Hybrid Democracy: How Initiatives and Referendums Affect the Selection and Incentives of Elected Officials*

Carlo Prato Bruno Strulovici
Georgetown University Northwestern University
First draft: Dec. 9, 2010; this version: March 26, 2014

Abstract

Citizen initiatives and referendums play an important role in modern democracies, from treaty ratification in the European Union to gay marriage in California, to the control of foreign workers in Switzerland. How does that affect the role of elected officials? This paper explores the trade-off between the virtue of direct democracy as a counteracting power to politicians with its substitutive impact on the institution of representative democracy. Departing from the comparison, at the heart of existing research, between pure direct democracy and pure representative democracy, the paper analyzes the equilibrium effects of direct democracy on the incentives and selection of elected officials. The main finding is that facilitating direct democracy induces a negative spiral on politicians’ role and contribution to society, which may reduce voters’ welfare. The theory offers predictions on reelection probabilities and politicians’ performance which are consistent with recent empirical evidence from U.S. states.

Keywords: Direct Democracy, Initiatives, Political Agency, Referendum.

*We thank two anonymous referees, David Austen-Smith, Eddie Dekel, Georgy Egorov, Jon Eguia, Tim Feddersen, Simone Galperti, Bård Harstad, Alex Hirsch, Navin Kartik, Patrick Le Bihan, Antoine Loeper, Arthur Lupia, Eric Maskin, Meg Meyer, Kris Ramsay, Debraj Ray, Stephane Wolton and seminar audiences at Northwestern, the 2011 SPSA and 2011 MPSA Annual National Conference, for helpful comments and suggestions. We are especially grateful to Elizabeth Maggie Penn and Justin Fox for their comments. Email addresses: carlo.prato@u.northwestern.edu and b-strulovici@northwestern.edu
1 Introduction

Direct democracy institutions, such as citizen initiatives and referendums, play an important role in regimes otherwise based on representative democracy. According to the Initiative and Referendum Institute in the period 1990-2010, almost 800 ballot initiatives have been proposed in the U.S. states (against the 433 of the two preceding decades), of which roughly 45% have been approved. Between 1990 and 2003, 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, and 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), 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 of policy expertise (Maskin and Tirole, 2004; Campbell et al., 1960). These arguments have been countered by several authors (Lupia, 1994; Bowler and Donovan, 1998; Garrett and McCubbins, 2008).

Crucially, both arguments focus on a single policy decision: They do not study how direct democracy affects the role of politicians and the overall democratic process. Clearly, not all decisions are made through initiatives and referendums: In large democracies, citizens must delegate most daily decisions as well as complex and strategic decisions to elected officials. Some decisions involve expertise or sensitive information unavailable to the lay person.

It is thus unescapable that some policy decisions should be left to elected politicians. This paper departs from the existing literature by studying how direct democracy on some issues affects equilibrium incentives, and selection and politicians, and policies on all issues. This question, and the theoretical framework developed here, seem especially important in light of the recent wave of empirical work.

---

1 Data available at www.iandrinstitute.org/data.html.
2 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.”
on direct democracy (Bali and Davis, 2007; Dyck, 2009; Kelleher and Wolak, 2007; Rydberg, 2010; Matsusaka, 2008 and 2010; Funk and Gathmann, 2011 and 2013; Hinnerich and Pettersson-Lidbom, 2014), as well as the growing importance of the question among legal scholars (Garrett, 2005).

Most theoretical studies of the interaction between elected officials’ behavior and direct democracy (Matsusaka, 1992; Gerber, 1996; Matsusaka and McCarty, 2001) have focused on the following intuitive point: allowing direct democracy typically improves the congruence between policies and the electorate’s preferences. This intuition is in line with both recent empirical evidence on issues such as abortion rights and death penalty (Matsusaka, 2010), as well as the historical origins of modern direct democracy institutions in United States (Cronin, 1989), where citizens’ demand for increased ability to control potential abuses and failures of representative democracy played a key role.\(^3\)

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 it turns out, addressing the equilibrium effects of direct democracy on the role of elected officials permits to explain recent empirical evidence (Bali and Davis, 2007; Dyck, 2009; Kelleher and Wolak, 2007; Rydberg, 2010) which is hard to reconcile with previous theories.

The framework presented here is based on 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 of this paper is that increasing voters’ amendment power undermines their ability to credibly reward expertise acquisition by an incumbent and learn about her competence.

\(^3\)Kessler (2005) is a notable exception. See Section\(^2\).

\(^4\)This 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.”
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 concerning gay marriage or immigration[^5] or forbidding the construction of minarets, and political ones like the abolition of proportional rule[^6] The second dimension, by contrast, is not amenable to direct democracy. Some examples include the management of international conflicts[^7] the ability of attracting public funds and, more generally, any policy which requires access to sensitive information, expert knowledge, or high-level interactions with elected officials of other countries.

The quality of a policy typically depends on the intrinsic ability of the policymaker, as well as on her effort to acquire expertise on that particular policy. These two components are modeled separately: a politician’s competence is her type, and is assumed to be constant. In contrast, expertise is modeled as the result of an information acquisition task. The effort devoted to this task is endogenous, which implies that equilibrium expertise results depend on politicians’ election incentives and selection.

Citizens cannot directly observe competence. Instead, their ability to assess the competence of an incumbent politician is typically 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 more correlated with her abilities.

Suppose, in that context, that direct democracy becomes feasible, or more easily implementable, along one policy dimension. Specifically, suppose that citizens can avoid 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: Since citizens can “fix” the incumbent’s mistakes along this dimension, their ability to credibly commit to reward effort (via equilibrium reelection incentives) is lowered as a result. This increased leniency

[^5]: Switzerland’s recent decision to limit immigration might force the country to quit the Schengen area.
[^6]: Pure PR was abolished in Italy with two referendums 1991 and 1993.
[^7]: The use of direct democracy in ratifying international treaties is explicitly prohibited by the Italian Constitution.
reduces the incumbent’s incentives for effort, and citizens are more likely to end up with a bad policy (before amendment). Moreover, this reduces citizens’ ability to discern the competence of the incumbent.

While already damaging, these effects are just the first iteration of an even more damaging 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 poor policy, to reelect her. This is only possible, however, if poor policies are less revealing of incompetence than before. This, in turn, implies that the equilibrium level of effort of the incumbent must go down. While the precise mechanics of this disincentive multiplier 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. This result conveys the following idea: 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’s expected payoff 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 negative effect of direct democracy uncovered here is worth emphasizing for several reasons. Firstly, it is derived in a framework where earlier criticisms of direct democracy 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 criticism (Broder, 2000; Gerber, 1999; Cronin, 1989) that these institutions have received in recent years in United States.

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, in such context, 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 multiplier 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.2) may be applied to other economic settings.

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 multiplier 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).

---

10For 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 Majority” in The Economist, December 17th 2009, or David Brooks’ “The Bloody Crossroads,” published on September 7, 2009 in The New York Times, where direct democracy is believed to be a cause of the recent fiscal crisis in California – see Matsusaka, 2005a, for a description of the debate.)

11One application concerns the relational contract between a manager and an employee when the manager can observe and correct the actions of the employee at some cost and only in some dimensions.
2 Related Literature

Most theories of direct democracy have focused on its comparison with representative democracy.\footnote{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).} 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 (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).\footnote{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.} 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 task by bureaucrats. Similarly, the judicial review of elected officials (Fox and Stephenson, 2011) can exacerbate a politician’s incentives to posture.
Kessler (2005) compares two examples in which direct democracy allows citizens to control ideologically incongruent politicians and reduces candidates’ incentives to acquire information before the election. She observes an endogenous deterioration in the quality of elected politicians, which is similar to the disincentive multiplier arising in the present model. This paper provides a framework, comparative statics, and empirical implications which complement and expand this observation. The 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.

3 Model

There are two periods. In the first one, the incumbent politician chooses how much expertise to acquire about a policy, at a cost that decreases in her competence and increases in the chosen level of expertise. She then observes a signal, whose accuracy depends on the chosen level of expertise, and decides on the policy $x$ for that period. For simplicity, both the state of the world and the policy are assumed to be binary, taking values in the set $\{S, R\}$. The incumbent’s choice is therefore either a success or a failure: Denoting by $h$ the state of the world, the policy is successful if $x = h$, and a failure otherwise.

At the end of the first period, citizens observe whether the policy was successful, and then decide whether to reelect the incumbent. Their vote is determined by their belief about the incumbent’s competence and acquired expertise, which are both relevant for their second-period utility (see below). The assumption, here, is that a politician’s competence affects two policies: The policy for which she acquires expertise, and a second policy which is not amenable to direct democracy (many such policies exist in practice, see the introduction and the discussion below).

In the second period, a new state of the world is drawn. If the incumbent is reelected, her expertise is used to match the new state. Challengers, having no expertise, are necessarily worse than the incumbent for this policy decision. However, their competence is relevant for the second policy: A competent challenger is more valuable than an incompetent incumbent along that second policy.
A politician is competent (C) with probability $q$ and incompetent (N) with probability $1 - q$. The competence type of the incumbent is denoted $\theta$ and is constant across periods. The cost to acquire expertise level $e$ is

$$\tau_\theta(e) = \frac{e^2}{2\tau_\theta}$$

where $\tau_\theta$ is a productivity parameter and is increasing in competence. The ratio $\frac{\tau_N}{\tau_C} < 1$ will be 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.

The technology that maps expertise into signals works as follows: If the incumbent has expertise $e$, then she learns that $h = R$ with probability $e$. Otherwise, she gets no signal. So, in particular, the incumbent gets no signal if $h = S$, or if $h = R$ but $e = 0$. In the second period, a reelected incumbent keeps the same expertise as in the first period, and learns about the new state of the world according to the same signal technology.

Voters are assumed to be perfectly homogenous (this avoids concerns about the tyranny of the majority, orthogonal to the analysis). Their utility has two separable components, corresponding to the two policy dimensions. Let $u(x_t, h_t)$ denote the utility they get in period $t \in \{1, 2\}$, if the state is $h_t$ and the policy is $x_t$. The first-policy component is simply given by $u(x_1, h_1) + u(x_2, h_2)$ (discounting across period is irrelevant to the analysis and omitted). The second-policy utility is equal to $V$ if the politician is competent and to 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: \( u \) is equal to zero if \( x = S \), regardless of the state; \( u(R, S) = -1 \) (bad reform), and \( u(R, R) = B > 0 \) (good reform). It is assumed throughout that reforms are 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.

**Elections**

 Voters’ strategy is defined by the reelection probability \( f(x, h) \) for each outcome of the first period, and is chosen optimally given voters’ belief about the incumbent’s expertise and competence. The incumbent gets a payoff of 1 if reelected in the second period. This reelection payoff is, of course, what motivates her to put effort in the first period.

In the second period, a challenger, having no expertise and no information, chooses policy \( S \), which is optimal given the prior. A reelected incumbent with expertise \( e \) learns that the state is \( R \) with probability \( e \), in which case she chooses \( x = R \). Otherwise, she chooses \( S \).

**Initiatives**

Direct democracy is modeled as voters’ ability to amend their elected official’s policy.\(^{18}\) After learning the state in the second period, citizens can modify the second-period policy, if needed, by implementing the reform \( R \) to match the state. This is the only relevant amendment in equilibrium: Politicians’ only mistakes in this model arise when they choose \( S \) by default, when \( R \) was in fact needed.\(^{19}\)

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. If \( c = 0 \), on the
to divert away public resources for personal reasons, tasks which are less visible and more difficult to assess in the short term.

\(^{18}\) Among existing direct democracy institutions, this would formally correspond to the notion of “popular referendum,” as discussed by the Founding Fathers of the US constitution. However, here the policy is amended by citizens only in favor of the reform, which is closer, in spirit, to a citizen initiative.

\(^{19}\) Allowing for amendments in the first period does not affect the strategic environment of the problem, and is omitted again for the sake of simplicity.
other hand, direct democracy bears no direct cost. (Remarkably, direct democracy may turn out to be inefficient, even in that case.) In between \( e \) may be conceived as a legal costs (e.g., signature requirements), technological and cognitive costs (e.g., the cost of producing, disseminating, and acquiring detailed policy information), or a combination of both. In any case, \( e \) will be restricted without loss of generality to the interval \([0, B]\).

Citizens’ expected utility in the second period is modified as follows: if the incumbent with type \( \theta \) is reelected,

\[
U(e) = Be\rho + (1 - e)(B - c)\rho + I_{\theta = C}V,
\]

where \( e \) is the expertise of the incumbent. If a challenger is elected, the expected utility is simply \( \underline{U} = \rho(B - c) + qV \). 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 \( \rho \) with a challenger.

Finally, it is assumed that \( B\rho < qV \), which guarantees that

\[
U(C) > \underline{U} > U(N). \tag{1}
\]

These inequalities mean 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. The assumption is made for convenience, as it avoids corner equilibria.

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 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.
- If the incumbent is reelected, she may observe a signal about \( h_2 \) (depending on expertise and \( h_2 \)).
- Whoever was elected 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.

4 Equilibrium and Political Contract

This model features several equilibria, including a “zero expertise” equilibrium, in which the incumbent, regardless of her type, never acquires expertise and is always reelected. It also includes a Pareto-dominated equilibrium 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\textsuperscript{20} 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

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

$$\bar{p}^r = \frac{qV - e_N \rho c}{V + c\rho(e_C - e_N)}.$$\[20\]

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

$$p[V + \rho e_C B + \rho(1 - e_C)(B - c)] + (1 - p)[\rho e_N B + \rho(1 - e_N)(B - c)],$$

\textsuperscript{20}The intuitive claim that this equilibrium is Pareto dominated is formally shown in Appendix B.1
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 \( f(S, S) = f(R, R) = 1, f(R, S) = 0 \).

### 4.2 Incumbent’s Choice

In the first period, the incumbent chooses the expertise level \( e_{\theta} \in [0, 1] \) which solves the optimization problem\(^{21}\)

\[
\max_{e \in [0, 1]} \left[ (1 - \rho) + \rho e + (1 - e)\rho f \right] - \tau(\theta, e),
\]

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 \( e_{\theta} = (1 - f)\tau(\theta, \rho) \).

### 4.3 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, when \( c < c_2 \), there are two natural equilibria: (i) a responsive equilibrium, uniquely characterized by the reelection probability \( f^* \in [0, 1] \), implicitly defined by the smallest solution of

\[
Z(f) = e_{N}(f^*)[1 + c\phi_1] - e_{N}^2(f^*)c\phi_1 - e_{C}(f^*)[1 - c\phi_2] - e_{C}^2(f^*)c\phi_2 = 0. \tag{2}
\]

\(^{21}\)The 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 \( h = S \).
(ii) an unresponsive equilibrium, where $f = 0$.

Moreover, 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.

### 4.4 Equilibrium Selection

The analysis has uncovered two equilibria: responsive and unresponsive ones. The unresponsive equilibrium is straightforward to study: any incumbent mistake is punished maximally ($f = 0$) and independently of all parameters of the model, and achieves a uniform, maximal incentive for politicians.

While it has the virtue of simplicity, the unresponsive equilibrium generates untenable predictions. In particular, it predicts that politicians’ behavior and voters’ reelection strategy are not affected whatsoever by the presence of direct democracy. Such invariance is violated in all empirical studies of direct democracy that we are aware of (Bali and Davis, 2007; Boehmke and Patty, 2007; Bowler and Donovan, 1998; Funk and Gathmann, 2011 and 2013; Gerber, 1996; Kelleher and Wolak, 2007; Matsusaka 1992 and 2008; Rydberg, 2010; Dyck, 2009).

From a theoretical standpoint too, there are several compelling reasons why voters might be unable to induce the highest theoretical level of expertise and screening. In particular, Appendix B.2 shows that when the quality of candidates varies endogenously with electoral incentives, a participation constraint arises, which sets an upper bound on the amount of pressure that voters are able to impose on the incumbent, and rules out the unresponsive equilibrium. While this richer model seems realistic and consistent with the literature on political selection (Caselli and Morelli, 2004; Mattozzi and Merlo, 2008; Messner and Polborn, 2004), explicitly modelling a participation constraint is not necessary for the existence of a responsive equilibrium, nor for the responsibility substitution effect. To present the theoretical mechanism as clearly as possible, the main text focuses on this more parsimonious setting.

---

22In that extension, the incumbent, after learning her signal, can choose some outside option instead of running for reelection.
The next sections focus on the responsive equilibrium, and refer to it as ‘the’ equilibrium.

5 Disincentive Multiplier 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

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

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 (3) 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 3), 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 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 (3), 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 displays the equilibrium reelection probability (derived from the voter’s indifference condition, corresponding to the zero of the function $Z(f)$ defined in (2) 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.

Let \( EU(c) \) 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 \( c \in [c_1, c_2] \).

The intuition for this result is as follows. An increase in \( c \) has two consequences:

- The direct policy loss associated with a politician’s mistake in the second period, which occurs with probability \( \rho[1 - E(e_\theta)] \), increases.
- The probability of a policy mistake decreases (responsibility substitution), because (a) the electoral incentive \( 1 - f \) increases (incentive gain), and (b) the average expertise is increasing in \( 1 - f \) (expertise gain).

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.

---

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.
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 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: \( E(e_0)|_{c=B} \geq 1/2 \).*

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, two key factors which modulate the impact of direct democracy are identified:

- The initial quality of representative democracy: when equilibrium expertise under representative democracy is high enough, voters’ welfare is lower even when direct democracy is costless

- The level of type separation: 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 by no means creates that quick interest in affairs which its originators had hoped to see it excite. It 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, in 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): dissatisfaction towards government and incumbents is higher in states with direct democracy.

As shown in Section 5, the model predicts that direct democracy institutions lead to a more lenient reelection probability _conditional on a policy mistake_. 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 also become more likely.

**Implication 2** The cost of direct democracy

(i) increases the incumbent’s *ex ante* reelection rate if and only if \(E(e_\theta) > 1/2\);

(ii) decreases the incumbent’s reelection rate *conditional on a policy mistake*.

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 they are statistically higher in times of bad economic performance.

### 6.2 Likelihood of Reform 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 \([0, 1] \times [c_1, c_2]\): \(\frac{\partial^2 f^*}{\partial c \partial \rho} > 0\).

To understand the second result, one should first observe that, 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 \(\tau_\theta \rho(1 - f)\) 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 arguable higher dissatisfaction with representative government). This prediction is supported by Boemhke (2005), who finds that the size of the budget deficit is indeed 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 (2), 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 the intuition contained in Figure 2.

**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
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 a higher potential to generate pernicious indirect effects for intermediate values of the politicians’ expected competence.

When initiatives were introduced in the United States, politicians were often perceived as weak and controlled by large corporations. At least some important aspects of politician’s competence have arguably improved over the last century.\(^{24}\) The model suggests that this increased average competence should come along with an increased relevance of the responsibility substitution effects.

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.

\(^{24}\)The improvement in average competence can be related to both better accountability and information mechanisms and to the improvements in the resources that politicians can access to deliver good policies to voters, especially relevant for complex issues where voters are likely not to be able to use initiatives (captured by the term \( V \)).
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

\[
    c_1 = 1 - \frac{\tau}{\phi_2 + \phi_1\tau}, \quad (4)
\]

\[
    c_2 = 1 - \frac{\tau}{\phi_2(1 - \rho) + \phi_1(1 - \tau \rho)}. \quad (5)
\]

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 to induce a positive expertise level, can only hold in equilibrium if voters weakly prefer the challenger to the incumbent, conditional on a policy mistake $(S, R)$:

\[
    \frac{q(1 - e_C(f))}{q(1 - e_C(f)) + (1 - q)(1 - e_N(f))} = \frac{qV - c\rho e_N(f)}{V + c\rho(e_C(f) - e_N(f))}. \tag{12}
\]
This condition may be re-expressed as

\[ a(f) = b(f), \]  

(6)

where \( a(f) = e_N(f)[1 + c\phi_1] - e_N'(f)c\phi_1 \) and \( b(f) = e_C(f)[1 - c\phi_2] + e_C'(f)c\phi_2 \). 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.\(^{26}\) An equilibrium probability \( f \) is a root of the function \( a - b \). It 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 \( f < f_b \), \( a'(f) \) and \( b'(f) \) 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 (6) 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 \( c \leq c_2 \).

\(^{26}\) 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 \( f > f_b \).
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 \(^{27}\) 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 (6) is satisfied for some positive expertise obtains if and only if (6) is satisfied for some \(f \in (0, 1)\). Since \(a(1) = b(1) = 0\) and \(\frac{a'(f)}{b'(f)}\) 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 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:

\[
\frac{dEU}{dc} = \left(\frac{df^*/dc}{Incentive gain}\right) \times (B + c)((1 - q)e_N' + qe_C') - ((1 - q)(1 - e_N) + q(1 - e_C)) \quad \text{Expertise gain} \quad \text{Direct policy loss}
\]

The incentive gain can be computed by applying the Implicit Function Theorem to the equation that defines \(f^*\), which is \(e_C(1 - \phi_2 c + e_C \phi_2 c) - e_N(1 + \phi_1 c - e_N \phi_1 c) = 0\). This yields

\[
\frac{df^*}{dc} = \frac{1}{c e_C'} \frac{(1 - e_N)e_N \phi_1 c + (1 - e_C)e_C \phi_2 c}{e_N' \phi_1 c - 1 - \phi_1 c + e_C'(2e_C \phi_2 c + 1 - \phi_2 c)}.
\]

Using the fact that \(e_N' = e_C' \tau\) and substituting the equation defining \(f^*\) into the denominator of \(\frac{df^*}{dc}\) (also the fact that \(e_N = \tau e_C\)), the following obtains:

\[
\frac{df^*}{dc} = \frac{1}{ce_C'} \frac{(1 - e_N)e_N \phi_1 c + (1 - e_C)e_C \phi_2 c}{\tau e_N \phi_1 c + e_C \phi_2 c} = \frac{1}{ce_C'} \frac{(1 - e_N)\tau(1 - q) + (1 - e_C)q}{\tau^2(1 - q) + q}
\]

\(^{27}\)Notice that \(f = 1\) is always a solution of \(Z(f) = 0\).
As a consequence, the derivative becomes

\[
B + c \frac{(1 - q)\tau + q}{\tau^2(1 - q) + q}[(1 - e_N)\tau(1 - q) + (1 - e_C)q] - ((1 - q)(1 - e_N) + q(1 - e_C))
\]

Letting \(\Lambda(\tau) = B + c \frac{(1 - q)\tau + q}{\tau^2(1 - q) + q} > 1\), the derivative can be re-expressed as

\[
D = (1 - e_C)q(\Lambda(\tau) - 1) - (1 - e_N)(1 - q)(1 - \Lambda(\tau)\tau).
\]

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\(^{28}\)) such that \(\forall \tau > \tau^*, D > 0\). ■

**Proof of Proposition 5.** i) 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 2 (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

\[
(1 - f)\rho = \left[\frac{\tau}{q} + \frac{1}{1 - q} - \frac{V}{\rho C}(1 - \tau)\right]^{-1} \left[\tau \tau_N \frac{q}{q} + \frac{\tau_C}{1 - q}\right]
\]

\(^{28}\)When \(\tau = 1/2\),

\[
\Lambda\tau = 1 + c \frac{(1 - q)\tau^2 + \tau q}{(1 - q)\tau^2 + q} \geq \frac{2(1 - q)\tau^2 + 2\tau q}{(1 - q)\tau^2 + q} \geq \frac{(1 - q)\tau + q}{(1 - q)\tau^2 + q} > 1.
\]

Moreover, when \(qe_C > (1 + (1 - q)\Lambda)(1 + \Lambda)^{-1}\), it is possible to show that \(\tau^* = 0\).
(ii) Using the formula above,

\[
\frac{\partial f}{\partial c} = -\frac{V}{\rho c^2}(1 - \tau) \left[ \frac{\tau \tau_N}{q} + \frac{\tau_C}{1 - q} \right]^{-1}
\]

\[
\frac{\partial^2 f}{\partial c \partial \rho} = \frac{V}{\rho^2 c^2}(1 - \tau) \left[ \frac{\tau \tau_N}{q} + \frac{\tau_C}{1 - q} \right]^{-1} > 0
\]

Proof of Proposition 7

The thresholds can be expressed as

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

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.\footnote{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.}

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

28
The above analysis also applies to the more general model presented in the following section.

**B.2 General Model**

This section studies a general model in which there is a unique natural equilibrium, corresponding to the *responsive* equilibrium of the model presented in the main text. This version of the model also features weaker assumptions on the politicians’ cost of expertise acquisition, and 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).

It has been already argued that the unresponsive equilibrium is both intuitively implausible and empirically problematic. This section argues against it on a theoretical level, by showing that the unresponsive equilibrium is not robust to the introduction of an alternative career choice for the incumbent.

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

---

$^{31}$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.
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

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

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 \( \lambda_C > 0 \).

Convexity of the cost function implies that the optimal expertise is characterized by

\[
e_\theta = \left[ \lambda' f(y'_\theta, R) + (1 - \lambda') O_\theta - \lambda f(y_\theta, R) - (1 - \lambda) O_\theta \right] \rho \tau_\theta.
\]

**Proposition 9** 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 \( \frac{1 - \rho}{1 - \rho e} + \frac{\rho (1 - e)}{1 - \rho e} f \), 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 \( \partial^2 L(I)/\partial^2 I = 2\tau_C \rho^2 (1 - \rho)(1 - \rho^2 \tau_C I)^{-3} > 0 \), \( L(I) \) is strictly convex. Moreover, it is decreasing for \( I \) close enough to 0. As a consequence, \( L \) must

\[32\] Otherwise, an incumbent running for reelection would reveal himself as an incompetent type. Voters would, therefore, never reelect her (see (1)), and the incumbent would optimally choose a zero level of expertise.
be monotonic on the range values in \((L(0), L(1))\). □

In light of this result, it is possible to show\(^{33}\) 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\(^{34}\).

**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)**

\[
\kappa(z) = \frac{T''_C(T'_{C}^{-1}(z))}{T''_N(T'_{N}^{-1}(z))} \text{ is weakly decreasing in } z.
\]

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.3 Robustness**

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

\(^{33}\)Details are available under request. The argument closely following the proof of Proposition\(^2\).

\(^{34}\)The proofs of Lemma\(^1\) and of Propositions\(^2, 3, 5\) and\(^4\) apply to this setting with minimal modifications.
**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 direct democracy.** Assuming that citizens can cancel a reform (i.e., move from $R$ to $S$) does not affect the analysis, given the signal structure, since in any natural equilibrium politicians only implement the reform if it is optimal to do so (i.e., $(R, S)$ does not arise in equilibrium).

**Symmetric signal structure.** The model can accommodate a symmetric structure in the incumbent’s signal about the state: given expertise $e$, $\Pr(\text{Learn } R \mid h = R, e) = \Pr(\text{Learn } S \mid h = S, e) = \frac{1+e}{2}$. 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.

**References**


