The Hidden Cost of Direct Democracy:
How Ballot Initiatives Affect Politicians’ Selection and Incentives

July 1, 2015

Abstract

Citizen initiatives and referendums play an important role in modern democracies, from treaty ratifications in the European Union to gay marriage in California, to the control of foreign workers in Switzerland. Departing from the classic opposition between direct and representative democracy, we study the equilibrium effects of direct democracy institutions on the incentives and selection of elected officials. We find that facilitating direct democracy induces a negative spiral on politicians’ role and contribution to society, which may dominate any direct benefit. The theory offers predictions on reelection probabilities and politicians’ performance consistent with recent evidence from the U.S. states.
1 Introduction

Direct democracy institutions, such as citizen initiatives and referendums, play an important role in regimes otherwise based on representative democracy: In the period 1990-2010, almost 800 ballot initiatives were proposed in U.S. states (against the 433 of the two preceding decades), of which roughly 45% were approved.\footnote{Data available at www.iandrinstitute.org/data.html.} Since 1990, referendums took place in 91 countries, including 30 European ones (Kaufmann and Waters, 2004). Following the collapse of the Soviet Union, 27 of the 30 newly established democracies adopted direct democracy institutions. Since 2012, direct democracy is also available at the European Union level via the European Citizens’ Initiative.

Understanding the impact of these institutions on the democratic process is thus of prime importance. Scholars of direct democracy, however, have traditionally focused on two comparatively narrow themes. One, dating back to Plato’s Republic (Book VIII),\footnote{The Federalist Papers (1788) feature another eminent exposition of this theory: Madison highlights the risk of collective decision made “not according to the rules of justice and the rights of the minor party, but by the superior force of an interested and overbearing majority.”} is the tyranny of the majority: Larger social groups can systematically impose their preferences on minorities. The other theme emphasizes voters’ lack of wisdom and expertise (Campbell et al., 1960; Maskin and Tirole, 2004).

Crucially, both arguments – as well as important counter-arguments (Lupia, 1994; Bowler and Donovan, 1998; Garrett and McCubbins, 2008) – focus on a single policy decision: They do not study how direct democracy affects politicians’ incentives and the overall democratic process. Clearly, not all decisions can be made through initiatives and referendums. In modern, hybrid democracies (Garrett, 2005), citizens still delegate many decisions to elected representatives. This paper departs from the existing literature by studying how direct democracy on some issues affects the quality of politicians’ decisions on all issues. This question seems especially important in light of the growing body of empirical work on the topic.\footnote{See, for example, Bali and Davis (2007), Kelleher and Wolak (2007), Matsusaka (2008, 2010), Dyck (2009), Dyck and Lascher (2009), Rydberg (2010), Funk and Gathmann (2011, 2013), Hinnerich and Pettersson-Lidbom (2014).}

Most theoretical studies of the interaction between elected officials’ behavior and direct democracy (Matsusaka, 1992; Gerber, 1996; Matsusaka and McCarty, 2001) focus on the following intuitive point: Allowing direct democracy typically improves the congruence between policies and the elec-
torate’s preferences. This intuition is in line with recent evidence on issues such as abortion rights and death penalty (Matsusaka, 2010), as well as with historical accounts of the origins of direct democracy in the U.S. (Cronin, 1989), where citizens’ demand for increased ability to control potential abuses and failures of representative democracy played a key role.

Despite its intuitive appeal, the previous argument is incomplete: To assess how efficient direct democracy is at preventing political failures, one should also consider how direct democracy affects politicians’ equilibrium incentives, expertise, and selection. If direct democracy is effective at correcting politicians’ mistakes, how does that affect the frequency of such mistakes? And how does it affect the risk of mistakes along dimensions which are not amenable to direct democracy?

As a first step towards answering these questions, this paper analyzes a principal-agent model of electoral control. The expertise (built on costly information acquisition) and competence of an elected official are endogenously determined through electoral selection and incentives. Direct democracy is modeled as citizens’ ability to amend some policies chosen by their elected official. The main finding is that increasing voters’ amendment power undermines their ability to credibly reward expertise acquisition by an incumbent and to learn about her competence. Besides its normative implications, addressing the equilibrium effects of direct democracy on elected officials’ incentives and behavior also helps to explain recent empirical evidence (Bali and Davis, 2007; Dyck, 2009; Dyck and Lascher; Kelleher and Wolak, 2007; Rydberg, 2010) that is hard to reconcile with previous theories.

**Direct Democracy and Responsibility Substitution**

In the model, a politician’s competence level affects two dimensions of policymaking. One dimension can be amended through direct democracy. Examples include economic decisions, such as banning golden parachutes for CEOs or limiting property taxes (California’s Proposition 13), social decisions, such as Switzerland’s recent decision to limit immigration or forbidding the construction of minarets, and political ones, such as Italy’s repeal of proportional rule in 1991 and 1993. The second dimen-

---

4This motivation, put forth by the Progressive Movement, is summarized in Theodore Roosevelt’s “Charter of Democracy” speech (1912): “I believe in the initiative and the referendum, which should be used not to destroy representative government, but to correct it when ever it becomes misrepresentative. ... In the great majority of cases it is far better that action on legislative matters should be taken by those specially delegated to perform the task; in other words, that the work should be done by the experts chosen to perform it. But where the men thus delegated fail to perform their duty, then it should be in the power of the people themselves to perform the duty.”
sion, by contrast, is not amenable to direct democracy. Some examples include the management of international conflicts, the ability of attracting public funds and, more generally, any policy which requires access to sensitive information or high-level interactions with foreign elected officials.

The quality of a policy typically depends on the intrinsic ability of the policymaker, as well as her effort to acquire expertise on that particular policy. These two components are modeled separately: a politician’s competence is viewed as her intrinsic type, and is assumed to be exogenous and constant throughout the game. Expertise, by contrast, is the result of an information-acquisition task. The effort devoted to this task is endogenous: in equilibrium, expertise depends on politicians’ electoral incentives as well as innate ability.

Citizens cannot directly observe competence. However, their ability to assess the competence of an incumbent politician is increasing in the effort that she exerts: If voters expect a politician to spend little time working on a given policy issue, they will hardly learn anything about her competence by observing outcomes along that issue. If, on the other hand, they expect her to invest a great deal of effort and resources, outcomes should be highly correlated with her abilities.

Suppose, in that context, that direct democracy becomes feasible along one policy dimension. Specifically, suppose that citizens can avoid at least some of the negative consequences of a bad policy along that dimension by amending it. While this ability clearly carries positive value, it also entails indirect effects. The first negative consequence is a responsibility substitution effect: Citizens’ improved ability to “fix” the incumbent’s mistakes reduces their ability to credibly commit to reward effort (via equilibrium reelection incentives). This increased leniency reduces the incumbent’s incentives to exert effort, and thus increases the probability that she chooses a bad policy.

This effect is just the first iteration of a disincentive spiral. Lower effort reduces citizens’ ability to assess the politician’s competence, and therefore their expected utility along the second dimension, which cannot be amended. Since citizens are now more concerned about the incumbent’s competence than her effort, they require a higher posterior belief about her competence, conditional on observing a bad policy, to reelect her. This is only possible, however, if bad policies are less revealing of

---

5 The use of direct democracy in ratifying international treaties is explicitly prohibited by the Italian Constitution.
6 The entire analysis concerns the probability of reelection following a poor policy. It is easy to show, and intuitive, that successful policies should always be rewarded by reelection of the incumbent.
incompetence than before. This, in turn, implies that the incumbent’s equilibrium effort must go down.

While the precise mechanics of this disincentive spiral are beyond the scope of this introduction, its implications are clear: The introduction (or facilitation) of direct democracy have negative indirect effects on citizens’ welfare. As it turns out, this negative spiral can dominate the initial benefit of direct democracy. By weakening the role of elected officials, direct democracy can have perverse effects on citizens’ ability to select good politicians. These effects can prove extremely damaging, as they not only increase the risk of bad policies along dimensions which citizens can correct, but also along dimensions that citizens cannot directly affect. As a result, voters’ expected utility is non-monotonic in the cost of direct democracy and, under certain circumstances, citizens might be better off without direct democracy. Strikingly, this latter result may occur even when direct democracy entails no legal, “technological,” or cognitive cost whatsoever.

Implications

The indirect effect of direct democracy uncovered here is worth emphasizing for several reasons. First, it is derived in a framework where earlier criticisms of direct democracy (such the tyranny of the majority or voters’ lack of knowledge) do not apply. It is a pure and novel equilibrium phenomenon, completely independent of other considerations. Secondly, recent theories have cast a homogeneously favorable light on direct democracy institutions (Matsusaka, 1992; Matusaka and McCarty, 2001; Boehmke and Patty, 2007; Besley and Coate, 2008; Le Bihan, 2013). Ignoring the equilibrium effects of such institutions on the role of politicians seems dangerous, as this paper illustrates.

The analysis also sheds light on key empirical observations about direct democracy. Firstly, the theory explains why incumbents in U.S. states with initiatives seem to enjoy a lower level of support (Dyck, 2009), but display higher reelection rates, especially conditional on poor performance (Kelleher and Wolak, 2007; Bali and Davis, 2007). Secondly, the model provides an interpretation for the recent increase in the frequency of initiatives, as well as the recent criticism that these institutions have received (Cronin, 1989; Gerber, 1999; Broder, 2000).²

---

²For example, in Arizona, where initiatives that restricted bilingual education in 2000, or, in 2004 and 2006, denied benefits and access to state education to undocumented immigrants, or in California (see, for example, “Tyranny of the
The model also makes a contribution to contract theory: It formalizes relational contracts (Macaulay, 1963; Macneil, 1978; Levin, 2003) with both adverse selection and moral hazard and studies how the principal’s ability to correct the agent’s actions affect equilibrium incentives and selection. The present mechanisms and effects, such as the disincentive spiral and the responsibility substitution effect (see Section 5), the structure and uniqueness of a “natural” equilibrium (see Appendix B.1) and the technical condition uncovered here regarding the discerning power of incentives (Appendix B.3) may be applied to other economic settings, such as the interaction between an employee/contractor and a manager/owner with limited contractual instruments.

The paper is organized as follows. Section 2 contains the literature review. Section 3 introduces the model and discusses the key assumptions of the paper. Section 4 characterizes the equilibria of the model. Section 5 exposes the disincentive spiral induced by direct democracy and its implications for citizens’ welfare. Section 6 discusses the testable implications of the model, and relates them to recent empirical evidence. Appendix A contains the proofs omitted from the main text. Appendix B shows that the class of equilibria studied in the paper are optimal from the voters’ viewpoint and are the only ones surviving in a generalized version of the model (where a few results of technical interest are also established, and the consequences of relaxing some other key assumptions are discussed).

2 Related literature

Most theories of direct democracy have focused on its comparison with representative democracy. Closer to the present paper, some theories consider the impact of direct democracy institutions on representative democracy. These works have typically studied distributional or ideological issues in the context of a spatial model, showing that direct democracy can improve policies by introducing competition in policymaking. Direct democracy may help alleviate political inertia (Matsusaka, 1992) and can allow a better expression of citizen preferences by unbundling multidimensional issues.

---

8 Persson and Tabellini (1994) show how electing a representative agent can help solve the credibility problem associated with dynamic capital taxation. In Redoano and Scharf (2004) and Loeper (2011), direct and representative democracy are compared in a setting with multiple districts and inter-jurisdictional externalities (see also Maskin and Tirole, 2004).
(Besley and Coate, 2008; Matsusaka, 2008). Initiatives and referendums may also benefit the electorate by creating informational cues that disseminate socially beneficial knowledge (Boehmke and Patty, 2007).

In contrast to these works, the present paper focuses on the equilibrium effects of direct democracy institutions on the role of elected officials, including their incentives, selection, and policies. Similar questions have been studied in a static setting. In Matsusaka and McCarty (2001), direct democracy translates into giving amendment power to an interest group over the policy chosen by an elected official. In Gerber (1996), the threat of an initiative endogenously increases the alignment between voters’ preferences and implemented policies. A similar idea is featured in Le Bihan (2013) and Hugh-Jones (2012). Those works focus on static agency problems and ignore the connection between political delegation, information acquisition, and political selection.

The responsibility substitution effect identified in this paper is related to recent work identifying indirect drawbacks of ex post control in other political arenas. For instance, the legislative oversight of bureaucracy (Bueno de Mesquita and Stephenson, 2007) can distort the allocation of effort across observable and non-observable tasks by bureaucrats. Similarly, the judicial review of elected officials (Fox and Stephenson, 2011) can exacerbate a politician’s incentives to posture.

Like the present paper, Kessler (2005) studies a model where agents can acquire costly information about the mapping between policies and outcomes. Unlike the present paper, she contrasts pure direct legislation to representative democracy. However, the observation that representative democracy, by concentrating rewards from policymaking to a single agent, provides stronger incentives for acquiring information is an intellectual precursor of the disincentive spiral described here. The present paper also shares some characteristics with Besley and Smart (2007), in which voters use elections to discipline the incumbent and select honest politicians who do not act on their hidden information for the second period. In contrast to that work, here politicians have heterogeneous abilities and acquire information endogenously to produce good policies.

---

9 In Le Bihan (2013) initiatives reduce the opportunity cost for non-congruent politicians to act on their policy preferences, rather than on voters’, even on issues not amenable to direct democracy. In High-Jones (2012), non-congruent politicians choose to adopt direct democracy to “tie their hands” and ultimately improve their long term electoral chances. The model provides an explanation for the higher electoral survival of incumbents under direct democracy, but not for the lower approval rate that they typically enjoy.
3 Model

The model has two periods, and features two politicians—an incumbent and a challenger—and a polity of voters which acts as a unitary agent. The incumbent is in office in the first period, and the voters decide whether or not to replace her with a challenger. Incumbent and challenger are drawn from the same pool of politicians. A politician privately observes his type $\theta$, which can either be competent ($\theta = C$) or incompetent ($\theta = N$) and is constant over time. It is common knowledge that the share of competent types is $q$.

In each period, the politician in office chooses a policy $x$. The mapping between policies and voter payoff depends on a state $h$. For simplicity, both the state and the policy are assumed to be binary, taking values in $\{R, S\}$. To make the model more concrete, $S$ is interpreted as a default policy, and $R$ as a policy change (see Appendix B.4 for the alternative scenario).

While unable to directly observe the state, the representative can acquire costly expertise ($e \in [0, 1]$) that enhances her ability to learn the state. Specifically, when $h = R$ a politician with expertise $e$ gets no signal with probability $1 - e$ and learns the state with probability $e$, while she gets no signal whenever $h = S$.[10] The cost of acquiring expertise level $e$ for a politician of type $\theta$ is $\tau_\theta(e) = e^2 / 2\tau_\theta$, where $\tau_\theta$ is (with a slight abuse of notation) a productivity parameter that is increasing in competence.[11] The ratio $\tau_N / \tau_C < 1$ is denoted by $\tau$. The state $h$ is i.i.d. across periods, equal to $S$ with probability $\rho$ and to $R$ with probability $1 - \rho$. To avoid corner solutions, the condition $\rho\tau_C < 1$ is imposed throughout the analysis.

At the end of the first period, voters observe both the policy and the state, and decide whether to reelect the incumbent or to replace her with a challenger randomly drawn from the pool of politicians. Voters are unable to directly observe politicians’ competence and expertise, which are both relevant to their second-period utility (see below). The assumption, here, is that a politician’s competence affects two policy dimensions: The policy for which she acquires expertise, and a second policy dimension which is not amenable to direct democracy (many such policies exist in practice, see the introduction and the discussion below).

[10]The signal structure here is chosen for simplicity (see Appendix B.4), and follows Bénabou and Tirole (2006) and Tirole (2009).
[11]The quadratic formulation is chosen for simplicity, and is generalized in Appendix B.
In the second period, a new state is drawn. An elected challenger can only acquire new expertise. A reelected incumbent, instead, can also rely on past expertise to detect the state. In other words, policy expertise is assumed to be persistent.

Voters’ utility has two separable components, corresponding to the two policy dimensions. Let \( u(x_t, h_t) \) denote voters’ utility in period \( t \in \{1, 2\} \) when the state is \( h_t \) and the policy is \( x_t \). The first-policy component simply yields \( u(x_1, h_1) + u(x_2, h_2) \) (discounting across period is irrelevant to the analysis and omitted). The utility from the second dimension equals \( V \) if the politician is competent and 0 otherwise, and is received at the end of the second period. Thus, the overall utility is equal to

\[
 u(x_1, h_1) + u(x_2, h_2) + I_{\theta=C} V
\]

The utility is normalized as follows: \( u \) equals zero if \( x = S \), regardless of the state; \( u(R, S) = -1 \) (bad reform), and \( u(R, R) = B > 0 \) (good reform). In the baseline model, it is assumed that \( x = R \) is undesirable in the absence of additional information beyond the prior: \( \rho B < 1 - \rho \). Therefore, choosing \( x = R \) is efficient only after receiving the signal.

A politician gets a payoff of 1 from being in office. This reelection payoff is, of course, what motivates the incumbent to put effort in the first period. Notice that, once in office in the second and final period, no politician has incentive to acquire additional expertise. As such, it is assumed that in the second period politicians in office choose the optimal policy conditional on their information. As a consequence, a challenger who cannot rely on previous expertise chooses policy \( S \). A reelected incumbent with expertise \( e \), instead, learns that the state is \( R \) with probability \( e \), in which case she chooses \( x = R \), and chooses \( S \) otherwise.

Voters’ strategy is defined by the reelection probability \( f(x, h) \) for each first period outcome, chosen optimally given their belief about the incumbent’s expertise and competence, and by whether or not an initiative is proposed.

**Initiatives**

In this environment, it seems natural to model direct democracy as a policy amendment to the elected official’s decision. In particular, it is assumed that citizens can modify the second-period policy after
learning the state. In practice (this is the only relevant amendment under the assumptions), citizens can implement the reform $R$ to match a realization of state $h = R$ that was undetected by the politician in office\textsuperscript{12}

The cost of the amendment is denoted by $c$, and has several interpretations. First, if $c$ is above $B$, it makes the exercise of direct democracy so costly that it amounts to ruling it out. As a consequence, $c$ is restricted without loss of generality to the interval $[0, B]$. If $c = 0$, on the other hand, direct democracy bears no direct cost. In between, $c$ captures (i) voters’ inability to fully undo the effects of an unwelcome policy, due to well documented lags or imperfections (Gerber et al., 2004) in the implementation of direct legislation, (ii) legal costs (e.g., signature requirements), and (iii) technological and cognitive costs (e.g., the cost of producing, disseminating, and acquiring detailed policy information).

Citizens’ expected utility in the second period can be then written as follows: if the incumbent with type $\theta$ is reelected, $U(\theta, e) = B\epsilon + (1 - e)(B - c)\rho + I_{\theta = C}V$, where $e$ is the expertise of the incumbent, and $\theta$ is (with a slight abuse of notation) her type. If a challenger is elected, the expected utility is simply $U = \rho(B - c) + qV$. The difference $(B - c)$ is the direct benefit of direct democracy in case a useful reform has not been undertaken by the elected official (which happens with probability $\rho(1 - e)$ when the incumbent is reelected, and with probability $\rho$ when a challenger is in office).

Finally, it is assumed that $B\rho < qV$, which guarantees that, if voters could observe the type of the incumbent, they would always reelect a competent type and always oust an incompetent type, regardless of her expertise:\textsuperscript{13} $\forall e \in [0, 1], U(C, e) > U > U(N, e)$.

In summary, the timing of the game is as follows:

- Nature draws the type $\theta$ of the incumbent and the state $h_1$ in the first period.
- The incumbent chooses expertise $e$, possibly observes a signal about $h_1$, and then chooses $x_1$.
- Citizens observe $h_1$ and $x_1$ and vote between the incumbent and the challenger.
- Nature draws $h_2$ and, if applicable, the type of the challenger.

\textsuperscript{12}Among existing direct democracy institutions, an ex post policy amendment would formally correspond to the notion of “popular referendum,” as discussed by the Founding Fathers of the U.S. constitution. However, here the policy is amended by citizens only in favor of the reform, which is closer, in spirit, to a citizen initiative.

\textsuperscript{13}The assumption is made for convenience, as it avoids corner equilibria.
• If the incumbent is reelected, she may observe a signal about $h_2$ (depending on $e$ and $h_2$).
• The politician elected to the office chooses $x_2$.
• $h_2$ and $x_2$ are observed. If $h_2 = R$ and $x_2 = S$, citizens may amend $x_2$ to $R$ at cost $c$.
• Everyone gets their utility and the game ends.

**Discussion**

While Appendix B shows that the insights of this baseline model are robust to several alternative or more general specifications, a few observations are in order.

First, assuming that voters are perfectly homogeneous avoids concerns about the tyranny of the majority (orthogonal to the analysis). However, it also generates important implications for the interpretation of the policy dimension $x$. While direct democracy sometimes affects divisive ethical issues, such as the death penalty, the available evidence (e.g., Matsusaka, 2010) shows that many ballot initiatives are about issues with important common value aspects, such as labor and product market regulation, term limits, and school funding.

Second, the model allows both initiatives and the realization of the utility from competence ($V$) only in the second period. The reason is that the strategic environment of the model is unaffected by first period quantities. As such, having initiatives and/or competence payoff (as long as competence is unobserved before the election) also in the first period has no consequence on the equilibrium analysis.

Third, the model assumes complementarity in expertise across policy dimensions. However, the second policy dimension can be interpreted more broadly: the competence payoff $V$ can represent unforeseen (or low-probability) contingencies, such as a terrorist attack or a financial crisis, or the gain from having a “dedicated” politician, who is more inclined to work in the interest of the population and less prone to divert away public resources for personal reasons.

Fourth, the model’s asymmetries do capture real phenomena. In particular, (i) incumbents have the opportunity of acquiring policy expertise that challengers cannot access, (ii) incumbents have many reasons to have voters fix their polices rather than attempting to do so themselves (interest groups, reputational risk, organizational rigidities), and (iii) assuming that voters are more informed *ex post*
about policy is not equivalent to assuming that they have more expertise in general (fixing a policy mistake ex post is not equivalent to getting it right from the start).

4 Equilibrium and political contract

This model features several equilibria, including one in which both types of incumbent acquire zero expertise and are always reelected. It also includes a Pareto dominated equilibrium (see Appendix B.1) in which the incumbent acquires expertise but does, in the first period, the opposite of what her signal tells her to do, as a way of signaling her competence. The present analysis focuses on the more relevant and realistic equilibria in which (1) the incumbent chooses the reform $R$ if and only if the need of a reform is revealed to her, and (2) the expertise acquisition is strictly positive for at least a subset of parameters’ space where the equilibrium is defined. Those equilibria are referred to as natural. The next section shows that there are two natural equilibria, and the paper will select among them on grounds of empirical plausibility and theoretical robustness.

4.1 Reelection probabilities and incumbent’s choice

Equilibria are characterized by backward induction. Let $e_\theta$ denote the expertise level of each type $\theta$. Taking these levels as given, voters reelect the incumbent conditional on observing the mismatch $(S, R)$ if and only if their posterior belief about her competence is above some threshold, denoted by $\bar{p}^r$, which is given by the following formula.

**Proposition 1** The posterior threshold for reelection of the incumbent is given by

$$\bar{p}^r = \frac{qV - e_N c \rho}{V + c \rho (e_C - e_N)}.$$

**Proof.** Voters’ expected second-period utility from reelecting the incumbent, given a posterior $p$, is

$$p[V + \rho e_C B + \rho (1 - e_C)(B - c)] + (1 - p)[\rho e_N B + \rho (1 - e_N)(B - c)],$$
while their expected second-period utility with a challenger is $qV + \rho(B - c)$. Comparing these expressions yields the threshold. ■

In Appendix B, it is also shown that it is interim and ex ante optimal for the voter to always reelect when the policy matches the state, an not to reelect when an unnecessary reform is implemented $f(S, S) = f(R, R) = 1, f(R, S) = 0$.

In the first period, the incumbent chooses the expertise level $e_0 \in [0, 1]$ which solves the optimization problem $\max_{e \in [0, 1]} \left[ (1 - \rho) + \rho e + (1 - e)\rho f \right] - \tau_0(e)$, where $f$ is her reelection probability in case of a mistake. (Recall that the reelection payoff has been normalized to 1). Convexity of the cost function implies that the optimal expertise level is characterized by the first-order condition $e_0 = (1 - f)\tau_0\rho$.

4.2 Equilibria

In this setting, there may exist two natural equilibria: a responsive equilibrium, where the incumbent is elected with strictly positive probability even after making a mistake (i.e., $f > 0$), and an unresponsive equilibrium, where failure of implementing $R$ when the state is $R$ leads to the demise of the incumbent with probability one ($f = 0$). Let $\phi_1 = \frac{\rho}{qV}$ and $\phi_2 = \frac{\rho}{(1-q)V}$.

**Proposition 2 (Equilibrium)** There exists a threshold $c_2 > 0$ such that no natural equilibrium exists when $c > c_2$. When $c \leq c_2$, there are two natural equilibria: (i) an unresponsive equilibrium, characterized by the reelection probability $f = 0$; (ii) a responsive equilibrium, where $f = f^* \in [0, 1]$, which is implicitly defined by

$$Z(f) = e_N(f^*)[1 + c\phi_1] - e^2_N(f^*)c\phi_1 - e_C(f^*)[1 - c\phi_2] - e^2_C(f^*)c\phi_2 = 0 \quad (1)$$

and such that there exists a threshold $c_1 \in [0, c_2]$ such that, $f^* < 1$ whenever $c \in [c_1, c_2]$ and $f^* = 1$ otherwise.

A direct implication of Proposition 2 is that the responsive equilibrium is associated with positive expertise only when the cost of direct democracy $c$ is large enough.

---

14The equation uses the result, established in Appendix B.1, that the incumbent is surely reelected if she manages to detect the reform state $R$ or if voters observe $h = S$. 

13
4.3 Equilibrium selection

The game’s two natural equilibria (responsive and unresponsive) have opposite normative implications for the value of direct democracy and make starkly different predictions for its impact on politicians’ and voters’ behavior.

In the unresponsive equilibrium, any incumbent’s mistake is punished maximally \((f = 0)\) and independently of all parameters of the model, including the cost of direct democracy. Compared to the responsive equilibrium, this maximal punishment results in a higher expected utility for the voter and a lower expected utility for the incumbent (regardless of his type).\(^{15}\) As a consequence, if voters’ and politicians’ behaviors empirically respond to the presence of direct democracy, then it must be that incumbents, not voters, are their highest payoff among natural equilibria. Indeed, a criterion based on the incumbent’s equilibrium payoff (as, for example, in Eguia and Shepsle, 2015) selects the responsive equilibrium.

Direct democracy does seem to influence various aspects of political outcomes. In particular, direct democracy systematically affects voters’ opinions and voting behavior (Bowler and Donovan, 1998; Dyck, 2009; Dyck and Lascher, 2009), as well as on politicians’ actions, characteristics, and electoral prospects (Bali and Davis, 2007; Funk and Gathmann, 2011 and 2013; Rydberg, 2010).

There are also theoretical reasons why voters might be unable to carry out this maximal punishment, developed in Appendix B. Firstly, incorporating probabilistic voting into our model would yield a unique equilibrium similar to our responsive equilibrium and featuring responsibility substitution, a central prediction of our analysis. Secondly, if the quality of candidates is allowed to vary endogenously with electoral incentives – for example, as a result of participation constraints (Caselli and Morelli, 2004; Messner and Polborn, 2004; Mattozzi and Merlo, 2008)– this creates an upper bound on the size of the punishment that voters are able to impose on the incumbent, ruling out the unresponsive equilibrium. While these richer models seem realistic and in line with previous literature, our equilibrium selection allows us to present the theoretical mechanism in its simplest form.

We thus focus on the responsive equilibrium, and refer to it as ‘the’ equilibrium.

\(^{15}\)To see why the latter claim must hold, notice that, by the Envelope Theorem, the incumbent’s equilibrium payoff is increasing in \(f\).
5 Disincentive spiral and welfare analysis

In equilibrium, the cost of initiatives affects voters’ ability to induce expertise acquisition and filter out incompetent politicians. The incumbent risks not being reelected if she made a mistake \((S,R)\). Ideally, voters would like such a mistake to reveal, as much as possible, an incompetent type. The *separating power* of the electoral process is its ability to weed out incompetent politicians. This is measured by the posterior probability \(p\) that the incumbent is competent, conditional on observing \((S,R)\): the lower that posterior, the more selective the electoral process.

The next proposition shows that, in equilibrium, this posterior is decreasing with electoral incentives: a lower probability \(f\) of reelection conditional on the mistake \((S,R)\) lowers the probability of ousting a competent type. Let \(p(f) = \frac{q(1-e_C(f))}{q(1-e_C(f)) + (1-q)(1-e_N(f))}\) denote the probability that the incumbent is competent, conditional on \((S,R)\).

**Proposition 3 (Responsibility Substitution)** *In equilibrium, (i) \(e_C(f) - e_N(f)\) is decreasing in \(f\), and (ii) \(p(f)\) is increasing in \(f\).*

The value of \(f\) that induces voters’ indifference is given by

\[
\frac{q}{q + (1-q)\eta(f)} = \frac{qV - cpe_N(f)}{V + cp(e_C(f) - e_N(f))},
\]

where \(e_C(f), e_N(f)\) are the equilibrium expertise levels of both types of incumbent given the ‘electoral disincentive’ \(f\), and \(\eta(f) = [1 - e_N(f)]/[1 - e_C(f)] > 1\). Indeed, the left-hand side of (2) is the equilibrium posterior probability that the incumbent is competent, conditional on making a mistake. The right-hand side is the posterior threshold above which citizens prefer the incumbent over the challenger (see Proposition [1]), and depends on (a) the equilibrium expertise levels \(e_N(f)\) and \(e_C(f)\) chosen by the incumbent, and (b) the cost \(c\) of correcting mistakes in the second period.

This posterior threshold is decreasing in \(c\), reflecting the fact that, when direct democracy is cheaper, voters care more about competence per se and less about the value of past expertise, since they can amend mistakes at lower cost.

As a result, a lower cost \(c\) weakens electoral incentives in a self-reinforcing pattern that is now ex-
explained. First, it raises the posterior threshold at which voters accept to reelect the incumbent. Other things equal, thus, $\eta$ must decrease: the equilibrium posterior conditional on a mistake must increase, in order to keep voters indifferent. The Proof of Proposition 3, part (ii), shows that $\eta(f)$ is decreasing in $f$. Therefore, $f$ must increase in equilibrium. The intuition is that, to achieve a higher posterior conditional on a mistake, the incumbent must have acquired a lower level of expertise, which is possible only if the “electoral disincentive” $f$ is higher. From part (i) of Proposition 3, a higher $f$ implies that $e_N(f)$ and $e_C(f) - e_N(f)$ decrease. Other things equal, this increases the right-hand side of (2), which must again be offset by a lower $\eta$, hence a higher $f$, and so forth. Therefore, the impact of a lower direct democracy cost has a spiraling negative effect on electoral incentives.

![Figure 1: Responsibility substitution effect. The zero of $Z(f)$, which implicitly defines $f^*$, shifts to the left as $c$ declines. Parameter values: $c = 0.4$ (black line) and $0.3$ (grey line); $V = 1$, $\rho = 0.8$, $q = 0.44$, $\tau_C = 1$, $\tau_N = 0.6$.](image)

Figure 1 displays the equilibrium reelection probability (derived from the voter’s indifference condition, corresponding to the zero of the function $Z(f)$ defined in (1) for two values of $c$. As the cost of direct democracy decreases, the responsibility substitution effect makes voters more willing to reelect an incumbent who mistakenly chose the status quo policy over the reform, which in turns dampens her incentive to acquire expertise to avoid these mistakes. The impact of direct democracy on welfare is a priori ambiguous: Policy mistakes harm voters less than before, but they are more likely to occur. The next section compares these direct and indirect effects.
5.1 Welfare and direct democracy

This section studies the impact of direct democracy on citizens’ welfare. The benchmark corresponds to \( c = B \), which is equivalent to the absence of direct democracy.\(^{16}\) Let \( EU(c) \) be the voter’s expected utility when the cost of direct democracy is equal to \( c \)

\[
EU(c) = qV + \rho(B - c) + \rho(c + B)E(e_{\theta})
\]

where \( E(e_{\theta}) = (1 - q)e_N + qe_C \) is the expected expertise. If \( EU(B) \geq EU(c) \), direct democracy at cost \( c \) reduces citizens’ ex ante welfare with respect to pure representative democracy.

**PROPOSITION 4** There exists \( \tau^* \in (0, 1/2) \) such that for all \( \tau \geq \tau^* \), \( EU(c) \) is strictly increasing in \( c \in [c_1, c_2] \).

Intuitively, increasing \( c \) has two consequences: First, the direct loss associated with a politician’s mistake in the second period, which occurs with probability \( \rho[1 - E(e_{\theta})] \), increases. Second, the probability of a policy mistake decreases because (a) the electoral incentive \( 1 - f \) increases (incentive gain), and (b) the average expertise is increasing in \( 1 - f \) (expertise gain). This second consequence embodies the responsibility substitution effect mentioned earlier in this paper.

When the productivity ratio \( \tau \) increases (i.e., competence types get closer, and the screening problem voters face is more severe), the direct policy loss decreases (keeping the electoral incentive constant). Moreover, the equilibrium electoral incentive, which requires the posterior to take a specific value, must be larger: if it was to remain constant, then type separation would be lower, and the posterior conditional on a mistake would be too high. Finally, an increase in \( \tau \) also increases the expected responsiveness to electoral incentives, thereby increasing the incentive gain.

As a consequence, when \( \tau \) is large enough (a sufficient condition being \( \tau \geq 1/2 \)), responsibility substitution dominates the direct effect of “cheaper” policy amendments on the whole domain \([c_1, c_2] \): although initiatives are ex post beneficial for voters, as long as the incumbent is acquiring some expertise in equilibrium, a marginal reduction in \( c \) hurts voters, ex ante, through worse behavior and

\(^{16}\)Notice the implicit assumption that \( B \leq c_2 \). Since \( B > c_2 \) implies that representative democracy features zero expertise acquisition, assuming \( B \leq c_2 \) rules out the case of a fully dysfunctional representative democracy, where any institutional change can only be welfare improving.
poorer selection of elected officials.

The next result compares welfare levels, and focuses on the extreme case in which the cost of initiatives is zero: this situation is referred to as costless amendments. The goal is to study under which conditions pure representative democracy dominates costless amendments.

**Proposition 5** The introduction of costless amendment reduces citizens welfare when the expected expertise is high enough: \( E(e_\theta) \big|_{c=B} \geq 1/2 \).

For parameters such that the equilibrium level of expertise under representative democracy is high enough (i.e., \( \rho \) large enough or \( V \) small enough), Proposition 5 implies that introducing costless direct democracy reduces expected welfare. The result is driven by responsibility substitution.

### 5.2 Discussion

The analysis shows that the introduction or facilitation of direct democracy along a single policy dimension can reduce voters’ ex ante welfare, even when the ex post exercise of direct democracy always improves welfare. In particular, there are two key factors which modulate the impact of direct democracy: the initial quality of representative democracy and the degree of type separation. On the one hand, when equilibrium expertise under representative democracy is high enough, voters’ welfare is lower even when direct democracy is costless. On the other, when the screening problem faced by voters is difficult enough (the productivity ratio \( \tau \) is close enough to 1) voters’ welfare is strictly increasing in \( c \).

The observation that direct democracy can reduce citizens’ welfare because of its indirect effect on politicians behavior is consistent with Woodrow Wilson’s appraisal of direct democracy in Switzerland ("The State," 1898, Chapter VIII):

[...] the Referendum [...] has dulled the sense of responsibility among legislators without in fact quickening the people to the exercise of any real control in affairs.

The same idea appears, a few years earlier, in the work of James Bryce (1888), an important scholar of the American political system, who highlights the lowered “authority and sense of responsibility of the legislature” induced by direct legislation. This paper not only formalizes these ideas, but also
provides an explanation for the dulled responsibility among legislators: the reduced effectiveness of the electoral process at screening competence and rewarding expertise.

6 Testable implications and existing evidence

This section derives the model’s main comparative statics and predictions and discusses their empirical support. In particular, the model can explain the puzzling fact that elected officials in states with direct democracy, while enjoying lower job approval rates than other states, are more likely to be reelected, especially conditional on bad economic performance.

6.1 Job approval, reelection rates, and cost of initiatives

The responsibility substitution effect predicts that direct democracy should be associated with a lower probability that the incumbent chooses the optimal policy for the voters. Direct democracy should then be correlated with higher dissatisfaction towards elected officials (and, more generally, representative government), and that this effect should be stronger when direct democracy is cheaper (low $c$):

**Implication 1** The incumbents’ approval rate should be lower with direct democracy, and increasing in the cost of proposing an initiative.

This predictions is consistent with the evidence documented by Dyck (2009) and Dyck and Lascher (2009): Dissatisfaction towards government and incumbents is higher in states with direct democracy.

Despite being relatively less satisfied about their elected officials’ performance, Section 5 shows that citizens under direct democracy do not necessarily punish them with lower reelection rates. In fact, the analysis shows that the reelection rate conditional on a policy mistake decreases in the cost of direct democracy. The prediction is more subtle with respect to the unconditional reelection probability, $1 - \rho(1 - E(e_\theta))(1 - f)$: although voters are more forgiving with policy mistakes, those mistakes

\[17\] While the possibility that causation runs the other way—that is states with systematically worse elected officials are more prone to the adoption of direct democracy—cannot be excluded, the fact that most of the initiatives states adopted direct democracy over a century ago should significantly mitigate these concerns.
also become more likely, so the difference across regimes in overall reelection rates should be less pronounced.

**Implication 2** The difference between direct and pure representative democracy in incumbent’s reelection rates *conditional on a policy mistake* should be positive, and larger than the overall difference in incumbent’s reelection rates.

These predictions are consistent with the evidence documented by Kelleher and Wolak (2007), and Bali and Davis (2007): In states that allow initiatives, legislators’ reelection rates do not statistically differ from non-initiative states in normal times, but are statistically higher in times of bad economic performance.

### 6.2 Likelihood of policy optimality and expected competence

The equilibrium probability that an initiative occurs can be interpreted, for empirical purposes, as the frequency at which initiatives occur in the data. The unprecedented increase in the frequency of US initiatives (Matsusaka, 2005b) over the last three decades, together with the criticism (Broder, 2000; Gerber, 1999; Cronin, 1989) that these institutions have received, can be interpreted through the lenses of responsibility substitution: While direct democracy enables a “reaction to a legislature’s inactivity on issues of importance to the voters,” as argued by California State Senator Tom Harman, such inactivity might indeed be endogenous to the presence of direct democracy.

The next result describes how the expected need for a policy change (i.e., the probability \( \rho \) that a reform is optimal) affects the equilibrium.

**Proposition 6** *The incumbent’s equilibrium expertise level is increasing in \( \rho \), and the equilibrium reelection probability \( f^* \), as a function of \( (\rho, c) \), is supermodular in \([0, 1] \times [c_1, c_2] \):* \[
\frac{\partial^2 f^*}{\partial c \partial \rho} > 0.
\]

Since \( \frac{\partial f}{\partial c} < 0 \), the sensitivity of the electoral incentive to the cost of direct democracy is *lower* when reforms are more likely to be needed, because the derivative of the equilibrium expertise \( \tau_0 \rho (1 - f) \) to the electoral incentive \( (1 - f) \) is proportional to \( \rho \). As a consequence, the adjustment in \( f \) that is needed to offset a change in \( c \) and to maintain voters’ indifference is decreasing in \( \rho \).

The result may be interpreted as follows: The indirect effect of direct democracy on a politician’s
performance is stronger when the expected optimality of the reform is relatively low. During policy emergencies or economic turmoil, by contrast, the incumbent’s behavior should be less elastic to changes in the cost of direct democracy. As a result, a weaker responsibility substitution effect should result in a lower frequency of initiatives (despite an arguably higher dissatisfaction with representative government). This prediction is indirectly supported by Boemhke (2005), who finds that the size of the budget deficit is associated with a lower frequency of initiatives.

Finally, this paper studies how the equilibrium depends on the probability \( q \) that politicians (both incumbent and challengers) are competent. Inspection of equation (1), which defines \( f \), suggests that \( q \) has an ambiguous effect on this variable. This point is illustrated by Figure 2, which displays the relationship between \( q \) and \( c_1, c_2 \) (the threshold values for \( c \) that define equilibrium expertise in the model). \( c_1 \) is the smallest cost of initiative for which expertise matters to voters and \( c_2 \) is the largest cost of initiatives for which nonzero expertise acquisition can arise in equilibrium. The following proposition formalizes this intuition.

**PROPOSITION 7** \( c_1 \) and \( c_2 \) are strictly quasiconcave in \( q \), with an interior maximizer.

In equilibrium, expertise is maximized when \( c = c_2 \) (that is, when \( f = 0 \)), and is equal to zero for \( c \leq c_1 \). Therefore, fixing a value for \( c \) (e.g., low enough to be below \( c_1 \) for some \( q \)) and letting \( q \) vary, the analysis shows a non-monotonic relationship between \( q \) and the equilibrium level of expertise. As the uncertainty surrounding the incumbent’s type vanishes (i.e., \( q \) goes to 0 or 1), the range of costs inducing a positive equilibrium level of expertise shrinks to zero.

As a consequence, initiatives have more likely to generate pernicious indirect effects when voters

---

**Figure 2: Expertise acquisition, as a function of competence distribution and direct democracy cost.**
face significant uncertainty about politicians’ competence. By contrast, when voters are more concerned about disciplining politicians (as in the Progressive Era) rather than selecting them, then direct democracy is less likely to negatively affect elected official’s incentives and selection.

7 Conclusion

Direct democracy institutions have a broader impact than suggested by merely looking at the individual issues which they have successfully influenced. In the United States and elsewhere, the growing importance of citizen initiatives concerns not only the policies that are decided by those initiatives, but also the behavior and the selection of elected officials. While citizen initiatives, referendums, and petitions provide a vital energy to democracies, they also reduce the accountability of politicians and narrow their role in the political landscape, giving them a perverse incentive to default on their responsibilities. This paper provides the first formalization of this idea, which may be traced back to the century-old concerns expressed by James Bryce (1888) and Woodrow Wilson (1898).

More broadly, the paper provides a new framework to think about the dynamic equilibrium effects of direct democracy on citizens’ welfare and politicians’ selection, incentives, and electoral survival. Even when it is immune from its standard criticisms, direct democracy creates a responsibility substitution effect and ensuing disincentive spiral which can more than offset the direct welfare gain from initiatives. The increased alignment between policies and voters’ preferences, which has been predicted by the existing literature on direct democracy, is no longer guaranteed.

The model also generates comparative statics that are consistent with existing and novel evidence from U.S. states. In particular, the theory can explain why (1) incumbents are more likely to be reelected, conditional on poor economic performance, when direct democracy is present, and (2) voters are overall less satisfied with their elected officials in states where initiatives are present and/or subject to lower signature requirements. As explained in Section 6.2, there is indirect evidence that the responsibility substitution has become more salient in the United States, which helps explain the criticism that direct democracy institutions have received there in recent years. 

\footnote{In 2011, less than 18\% of Californians thought that the initiative process did not need to be reformed (source: Public Policy Institute of California Statewide Survey, May 2011).}
Appendices

A  Proofs

Proof of Proposition 2. First, let’s define
\[ c_1 = \frac{1 - \tau}{\phi_2 + \phi_1 \tau}; \quad c_2 = \frac{1 - \tau}{\phi_2 (1 - \rho) + \phi_1 \tau (1 - \tau \rho)} \]

The proof is based on the following lemma.

Lemma 1 For \( c \leq c_2 \), there exists a natural equilibrium with \( f = 0 \) and at most one other natural equilibrium with \( f \in (0, 1) \).

Proof. The condition \( f < 1 \), which is necessary for a positive expertise level, can only hold in equilibrium if voters weakly prefer the challenger to the incumbent, conditional on a policy mistake \((S, R)\):
\[ \frac{q(1 - e_C(f))}{q(1 - e_C(f)) + (1 - q)(1 - e_N(f))} = \frac{qV - c\rho e_N(f)}{V + c\rho (e_C(f) - e_N(f))}. \]

This condition may be re-expressed as
\[ a(f) = b(f), \quad (3) \]

where \( a(f) = e_N(f)[1 + c\phi_1] - e_N^2(f)c\phi_1 \) and \( b(f) = e_C(f)[1 - c\phi_2] + e_C^2(f)c\phi_2 \). The case \( f = 0 \) corresponds to \( a(0) \leq b(0) \). Using that \( e_C = \rho \tau C \) and \( e_N = \rho \tau N \), this yields \( c \leq c_2 \).

If \( f \in (0, 1) \), equilibrium expertise must be such that voters are, conditional on a mistake, exactly indifferent between challenger and incumbent, i.e., such that \( a(f) = b(f) \).

It is easy to show that \( a \) is decreasing in \( f \), and that \( b \) is either everywhere decreasing in \( f \) or first decreasing and then increasing.\(^{19}\) An equilibrium probability \( f \) is a root of the function \( a - b \). It

\(^{19}\)Seen as a function of \( e_N \), \( a \) is a quadratic concave polynomial that reaches its peak at \( (1 + \phi_1)/(2\phi_1) > 1 \). Therefore, \( a \) is increasing in \( e_N \) on \([0, 1]\) and, therefore, decreasing in \( f \). Similarly, \( b \) is a quadratic convex polynomial in \( e_C \), whose minimizer \( e_b \) is always strictly less than 1, and is negative if \( \phi_2 < 1 \). Therefore there exists \( f_b \) such that \( e_C(f) \) is in the increasing range of \( b \) if and only if \( f > f_b \).
is now shown that the function \( a - b \) is strictly quasiconcave (i.e., strictly increasing, then strictly decreasing). Equivalently, it suffices to show that \( b' - a' \) satisfies the single-crossing property (i.e., crosses the \( x \)-axis only once, from negative to positive. Let \( f_b \) denote the (possibly, equal to 1) value of \( f \) such that \( b \) is decreasing on \([0, f_b]\) and increasing on \([f_b, 1]\). Since \( a \) is decreasing, the function \( a - b \) is decreasing on \([f_b, 1]\). To conclude the proof of quasiconcavity, it suffices to show that if \( a - b \) is decreasing around any value \( f < f_b \), then it is decreasing on \([f, 1]\). For \( f < f_b \), \( a'(f), b'(f) \) are negative, and

\[
\frac{a'(f)}{b'(f)} = \frac{1 + c\phi_1 - 2e_N(f)c\phi_1}{1 - c\phi_2 + 2e_C(f)c\phi_2},
\]

which is increasing on \( f \), since \( 1 - c\phi_2 + 2e_C(f)c\phi_2 \) is positive (since \( b' \) is negative). Now suppose that \((a - b)'(\hat{f}) \leq 0 \) for any \( \hat{f} < f_b \). Then, \( \frac{a'(\hat{f})}{b'(\hat{f})} > 1 \) and hence \( \frac{a'(f)}{b'(f)} > 1 \) and \((a - b)'(f) < 0 \) for all \( f \in (\hat{f}, f_b) \). In particular, given that \( a(1) - b(1) = 0 \), the function \( a - b \) can only cross the \( x \)-axis once, and it does so from below. This shows that there exists at most one level of \( f \) for which (3) holds and, therefore, at most one intermediate incentives equilibrium. Now suppose that \( a(0) > b(0) \). If \( a(\hat{f}) = b(\hat{f}) \) for some \( \hat{f} \in (0, 1) \), strict quasi-concavity implies that \( a(f) < b(f) \) for all \( f > \hat{f} \), which contradicts the fact that \( a(1) = b(1) \). As a consequence, it is required that \( a(0) \leq b(0) \), which is equivalent to \( c \leq c_2 \).

Lemma 1 shows that there are three possible cases: (i) no natural equilibrium \((c > c_2)\), (ii) an unresponsive equilibrium and a responsive equilibrium with \( f = 1 \) and zero expertise\(^{20}\) and (iii) a responsive equilibrium with positive expertise \((f \in (0, 1))\) and an unresponsive equilibrium. There remains to identify under which condition the last two cases hold. A responsive equilibrium with positive expertise obtains if and only if (3) is satisfied for some \( f \in (0, 1) \). Since \( a(1) = b(1) = 0 \) and \( a'(f)/b'(f) \) is increasing in \( f \), a necessary condition for the existence of \( f < 1 \) such that \( a(f) = b(f) \) is that \( \frac{a'(1)}{b'(1)} \geq 1 \). Otherwise, \( b'(f) < a'(f) < 0 \) for all \( f \) (recalling that \( a, b \) are decreasing) and, therefore, \( a(f) > b(f) \) for all \( f < 1 \). Using that \( \frac{a'(1)}{b'(1)} = \frac{\tau(1 + c\phi_1)}{(1 - c\phi_2)} \), one obtains the desired condition \( c \geq c_1 \).

**Proof of Proposition 3** (i) immediately follows from \( e_C(f) - e_N(f) = (1 - f)\rho(\tau_C - \tau_N) \geq 0 \). (ii) It suffices to show that the ratio \( \frac{1 - e_N(f)}{1 - e_C(f)} \) is decreasing in \( f \). Since \( e_\theta(f) = (1 - f)\rho\tau_\theta \), the derivative of the ratio \( \frac{1 - e_N(f)}{1 - e_C(f)} \) has the same sign as \( \rho[\tau_N(1 - e_C(f)) - \tau_C(1 - e_N(f))] < 0 \), where the inequality

\(^{20}\)Notice that \( \hat{f} = 1 \) is always a solution of \( Z(f) = 0 \).
follows from $e_C(f) \geq e_N(f) \forall f$. 

**Proof of Proposition 4**

Denote by $EU$ the expected utility to the voter as a function of the cost of proposing an initiative $c$. Then $EU(c) = qV + \rho(B - c) + \rho(c + B)((1 - q)e_N + qe_C)$. Differentiating it with respect to $c$ yields:

$$\frac{dEU}{dc} = \frac{df^*/dc}{\text{Incentive gain}} \times (B + c)((1 - q)e'_N + qe'_C) - (1 - q)(1 - e_N) + q(1 - e_C)$$

The incentive gain can be computed by applying the Implicit Function Theorem to the equation that defines $f^*$, which is $e_C(1 - \phi_2c + e_C\phi_2c) - e_N(1 + \phi_1c - e_N\phi_1c) = 0$. This yields

$$\frac{df^*}{dc} = \frac{1}{ce'_C} \frac{(1 - e_N)e_N\phi_1c + (1 - e_C)e_C\phi_2c}{e_N\phi_1c + e_C\phi_2c}.$$ 

Using the fact that $e'_N = e'_C\tau$ and substituting the equation defining $f^*$ into the denominator of $df^*/dc$ (also the fact that $e_N = \tau e_C$), the following obtains:

$$\frac{df^*}{dc} = \frac{1}{ce'_C} \frac{(1 - e_N)e_N\phi_1c + (1 - e_C)e_C\phi_2c}{\tau e_N\phi_1c + e_C\phi_2c} = \frac{1}{ce'_C} \frac{(1 - e_N)\tau(1 - q) + (1 - e_C)q}{\tau^2(1 - q) + q}$$

As a consequence, the derivative becomes

$$D = (1 - e_C)q(\Lambda(\tau) - 1) - (1 - e_N)(1 - q)(1 - \Lambda(\tau)\tau),$$

where $\Lambda(\tau) = \frac{B + c}{c} \frac{(1 - q)\tau + q}{(1 - q)\tau^2 + q} > 1$. The necessary and sufficient condition for $D > 0$ can be expressed as

$$(1 + \Lambda)(qe_C + (1 - q)e_C\tau) \geq 1 + \Lambda(1 - e_C\tau)(1 - q)(1 - \tau)$$

which implicitly defines a $\tau^* < 1/2$ (possibly equal to zero\footnote{When $\tau = 1/2$, $\Lambda\tau = \frac{1 + c(1 - q)\tau^2 + \tau q}{(1 - q)\tau^2 + q} \geq \frac{2(1 - q)\tau^2 + 2\tau q}{(1 - q)\tau^2 + q} \geq \frac{(1 - q)\tau + q}{(1 - q)\tau^2 + q} > 1$. Moreover, when $qe_C > (1 + (1 - q)\Lambda)(1 + \Lambda)^{-1}$, it is possible to show that $\tau^* = 0.$}) such that $\forall \tau > \tau^*, D > 0$. 

\footnote{When $\tau = 1/2$, $\Lambda\tau = \frac{1 + c(1 - q)\tau^2 + \tau q}{(1 - q)\tau^2 + q} \geq \frac{2(1 - q)\tau^2 + 2\tau q}{(1 - q)\tau^2 + q} \geq \frac{(1 - q)\tau + q}{(1 - q)\tau^2 + q} > 1$. Moreover, when $qe_C > (1 + (1 - q)\Lambda)(1 + \Lambda)^{-1}$, it is possible to show that $\tau^* = 0.$}
Proof of Proposition 5. The difference between $EU(B)$ and $EU(c)$ in equilibrium is given by

$$d(c) = qV + \rho(c + B)E(e(1)) - qV - \rho(B - c) - \rho(c + B)E(e(c))$$

and the condition follows from computing that difference for $c = 0$: $d(c) = 2\rho E(e(1)) - \rho$.

Proof of Proposition 6. (i) Since $e_\theta = (1 - f)\rho \tau \theta$, $\partial e_\theta / \partial \rho = \partial [\rho(1 - f)] / \partial \rho$, equation [1] (which implicitly defines $f$) can be re-expressed as

$$(1 - f)\rho \tau N[1 + c\phi_1] - [(1 - f)\rho \tau N]^2c\phi_1 - (1 - f)\tau C\rho[1 - c\phi_2] - [(1 - f)\tau C\rho]^2c\phi_2 = 0$$

After dividing both sides by $(1 - f)\rho \tau C$ and rearranging terms, one obtains

$$(1 - f)\rho = \left[\frac{\tau}{q} + \frac{1}{1 - q} - \frac{V}{\rho c}(1 - \tau)\right] \left[\frac{\tau \tau N}{q} + \frac{\tau C}{1 - q}\right]^{-1} \tag{4}$$

(ii) Using (4) yields $\frac{\partial f}{\partial c} = -\frac{V}{\rho c^2}(1 - \tau) \left[\frac{\tau \tau N}{q} + \frac{\tau C}{1 - q}\right]^{-1}$ and $\frac{\partial^2 f}{\partial c \partial \rho} = \frac{V}{\rho^2 c^2}(1 - \tau) \left[\frac{\tau \tau N}{q} + \frac{\tau C}{1 - q}\right]^{-1} > 0$.

Proof of Proposition 7. The thresholds can be expressed as

$$c_1 = \frac{V}{\rho} [1 - \tau] \left[\frac{1}{1 - q} + \frac{\tau}{q}\right]^{-1}; \quad c_2 = \frac{V}{\rho} [1 - \tau] \left[\frac{1 - \rho}{1 - q} + \frac{1 - \rho \tau}{q}\right]^{-1},$$

their denominators are strictly quasiconvex in the relevant domain: therefore, they are both decreasing between 0 and their unique interior minimizer, and then increasing from there onwards. Moreover, these denominators are also strictly positive in the whole domain of $q$. Since the reciprocal of a strictly quasiconvex function that is positive on a subset of $\mathbb{R}^n$ is strictly quasiconcave on that subdomain, then $c_1$ and $c_2$ are both strictly quasiconcave functions, and their maximizer must be the minimizer of their denominators.
B General model and robustness

B.1 Self-enforcing political contracts

This section establishes that the “natural” policy of implementing the reform if and only if the incumbent learns that the state is equal to $R$ is optimal from voters’ viewpoint, and optimally achieved by surely reelecting the incumbent if she made no mistake, reelecting her with lower probability (and ousting her otherwise) if she chose the default option when the reform was needed, and punishing her with certain ousting if she implemented a bad reform.

An equilibrium of the kind described above, with $f(S, S) = f(R, R) = 1$, $f(R, S) = 0$, and $f(S, R) \in [0, 1)$, clearly dominates any no-expertise equilibrium. It is now established that it is also optimal among all possible equilibria inducing a positive expertise level.

First, consider voters’ reelection strategy conditional on observing $h = S$. In that case, the incumbent cannot have received the signal revealing $h = R$, irrespective of her type and expertise level. Type selection cannot occur in that case, and the best voters can hope for, given $h = S$, is that the incumbent chooses $x = S$, which is ex ante optimal, and a fortiori optimal conditional on not receiving any signal. Moreover, voters want to reelect the incumbent, other things equal, for two reasons. First, their posterior about her competence is the same as their belief about the competence of the challengers. Second, the incumbent has the strict advantage, compared to her challengers, of having acquired some expertise, which is useful for the second period. Therefore, an optimal strategy is to set $f(S, S) = 1$ and $f(R, S) = 0$, as it achieves exactly this.$^{22}$

Suppose, instead, that the state is $h = R$. In this case, voters want the incumbent to choose the reform $x = R$ when she learns the state. Thus, abstracting for now from the adverse selection problem, voters want the incumbent to choose $R$ if and only if she received the signal. Taking now into account adverse selection, in any equilibrium with some expertise acquisition, the competent type acquires more expertise than the incompetent type because it is less costly for her to do so. Therefore, the natural policy not only maximizes voters’ first-period utility, but it also increases the probability of selecting a more competent politician. Thus, setting $f(R, R) = 1$ is optimal. The remaining

$^{22}$Setting $f(R, S)$ small enough would also yield an optimal strategy, provided that the incumbent follows the “natural policy,” since $(R, S)$ does not arise on the equilibrium path.
probability $f(S, R)$ is pinned down by equilibrium conditions. If $f(S, R) = 1$, the incumbent is always reelected if she chooses the default action $S$; as a consequence, she has no incentive for acquiring expertise.\footnote{In that equilibrium, voters are indifferent between reelecting the incumbent and the challenger, as both have the same expected competence and the incumbent has no expertise.} This establishes the following result.

**Proposition 8 (Natural Positive-Expertise Equilibrium)** Any optimal equilibrium can be implemented by setting $f(S, S) = f(R, R) = 1$ and $f(R, S) = 0$. The remaining reelection probability $f = f(S, R) \in [0, 1]$ is determined by equilibrium conditions.

In this equilibrium, the expertise level is given by $e(\theta) = (1 - f)\tau \rho$. Because the cost function $T$ is convex, $e(\theta)$ is decreasing in $f$. Since $\tau(\theta)$ has increasing differences in $(e, \theta)$, the expertise levels are also ranked: $e_C(f) \geq e_N(f)$ for all $f$.

The above analysis also applies to the more general model presented in the Section B.3.

### B.2 Probabilistic voting

Suppose that, instead of being fully rational, voting behavior was probabilistic. In that case, voters’ would elect the incumbent if and only if the net gain in expected utility vis à vis the challenger was above the realization of a preference shock $\xi$: given a posterior $p$ about the incumbent, her reelection probability is

$$Pr[\xi \leq p[V + \rho e_C(B + \rho(1 - e_C)(B - c)) + (1 - p)[\rho e_N(B + \rho(1 - e_N)(B - c)] - qV - \rho(B - c)].$$

which simplifies to

$$Pr[\xi \leq (p - q)V + \rho c(1 - \tau)][\xi \leq (p - q)V + \rho c(1 - \tau).$$

For simplicity,\footnote{What matters is the supermodularity of $5$ in $(p, c)$.} let’s assume (as often the case in the literature) that $\xi$ is drawn from a uniform with support $[-1/2\psi, 1/2\psi]$, where, to guarantee that reelection probabilities are interior, it is also assumed that $1/2 - \psi qV > 0$ and $1/2 + \psi[(1 - q)V + \rho c(1 - \tau)] < 1$. In this case, it immediately follows that $f(S, S) = 1/2$. Moreover, an incumbent of type $\theta$ chooses expertise $e(\theta) = \tau(\theta)\rho[f(R, R) - f(S, R)$.}
In a natural equilibrium, where politicians act optimally on their information, the following obtains (after a few steps):

\[
f(R, R) - f(S, R) = \psi q [V + c \rho (e_C - e_N)] \left[ \frac{qe_C}{qe_C + (1 - q)e_N} - \frac{q(1 - e_C)}{q(1 - e_C) - (1 - q)(1 - e_N)} \right]
\]

As a consequence, using the fact that \( e_c = \tau e_N \), and the equilibrium is identified by the solution of \( G(e; c) = e \), where

\[
G(e; c) = \psi \tau C \rho q [V + c e \rho (1 - \tau)] \frac{1 - q + (1 - q)e}{q + (1 - q)\tau}(1 - q + [(1 - q)\tau]e).
\]

**Proposition 9 (Responsibility Substitution)** There is a unique equilibrium, where expertise is increasing in \( c \).

**Proof.** To see that the fixed point must be unique, notice that \( G(0; c) > 0 \). Combining that with the fact that \( G \) is increasing and convex in \( e \) with the boundary conditions \( \psi q V < 1/2 \) and \( \psi [(1 - q) V + \rho c (1 - \tau)] \) yields

\[
1 > \frac{q}{q + (1 - q)\tau} \tau C \rho \psi [V + c e \rho (1 - \tau)] = \frac{\partial G(e; c)}{\partial e} \bigg|_{e=1} > \frac{\partial G(e; c)}{\partial e}.
\]

As a consequence, Banach’s theorem holds. Since \( G \) is also increasing in \( c \), then it must be that supermodular in \((c, e)\), which implies that the zero of \( G - e \) must also increase in \( c \). This completes the proof. ■

### B.3 General model

This section studies a general model whose unique natural equilibrium corresponding to the responsive equilibrium of the model presented in the main text. This version of the model also shows that the results derived in the main text generalize to a large class of cost functions.

As discussed in Section 4, the model presented in the body of this paper allows us to provide the simplest formal description of the responsibility substitution effect. In this environment, however, responsibility substitution is only associated with one – the responsive equilibrium – of the two natural
equilibria that arise. In the other, the *unresponsive* equilibrium, voters are able to maximize the strength of the electoral incentives for the incumbent to avoid policy mistakes. Moreover, this ability is not affected by the cost of correcting these mistakes *ex post* (i.e., the cost of proposing initiatives).

**Outside option.** After observing the outcome of her expertise acquisition, the incumbent chooses whether or not to run for reelection and which policy to implement. If she chooses not to run for reelection, she obtains an outside option $O_\theta$, which is assumed to depend on her competence. The outside option can be interpreted as the choice of a private sector job, or the choice of running for a higher level political office.\(^{25}\) In line with the literature (Caselli and Morelli, 2004; Mattozzi and Merlo, 2008; Messner and Polborn, 2004), it is assumed that political competence is correlated with the value of the outside option ($O_C > O_N$). For simplicity, $O_N$ is set to zero. The following assumptions also requires that, at least in some circumstances (depending on voters behavior), the $C$-type might actually prefer $O_C$ over running for reelection.

**Assumption 1 (Participation Constraint)** $O_C \in ((1 - \rho)(1 - \rho^2)^{-1}, 1)$.

This assumption implies that a competent incumbent prefers her outside option over a regime in which voters reelect her only when she chooses the optimal policy. It therefore imposes an upper bound on the reelection pressure that voters can put on the incumbent.

In this modified environment, $\lambda_\theta$ (resp. $\lambda'_\theta$) denotes type $\theta$'s probability of running for reelection upon receiving no signal. $\theta$'s optimal strategy then solves

\[
\max_{e,\lambda,\lambda',y,y'} \left\{ \rho e \left[ \lambda' f(y', R) + (1 - \lambda') O_\theta \right] + (1 - \rho e) \left[ \lambda \left( \frac{1 - \rho f(y, S)}{1 - \rho e} + \frac{\rho(1 - e) f(y, R)}{1 - \rho e} \right) + (1 - \lambda) O_\theta \right] - \tau_\theta(e) \right\}
\]

The incompetent type always runs for reelection since her outside option is zero. In any equilibrium with a positive expertise level, the competent type must also run with positive probability $\lambda_C > 0$.\(^{26}\)

\(^{25}\)Mitt Romney constitutes a recent example of a politician who chose not to run for reelection in order to try to compete for a higher political office.

\(^{26}\)Otherwise, an incumbent running for reelection would reveal himself as an incompetent type. Voters would, therefore, never reelect her ($U(C, e) > U > U(N, e)$), and the incumbent would optimally choose a zero level of expertise.
Convexity of the cost function implies that the optimal expertise is characterized by

\[ e_\theta = [\lambda' f(y'_\theta, R) + (1 - \lambda') O_\theta - \lambda f(y_\theta, R) - (1 - \lambda) O_\theta] \rho_{\tau_\theta}. \]

**Proposition 10** The mixing probability \( \lambda \) satisfies the following conditions:

(i) When \( f(y', S) = 0 \), \( \lambda_C = 0 \);

(ii) When \( f(y', R) = f(y, S) = 1 \), there exists a unique lower bound \( f \) for \( f(y', S) \) above which \( \lambda_C > 0 \).

**Proof.** (i) The maintained assumptions imply that \( \frac{1 - \rho}{1 - \rho e} f(y, S) < \frac{1 - \rho}{1 - \rho^2} < O_C \), from which the result follows. For (ii), first observe that the expected payoff from running for reelection is equal to \( \frac{1 - \rho}{1 - \rho e} + \rho e f(y', R) \), where \( f = f(y, R) \). Substituting into this equation the search level \( e = (1 - f) \tau_C \rho \) that permits a strictly positive value of \( \lambda \), one obtains

\[ \frac{1 - \rho + \rho (1 - (1 - f) \tau_C \rho) f}{1 - \rho^2 (1 - f) \tau_C} = 1 - I \frac{\rho - \rho^2 e \tau_C}{1 - \rho^2 e \tau_C} := L(I) \]

where \( I = (1 - f) \) is the electoral incentive. Since \( \frac{\partial^2 L(I)}{\partial^2 I} = 2 \tau_C \rho (1 - \rho)(1 - \rho^2 \tau_C I)^{-3} > 0 \), \( L(I) \) is strictly convex. Moreover, it is decreasing for \( I \) close enough to 0. As a consequence, \( L \) must be monotonic on the range values in \((L(0), L(1))\).

In light of this result, it is possible to show\(^{27}\) that there exists a threshold \( \hat{c} > c_1 \) (whose definition is similar to \( c_2 \)) below which the unique natural equilibrium with a positive level of expertise is analogous to the responsive equilibrium of the baseline model (i.e., \( \lambda_C = 1, f = f^* \)). It must be also noted that above \( \hat{c} \) the unique natural positive expertise equilibrium features a fixed level of expertise \( (e_\theta(f)) \) and the competent type randomizes between entry and outside option. In that equilibrium, a lower cost of direct democracy increases \( \lambda_C \), thereby improving voters’ welfare.

**General functional form for the cost of effort.** If one assumes that \( \{\tau_\theta(e)\}_{\theta \in \{C, N\}} \) are increasing, continuous, and convex functions, an interior expertise acquisition requires \( \tau_\theta(0) = \tau'_\theta(0) = \tau'_\theta(1) < \rho \). Assuming that condition is also satisfied, a sufficient condition for the model to generate the same set of results presented in the paper is that the two assumptions below hold\(^{28}\)

\(^{27}\)Details are available under request. The argument closely following the proof of Proposition 2.

\(^{28}\)The proofs of Lemma 1 and of Propositions 2, 3, 5, and 4 apply to this setting with minimal modifications.
**Assumption 2 (Increasing Differences)** \( T^*_N(e) > T^*_C(e) \) for all \( e > 0 \).

This assumption, which is very standard in signaling models, ensures that the politician’s effort is increasing in her competence, in the presence of moral hazard and adverse selection.

**Assumption 3 (Discerning Power of Electoral Incentives)**

\[
\kappa(z) = \frac{T^*_C(T^{r-1}_C(z))}{T^*_N(T^{r-1}_N(z))} \text{ is weakly decreasing in } z.
\]

This second assumption states that the differential in responsiveness across types does not decrease with the strength of the electoral incentives, and is essentially equivalent to the signed ratio monotonicity condition introduced by Quah and Strulovici (2012): stronger electoral incentive lead to higher type separation.

### B.4 Robustness

This section discusses alternative assumptions, which could easily be accommodated by the model.

**Discount factor.** If voters discount second-period utility by a factor \( \beta < 1 \), their posterior threshold, which is only based on second-period utility, is not affected. If politicians discount their second-period payoff, this is equivalent to reducing their rent from being reelected, which does not affect qualitatively the analysis. Overall, discounting second-period payoffs reduces equilibrium effort but does not qualitatively affect the impact on welfare of direct democracy.

**Initiatives in both periods.** Allowing initiatives in the first period has no impact on the equilibrium: it does not affect voters’ threshold or politicians’ reelection incentives. However, it does mitigate the impact of second-period direct democracy on politicians’ first-period effort.

**Symmetric policy amendment.** Assuming that citizens can undo a reform (i.e., move from \( R \) to \( S \)) does not affect the analysis. Given the signal structure, in any natural equilibrium politicians only implement the reform if it is optimal (i.e., \( (R, S) \) does not arise in equilibrium).

**Symmetric signal structure.** The model can accommodate a symmetric signal structure: given expertise \( e \), \( \Pr(Learn\ R \mid h = R, e) = \Pr(Learn\ S \mid h = S, e) = \frac{1+e}{2} \). This framework is considerably
less tractable, but does not affect the qualitative nature of the results studied in this paper.

**Weaker form of persistence of expertise.** Assuming that a reelected incumbent inherits only a (possibly type–dependent) fraction $B_\theta$ of her first period effort does not affect the results.

**Partially observable competence.** In general, one could allow citizens to get a signal linked to the incumbent’s competence in the first period, as long as the signal is imperfect, at the expense of expositional simplicity.

**Interpretation of the policy space.** The interpretation of the policy space can be reverted without altering the basic message of the paper: Rather than the reform, $x = R$ can represent a safe policy to which voters can return by proposing an initiative. Under this interpretation, representatives acquire expertise in order to learn when and when not to implement a change. While the baseline interpretation seems more appropriate as a model of initiative, the latter can be better suited to model referendums in Europe.
References


36