such that the previous condition fails. Hence, for $\mathfrak{C}_L^1 > a$ there are multiple equilibria where candidate $R$ plays $p_R^1 \in [a - \mathfrak{C}_L^1, \mathfrak{C}_L^1 - a]$. As candidate $L$ can not choose any promise to win, she is indifferent to any promise that she can choose. In equilibrium, $p_L^1 \in \mathbb{R}$.

If $a > \mathfrak{C}_L^1 \geq 0$, candidate $L$ can guarantee herself a positive probability of winning choosing $p_L^1 = 0$, as

$$\mathfrak{C}_L^t - a < 0 \text{ and } 0 \geq |p_R^1|$$

Both candidates, choosing a promise close to the voter's ideal policy, can have a strictly positive probability of winning. With this, candidate's $R$ utility maximization is equivalent to maximize

$$\frac{-1}{a}\left(b + |p_R^1| - |p_L^1| - \mathfrak{C}_L^t\right)$$

which is decreasing for $|p_R^1|$. Reasoning equivalently for candidate $L$, I can state that both candidates best reply to each other with $p_R^1 = p_L^1 = 0$.

□