POLITICAL REPUTATIONS AND CAMPAIGN PROMISES

Enriqueta Aragonès
Institut d’Anàlisi Econòmica, CSIC

Andrew Postlewaite
University of Pennsylvania

Thomas Palfrey
California Institute of Technology

Abstract
We analyze conditions under which candidates’ reputations may affect voters’ beliefs over what policy will be implemented by the winning candidate of an election. We develop a model of repeated elections with complete information in which candidates are purely ideological. We analyze an equilibrium in which voters’ strategies involve a credible threat to punish candidates who renege on their campaign promises and in which all campaign promises are believed by voters and honored by candidates. We characterize the maximal credible campaign promises and find that the degree to which promises are credible in equilibrium is an increasing function of the value of a candidate’s reputation. (JEL: D8)

1. Introduction

Politicians seeking office make promises. This is presumably done in the belief that such promises will alter voters’ beliefs about the policies the politician will implement if elected and about the capabilities of the politician. The flip side of the coin is that these promises may later come back to haunt an office holder seeking reelection, so candidates must temper their promises in anticipation of future elections. This paper presents a model in which these effects arise as equilibrium phenomena.

We focus on one aspect of political campaigns, which we refer to as credible commitment, and study it using an infinitely repeated version of the
one-dimensional spatial model where candidates have policy preferences that change over time. With sufficiently patient voters and candidates, there are many equilibria. We characterize the range of credible promises to which candidates can commit. Campaign promises provide a mechanism for voters to select among multiple equilibria in a repeated game, much like a focal point. In this sense, credible promises solve a coordination problem that arises naturally in the context of multi-principal agency problems, where the many principals must somehow converge on a common rule in order to effectively control the agent. Campaign promises affect voters’ expectations about what policies will be chosen by an elected official, and they provide a benchmark for voters to link policy decisions with future re-election. In the absence of such public announcement, it is difficult to imagine how voters could agree about what constitutes acceptable performance by an elected official.

The difficulty with the argument that campaign statements are a mere act of promising, or pledging, to carry out a particular policy is that they are cheap talk. That is, fixing all actions of all participants, no payoffs differ when messages alone are changed. Consider, then, a problem in which there is a single election in which candidates vie for office. Suppose candidates are purely ideological: they receive no direct payoff from holding office and care only about the policy chosen. In this environment, any candidate who is elected will choose the policy alternative that he most prefers, regardless of any campaign promise that might have been made. Consequently, if voters have rational expectations, then no campaign promise can alter voters’ beliefs about what action will be taken by a candidate once elected. If there were any statement that would alter beliefs in a way that increased the candidates’ probability of election, then candidates would make such statements regardless of what they intended to do if elected. Hence, no campaign statement can convey information that alters the chance of election.\footnote{See Harrington (1992) for an elaboration of this argument.}

When we move from the case of a single election to multiple elections, campaign promises may be costly because voters can condition their strategies on these promises in the repeated game. Voters may vote differently in future elections if a candidate promises to do something if elected but subsequently reneges on that promise. Simply put, voters may punish a candidate for reneging on campaign promises by voting him out of office. In this way the promises serve a coordinating role for voters. Under certain conditions, threats of such punishment can support an equilibrium in which campaign promises are kept and in which voters’ beliefs about what a candidate will do if elected are affected by campaign promises. There is a potential problem, however, with voters behaving on the basis of “retrospective” assessments of candidates: at the time of the next election, the future choices that the candidate might make could look far better than those of his opponent. Threats to vote candidates out of office regardless of
the circumstances may not be credible; in other words, strategies employing such threats are dominated. Despite the fact that these strategies are dominated, they are often used to justify the assumption that politicians can commit to platforms or policies prior to an election.

We present and analyze a dynamic model in which candidates make campaign promises and then voters use those promises to form beliefs about the policies the candidate will choose if elected. We analyze equilibria of the model in which some promises will be kept—even when the promised policy differs from the elected candidate’s ideal point owing to fear of voter reprisal. However, unlike the retrospective punishments just described, punishment in our model is prospective. Voters discipline candidates by believing some promises a candidate makes as long as that candidate has never reneged on a promise in the past. Once he reneges, no future promises will be believed. Candidates only make promises they intend to keep, and they keep those promises if elected. In other words, we consider only subgame perfect equilibria.

Modeling campaign promises in this way has advantages beyond simply avoiding dominated strategies. The incentive to fulfill campaign promises is based on the threat that future promises will not be believed; the cost to a candidate of this punishment is finite. Consequently, promises to carry out policies that are known to be anathema to the candidate will not be believed, because it will be understood that the gain from reneging will outweigh the cost in lost credibility.\(^2\) Thus, unlike models that simply assume that candidates can commit, we find that there typically will be some policies to which candidates can (credibly) commit, but other policies to which they cannot commit. In addition, the precise modeling of the source of a candidate’s ability to alter voters’ beliefs about what he will do, once elected, permits an analysis of how the magnitude of his credibility is affected by circumstances such as the probability of being elected, the expected duration of his political career, the reputation of his opponent, and so forth.

Related literature. As mentioned previously, much of the work on campaigns has followed Downs (1957) in assuming, implicitly or explicitly, that candidates could commit to platforms or policies that they would implement if elected. Ferejohn (1986) and Barro (1973) consider a repeated principal-agent model of sequential elections in which the threat of being thrown out of office reduces the incentives for shirking while in office. Candidates are identical and have no policy preferences, and they are judged on their past performance rather than on any campaign promises or commitments they might make. Austen-Smith and Banks (1989) explore a two-period variation of this principal-agent model. Candidates propose performance goals during the election, and achievement of these goals

---

\(^2\) Consider, for example, the skepticism that greeted 1996 US presidential candidate Bob Dole’s promise to cut taxes after a long history of arguing against the wisdom of such cuts.
depends on a combination of effort and luck. The authors look at the subset of implicit contracts where voters discipline the incumbent by a quadratic scoring rule that compares actual performance to the incumbent’s performance goal.

Wittman (1990) analyzes a model with politicians facing an infinite sequence of elections with unchanging ideal points. He characterizes the equilibrium between the candidates when they are restricted to choosing the same policy each period. This differs from our model in two ways: Voters play no active role in that model, and candidates never compare the costs and benefits of carrying out the policies, so issues of credibility do not enter the model. Banks and Duggan (2002) analyze a dynamic, multi-dimensional policy model without campaigns and then characterize equilibria in terms of simple strategies. In each period, the incumbent faces a random opponent; the authors show existence of an equilibrium in which an individual votes for the incumbent if the individual’s utility meets a critical threshold, which is determined endogenously. There is no consideration of prospective evaluations of candidates. Duggan and Fey (2006) investigate properties of the set of equilibria with infinitely repeated elections, complete information, and office-motivated candidates but no campaign promises. In their model there is no issue of candidate credibility or retrospective voting, because candidates are purely office-motivated and thus are indifferent concerning which policy they actually implement if elected.

This earlier work either ignored the effect of a politician’s performance in office on the chances of reelection or considered only office-motivated candidates. Most of the work that embodies retrospective assessment leaves out any possibility of campaign credibility. Our contribution is to model political campaigns by ideological candidates who make campaign promises, given voters who are fully rational in the degree to which the promises can be believed.

This paper is organized as follows. In section 2, we focus on the case in which there are (potentially) infinite elections and complete information. In this case we show how candidates may (rationally) choose to maintain a reputation for fulfilling campaign promises. We do this initially for the case in which candidates have linear utility functions; we then analyze several extensions, including the effect that concavity in utility has on the set of believable promises. Section 3 comprises a brief discussion of our results, and our formal propositions are proved in section 4.

2. Sequential Elections

2.1. Overview

With an infinite horizon, promises can be credible in equilibrium as long as reputation has a value. Of course, promises can always be broken—and will be
broken if it is in the interest of the candidate to do so. Likewise, promises are kept only if it is in the interest of the candidate to do so, because future payoffs are different for the candidate when he keeps his promise than when he does not. Promises may change voters’ beliefs about the choices that candidates will make if elected, because voters understand that it is sometimes in a candidate’s selfish interest to fulfill his promises even when there is a short-term gain from reneging. Voters also understand that the threat of future punishment is not sufficient to deter all reneging: some promises may be so far from a candidate’s preferred outcome that the short-term gain from reneging is sufficiently high that a candidate will relinquish his electoral future. In short, the ability of a candidate to alter voters’ beliefs is not a “technological” given but rather is an equilibrium phenomenon.

We assume complete information: voters know candidates’ preferences over policies perfectly at the time they vote. We assume that, at each election, candidates’ reputations may be either good or bad: candidates with good reputations are candidates who have never reneged in the past whereas candidates with bad reputations are those who have reneged on a promise sometime in the past. Voters believe only promises of candidates who have a good reputation and never believe any promise of candidates who have a bad reputation. After each election, a winning candidate with a good reputation compares the one-time benefit of reneging on any promise he may have made with the value of maintaining his reputation by fulfilling the promise. Candidates with a bad reputation choose their optimal policy independent of their promises. Voters predict that (a) candidates with a bad reputation will implement their ideal policy regardless of any promises, and (b) candidates with a good reputation will fulfill any promise that is not too costly to carry out—that is, for which the benefit of reneging is less than the decrease in their continuation payoffs if they renge. These strategies constitute a subgame perfect equilibrium. If there is no uncertainty, then candidates do not make promises they do not intend to keep because, with complete information, voters can predict that candidates will renge and so the promises will not influence the voters’ choices.

Candidates will be able to change voters’ beliefs about the policy they will undertake as long as the discount factor is large enough. That is, as long as the future has sufficient value, candidates will carry out their promises when it is not too costly to do so. If there is a positive (expected) value to being elected in each

---

3. We will discuss later a variant of the model in which candidates’ preferences are not known with certainty at the time of the election.

4. Reputations need not have this “all-or-nothing” property; we shall also discuss richer possibilities of how past behavior can affect reputation.

5. Uncertainty (symmetric between voters and candidates) about what alternatives will arise between the time of voting and the time at which the alternatives to the promised action are known would change this. In that case one would expect, in equilibrium, that some promises will not be kept when the benefits of reneging outweigh the value of reputation.
of the future periods, then the value to retaining a good reputation goes to infinity as the discount factor goes to one. For high enough value to retaining a good reputation, all promises will be kept (and, hence, believed by voters).

In these models there will always be one equilibrium in which campaign promises are irrelevant: all candidates make random promises and voters, despite any message they hear, do not alter their beliefs about a candidate’s type or the choices he will make if elected. Candidates choose their most preferred policy if elected. Here, the only information relevant to voters is the candidate’s choice: Their predictions of choice in the second period are independent of any campaign promises, so reneging on campaign promises cannot affect voting in the second election.

What is interesting, however, is that—in addition to this uninformative equilibrium—there may be equilibria in which voters do change their beliefs about candidates, policies and their voting behavior on the basis of campaign promises.

Campaign promises matter if and only if candidates’ payoffs from reneging on their campaign promises are different from their payoffs from fulfilling those promises. That is, we obtain different election outcomes following a failure to fulfill a promise than after a promise has been fulfilled. For the outcome of future elections to differ following fulfillment or nonfulfillment of promises, voters’ strategies must depend on the relationship between a campaign promise and the policy choice of a candidate: voters’ actions must depend on the candidates’ promises.

The set of payoff-relevant equilibrium outcomes (i.e., sequences of elections and policies) of the repeated game analyzed in this paper is the same as the set of equilibrium outcomes that would obtain in a model without candidates making any campaign promises. Complete information implies that there is also an equilibrium without promises that duplicates this sequence of outcomes. In that other equilibrium, no candidate makes an announcement, and the candidate who wins the elections is the one whose promise (in the equilibrium with cheap talk) would have been closer to the median voter. The elected candidate then implements the policy that corresponds to the announcement he would have made in the corresponding cheap-talk equilibrium. If a candidate is elected then the voters expect him to implement it, and he always does this. If not, the candidate would be punished just as in the equilibrium with cheap talk. Because it was a credible equilibrium with cheap talk, this means the elected candidate would prefer to implement what would have been his campaign promise and so maintain a good reputation, than to implement his ideal point and thereby acquire a bad reputation. Thus, the role of promises in the equilibrium analyzed here is primarily to coordinate on an equilibrium strategy for voters.

In general, candidates will not be able to induce all possible beliefs in voters. We consider this an important feature of our approach. It is endogenous which promises will be made, believed, and fulfilled when both candidates and voters
are fully rational. Each candidate will have available to him only a subset of the possible beliefs that voters might have about his policy choices if elected. It is important to note that the sets of beliefs that candidates can induce in voters are typically quite different; they depend on voters’ initial beliefs about the candidates, including their discount factors, utility functions, and so forth.\footnote{This construction provides a rational explanation for the exogenous cost of commitment assumed in Banks (1990), for example.}

2.2. The Model

There are two candidates, $L$ and $R$, who compete in all elections. At each election, the structure of the game is as follows.

\textit{Campaign stage.} Both candidates simultaneously make an announcement. Each candidate must decide between making a promise about the policy he will implement in case he wins the election or sending a message devoid of promises.

\textit{Voting stage.} Each voter votes for the candidate who maximizes the voter’s expected utility, which depends on the policy that he or she believes will be implemented after the election.

\textit{Office stage.} The winner of the election implements a policy.

Candidates and voters derive utility only from the policy implemented. We assume that the utility an agent obtains from each election is represented by

$$u_i(x) = -|x - x_i|,$$

where $x_i$ represents the ideal point of agent $i$.

The policy space is represented by the interval $[-1, 1]$. We assume that the ideal point of the median voter is the same at all elections; it is normalized to be $x_m = 0$.

Elections take place over time. Voters simply vote in each election for the candidate whose predicted policy choice is most preferred.\footnote{We rule out the possibility that voters will “punish” candidates when it is not in their interest to do so for the same reason that attention is restricted to subgame perfect equilibria.} Candidates discount future payoffs with a discount factor $\delta \in [0, 1)$. The discount factor represents the weight that future payoffs have on candidates’ total utility. We have in mind an interpretation of $\delta$ that combines both time preference and the probability that a candidate will run for office in the future. For example, we can think of it as $\delta = \lambda \beta$, where $\lambda \in [0, 1]$ represents the probability that the candidate will run for office in any period and $\beta \in [0, 1)$ represents time preference. Because the value of $\delta$ is less than 1, elections that are further in the future have less effect on the total utility of a candidate than do earlier elections.
We assume that the policy preferences of the two candidates change at each election. In particular, we assume that at each election the ideal point of candidate $L$ is $x_L \in [-1, 0]$, which is given by a random draw from a uniform probability distribution over $[-1, 0]$. At each election the ideal point of candidate $R$ is $x_R \in [0, 1]$, which is similarly given by an independent random draw from a uniform probability distribution over $[0, 1]$. Candidates’ ideal points are drawn independently of each other and of past draws before each election.

Candidates know the preferences of the median voter. At the beginning of each electoral period, voters and candidates learn the ideal points of both candidates for that period.

For each one-period game, a candidate’s strategy selects a pair $(p, x)$, where $x \in [-1, 1]$ represents the policy the candidate will implement if he wins the election and $p \in [-1, 1] \cup \{Ø\}$ represents the announcement (either a promised policy or nothing) the candidate makes at the campaign stage. Formally, we may define a promise by the exact policy that will be implemented; hence if a candidate promises policy $x \in [0, 1]$, then he will break his promise only if he implements $x' \neq x$. We may also think of a promise as the worst policy in terms of the median voter’s preferences, that is, if a candidate promises policy $x \in [0, 1]$, then he will break his promise only if he implements $x' \in (x, 1]$. In our model these definitions are equivalent.

Before deciding their vote, voters may update their beliefs about the candidates’ policy choices (in case they win the election) based on the announcements made at the campaign stage. Given their beliefs, voters decide to vote for the candidate who maximizes their expected utility.

Because voters know the candidates’ ideal points, we assume that—in the absence of promises—voters believe that a candidate will choose his ideal point if elected. After the campaign stage, voters may update their beliefs about the policy choices the candidates would make if elected. Voters decide rationally whether to believe the campaign promises or not. Voters will believe a promise only if honoring it is compatible with his incentives after the election. Thus, even though campaign promises do not affect the payoffs of any of the agents, such promises may affect their decisions.

### 2.3. Credible Promises

We describe an equilibrium of this repeated game in which campaign promises matter in the sense that different promises imply different strategy choices and therefore lead to different payoffs. In this equilibrium, voters will believe the maximal set of incentive compatible promises—that is, promises that the candidate would have an incentive to fulfill should he be elected. For a candidate with discount factor $\delta$, we will show that there is a number $d(\delta)$ such that voters will
believe promises made by the candidate if and only if the distance between the candidate’s ideal point and his promise is not greater than \(d(\delta)\). In the equilibrium we describe, voters will believe all promises from a candidate for which the distance from the candidate’s ideal point is not greater than \(d(\delta)\) if the candidate has never reneged on a promise but will believe no promise if he has ever reneged (i.e., has implemented a policy other than a promised policy).\(^8\) If the candidate makes a promise that is not incentive-compatible or if he makes no promise, then voters believe that he will implement his ideal point.

These strategies essentially treat candidates as one of two types. At each election we may have candidates with a good reputation, who have never reneged on any promises and whose (incentive-compatible) promises will be believed by voters, as well as candidates with a bad reputation. These latter candidates have reneged on a promise at some time in the past and regardless of what promises are made at the campaign stage, a voter will believe that they will implement their ideal points if elected.

After the election, the winner implements the policy that maximizes his expected payoffs, taking into account that the voters’ strategies for future elections might depend on the candidate’s promises and choice. Thus, candidates at this stage will compare the gains and costs of reneging. The gains from reneging are represented by the instantaneous increase in the candidate’s utility produced by deviating from his promised policy and choosing instead his ideal point. The costs of reneging are reflected in his expected payoffs from future elections: the difference between the future expected payoffs for a candidate with a good reputation and a candidate with a bad reputation. A candidate will reneg on a promise only if the instantaneous gain is larger than his future expected loss.

In the equilibrium we describe, candidates will make only incentive-compatible promises and will fulfill the promises they make. Therefore, voters will believe the promises that are made and the winner will be the candidate who is able to promise a policy closer to the median voter’s ideal point. The winning candidate must promise a policy that is at least as attractive to the median voter as his opponent’s policy. If the losing candidate promises a policy that is consistent with incentive compatibility and as close as possible to the median voter’s ideal point, then the winning candidate will have to promise a policy that is at least as close to the median voter’s ideal point.\(^9\) Because we assume that the candidates’

---

\(^8\) There are other equilibria that can be thought of as intermediate cases in which voters believe some, but not all, promises that are incentive compatible. The equilibria in these cases will look like the equilibrium we describe but with a smaller \(d\); that is, voters will believe fewer promises.

\(^9\) If candidate A is promising an incentive-compatible policy that is as close as possible to the median voter’s ideal point and if candidate B has an incentive-compatible promise that is closer, then candidate B will win the election. However, the set of incentive compatible promises that are strictly preferred by the median voter is open. We assume that candidate B is the winning candidate in this
ideal points are on opposite sides of the median voter’s ideal point, it follows that the losing candidate’s utility increases when the winner makes promises that are closer to the median voter’s ideal point. The candidates’ strategies in the equilibrium we describe have (a) the losing candidate promising the policy closest to the median voter’s ideal point that is consistent with incentive compatibility and (b) the winning candidate making a promise that is equally close.

Formally, the strategies for the equilibrium described are as follows.

**Candidates’ strategies**

1. If neither candidate has ever reneged on a promise, then the candidate whose ideal point is farther from the median voter’s ideal point promises the policy that is closest to the median voter’s ideal point consistent with incentive compatibility. The candidate whose ideal point is closer to the median voter’s ideal point promises a policy that is equally attractive to the median voter. If elected, both candidates fulfill their promise.
2. If both candidates have reneged on a promise in the past, then both candidates promise to implement the median voter’s ideal point. If elected, they implement their own ideal point.
3. If one candidate has reneged on a promise but the other candidate has never reneged, the candidate who has reneged promises to implement the median voter’s ideal point. If elected, he implements his own ideal point. The candidate who has not reneged promises a policy that is as attractive to the median voter as the opponent’s ideal point, if such a promise is incentive compatible. If that policy is not incentive-compatible, then he promises his ideal point. If elected, he fulfills his promise.

**Voters’ strategies.** Each voter casts his or her vote for the candidate whose expected policy, if elected, maximizes the voter’s utility. Voters’ beliefs are as follows.

1. Voters believe that incentive-compatible promises of candidates who have never reneged on a promise will be fulfilled.
2. Voters believe that a candidate who makes a promise that is not incentive-compatible will implement his ideal point.
3. Voters believe that a candidate who has reneged on a promise in the past will implement his ideal point.

**Proposition 1.** The strategies just described constitute an equilibrium. The promises believed and fulfilled in equilibrium with linear utility functions are
those within a distance \( d^*(\delta) \) of the candidates’ ideal points, where

\[
d^*(\delta) = \begin{cases} 
0 & \text{if } 0 \leq \delta \leq \frac{1}{2}, \\
\frac{3}{2} \left( 1 - \sqrt{(4 - 5\delta)/(3\delta)} \right) & \text{if } \frac{1}{2} \leq \delta \leq \frac{3}{4}, \\
1 & \text{if } \frac{3}{4} \leq \delta \leq 1.
\end{cases}
\]

The proof is in Appendix A.2.

The distance \( d^*(\delta) \) characterizes an equilibrium with the maximal range of incentive compatible promises. We obtain that, in the equilibrium we have analyzed, candidates who have never reneged on a promise fulfill all the promises they make and voters believe these promises; both candidates maintain a good reputation over time. There is a continuum of equilibria with similar characteristics: For all \( d \leq d^*(\delta) \), there is an equilibrium in which voters believe promises up to a distance \( d \) away from the candidate’s ideal point.

Our analysis yields some simple but interesting comparative statics. Observe that the maximal promise believed in equilibrium is an increasing function of the discount factor, because

\[
\frac{\partial d^*(\delta)}{\partial \delta} = \frac{1}{\delta^2 \sqrt{4 - 5\delta}} \geq 0.
\]

Thus, as the discount factor increases, the value of reputation (the cost of reneging) increases, which implies that larger promises will be kept and believed in equilibrium.

In general, we should expect to see that candidates with high probability of running for office in the future are more likely to fulfill their promises and that voters are more likely to believe promises from these candidates. Hence promises are more likely to be believed at the same time that candidates are more likely to make them.

A particularly interesting consequence of this is that, ceteris paribus, two-candidate systems have an advantage over multi-candidate systems. In the latter, the average candidate clearly has a lower chance of being elected in future elections and thus has a lower value for maintaining a reputation. This lower value of reputation makes fewer promises credible, with the result that there will be less mediating effect of credible promises and hence fewer implemented policies with more candidates.\(^{10}\)

Similarly, all else equal, younger candidates are more likely to fulfill their promises because they have a longer time horizon to consider and so their reputation is more valuable. However, there may be factors such as seniority effects that

\(^{10}\) We thank Abhijit Banerjee for this observation.
cause younger candidates to have smaller chances of being elected in the future; these factors would work in the opposite direction.

Note that a candidate’s expected value $v_G$ of maintaining a good reputation is the same regardless of whether his opponent has a good or a bad reputation:

$$v_{GG}(d^*(\delta)) - v_{BG}(d^*(\delta)) = v_{GB}(d^*(\delta)) - v_{BB}(d^*(\delta)),$$

where the first (respectively, second) subscript denotes the good or bad, G or B, reputation of the candidate (respectively, opponent). That the value of a good reputation is independent of the opponent’s reputation is due to the assumed linearity of the utility functions.

We also analyze the effects of maintaining a good reputation on the welfare of the median voter. The median voter’s expected utility $u$ from each election (as a function of credible promises) in equilibrium is given by

$$u_{GG}(d) = -\frac{1}{3} + d^2 \left( 1 - \frac{2}{3} d \right) > -\frac{1}{3},$$

and

$$u_{BB}(d) = u_{GB}(d) = u_{BG}(d) = -\frac{1}{3},$$

where $\partial u_{GG}(d)/\partial d > 0$ for $0 \leq d \leq 1$. Thus, the median voter is better off when both candidates have a good reputation because all promises that are made tend toward the median voter’s ideal point. In equilibrium, both candidates have a good reputation and the utility of the median voter increases with the size of the set of credible promises.

The probability that a voter is better off when candidates can make credible promises decreases with the absolute value of the ideal point of the voter. In particular, this implies that the voter most favored by the credibility of promises is the median voter ($x_m = 0$). Voters with ideal points at the extremes of the policy space obtain the same expected utility when both candidates have a good reputation as when both candidates have a bad reputation. The reason is that, for each realization of the candidates’ ideal points such that a voter’s utility decreases when some promises are credible, there is another (symmetric) realization of the candidates’ ideal points such that the voters’ utility increases by the same amount when promises are credible.

**2.4. Extension to Concave Utility Functions**

Until now we have assumed that the utility function of each candidate is linear with respect to the distance between his ideal point and the implemented policy. In this section, we will assume that this function is concave. Formally, we assume that

$$U_i(x) = -|x_i - x|^k$$

for all $i$, where
where \( k \geq 1 \) measures the degree of concavity (i.e., the larger the value of \( k \), the larger the degree of concavity). A candidate with a strictly concave utility function \( (k > 1) \) suffers more than a candidate with a linear utility function \( (k = 1) \) from the implementation of policies that are far away from his ideal point. In a sense, the utility function’s degree of concavity is a measure of the intensity of the candidate’s political preferences.

We should expect that the value of maintaining a good reputation for a candidate is increasing in the degree of concavity of his preferences, because his utility loss from losing an election increases with the degree of concavity whereas his utility when winning (even with a promise different from his ideal point) is affected less. In this section, we replicate the preceding equilibrium analysis when candidates’ utility functions are concave, assuming that both candidates’ utilities exhibit the same degree of concavity. We find that the set of credible campaign promises is larger when the degree of concavity of the candidates’ utility functions is higher. In the following proposition we denote by \( d^S \) the maximal promise from a candidate with a good reputation that is credible when the other candidate has a bad reputation.

**Proposition 2.** The strategies described in section 2.3 constitute an equilibrium. The promises believed and fulfilled in equilibrium with concave utility functions are those within a distance \( \tilde{d}^D(\delta, k) \) of the candidates’ ideal points, where

\[
\tilde{d}^D(\delta, k) = \begin{cases} 
0 & \text{if } \delta = 0, \\
0 < \tilde{d}^D(\delta, k) < 1 & \text{if } 0 < \delta < \left(1 + \frac{2k\left[3-(1-d^S)^{k+2}\right]-k-3}{(k+1)(k+2)}\right)^{-1}, \\
1 & \text{if } \left(1 + \frac{2k\left[3-(1-d^S)^{k+2}\right]-k-3}{(k+1)(k+2)}\right)^{-1} \leq \delta,
\end{cases}
\]

and

\[
\frac{\partial (\tilde{d}^D(\delta, k, d^S))}{\partial k} \geq 0.
\]

The proof is in Appendix A.3.

When we assumed that the candidates’ utility functions were linear, we saw that their expected utilities were unaffected by the kind of reputation that they had as long as both candidates had the same kind of reputation. When both candidates have a good reputation, it is equally likely that a given candidate will be helped or hurt by his reputation. When a candidate’s ideal point is closer to the median voter, he will win regardless of whether both candidates have a good or a bad reputation. When both candidates have a good reputation, in equilibrium he will
make a promise and hence be worse off than if both have had a bad reputation, in which case he could have won by promising his ideal point. On the other hand, if his opponent’s ideal point is closer to the median voter then this candidate benefits from having a good reputation. With linear utility functions, these effects exactly offset and the candidates’ expected utilities when both candidates have a good reputation are the same as when neither does.

With concave utility functions, this is no longer the case. When both candidates have a good reputation, the equilibrium policies enacted will be closer to the median voter than they would be if both candidates had a bad reputation. However, this convergence toward the median voter is beneficial to candidates with strictly concave utility functions. When a candidate is forced to move his policy choice toward the median voter’s ideal point because both candidates have a good reputation, his loss is not as large as his gain from his opponent doing the same thing. Hence, with concave utility functions, candidates’ expected utilities are larger when both candidates have a good reputation, than when both candidates have a bad reputation, and the greater the degree of concavity, the greater the difference between the two candidates’ utilities.

Candidates’ welfare increases in our model because of the policy convergence that a good reputation generates. The effect is similar to the welfare increase that results from policy convergence in Alesina (1988) and Dixit, Grossman, and Gul (2000). In these papers, policy convergence arises through tacit cooperation between two parties that moderate their policies when in office. Although the welfare benefits in these papers (as in our paper) are due to policy convergence, the policy convergence that we obtain when we assume linear utility functions stems from the interactions between the voters and the candidates, not from those between the candidates themselves.

2.5. Extension to Random Median Voters

In the model analyzed in the previous sections of this paper, we assume that the ideal points of the candidates change from election to election and that the ideal point of the median voter does not change over time. These assumptions can be interpreted as if voters had stable preferences but the issues changed from election to election. For instance, in one period the main campaign issue (and thus the candidates’ promises) is tax reform, the next election’s issue is abortion, and so on. At each election the ideal point of the median voters is normalized to be zero, and the candidates’ ideal points differ because they reflect the different relative positions of all agents on each specific issue. In one sense, this can be thought of as a model of short-term policies.

In this section, we describe an alternative model in which the policies can be thought of as long-term policies. Here we assume that candidates’ ideal points are fixed at all elections while the ideal point of the median voter changes across
elections. This variation of the model can be interpreted as the candidates having long-run, stable ideal points regarding some policy—say, income distribution. The assumption that the ideal point of the median voter is random captures the idea that the median voter may change over time owing to demographic changes or that individual voters’ preferences may change because of changes in the economy.

Consider the following variant of the model described previously, where the ideal points of the candidates are $x_L = 0$ and $x_R = 1$ at each election and where the ideal point of the median voter $m$ at each election is an independent realization of a uniform random variable on the interval $[0, 1]$. Note that here the ideal points of the candidates are not independent, in contrast to what was assumed in the previous sections.

**Proposition 3.** The strategies described in section 2.3 constitute an equilibrium. The promises believed and fulfilled in equilibrium with a random median voter are those within a distance $d^*(\delta)$ of the candidates’ ideal points, where

$$d^*(\delta) = \begin{cases} 
0 & \text{if } \delta \leq \frac{2}{3}, \\
2(3\delta - 2)/\delta & \text{if } \frac{2}{3} \leq \delta \leq \frac{4}{5}, \\
1 & \text{if } \frac{4}{5} \leq \delta.
\end{cases}$$

The proof is in Appendix A.1.

In this extension of the basic model, we again obtain that the maximal promise depends on the discount factor in a very natural way: When the discount factor is very small, no promises are believed in equilibrium; for larger values of the discount factor, more promises are believed in equilibrium; and when the discount factor is sufficiently large, all promises are believed.

Thus, the results obtained with this alternative formalization of the two-candidate electoral competition are qualitatively the same as the results found when we assumed that the candidates’ ideal points were randomly determined at each election and the median voter’s ideal point was fixed at all elections.

The welfare effects in this case are similar to those in the previous section. As in that case, the median voter is strictly better off when candidates have reputations. When the candidates have linear utility functions, they are equally well off when both or neither have reputations; with strictly concave utility functions, they will be better off when both have reputations than when neither does.

### 2.6. Extension to Multi-Dimensional Policy Space

The logic and intuition of the foregoing results carry over to the multi-dimensional policy case. We discuss here the two-dimensional case, where voters and candidates have Euclidean preferences, each voter or candidate has an ideal point and
utility is a decreasing function of Euclidean distance from these ideal points. The voters’ ideal points are distributed such that there is a global median. For concreteness, suppose there is a continuum of voters whose ideal points are uniformly distributed on a rectangle; hence the global median is simply the individual whose ideal point is the center of the rectangle. For any two policies chosen by the candidates, each voter will prefer the policy that is closer to his or her ideal point. The policy that is preferred by the median voter will be preferred by a majority of the voters (i.e., by more than half of the voters).

In each period, the candidates’ ideal points are drawn from the uniform distribution on the rectangle; these points are independent of each other and are independent across elections. As before, suppose that each candidate can make promises and that those promises will be believed if (they are incentive-compatible and) the candidate has never reneged on a promise in the past but will not be believed if he has ever reneged. It is straightforward to see that, for a candidate who has never reneged, there is a maximum distance from the incentive compatible ideal point that he can promise; that distance is determined by the net present value of being able to make promises in future periods.

The Nash equilibrium of the repeated game of promises between candidates follows the same logic as in one dimension, where generally one of the candidates is closer to the median and the other is farther away. Call the latter candidate the loser. There can’t be an equilibrium with the loser making a less than maximal promise when they both have good reputations, and this follows from exactly the same reasoning as in the one-dimensional case. The equilibrium is illustrated in Figure 1. On the equilibrium path, the loser adopts the closest possible credible policy to the median voter—along the line segment between the loser’s ideal point $B$ and the median voter’s ideal point $V$. The policy is labeled $B'$. The other candidate matches by adopting the (unique) policy on the line segment between that candidate’s ideal point, $A$, and $V$. This policy is labeled $A'$. Although indifferent, the voter votes for $A'$ by the same reasoning as in the one-dimensional case.11

The two-dimensional case has an interesting feature that is absent in one dimension. Namely, if there were a different timing structure in the game, then the losing candidate might gain by making a promise that induces his opponent to counter-promise something other than $A'$ that the loser prefers to $A'$. With simultaneous moves such manipulation is not possible, but it would be if $B$ could irrevocably make a promise before $A$.

Figure 2 illustrates a case in which the optimal “manipulative” promise for the losing candidate is not the maximal incentive compatible promise that he

---

11. That is, $A$ could make an arbitrarily small move toward $V$ in his promise that $B$ could not credibly match. This would make the voter strictly prefer candidate $A$ to candidate $B$. Rather than model this formally, we simply assume that the better candidate wins in case of a tie.
could make. Suppose $B'$ is the largest incentive compatible promise that $B$ can make, inducing $A$ to make promise $A'$. But inducing $A$ to make promise $A'$ is not the best that $B$ can do. As $A$ makes promises closer to $V$ in attempting to win the election, the outcome moves from point $A$ (if $A$ could win without making any promise) toward $V$’s ideal point. The move along this line initially increases $B$’s utility, because movements along the line $AV$ initially lead to outcomes that are closer to $B$. However, once $A$ makes promises greater than $M$, the move along the line $AV$ leads to outcomes that are farther from $B$, thereby decreasing $B$’s utility. If $B$ were to scale down his promise to $B''$ (prior to $A$ making his promise), then $A$’s response would be to scale back his promise to $M$, which is the point on the line $AV$ that gives $B$ the highest utility.

It is thus easy to see that, with this kind of manipulation, there may be no promises at all. If the lines $AB$ and $AV$ form an obtuse angle, then any movement

\[\text{Figure 1. Equilibrium with simultaneous moves in a two-dimensional policy space.}\]
along the line $AV$ reduces $B$’s utility; hence the loser would be better off making no promise than making the equilibrium promise. The analogue of this in the one-dimensional case is when both candidates are on the same side of the median voter. In the absence of any promises, the candidate whose ideal point is closer to the median will win the election. The only effective promises that the other candidate can make are promises of outcomes that are closer to the median voter’s ideal point than his rival’s ideal point, but movements in this direction decrease the candidate’s utility.

These arguments carry over generally to any number of dimensions, provided a global median exists. If a global median does not exist then there is still always an equilibrium in which no promises are made (and none believed) and the winning candidate simply implements his ideal point. This suggests that equilibria with small but positive promises may exist when $\delta$ is positive.

**Asymmetry of Issues.** The discussion so far treats the case in which all parties have circular indifference curves. Suppose the issues are not symmetric—that is, suppose the voters’ utility functions are given by $U(x, y) = -a^2(x - \bar{x})^2 -$
\[ b^2 (y - \bar{y})^2. \] One can make a change of variables with \( \tilde{x} = x/a \) and \( \tilde{y} = y/b \) so that \( U(\tilde{x}, \tilde{y}) = -(\tilde{x} - \bar{x})^2 - (\tilde{y} - \bar{y})^2. \) If we assume that the voters’ ideal points in the space of changed variables are uniform, then the foregoing analysis carries over more or less intact. As before, voters will choose the candidate whose (incentive compatible) promise is closer to their ideal point. There is one difference, however. If the candidates do not have the same utility function as the voters (i.e., if the \( a \) and \( b \) parameters in their utility function are not the same as in the voters’ utility functions), then the voters’ indifference curves after the change of variables will not be circles but rather ellipses. The promises that will be made by a candidate will still be on the locus of tangencies of the voter and candidate indifference curves but this locus will no longer be a straight line.

3. Discussion

There are several features of this model that deserve further discussion.

The Effect of Candidate Ideology on Credibility. How does the intensity of a candidate’s ideology affect his credibility? Our results have assumed that candidates’ ideal points were uniformly distributed on the unit interval. Imagine instead environments in which there is more polarization between the candidates, as captured by distributions of ideal points putting greater weight on points that are farther from the median voter. The value \( d \) measured the magnitude of candidates’ credibility in section 2.3; we are interested in whether this value would increase or decrease when there is greater polarization.

Suppose we symmetrically change the distributions of the candidates’ ideal points, putting greater weight on points farther from the median voter and less weight on points that are nearer. As before, it will still be the case that a candidate is more likely to win an election when his reputation is intact than when he has lost his reputation. The candidate whose reputation is intact benefits from this. That benefit will sometimes come about when the candidates’ ideal points are relatively close to the median voter’s and sometimes when those points are farther away from the median voter’s. The magnitude of the benefit of a good reputation will be greater when the ideal points are farther away—simply because the distance between the ideal points is larger in this case. But then the effect of an increase in ideological intensity is to put greater probability on those cases where the benefit is larger, so the value of having a good reputation is greater with the increase.

The increased value of having a good reputation when there is greater ideological intensity translates into an increase in potential credibility. Not all promises are typically believed by voters; what voters will (can) believe is limited by what the candidate has to lose by reneging after being elected. Anything that increases the value of maintaining one’s reputation also increases the loss to the candidate
should he renege and thereby increases the magnitude of the promises that he will have an incentive to keep.

**Uncertainty.** Suppose that between the voting stage and the office stage, the policy preferences of the winner suffer a shock that changes the candidate’s ideal point with some positive probability. In the case analyzed in the previous section, all promises made by a candidate during the campaign were fulfilled in equilibrium. Adding uncertainty about the candidates’ preferences alters this: now some promises that are believed in equilibrium will not be fulfilled. Furthermore, a greater probability of shocks on candidates’ preferences should also imply a lower future expected value from maintaining a good reputation (because with positive probability it will be lost in any case), and thus a lower value of reputation (lower cost of reneging). Therefore, in equilibrium we will obtain a smaller \( d \): fewer promises will be credible.

**Alternative Punishment Strategies.** We have assumed that voters’ punishment of candidates who renege is extreme: after a candidate reneges once, voters continue to punish not believing any of his promises for all future elections. There are other equilibria in which voters’ punishment is less extreme. It might be possible that, after a candidate reneges once, voters punish the candidate for a finite number of periods but afterward again believe his incentive compatible promises. Because the future expected payoffs if he reneges will be higher in equilibrium, we will obtain a lower value for maintaining reputation and thus a smaller \( d \) that is, fewer promises will be credible.

**Endogenous Entry.** We have assumed that candidate ideal points are exogenous. A natural extension of the model would be to incorporate endogenous entry of candidates. One possible approach would be to combine the random median voter model of section 2.4 with the citizen candidate approach of Besley and Coate (1997) and Osborne and Slivinski (1996). In the citizen candidate model, there is a cost \( c \) to entry, and each citizen simultaneously makes a choice to either pay \( c \) or not. Then an election occurs among all the candidates. As in our model, the citizen candidate approach assumes that candidates have policy preferences so that, in a single-election world, no promises are credible. In principle, one could extend the citizen candidate approach to a repeated elections setting that includes a “campaign promise” stage in each election. This could lead to interesting dynamic interactions between current promises by incumbents and endogenous entry of challengers in later elections. Such an extension, however, is not straightforward. The benefit of making a promise in a given period and of subsequently carrying out that promise depends on the past history of promises by all other agents. Thus, the resulting interaction is a dynamic game where the state variable is a vector
summarizing all agents’ promise histories. Restricting attention to exogenous entry, as we have done, simplifies the analysis.

Forward-Looking Voters. We have assumed that voters are myopic and do not take the future into account. If voters, like candidates, discount the future, then the equilibria we have characterized continue to be equilibria. Voters will in each period continue to choose the candidate who offers the preferred platform, because there is no link between what platform is promised and/or enacted except the extent to which promises are kept. Some additional equilibria are possible if the voters not only discounted the future but also were more strategic than in our model. Although we do not carry out the analysis, we shall outline the logic as follows. When voters are strategic, they can induce candidates to make even better promises. In our model, the winning (more moderate) candidate need only offer the median voter a platform with the same utility as that offered by the losing (more extreme) candidate. A strategic forward-looking voter might adopt a strategy that requires the more moderate candidate to make an even better (credible) promise or suffer a bad reputation, with the losing (more extreme) candidate still always making the maximum promise as in our analysis. In equilibrium, this would reduce the value to a candidate of maintaining a reputation and would therefore decrease the maximal promise that can be sustained, so the voters would face a trade-off. On the one hand, they would receive the maximal promise from the more moderate candidate every period, whereas in our model they get strictly less than that each period because the moderate candidate only matches the extreme candidate’s promise. On the other hand, the maximal promise is less and so, when the two candidates have ideal points that are nearly the same distance from the median voter, the median voter is worse off. There are many subgame perfect equilibria of this sort, and in theory one could compute the best equilibrium from the standpoint of the (strategic) median voter.

Standard Preferences. Voters and candidates in our model have standard preferences; that is, they care about the policies that will be chosen. One might argue that there are honest politicians and dishonest politicians, and that voters have a preference for honest politicians. It is certainly plausible that politicians differ in the degree to which they prefer to keep their word and that voters care about this. Candidates’ behavior in a model that incorporated these ingredients would be similar to their behavior in our model: they would hesitate to make promises that they did not intend to keep, and voters would be less likely to vote for candidates who have reneged on promises in the past.

One disadvantage of such a model is the introduction of additional arguments in candidates’ and voters’ utility functions. It is relatively easy to explain a particular phenomenon by adding parameters to a model, but formal modeling is successful when a single parsimonious model is able to account for a wide
range of phenomena. Even if one believes that candidates do have preferences for keeping their word and that voters do have a preference for honest politicians, it is valuable to know the extent to which the campaign promises that politicians make and fulfill can be understood within the standard model (i.e., without adding these parameters).

**Appendix**

**A.1. Proof of Proposition 1**

In order to find equilibrium strategies for the two candidates, we will consider three different cases: when both candidates have a bad reputation, when only one of the candidates has a good reputation, and when both have a good reputation.

Suppose that both candidates have a bad reputation. Then, given that voters do not believe any promises (other than the candidates’ ideal points), the cost of reneging is zero because no promises will be believed in any event. Therefore at the “office stage”, all candidates will always implement their ideal points. Similarly, given that the only promise that is incentive-compatible for the candidates is their own ideal point, it is optimal for the voters not to believe any other promise. Thus, at each election the winner will be the candidate whose ideal point is closer to the ideal point (zero) of the median voter and the policy implemented after the election will be the candidate’s ideal point. In this case, the expected payoff (prior to the realization of the candidates’ ideal points) for each candidate at each election is given by

\[
v_{BB} = \int_0^1 \int_{-1}^{-x_R} u_L(x_R) \, dx_L \, dx_R + \int_0^1 \int_{-x_R}^0 u_L(x_L) \, dx_L \, dx_R = -\frac{1}{2}.
\]

(See Figure A.1.)

Now suppose that candidate \( R \) has a bad reputation, which means that voters will believe that he will implement his ideal point, and candidate \( L \) has a good reputation, which means that voters believe all his promises that are consistent with the incentive compatibility constraints.

We start by assuming that voters believe all promises made by candidate \( L \) that are less than a distance \( d \) from his ideal point. Then, solving for the equilibrium strategies, we must find the maximal \( d \) that is consistent with incentive compatibility.

If \(-x_L < x_R\), then candidate \( L \) wins by promising his ideal point. In this case, he does not need to make any promises and obtains the maximal possible utility.
If $-x_L > x_R$, then candidate $L$ loses unless he can credibly promise a policy that is closer to the ideal point of the median voter than $x_R$ is. In this case candidate $R$ wins the election and implements $x_R$, otherwise, candidate $L$ may credibly promise a policy $-x_R$ that, for the median voter, is at least as good as $x_R$. Making a promise that allows him to win the election is a better strategy for $L$ than allowing $R$ to win, since $L$ gets a higher utility even if he decides to fulfill his promise:

$$u_L(-x_R) = x_L + x_R > u_L(x_R) = x_L - x_R.$$ 

In equilibrium, then, candidate $L$ promises policy $-x_R$. Voters will believe him only if he has a good reputation and if implementing $-x_R$ is incentive compatible for candidate $L$—that is, if the gain he obtains from fulfilling his promise

![Figure A.1. Winner as a function of the candidates’ ideal points when both have bad reputation.](image)
in terms of future expected payoffs is larger than the cost of reneging. In this case, candidate \( L \) wins the election.\(^{13}\)

The cost of reneging is the difference between \( L \)'s future expected payoff if he maintains a good reputation and his future expected payoff if he loses his reputation, given that candidate \( R \) does not have a good reputation.

Let \( v_{GB}(d) \) denote the one-election expected utility for a candidate with a good reputation when his opponent has a bad reputation. Similarly, let \( v_{BB}(d) \) denote the one-election expected utility for each candidate when both have a bad reputation. Then, given the assumptions of our model we have the following:

\[
v_{GB}(d) = \int_{0}^{1-d} \int_{-x_R-d}^{0} u_L(x_R) \, dx_R \, dx_L + \int_{1-d}^{1} \int_{-x_R}^{0} u_L(-x_R) \, dx_L \, dx_R \\
+ \int_{0}^{1} \int_{-x_R}^{0} u_L(-x_R) \, dx_L \, dx_R + \int_{0}^{0} \int_{-x_R}^{x_L} u_L(x_L) \, dx_L \, dx_R \\
= -\frac{1}{6} - \frac{(1 - d)^3}{3}
\]

and

\[
v_{BB}(d) = \int_{0}^{1} \int_{-1}^{-x_R} u_L(x_R) \, dx_R \, dx_L + \int_{0}^{0} \int_{-x_R}^{x_L} u_L(x_L) \, dx_L \, dx_R = -\frac{1}{2}.
\]

(See Figure A.2.)

Given the one-election expected payoffs, we can compute the expected future payoffs \( v_{GB} \) for a candidate with a good reputation given that his opponent has a bad reputation:

\[
V_{GB}(d; \delta) = \sum_{t=1}^{\infty} \delta^t v_{GB}(d).
\]

Similarly, the future expected payoffs for a candidate with a bad reputation given that his opponent also has a bad reputation are

\[
V_{BB}(d; \delta) = \sum_{t=1}^{\infty} \delta^t v_{BB}(d).
\]

Thus we obtain the cost of reneging as a function of the maximal promise believed by voters and the discount factor. Let \( C^S(d; \delta) \) denote the cost of reneging. Then

\[
C^S(d; \delta) = V_{GB}(d; \delta) - V_{BB}(d; \delta) = \frac{\delta}{1 - \delta} \frac{1}{3} (1 - (1 - d)^3).
\]

\(^{13}\) Observe that, when candidate \( L \) promises \(-x_R\), the median voter is indifferent between the two candidates. We assume that a voter who is indifferent between the two candidates will vote for the unconstrained candidate.
The gain from reneging is $d$, the maximal gain a candidate may obtain from reneging of a promise, that is, the maximal difference in utility between implementing the policy he promised and implementing his ideal point. Therefore, it is an optimal strategy for candidate $L$ to fulfill all promises that are a distance of at most $d$ from his ideal point, where $d$ satisfies $d \leq C^S(d; \delta)$.

It is also an optimal strategy for voters to believe all promises that are a distance of at most $d$ from the candidate’s ideal point, with $d$ again satisfying $d \leq C^S(d; \delta)$, because in equilibrium those promises will be fulfilled.

We denote by $d^S$ the value of $d$ that solves

$$d = C^S(d; \delta).$$
$d^S$ is the maximal promise that a candidate will always fulfill and it is also the maximal promise that voters will believe.

Because

$$\frac{\partial C^S(d)}{\partial d} = \frac{\delta}{1 - \delta} (1 - d)^2 \geq 0 \quad \text{and} \quad \frac{\partial C^S(0)}{\partial d} = \frac{\delta}{1 - \delta},$$

it follows that, in equilibrium:

1. for $\delta \leq 1/2$ we must have $d^S = 0$, so no promises are believed;
2. for $1/2 < \delta < 3/4$ we must have $0 < d^S < 1$, so some promises may be believed; and
3. for $3/4 \leq \delta \leq 1$ we must have $d^S = 1$, so all promises may be believed.

(See Figure A.3.)

**Figure A.3.** Cost of reneging to candidate 1 when only he has good reputation.
Therefore, the promises that (in equilibrium) may be believed and fulfilled are

\[
d^S(\delta) = \begin{cases} 
0 & \text{if } 0 \leq \delta \leq \frac{1}{2}, \\
\frac{3}{2} \left(1 - \sqrt{\frac{4 - 5\delta}{3\delta}}\right) & \text{if } \frac{1}{2} \leq \delta \leq \frac{3}{4}, \\
1 & \text{if } \frac{3}{4} \leq \delta \leq 1.
\end{cases}
\]

Observe that, because

\[
\frac{\partial^2 C^S(d)}{\partial d^2} = -2 \frac{\delta}{1 - \delta} (1 - d) \leq 0,
\]

the cost of reneging must be a concave function. This is intuitively plausible since a candidate benefits from an increase in the set of credible promises (i.e., an increase in \(d^S(\delta)\)) only when his ideal point is a distance of more than \(d^S(\delta)\) from the median voter’s ideal point, and the probability of this event is lower the larger is the value of \(d^S(\delta)\).

Now consider the case in which both candidates have a good reputation. Let \(v_{GG}(d)\) denote the one-election expected utility for a candidate with a good reputation when both candidates have a good reputation, and let \(v_{BG}(d)\) denote the one-election expected utility for a candidate with a bad reputation when his opponent has a good reputation. As before, we start by assuming that voters believe all promises that are a distance of at most \(d\) from the ideal point of the candidate. We then look for a function \(d^D(\delta)\) that characterizes the maximal promise that candidates will fulfill (and voters will believe) if both candidates have a good reputation. If both candidates have a good reputation—that is, if both candidates can make credible promises—then the maximal promise that is incentive-compatible could be different than the one we found in the case where only one candidate can make credible promises. Given the assumptions of our model, we now have

\[
v_{GG}(d) = \int_0^{1-d} \int_{-x_R-d}^{-x_R} u_L(x_R) \, dx_L \, dx_R + \int_d^1 \int_{-x_R+d}^0 u_L(x_L) \, dx_L \, dx_R \\
+ \int_0^d \int_{-d}^0 u_L(0) \, dx_L \, dx_R + \int_1^{-x_R+d} \int_{-x_R}^0 u_L(-x_R + d) \, dx_L \, dx_R \\
+ \int_{-d}^0 \int_{-x_L}^{x_R-d} u_L(-x_L - d) \, dx_R \, dx_L = \frac{1}{2}
\]
and

\[ v_{BG}(d) = \int_0^1 \int_{-x^R}^{-x^R + d^S} u_L(x_R) \, dx_L \, dx_R + \int_{-x^R}^1 \int_{-x^R}^{-x^R + d^S} u_L(-x_L) \, dx_L \, dx_R \\
+ \int_{-x^R}^{d^S} \int_0^{-x^R} u_L(-x_L) \, dx_L \, dx_R + \int_{-x^R}^{d^S} \int_0^{-x^R + d^S} u_L(x_L) \, dx_L \, dx_R \]

\[ = -\frac{5}{6} + \frac{(1 - d^S)^3}{3}. \]

(See Figure A.4.)

In this case the future expected payoff for a candidate who has a good reputation when the other candidate also has a good reputation is

\[ V_{GG}(d; \delta) = \delta \frac{1}{1 - \delta} v_{GG}(d) = -\frac{1}{2} \frac{\delta}{1 - \delta}. \]

Note that, when both candidates have a good reputation, their payoffs are independent of the size of the set of credible promises. This is due to the linearity of the candidates’ utility functions: The expected increase in utility for a candidate whose opponent can make promises compensates for his loss in utility that results from fulfilling his own promises. Similarly, the future expected payoff for a candidate who has a bad reputation when his opponent has a good reputation is

\[ V_{BG}(d; \delta) = \delta \frac{1}{1 - \delta} v_{BG}(d^S) = \frac{\delta}{1 - \delta} \left( -\frac{5}{6} + \frac{(1 - d^S)^3}{3} \right). \]

Observe that the expected future payoff for a candidate with a bad reputation when his opponent has a good reputation is a function of the maximal promise that voters believe when only one candidate can make promises—that is, the value \( d^S(\delta) \) that we found for the previous case—whereas the expected future payoff for a candidate with a good reputation when his opponent also has a good reputation is independent of \( d \). As a consequence when both candidates have a good reputation, the candidate’s cost of reneging is given by

\[ C^D(d; \delta) = V_{GG}(d; \delta) - V_{BG}(d; \delta) = \frac{\delta}{1 - \delta} \frac{1}{3} \left( 1 - (1 - d^S)^3 \right). \]

Comparing this cost with the results found for the case in which only one candidate has a good reputation, we conclude that

\[ C^D(d; \delta) = C^S(d^S; \delta) = d^S(\delta). \]

That is, the cost of reputation when both candidates have a good reputation equals the value of maintaining a good reputation for a candidate when his
opponent has a bad reputation, so it is equal to the maximal promise that voters believe when only one candidate has a good reputation. This implies that necessarily $d^D(\delta) = d^S(\delta)$. In other words, if both candidates have a good reputation, then the maximal promises to be fulfilled by candidates and believed by voters in

Figure A.4. Winner as a function of the candidates’ ideal points when both have good reputation.
equilibrium are the same as when only one candidate has a good reputation. We set \( d^* \) in the statement of Proposition 1 equal to this common value.

**A.2. Proof of Proposition 2**

We first consider the case in which both candidates have a bad reputation. As before, because no promises are ever believed by voters, the cost of reneging is zero and thus all candidates always implement their ideal point at the office stage. In each election the winner will be the candidate whose ideal point is closer to the median voter’s ideal point. The expected payoff (prior to the realization of the candidates’ ideal points) for each candidate at each election is given by

\[
\tilde{v}_{BB}(k) = \int_0^1 \int_{-1}^{-x_R} -(x_R - x_L)^k d x_L d x_R = \frac{1 - 2^{k+1}}{(k+1)(k+2)}.
\]

Observe that the expected payoff in this case is strictly decreasing with the degree of concavity of the candidates’ utility function:

\[
\frac{\partial \tilde{v}_{BB}(k)}{\partial k} = \frac{2^{k+1}[2k + 3 - (k + 1)(k + 2) \ln 2] - (2k + 3)}{(k+1)^2(k+2)^2} < 0.
\]

Now suppose that candidate \( L \) has a good reputation and candidate \( R \) has a bad reputation. As before, we first assume that voters believe all promises made by candidate \( L \) that are a distance of less than \( d \) from his ideal point, and we then determine the maximal \( d \) that is consistent with incentive compatibility.

The **gain from reneging** is \( d^k \), the maximal gain that a candidate may obtain from reneging on a promise—that is, the maximal difference in utility between implementing the promised policy and implementing his ideal point.

The **cost of reneging** is the difference between \( L \)'s future expected payoff if he maintains a good reputation and his future expected payoff if he loses his reputation, given that candidate \( R \) has a bad reputation. In this case, the one-election expected utility for candidate \( L \) is

\[
\tilde{v}_{GB}(d; k) = \int_0^{1-d} \int_{-1}^{-x_R-d} -(x_R - x_L)^k d x_L d x_R \\
+ \int_0^{1-d} \int_{-x_R}^{-x_R-d} -(x_R - x_L)^k d x_L d x_R \\
+ \int_{1-d}^1 \int_{-1}^{-x_R} -(x_R - x_L)^k d x_L d x_R \\
= \frac{1}{2}(2-d)^{k+2} - \frac{3}{2} d^{k+2} + 1 - \frac{d^{k+1}(1-d)}{k+1}.
\]
Given the one-election expected payoffs, we can once again compute the expected future payoffs for a candidate with a good reputation, given his opponent’s reputation, and then compute the cost of reneging as the difference between them:

\[
\tilde{C}^S(d; \delta, k) = \tilde{V}_{GB}(d; \delta, k) - \tilde{V}_{BB}(d; \delta, k) = \sum_{t=1}^{\infty} \delta^t [\tilde{v}_{GB}(d; k) - \tilde{v}_{BB}(k)].
\]

When both candidates’ utility functions are concave, the cost of reneging is given by the expression

\[
\tilde{C}^S(d; \delta, k) = \frac{\delta}{1 - \delta} (\tilde{v}_{GB}(d; k) - \tilde{v}_{BB}(k))
\]

\[
= \frac{\delta}{1 - \delta} \left[ \frac{1}{2} \left( 2^{k+2} - (2 - d)^{k+2} - 3d^{k+2} \right) \right] \frac{(k + 1)(k + 2)}{k+1} - \frac{d^{k+1}(1 - d)}{k+1}.
\]

Therefore, it is optimal for candidate \( L \) to fulfill all promises that are a distance of at most \( d \) from his ideal point, where \( d \) satisfies

\[
d^k \leq \tilde{C}^S(d; \delta, k).
\]

It is also optimal for the voters to believe all promises that are a distance of at most \( d \) satisfying this inequality, because in equilibrium these promises will be fulfilled.

Notice that the cost of reneging is increasing with the amount of promises believed by voters:

\[
\frac{\partial \tilde{C}^S(d; \delta, k)}{\partial d} = \frac{\delta}{1 - \delta} \left[ \frac{1}{2} \left( 2^{k+2} - (2 - d)^{k+2} - 3d^{k+2} \right) \right] \frac{(k + 1)(k + 2)}{k+1} - \frac{d^{k+1}(1 - d)}{k+1} \geq 0.
\]

The cost of reneging is also a concave function of the amount of promises believed by voters:

\[
\frac{\partial^2 \tilde{C}^S(d; \delta, k)}{\partial d^2} = \frac{\delta}{1 - \delta} \left[ - \frac{1}{2} (2 - d)^k + \frac{1}{2} d^k - kd^{k-1}(1 - d) \right] \leq 0.
\]

On the other hand, the gains from reneging, \( d^k \), are an increasing and convex function of the amount of promises believed by voters.

---

14. This is the case because \( \frac{\partial \tilde{C}^S(d; \delta, 1)}{\partial d} = \frac{\delta}{(1 - \delta)(1 - d)^2} \geq 0 \) and \( \frac{\partial (\partial \tilde{C}^S(d; \delta, k) / \partial d) / \partial k}{\partial d} / \partial k \) equals

\[
\frac{\delta}{1 - \delta} \left[ \frac{\frac{1}{2} (2 - d)^{k+1} [(k + 1) \ln(2 - d) - 1] + \frac{1}{2} d^{k+1} [1 - (k + 1) \ln d]}{(k + 1)^2} - d^k (1 - d) \ln d \right] > 0.
\]
We also have that \( \tilde{C}^S(0; \delta, k) = 0 \) and
\[
\tilde{C}^S(1; \delta, k) = \frac{\delta}{1 - \delta} \frac{2^{k+1} - 2}{(k+1)(k+2)} \leq 1 \iff \delta \leq \left(1 + \frac{2^{k+1} - 2}{(k+1)(k+2)}\right)^{-1}.
\]

It follows that the cost of reneging and the gains from reneging intersect at no more than one single point when \( d \in [0, 1] \). Thus, there is a value of \( d \) for which \( d^k = C^S(d; \delta, k) \), which determines the maximal promise believed by voters. Let \( \tilde{d}^S \) denote this value. Then
\[
\tilde{d}^S(\delta, k) = 0 \quad \text{if } \delta = 0,
\]
\[
0 < \tilde{d}^S(\delta, k) < 1 \quad \text{if } 0 < \delta < \left(1 + \frac{2^{k+1} - 2}{(k+1)(k+2)}\right)^{-1},
\]
\[
\tilde{d}^S(\delta, k) = 1 \quad \text{if } \delta \geq \left(1 + \frac{2^{k+1} - 2}{(k+1)(k+2)}\right)^{-1}.
\]
(See Figure A.5.)

Observe that when candidates’ utility functions are strictly concave, there are always some promises different from the candidates’ ideal points that are believable by voters—as long as the discount factor is greater than 0. And as in the linear case, when the discount factor increases the set of believable promises also increases, because
\[
\frac{\partial \tilde{C}^S(d; \delta, k)}{\partial \delta} = \tilde{C}^S(d; \delta, k)/(1 - \delta) \geq 0.
\]
Finally, if the discount factor is sufficiently large, then all promises are incentive-compatible.

We can also show that the maximal promise believed by voters increases with the degree of concavity of the candidates’ utility function:
\[
\frac{\partial \tilde{d}^S(\delta, k)}{\partial k} \geq 0.
\]
Because the cost of reneging for each value of \( d \) increases with the degree of concavity, we have\(^{15}\)
\[
\frac{\partial \tilde{C}^S(d; \delta, k)}{\partial k} \geq 0;
\]
on the other hand, the gain from reneging decreases with the degree of concavity:
\[
\frac{\partial (d^k)}{\partial k} = d^k \ln d \leq 0.
\]

\(^{15}\) This is true because: (1) \( \partial \tilde{C}^S(0; \delta, k)/\partial d \) increases with \( k \); (2) \( \tilde{C}^S(1; \delta, k) \) increases with \( k \); and (3) \( \partial \tilde{C}^S(d; \delta, k)/\partial d \) increases with \( k \).
Now consider the case in which both candidates have a good reputation. We first compute the one-election expected payoffs for a candidate with a good reputation and then for one with a bad reputation, given that the opponent has a good reputation. We assume that voters believe promises from either candidate that are a distance of at most $d$ from the candidate’s ideal point, and we look for a function $\tilde{d}(\delta, k)$ that characterizes the maximal promise that candidates will fulfill (and hence that voters will believe) given that both candidates have a good reputation. The one-election expected payoff for a candidate with a good reputation when his opponent has also a good
reputation is

\[ \tilde{v}_{GG}(d; k) = \int_{0}^{1-d} \int_{-d}^{-x_R} -(x_R - x_L)^k dx_L dx_R + \int_{-d}^{d} \int_{-d}^{0} -(x_L)^k dx_L dx_R \]

\[ + \int_{0}^{1} \int_{d}^{-x_R+d} -(x_R - x_L + d)^k dx_L dx_R \]

\[ + \int_{-1}^{-d} \int_{-x_L}^{-x_L+d} -(2x_L - d)^k dx_L dx_R \]

\[ = -\frac{1}{4} \left[ (2 - d)^{k+2} + d^{k+2} \right] + 1 + \left( \frac{d}{2} - 1 \right) \frac{d^{k+1}}{k+1} - \frac{d (2 - d)^{k+1}}{2} \]

When computing the expected utility for a candidate with a bad reputation when his opponent has a good reputation, we need to take into account that in this case the set of promises that voters believe is given by the function \( \tilde{d}^S(\delta, k) \) found previously; thus,

\[ \tilde{v}_{BG}(\tilde{d}^S(\delta, k); k) = \int_{0}^{1} \int_{-x_R}^{-} -(x_R - x_L)^k dx_L dx_R + \int_{-x_R+\tilde{d}^S}^{0} -(2x_L)^k dx_L dx_R \]

\[ - \int_{x_R}^{0} -(2x_L)^k dx_L dx_R \]

\[ = \frac{1}{4} 2^{k+2} \left[ (1 - \tilde{d}^S)^{k+2} - 3 \right] + 1. \]

Given the one-election expected utilities, we again find the value of the future expected payoffs and the cost of reneging as

\[ \tilde{C}^D(d; \delta, k) = \tilde{V}_{GG}(d; \delta, k) - \tilde{V}_{BG}(d; \delta, k) \]

\[ = \frac{\delta}{1 - \delta} \left[ \tilde{v}_{GG}(d; k) - \tilde{v}_{BG}(\tilde{d}^S(\delta; k); k) \right]. \]

Using the previous expressions, we can obtain the cost of reneging as a function of the size of the set of credible promises when both candidates have a good reputation compared to the maximal amount of credible promises when only one candidate has a good reputation:

\[ \tilde{C}^D(d; \delta, k, \tilde{d}^S) = \frac{\delta}{1 - \delta} \left[ -\frac{1}{4} \left[ (2 - d)^{k+2} + d^{k+2} \right] - \frac{1}{4} 2^{k+2} \left[ (1 - \tilde{d}^S)^{k+2} - 3 \right] \right. \]

\[ + \left. \left( \frac{d}{2} - 1 \right) \frac{d^{k+1}}{k+1} - \frac{d (2 - d)^{k+1}}{2} \right]. \]
First notice that, for all $\tilde{d}^S(\delta, k) > 0$, if voters believe no promises other than the candidates’ ideal points (when both candidates have a good reputation), then the cost of reneging is still positive (and recall that $\tilde{d}^S = 0$ only when $\delta = 0$):

$$
\tilde{C}^D(0; \delta, k, \tilde{d}^S) = \frac{\delta}{1 - \delta} \left[ \frac{2^k \left[ 1 - (1 - \tilde{d}^S)^{k+2} \right]}{(k + 1)(k + 2)} \right] > 0.
$$

This implies that the cost of losing a good reputation for one of the candidates (when both have a good reputation) might be positive, even if no promises are believed by voters. This can happen if some promises are believed by voters only when a single candidate has a good reputation. The reason for this anomaly is that, if a candidate were to lose his reputation, then we would revert to the state in which only one candidate has a good reputation—that is, the state in which the amount of credible promises is given by $\tilde{d}^S > 0$. In that state the candidate with a bad reputation is worse off than when both have good reputations, even if no promises are believed in that case.

Furthermore, the cost of reneging in this case is increasing with the size of the set of believable promises:16

$$
\frac{\partial \tilde{C}^D(d; \delta, k)}{\partial d} = \frac{\delta}{1 - \delta} \left[ \frac{d(2 - d)^k}{2} - \frac{1}{2}(2 - d)d^k \right] \geq 0.
$$

We can also show that, for low values of $d$, $\tilde{C}^D(d; \delta, k)$ is a convex function of $d$ and, as $d$ increases, $\tilde{C}^D(d; \delta, k)$ becomes a concave function:

$$
\frac{\partial^2 (\tilde{C}^D(d; \delta, k))}{\partial d^2} = \frac{\delta}{1 - \delta} \left[ 1 - \frac{d}{2} (k + 1) \right] [(2 - d)^{k-1} - d^{k-1}].
$$

Moreover, $\partial^2 (\tilde{C}^D(d; \delta, k))/\partial d^2 \leq 0$ if and only if $d \geq 2/(k + 1)$.

For a given value of $k$, the maximal credible promise, denoted by $\tilde{d}^D(\delta, k)$, is given by the largest value of $d$ that satisfies $\tilde{C}^D(d; \delta, k) \geq d^k$ (see Figure A.6).

16. This is the case because $\partial (\tilde{C}^D(d; \delta, 1))/\partial d = 0$ and

$$
\frac{\partial (\tilde{C}^D(d; \delta, k))}{\partial d} = \frac{\delta}{1 - \delta} \left[ \frac{d}{2} (2 - d)^k \ln(2 - d) - \frac{1}{2}(2 - d)d^k \ln d \right] \geq 0.
$$
In this case we also have that the size of the set of credible promises increases with the value of the discount factor if $\tilde{d}^D(\delta, k) > 0$:

\[
\tilde{d}^D(\delta, k) = 0 \quad \text{if } \delta = 0,
\]

\[
0 < \tilde{d}^D(\delta, k) < 1 \quad \text{if } 0 < \delta < \left(1 + \frac{2^{k+2} - (1 - d^S)^{k+2}}{(k + 1)(k + 2)}\right)^{-1},
\]

\[
\tilde{d}^D(\delta, k) = 1 \quad \text{if } \delta \geq \left(1 + \frac{2^{k+2} - (1 - d^S)^{k+2}}{(k + 1)(k + 2)}\right)^{-1}.
\]
Finally, we have that the cost of reneging for all \( d \) is an increasing function of \( k \);\(^{17}\) that is,

\[
\frac{\partial C^D(d; \delta, k)}{\partial k} \geq 0 \quad \text{and} \quad \frac{\partial \tilde{d}^D(\delta, k, \tilde{d})}{\partial k} \geq 0.
\]

Because we have already shown that the gain from reneging for all \( d \) decreases with \( k \), it follows that the value of the maximal credible promise increases as \( k \) becomes larger.

### A.3. Proof of Proposition 3

Consider first the case in which one of the candidates \( (L) \) has a good reputation while the other candidate \( (R) \) has a bad reputation. Then, the expected payoff from one election for candidate \( L \) are

\[
v_{BB}(d) = \frac{1}{2}u(0) + \frac{1}{2}u(1) = -\frac{1}{2}
\]

and

\[
v_{GB}(d) = \frac{1}{2}u(0) + \int_{1/2}^{(1+d)/2} u(2m - 1)dm + \frac{1-d}{2}u(1)
\]

\[= -\frac{1}{2} + \frac{d}{2} \left( 1 - \frac{d}{2} \right).
\]

Hence the cost of reneging when the opponent has a bad reputation is

\[C^S(d) = V_{GB}(d; \delta) - V_{BB}(d; \delta) = \frac{\delta}{1-\delta} \frac{d}{2} \left( 1 - \frac{d}{2} \right).
\]

Because the maximal gain from reneging is \( d \), we have that the maximal promise that is incentive compatible is

\[
d^S(\delta) = \begin{cases} 
0 & \text{if } \delta \leq \frac{2}{3}, \\
\frac{3\delta - 2}{\delta} & \text{if } \frac{2}{3} < \delta \leq \frac{4}{5}, \\
1 & \text{if } \frac{4}{5} < \delta.
\end{cases}
\]

(See Figure A.7.)

---

\(^{17}\) This is true because: (1) \( \tilde{C}^D(0; \delta, k) \) increases with \( k \); (2) \( \tilde{C}^D(1; \delta, k) \) increases with \( k \); and (3) \( \frac{\partial \tilde{C}^D(d; \delta, k)}{\partial d} \) increases strictly with \( k \) for all \( d \in (0, 1) \). Therefore, if \( k < k' \) then, for all \( d < d' \),

\[
\tilde{C}^D(d; \delta, k') - \tilde{C}^D(d; \delta, k) < \tilde{C}^D(d'; \delta, k') - \tilde{C}^D(d'; \delta, k).
\]
As before, the maximal promise that is credible in equilibrium when only one candidate has a good reputation is an increasing function of the discount factor. For small values of the discount factor \( \delta \leq \frac{2}{3} \) no promises are believed, whereas for large values \( \delta \geq \frac{4}{5} \) all promises are believed.

Now consider the case in which both candidates have a good reputation. The expected payoffs from one election for candidate \( L \) are

\[
v_{GG}(\delta) = \frac{1 - d}{2} u_L(0) + \int_{(1-d)/2}^{1/2} u_L(2m - 1 + d) \, dm
+ \int_{1/2}^{(1+d)/2} u_L(2m - d) \, dm + \frac{1 - d}{2} u_L(1)
= -\frac{1}{2}
\]
and

\[ v_{BG}(\delta) = -1 - v_{GB}(\delta) = -\frac{1}{2} - \frac{d^S}{2} \left( 1 - \frac{d^S}{2} \right). \]

So the cost of reneging in this case is

\[ C^D(d; \delta) = V_{GG}(d; \delta) - V_{BG}(d; \delta) = \frac{\delta}{1 - \delta} \frac{d^S}{2} \left( 1 - \frac{d^S}{2} \right) = d^S(\delta). \]

Therefore, in this case also we will have that \( d^D(\delta) = d^S(\delta) \) —that is, the maximal credible promise when both candidates have a good reputation coincides with the maximal credible promise a candidate can make when his opponent has a bad reputation.

**References**


