Explaining Institutional Change:

Why Elected Politicians Implement Direct Democracy

by David Hugh-Jones

Email: hugh-jones@econ.mpg.de

January 12, 2009

Abstract

In existing models of direct democratic institutions, the median voter benefits, but representative politicians are harmed since their policy choices can be overridden. This is a puzzle, since representative politicians were instrumental in creating these institutions. This paper presents a model in which (1) voters are uncertain about their representative's preferences; (2) direct and representative elections are substitute methods for voters to control outcomes. Some politicians benefit from the introduction of direct democracy, since they are more likely to survive representative elections. Historical evidence from the introduction of the initiative, referendum and recall in America broadly supports the theory, which also explains two puzzling empirical results: legislators are trusted less, but reelected more, in US states with direct democracy.

JEL Classification: D72

Keywords: direct democracy, institutional change, referendums

1 Introduction

There are two schools of thought about direct democratic institutions like the US ballot initiative and the popular referendum. One school has it that direct democracy benefits the voters, at the expense of politicians, by allowing voters to override their representatives when they make bad or self-interested choices (Matsusaka 2004, Gerber 1996, Lupia and Matsusaka 2004). The other believes that direct democracy is bad for everybody. Politicians lose the ability to set policy, but the inexpert voters are likely to make bad choices for themselves (Broder 2000, Schrag 2004). The majority of recent political science work, both theoretical and empirical, supports the first school: voters are made better off. Whatever their relative merits, these views both give rise to a fundamental puzzle. Modern direct democratic institutions could not have been introduced without the consent of the politicians then in power. But on either of the accounts above, they would have no reason to do this, as they lose some political power and gain nothing in return.

More broadly, extant rational choice theories of institutional change also fail to explain the introduction of direct democracy. These theories assume (1) that political actors initiate change because they foresee gains from it; (2) complete information among all participants. The assumptions provide a well-defined range of explanations for change. Weak governments can be forced to introduce change by an opposition coalition (Bowler et al. 2002); relatively equal political partners may bargain their way to an efficiency-improving reform (Geddes 1991); electoral changes may be introduced to protect incumbents from new political entrants (Rokkan 1970, Boix 1999); or they may be concessions to those incumbents, to head off deeper change including revolution (Acemoglu and Robinson 2006, Boix 2003). These theories also implicitly explain why change mostly does not happen. Political outsiders want changes that make it easier for them to get power, but incumbents do not, if they can avoid it; also, the outsiders have a commitment problem, as if they win power within the existing system, they become incumbents and lose their incentive to make changes. For example, the British Labour Party supported Proportional Representation when in opposition in the 1980s, but changed its mind after it came to power.

The introduction of direct democracy in America is problematic for these theories. Under complete information, no political party in office gains from introducing direct democracy, as doing so will give the voters veto power over its decisions and effectively bind its hands. Political parties out of office might prefer to bind their opponents' hands; but introducing a constitutional reform which will also inevitably harm them when they come to power seems a bad way to do it. The most promising explanation would be that a weak incumbent was forced to make concessions to interest groups which were not political parties. In fact, though, direct democracy was sometimes introduced by political parties who had just won power and were in a position of strength – in the US, reformist Democrats and Republicans. Furthermore, these parties wanted quite radical change in society; nevertheless they gave the people a veto over these changes. (Section 4 supports these claims in more detail.)

This paper builds a theoretical model in which some elected politicians do gain from introducing direct democracy. This happens in two ways:

1. Politicians who foresee losing office may prefer to tie their opponents' hands, as well as their own, through direct democracy.

2. With politicians' hands tied by direct democratic institutions, voters have less incentive to remove them in representative elections – that is, direct democracy and representative democracy are to some extent substitutes. So politicians are more likely to get reelected, and can do so while implementing policies closer to what they themselves prefer.

The first point has already been discussed: politicians introduce change when they are going to lose office. The innovation is in the second point. When this is added, we get a new result: even some politicians who foresee being reelected will want to introduce the change. An essential condition for point 2 is that voters have incomplete information. Voters remove politicians because they fear what they might do if reelected; politicians' actions before the election change the voters' expectations about this, and so politicians are constrained by elections to curry favour with the electorate. A politician who lessens the electorate's worries, by tieing her hands in the second term, may then increase her freedom of action in the first term.

The model shows that elected politicians may have long-term reasons to favour direct democracy. "Long-term" means that the short-term electoral benefits of introducing direct democracy (and thus being seen as a reformer) are ignored; only the comparative statics of politicians' welfare with and without direct democratic constraints are examined. Also, politicians are purely policy-motivated. A strong office motivation would increase the incentive to introduce direct democracy, by making it more attractive to trade policy power for stable tenure in office.

The model is based on some key assumptions. First, as mentioned, voters have incomplete information – that is, they are uncertain about the best policy to pursue. Representatives, by contrast, are informed.[1] However, representatives may not share the preferences of the electorate. The electoral mechanism provides a way to remove representatives who stray too far from what the voters want. My second assumption is that this mechanism is imperfect. There are occasions when representatives will not be disciplined by representative elections. For example, politicians who intend to retire at the end of their term, or who cannot run for re-election due to term limits, will not be constrained by elections. Certain issues may be too complicated or boring for the electorate to

---

[1]. Exactly how this happens is not a trivial question but is not the focus here. Kessler (2005) endogenizes the choice to become informed and shows that representatives will do so if the cost of acquiring information is not too high. When they do not do so, there is no advantage to representative democracy.

understand. Because of these occasions, voters have an interest in the real preferences of those they elect, and they may also benefit from introducing a second kind of control of politicians – direct democratic restraints such as the citizen's initiative and referendum.

Viewing direct democratic institutions as a curb on the actions of the representative leads to the third assumption: these institutions interact with representative democracy. This happens in at least two ways. First, popular referendums and ballot initiatives are introduced in response to the perceived shortcomings of the policy implemented by representatives. For example, in 1978 California's voters backed Proposition 13 – a swingeing tax limitation which is still controversial 30 years later – not because they were convinced in the abstract of the need for direct popular control over taxation, but because they thought that existing taxes, set by the legislature, were too high. In game-theoretic terms, the representative's choice of policy provides information to voters about the state of the world. Second, the availability of direct democratic measures affects representative elections: voters will be less concerned about a representative's policy preferences if they know that these can be overridden, and representatives will behave differently in the light of that knowledge.

The final assumption is about the different ways direct and representative elections work. A ballot initiative or referendum involves overriding a political decision. That is, the voters change their representative's policy and put another one in its place. In the model this happens after the policy is revealed, but before its consequences are discovered. Representative elections, by contrast, are about judging a political decision after the consequences. The trade-off is between *ex ante* ignorance and learning from bitter experience.

[Table 1 about here]

The key idea of this paper, that the interaction between different kinds of elections provides a motive for elected representatives to support direct democracy, could be modeled in different ways. For example, suppose that politicians vary on some dimension of policy preference, which the voters learn about only after observing an incumbent's policy choices, and on another dimension of, e.g., competence. Under representative democracy, the median voter will face a trade-off between competence and extremism. Politicians above a particular indifference curve will be reelected. Under direct democracy, the indifference curve will shift. Policy extremism will matter less, since the voters can hold this in check with referendums or initiatives, and competence will matter more. So competent extremists will be more likely to be reelected. They then have a reason to favour direct democracy.

This paper focuses on a different trade-off, which exists even when politicians are defined only by their policy preferences. Initiatives and referendums substitute the uninformed choice of the median voter for the informed but perhaps non-centrist choice of the politician. They therefore work best when they are not used, i.e. when the median voter can trust the politician's choice. In effect, from being an elected dictator, the politician becomes a

kind of adviser (though with strong agenda power): the voters can choose to accept or reject his proposals. While an extreme dictator is always dangerous, an extreme adviser can be corrected for. Knowing a politician's preferences helps the voters decide when to trust him. For example, a left-winger who chooses a rightwing policy is less likely to be motivated by ideology than by a recognition of political circumstances. This in turn means that both the median voter, and an incumbent politician, can gain from the incumbent's reelection. On the other hand, direct democracy is not very useful with a newly elected politician whose policy preferences are unknown. So direct democratic restraints shift the voters' calculations in favour of the incumbent. Thus, it improves legislators' reelection chances – as empirical research confirms (Bali and Davis 2007).

Another insight follows from this approach. Since direct democracy allows a wider range of incumbents to be reelected, the average policy distance between voters and politicians increases. This suggests a theoretical basis for another finding that has puzzled empirical researchers: voters trust their politicians less in states with direct democratic institutions (Kelleher and Wolak 2007).

The next section reviews the relevant literature. Section 3 introduces the model. Section 4 discusses historical evidence from the development of direct democracy in the United States during the Progressive era. The conclusion assesses the prospects for incomplete information and signaling models of institutional change more generally.

2 Literature review

There is a large literature on endogenous institutional change. This review examines existing explanations of deliberate change, as opposed to unintended evolution (e.g. Greif and Laitin 2004), and focuses on rational choice theories, which assume that actors have a sense of their interests and form rational expectations about the future. Other factors such as myopia, chance and character certainly influence institutional change but it is extremely hard to weave them into a systematic theory. Finally, I examine reforms of an existing stable system, rather than "constitutional choice" in an institutional vacuum (e.g. O'Donnell et al. 1986; Przeworski 1991; Bawn 1993).

These reforms are often paradoxical, because they are made by powerful actors – the winners under the existing system. Why, then, would they risk changing it? We can categorize existing theories by how they answer this question. First, there are efficiency theories. In these, change takes place because it enables all the players to become better off. This includes the case when winners gain more than losers and can compensate them. To avoid vacuity these theories require an explanation for the times when change does *not* take place. For example,

it may be that under some circumstances players cannot make credible commitments to compensate losers. Geddes (1991) suggests that Latin American civil service reform takes place when political parties are roughly equal in strength, so that both lose more from patronage – in effect they are in a Prisoner's Dilemma, which reform can eliminate. Rogowski (1987) claims that states with open economies are more likely to have proportional representation as a way of mitigating the inefficient distributive politics more associated with first-past-the-post systems.

Second, powerful actors may make reforms under pressure from a rising opponent. Acemoglu and Robinson (2006) and Boix (2003) claim that the rich allow transitions to democracy so as to avoid revolution. Relatedly, those in power may change the rules so as to head off the rise of their competitors, following Di Lampedusa's advice that "everything must change so that everything can stay the same". For example, it has been suggested that ruling parties brought in proportional representation so as to avoid being wiped out electorally by rising socialist parties (Rokkan 1970, Boix 1999). Similarly, Diaz-Cayeros and Magaloni (2001) show how the Mexican PRI set electoral rules so as to grant some, but not too much, power to its opponents, thus sustaining its hegemony for most of the 20th century. Dunleavy and Margetts (1995) have a broadly similar theory of changes in pluralist democracy as being designed to preserve existing elites from the threat of change in a multi-dimensional policy setting, though they also give some implicit credence to the efficiency explanation.

Third, if a group has power, but does not expect to keep it for long, they may change institutions as a way to embed their interests. This is particularly relevant in democratic contexts which already have some institutionalized ways to change leadership. McCubbins et al. (1987) theorize that regulations on bureaucracies are designed to ensure that the political coalition which creates a new law is able to monitor its enforcement. Political changes can effect more permanent changes in the balance of power. Thus Benoit (2004) builds a general theory in which political institutions are changed when there exists a coalition with "fiat power" to change institutions (e.g. with enough of a majority to change the constitution), every member of which will benefit from the change.

A final group of explanations is closest to the spirit of this paper: that of institutions as "credible commitments". North and Weingast (1989) explain the institutions brought in by the Glorious Revolution as a way for the sovereign to commit not to expropriate his subjects. This then allowed greater economic growth, which in turn benefited the sovereign. Acemoglu and Robinson (2006) emphasize that when under pressure, the rich offer democracy as a way to credibly commit to redistributive policies. If they simply promised redistribution, this promise would be broken in times of political stability, and so would not be credible. These papers do not invoke incomplete information: it is assumed that the sovereign would always expropriate if he needed to (the rich would never redistribute if they could avoid it). The literature on the political economy of central bank independence suggests that a Left-wing party may grant central bank independence so as to neutralize voters' worries

about its monetary policy (Milesi-Ferretti 1995). This will make the party more electable, so that it can afford to introduce its preferred policies in other areas.

How do these theories fare at explaining the introduction of direct democracy in the United States? The issue is as follows. The constitutional amendment introducing direct democracy was introduced in most states by the legislature, before it was voted on by the people.[2] Before direct democracy was introduced, legislatures made the law. Afterwards, the people could veto those laws by popular referendum, or introduce laws the legislatures did not want by the ballot initiative. Also, the populace as a whole cannot be a bargaining partner: there is no way that the median voter can "commit" to reelecting legislatures in exchange for any possible gains (cf. Acemoglu 2003). Thus, efficiency explanations are unlikely to work. Nor does it seem plausible that these reforms were granted for fear of something worse. Indeed, handing over law-making power was often a gateway to further constitutional changes. For example, in Oregon, after the 1905 legislature tried to bring in a constitutional convention in order, it was rumored, to rescind the new direct democratic provisions, citizens initiated a constitutional amendment allowing a constitutional convention only after a popular vote. And unlike changes to electoral institutions, direct democracy provides no way to buy off particular opponents and preserve an incumbent's power.

The third explanation seems more plausible: reformists gained temporary power, then introduced direct democracy so as to give the people a veto in case their successors undid their reforms. This argument fits many cases quite well (see Section 4), and the model here formalizes the intuition. However, there remain problems. The people's veto was a double edged sword which could be – and was – used against reformers. In a complete information setting, only politicians who are very likely to lose power will wish to tie their successors' hands in this way. Furthermore, as de Figueiredo (2002) points out, politicians with a long-term perspective can bargain with political opponents to preserve each others' reforms. Finally, constitutional changes in most US states required a 2/3 supermajority in the legislature; and at this time, legislative constituencies were heavily biased towards more conservative rural areas and against the new industrial cities (David and Eisenberg 1961; Hamm and Moncrief 1999). These factors together made it hard to achieve a legislative supermajority for direct democratic reform made only out of extreme reformers who expected to lose office.

The model here combines this explanation with the "credible commitment" story. Some extreme candidates for office, who do not expect to get reelected, prefer direct democracy because it will bind their successors. But less extreme candidates also prefer direct democracy, because voters who can vote down extreme policies in a ballot initiative have less desire to forestall those policies by voting a candidate out. This in turn lets some candidates implement policies closer to their ideal point. Thus, just as a party might delegate power to a central bank in

---

[2]. In Oklahoma, Arizona and Massachusetts a specially elected constitutional convention brought in direct democratic measures as part of a wholesale rewrite of the state constitution.

exchange for more freedom in other areas, a party or candidate may return power to the voters in exchange for more freedom in the face of representative elections.

Incomplete information plays a key role in this story, as it does not in previous credible commitment models of institutional reform. Voters always prefer politicians whose preferences are closer to their own, so if politicians' preferences are visible, they will always prefer an incumbent if and only if he or she is closer than the average to the median voter, whatever the democratic institutions (see Section 3.2). When politicians' preferences are unknown, voters must learn about them from the policies they choose. The restraints of direct democracy make it less tempting for extreme politicians to choose centrist policies so as to get reelected, since they will not be able to do what they want after the election. This in turn means voters who observe a given policy choice will believe the incumbent to be relatively more centrist, and so they reelect after a wider range of observed policy outcomes.

Empirical research has not found a very compelling general explanation of the spread of direct democracy. Scarrow (1997) examines the expansion of direct democracy in Germany and finds that parties introduced the reforms in the quest for short-term electoral advantage. Thus, there was a breakdown of long-term cooperation between the parties, driven by strong inter-party competition. However, in the case which interests us – the introduction of statewide initiative and referendum by state parties, i.e. where parties choose to limit their own power – all her cases show collaboration between the major parties. She explains this as a reaction to declining turnout. This may be a reasonable explanation but it assumes that parties take a long-term perspective on the health of democracy as a whole, which sits oddly with the short-termist competition used to explain other changes. Bowler et al. (2002) survey legislators' support for direct democratic reforms in 3 nations. In their theoretical section they admit candidly that "legislators will find direct democracy an unappealing institution". They do find that ideology predicts support for reforms, which fits the theory presented here. However, predicting behaviour from surveyed opinions, at best a hopeful enterprise, is particularly problematic in this case. (The same applies to Bowler et al. (2006), which finds quite high levels of expressed support for direct democracy among legislators.) They conclude by suggesting that reforms will be forced on weak legislators by outsider coalitions. Smith and Fridkin (2008) provide supporting evidence, showing that during the progressive era, minority party strength predicted the legislature's likelihood of placing on the ballot a constitutional amendment granting the citizen's initiative.

This argument fits many of the states which introduced direct democracy, complementing the case in which reformers gain temporary power (again, see Section 4). And it is intuitively attractive: if legislators have no reason to bind their hands, they must have been forced to do so. However, there is a problem of commitment on both sides. First, legislators may promise to support reform but fail to deliver once in office. Indeed, Bowler et

al. write "[i]t is not surprising... to find examples of legislators supporting bills in favour of CIRs [Citizen Initiated Referendums] while in Opposition, yet opposing an identical bill, or letting authorizing legislation languish, once they win Government". Second, voters or interest groups who have threatened to punish these non-performing legislators may face the problem that the alternative is no better.

Assuming incomplete information – i.e. that voters do not know exactly what their representatives' political preferences are – could resolve these problems in two ways. First, a vote for institutional reform may convince voters that a particular politician is on their side, thus providing them with a forward-looking incentive to vote for him, and therefore giving him an incentive to support reform. This is not modelled here. Second, the institutional reform itself, by weakening politicians' power to implement extreme outcomes, may make voters relatively more tolerant of politicians, and so some politicians may actually prefer direct democratic institutions. Even some politicians who would be reelected under representative democracy can implement policies closer to what they really want under direct democracy and still get reelected, as Proposition 2 below shows. This second possibility is modelled here, and provides a new explanation for the introduction and survival of direct democratic institutions.

This paper also adds to the collection of formal models of direct democracy. The classic complete information models are Romer and Rosenthal (1978) and Gerber (1996). In the Romer-Rosenthal "setter model" the monopoly agenda-setting power of a resource-maximizing bureaucrat means the median voter does not achieve his ideal point in a direct democratic process, in contrast with the competitive model of representative elections. Gerber introduces a representative with an exogenous ideal point who may be threatened by a direct democratic agenda setter; now, this threat may force the representative to move policy towards the median voter, compared to the outcome of unrestrained representative democracy. Besley and Coate (2001) present a model in which direct democracy benefits citizens indirectly by "unbundling" some issues from representative elections. The aim is to explain why direct democracy may bring outcomes closer to the will of the majority, even when representatives compete for election – a process not modelled in the papers above. In a complete information framework, they put forward different reasons why representative democracy may fail to respect the will of the majority on a particular issue, and show that direct democracy ameliorates this. Parties, who choose candidates and whose members may have different preferences from the population as a whole, play a crucial role. Introducing uncertainty about politicians' preferences, as done here, means that representative democracy does not perfectly fulfil the median voter's preferences even without parties.

Some later models introduce incomplete information. Matsusaka and McCarty (2001) recognize the downside of the direct democratic threat: if representatives are unsure of voters' preferences, and have their own policy preferences, then extreme interest groups may move policy away from the median voter's preference by threatening

to introduce a ballot initiative. Maskin and Tirole (2004) use a very simple model, in which voters may not know the best policy for them, to examine the trade-offs between direct democracy, elected representatives, and unelected bureaucrats. Kessler (2005) does the same in a model with a continuous policy space, and shows that direct democracy may lower elected officials' incentive to become informed – a choice not examined here. The paper's framework is similar to this one's, but the interaction between direct and representative democracy is not fully modelled.

Uncertainty is pervasive, but the most important kind is surely the uncertainty of the voter about policy, as discussed in the Introduction. The model here uses a signalling framework to examine policy choice under direct and representative democracy when the voter is uncertain. The interaction between these two ways of controlling policy-makers generates a new insight: incumbent representatives may benefit electorally from direct democracy. Another insight is that direct democracy changes which representatives get elected, for the worse from the point of view of the median voter. Thus, the model gives a theoretical basis for two recent findings that have puzzled empirical researchers, and indeed seem contradictory at first sight. First, voters trust their politicians less in states with direct democratic institutions (Kelleher and Wolak 2007). More importantly, direct democracy appears to improve legislators' reelection chances (Bali and Davis 2007). This is a direct confirmation of the mechanism behind the model's central prediction, and a finding which existing theory fails to explain.

3 The model

There are two periods. In each period $t = 1, 2$ a politician observes a shock $\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$ and chooses a policy p_t. The outcome is $w_t = p_t - \varepsilon_t$. (Numbered subscripts will be dropped when the sense is obvious.) The voters do not observe the shock, only the outcome in each period. All players' utility is quadratic in each period: total utility for a player with ideal point i is $-(i - w_1)^2 - \delta(i - w_2)^2$ where δ is the common discount rate. Rather than simply referring to time, δ can be interpreted as representing the tradeoff between periods 1 and 2, that is, the likelihood, over a range of issues and times, that a politician will be free of the threat of reelection. When this is very low, all players are more concerned with the "pre-election" period 1 effect of direct democracy. When δ is high they are concerned with period 2, and the possibility of restraining the politician with a ballot proposal.

All players' ideal points are private information. After period 1, the voters vote to keep the incumbent or elect a challenger, by majority rule. The median voter's ideal point is 0; Banks and Duggan (2006) provide a median voter theorem for lotteries over outcomes when utility is quadratic. The incumbent and challenger's ideal points k are drawn from a normal distribution $N(0, \sigma_k^2)$ which is independent of ε. Write φ_k for the density function of k and Φ_k for its cumulative distribution function. Similarly φ_ε and Φ_ε for the shock.

In period 2 the politician is not facing reelection. However, under direct democracy, the voters observe p_2 before the policy is implemented, and vote to keep it or replace it by a default policy of 0. This is a rough analogue of a popular referendum process in which voters may choose to challenge the proposed policy and replace it with an existing default, which is assumed to be the optimal policy given the distribution of shocks. (This maximizes the voters' incentive to choose the default and thus makes life as hard as possible for the politician: if the default were less attractive, the politician would be better off.)[3] A more complete model would allow voters to do the same in period 1: we limit direct democracy to period 2 for simplicity's sake. We also allow the politician to step down after period 1 (for example, if the voters would never accept his policies in period 2, and he would do better under a challenger).

Voters learn the proposed policy p before voting on it in period 2. This can be thought of as the effect of an election campaign, in which those with extreme preferences (on either side of the issue) will make costly efforts to inform others about what is at stake, through advertisements, leaflets et cetera. Thus voters learn p (as opposed to the outcome they experience, w) only when there is a direct democratic campaign.

In any period, a politician who knows his ideal point k and the shock ε gets utility from a policy p of $-(p - \varepsilon - k)^2$. This policy utility is single peaked with the peak at $p = k + \varepsilon$. Call $q = k + \varepsilon$ the politician's induced policy ideal point. Single-peakedness implies that, if policies within an interval $[\underline{p}, \bar{p}]$ are available, and a point in the interval's interior $p \in (\underline{p}, \bar{p})$ is chosen, then $q = p$; otherwise a lower or higher policy would have been chosen instead.

3.1 Results

The setup gives us a game of incomplete information. Perfect Bayesian Equilibrium (PBE) is the standard equilibrium concept for this kind of games (see e.g. McCarty and Meirowitz 2007). Using it I find a set of sensible equilibria. Unfortunately other, less sensible ones also exist: these are ruled out with the Intuitive Criterion, a well-known refinement for signaling games (Cho and Kreps 1987). Essentially this requires that if an off-equilibrium action is observed, other players must assume that the deviating player hoped to gain something from the deviation. PBE imposes no restriction on beliefs off-equilibrium and this results in the odd equilibria.

[3.] An alternative assumption would be that there is an initiative process with some kind of open agenda-setting process in which the median voter can select the optimal initiative given p. If so, this essentially becomes a game of cheap talk in which the politician recommends p and the median voter chooses the ultimate policy. This assumption seems optimistic about the openness of direct democratic agenda-setting.

We use the model to compare the expected welfare of politicians under the two different systems. The first proposition replicates the findings of the existing literature: if we ignore the effect of representative elections, no politician wants direct democracy. This is trivial. In period 2, under representative democracy, a politician gets exactly what she wants, while under direct democracy in general she does not.

Proposition 1. *In period 2, any politician weakly prefers representative democracy to direct democracy, and an incumbent reelected from period 1 strictly prefers it.*

All proofs are in the appendix.

Figure 1 shows the basic form of equilibrium behaviour in the first period. If the outcome is in an interval around the median voter's ideal point of 0, then the politician is reelected. Politicians with ideal points in the interval implement their ideal outcome, taking into account the effect of the shock. On either side of this reelection region, there is an area where politicians choose the nearest endpoint of the region rather than their own ideal point: they accept a policy loss now, in order to get reelected and choose policy in period 2. Further out, politicians again implement their own most preferred outcome, and are not reelected.

[Figure 1 about here]

The reelection region is small enough that the voters prefer to reelect an incumbent politician who is at the endpoint – knowing that in fact, this politician would prefer a more extreme policy, and will try to implement it in the second period. If this condition is satisfied, then the voters will be happy to reelect the less extreme incumbents politicians who choose outcomes inside the reelection interval. On the other hand, the reelection region is large enough that voters will not reelect a politician who chooses something outside the reelection region. Since some policy outcomes are never observed in equilibrium – the dashed line in the picture – there is a range of reelection regions satisfying both this constraint and the endpoint reelection constraint. So we have multiple equilibria.

We start the analysis of direct democracy by examining period 2. (Period 2 in representative democracy is straightforward: the elected politician implements her ideal outcome.) There are two cases. Either a challenger, with an unknown ideal point, has been elected, or the incumbent has been reelected, and, as we have seen, the voters know something about his ideal point – at a minimum, they know which side of the median voter he is on, and for centrist politicians they know his ideal point exactly, because he chose it in the first period. If a challenger has been elected, because voters do not know the likely direction of his bias, the choice to trust him or not is the same whatever policy he proposes. As it turns out, the voters accept any policy proposed if $\sigma_\varepsilon^2 > \sigma_k^2$, and reject any policy proposed if $\sigma_\varepsilon^2 < \sigma_k^2$, replacing it with the default of 0. In effect, the choice is between two kinds of shocks, one caused by the unknown policy environment and one caused by an unknown politician's difference from the median voter.

The case with a reelected politician is quite different. Even if the politician is very extreme, because the direction of his bias is known, it can be accounted for. Indeed, whenever the proposed policy goes against the politician's bias, the voters know they will do better by accepting it. That is, if the politician proposes $p_2 < 0$ and the voters know $k_2 > 0$, they will always accept p_2 instead of the default policy of 0. I.e., if a rightwing politician prefers a more leftwing policy to the default, the median voter will have the same preferences. So, whenever the shock leads the politician to "go against" his prior belief, his policy proposal is accepted. As the saying runs, it takes a Nixon to go to China (Cukierman and Tommasi 1998).

Intuitively, we might expect a rightwing politician proposing an extremely rightwing policy to be voted down. This is not necessarily true. If the politician were really extreme, he would not have been reelected in period 1: a very rightwing policy means he must be countering a big leftwards shock. So there are possible equilibria in which very rightwing policies are accepted when a moderate rightist proposes them. However, for simplicity, we assume that when a rightwing politician proposes any rightwing policy, it is rejected. This is the smallest possible equilibrium in terms of the set of policies accepted in period 2, and therefore the worst possible equilibrium for politicians. If some politicians prefer direct democracy under this equilibrium, this will be true in all cases. Period 2 under direct democracy is illustrated in Figure 2.

[Figure 2 about here.]

We can now say a bit more about period 1. The different results when a challenger is reelected (his proposal is always accepted if $\sigma_k^2 \leqslant \sigma_\varepsilon^2$, and always rejected when $\sigma_\varepsilon^2 < \sigma_k^2$) make a large difference to the period 1 equilibria. In fact, if $\sigma_\varepsilon^2 < \sigma_k^2$, a newly elected challenger is a useless object: the voters do not know enough about his ideal point to trust him, whatever policies he proposes! On the other hand, even a very extreme incumbent, whose ideal point is known, can be trusted some of the time. For this reason, in equilibrium when $\sigma_\varepsilon^2 < \sigma_k^2$, *incumbents are always reelected*. This result contains the central intuition of this paper. Direct democracy turns politicians from dictators into advisers. A dictator whose preferences are known is still dangerous, but an adviser whose preferences are known can be corrected for. Thus, direct democracy raises the relative value of incumbency.

However, this extreme result only applies when $\sigma_\varepsilon^2 < \sigma_k^2$ – that is, when an unconstrained politician is actually worse for the median voter than choosing himself. When $\sigma_\varepsilon^2 \geqslant \sigma_k^2$, the result is less clear. (A really extreme politician would almost never propose a trustworthy policy in period 2, so the voters would usually choose the default policy of 0 and suffer the effects of the policy shock. So it is better for them to elect an unknown challenger if he is fairly trustworthy.)

Concretely, what does $\sigma_\varepsilon^2 < \sigma_k^2$ mean? The variance of ε measures the unpredictability of the policy environment – not necessarily objectively, but to the voters. A number of factors affect this. σ_ε^2 will be lower when voters have direct knowledge of outcomes, for example when they feel the effects of the economy on their pocketbooks,

whereas it will be higher for, say, the progress of a foreign war. σ_ε^2 will be low for "easy" policy areas such as abortion law, and high for abstruse ones such as financial crisis management. It will be low when there is widespread consensus on the best course of action, and high when experts disagree. Finally, σ_ε^2 will rise during times of social change and fall when social patterns are well-known, stable and predictable. On the other side of the inequality, σ_k^2 is rather easier to understand. When politicians are trustworthy, or when the campaign process provides a lot of information on candidates so that bad ones can be weeded out before taking office, σ_k^2 will be low. In times when politicians are untrustworthy, and when the campaign process is uninformative, σ_k^2 will be higher. It will also be low when parties are strong, and are able to commit to well-defined programs, and high when parties are weak and individual candidates rely on their local power base rather than on campaign promises. Section 4 argues that during the Progressive Era in America, σ_k^2 was high.

When $\sigma_\varepsilon^2 < \sigma_k^2$, before reelection, some incumbents prefer direct democracy (contrast this with Proposition 1). This does not only include extremists, who are not reelected and therefore want to bind their successors hands; there are always some politicians who would be reelected under representative democracy, but who nevertheless prefer direct democracy. Their loss from direct democracy in period 2 is balanced either by being reelected instead of removed from office, or by being reelected with a policy closer to one's own ideal point.

All these results are restated formally below. Proofs are in the appendix.

Lemma 1. *Under representative democracy, any equilibrium satisfying the Intuitive Criterion can be characterized by a pair of points $w_\star < 0$ and $w^\star > 0$.*

Equilibrium behaviour is as follows: in period 2 incumbents and challengers implement their ideal point $p_2 = q_2$. In period 1, the voters reelect the incumbent if and only if the outcome $w_1 \in [w_\star, w^\star]$. Incumbents with $w_\star \leqslant k \leqslant w^\star$ implement their ideal point $p_1 = q_1$. So do incumbents with $k \geqslant k^\star$, where $k^\star > w^\star$ and $k^\star$ is uniquely determined by $w^\star$; and incumbents with $k \leqslant k_\star < w_\star$ where $k_\star$ is uniquely determined by $w_\star$. Incumbents with $k \in (w^\star, k^\star)$ (respectively $k \in (k_\star, w_\star)$) set $p_1 = w^\star + \varepsilon_1$ (respectively $p_1 = -w_\star + \varepsilon_1$) so that the outcome is $w_1 = w^\star$ ($w_1 = w^\star$).

The set of possible equilibria is as follows: $w^\star$ may be any point satisfying $\underline{w} \leqslant w^\star \leqslant \bar{w}$; similarly $w_\star$ may be any point satisfying $-\underline{w} \geqslant w_\star \geqslant -\bar{w}$. The bounds $\underline{w}$ and $\bar{w}$ are non-negative and fixed by the parameters. When $\delta \geqslant 1/2$, $\underline{w} = 0$, i.e. there are equilibria in which politicians must implement the median voter's ideal point, $w_1 = 0$ to be reelected.

Lemma 2. *In any equilibrium satisfying the Intuitive Criterion, in period 2 under direct democracy, write A for the acceptance region such that $p_2 \in A$ is accepted, and write Φ_k for the voters' probability distribution over the politician's k at the start of period 2.*

1. Politicians with ideal policy $q_2 = k_2 + \varepsilon_2$ choose $p_2 = q_2$ if $q_2 \in A$. Otherwise they choose the closest point in A to q_2. (If q_2 lies closer to 0 than to any other point in A, then any point in A^C may be chosen; we assume they choose 0.)

2. $\mathbb{R}_- \subset A$ when the politician is known by the voters to have $k \geqslant 0$. When $k \leqslant 0$ for sure, $\mathbb{R}_+ \subset A$.

3. If the voters always know the sign of k and have bounds on k (e.g. $\Phi(k) = 0$ for $k < k'$, $\Phi(k) = 1$ for $k > k''$, $0 < k'$), at the start of period 2, then the following is a possible equilibrium: $A = \mathbb{R}_-$ when k is known positive, $A = \mathbb{R}$ when $k = 0$, $A = \mathbb{R}_+$ when k is known negative.

Lemma 3. *Under direct democracy, in period 2, if a challenger has been elected, and if $\sigma_k^2 > \sigma_\varepsilon^2$, the voters always reject the challenger's proposal and choose $p_2 = 0$ instead. If $\sigma_k^2 \leqslant \sigma_\varepsilon^2$ then the challenger proposes $p_2 = q_2$ and the voters always accept.*

Proposition 2. *Under direct democracy, if $\sigma_\varepsilon^2 < \sigma_k^2$, then all politicians are reelected in period 1. Furthermore, at least some politicians who would be reelected under representative democracy nevertheless strictly prefer direct democracy.*

Equilibria when $\sigma_k^2 < \sigma_\varepsilon^2$ can be complex. In particular, even politicians who would get reelected by choosing their ideal point in period 1 do not necessarily do so. There may be an incentive to move towards the centre so as to persuade the electorate that one is more centrist, in order to get more policies accepted in Period 2. (In the smallest equilibrium described in Lemma 2 point 3, this does not happen.) I hypothesize, but do not prove, that for some values of the parameters and some equilibria, there are politicians who prefer direct democracy even though they are reelected under representative democracy, as in Proposition 2.

3.2   A complete information benchmark

To emphasize the effect of voters' incomplete information in the model, we can compare it with a benchmark in which everything is unchanged except that the voters learn a politician's ideal point k with certainty after he is elected.

Under representative democracy, voters know that the incumbent will implement his ideal point in period 2. So after period 1, they ignore w_1 and simply reelect the incumbent iff $k_1^2 \leqslant \sigma_k^2$ i.e. if $|k_1| \leqslant \sigma^k$. Therefore, in period 1 the politician always chooses $w_1 = k_1$.

Under direct democracy, whether the incumbent or a challenger is reelected, the voters know k_2 in period 2, and set the acceptance region A accordingly. As the voters' loss in period 2 is increasing in $|k_2|$,[4] again in period 1 they reelect if and only if $k_1^2 \leqslant \sigma_k^2$. Once again, in period 1 the politician chooses $w_1 = k_1$. As politicians who will not be reelected wish to control their successors through direct democracy, and politicians who will be reelected do not wish to be controlled, the following proposition is now easy to see.

Proposition 3. *if voters learn k with certainty in each period, then only politicians with $|k_1| > \sigma_k$, who will not be reelected, strictly prefer direct democracy in period 2; all other politicians strictly prefer representative democracy.*

Compare this to Proposition 2. There, some politicians prefer direct democracy even though they would be reelected under either institution, because it strengthens their position as an incumbent and gives them more freedom of action in period 1. Full voter information removes this effect. Now only those in a position of political weakness will prefer direct democracy. This is in line with some conventional explanations of institutional change. But as we shall see when we discuss the history of US direct democratic institutions, this explanation does not always fit the facts: the initiative and referendum were not always granted by weak politicians who expected to lose office. The model suggests why. Suppose there is a legislature composed of some politicians who have been reelected, and others who are in their first period. (Ignore the strategic issues of policy choice in the legislature and just assume that only politicians who will not get reelected prefer direct democracy: the logic is generally quite plausible.) Then, even if all legislators are newly elected, all those with ideal points $|k| < \sigma_k$ – within one standard deviation of the mean, i.e. about two-thirds of legislators – will prefer representative democracy. And if some legislators were reelected the proportion of opponents would get even higher, making a pro-reform majority, let alone a supermajority, rare indeed.

Adding incomplete information changes the story, so that even politicians in a strong position may favour direct democracy. This improves the odds of a reform passing. It also seems to fit the historical facts, to which I now turn.

4   Historical Discussion

How did direct democracy come to the US?[5] The standard story is as follows. In the 1900s, the political system in most US states came to be seen as broken. Legislators were in hock to special interests from business. The electoral system had been corrupted in some municipalities; on the other hand, many new industrial cities were

---

4. This is shown by equation (21) in the Appendix.

5. This section draws on Allswang (2000), Piott (2003), Goebel (2002), Bowler and Donovan (2006), Smith and Lubinski (2002) and Haynes (1919), in particular on the detail in Piott's book.

grossly underrepresented by their allocation of seats in the legislature. Neither of the main parties was trusted: turnout and partisanship, high throughout the 19th century, declined. (In terms of the previous section, σ_k^2 was high.)

The ideas of direct democracy were first taken up by the Populists in the 1890s, seeking a way round the corrupted and unrepresentative system of state representative democracy. The Populists did not succeed in general, but they did pass the first state initiative law in South Dakota. In the 1900s a succession of reformers such as William U'Ren continued the fight and eventually forced legislatures to concede in half the US states. Supporters of direct democracy included the Progressive movement, Labor unions and farmers' groups, as well as some Prohibitionists, all of whom hoped to use the ballot box to circumvent obstructionist legislatures.

This story shows us two of the mechanisms for institutional change discussed in Section 2. First, a group that gains short-term power, but does not expect to keep it, may use reform to entrench the policy changes it wants. This is what the Populists did in South Dakota. Similarly, in Oklahoma, direct democracy was introduced by a constitutional convention which was 90% Democratic. Previously Republicans had dominated state politics; this was the turning point. Later the Democratic legislature restricted direct democracy. The change of heart suggests that at convention time, Democrats sought to protect their preferred policies from Republican legislators; they changed their mind only when they foresaw a stable period of Democratic power. Similarly, in Arizona, a constitutional convention introduced direct democracy.

The more important theory is simply that divided or weak legislators may be pressured by outside forces into accepting change. In Oregon, the indefatigable activist William U'Ren got himself got himself elected along with the Populists, and then used logrolling tactics to win over the divided major parties in the legislature, offering Populist legislators' votes in exchange for support for direct democracy. Before the bill introducing the initiative and referendum came up for a second vote in 1901, U'Ren helped legislators who had voted for him before get reelected; further lobbying ensured it passed again. By 1904 U'Ren, now a formidable political force in the State, could say of a measure threatening the new institutions: "The man who votes for the Brownell bill... is digging his own political grave and getting into it."[6] In some state elections, the Direct Legislation League extracted pledges to support direct democracy from candidates for the legislature. For instance, in the Maine election of 1906, many candidates who had not pledged their support lost their seats; both houses of the subsequent legislature unanimously passed a moderate bill to introduce the initiative and referendum.

These theories are powerful and explain many of the historical cases. They also have the advantage of explaining the overall pattern that direct democracy was introduced mostly in Western states. These states had weak party systems so that candidates cared more about their individual electoral fortunes than about preserving the legisla-

---

6. Quoted in Piott (2003) p. 43.

ture's power as a whole. Reformers had less success in the East where parties were strong, and in the one-party South.

However, the history also reveals other actors who were instrumental in introducing direct democracy in many states. First, there were "reform" or "insurgent" Democrat and Republican politicians, progressives who aimed to wrest control of their parties from the machine. They were disciplined political groups, not divided individuals, and rather than gaining temporary power, they aimed to take over the political mainstream in the long-term, and often succeeded. When, as often, they supported direct democratic measures, they did so out of principle rather than short-term expediency.

For example, in California, the Direct Legislation League had failed throughout the 1900s to get measures brought in by pressure on individual legislators: the Southern Pacific Railroad, exerted its great influence over parties to prevent a vote. Direct democracy was introduced in 1911 after the Lincoln-Roosevelt League of Progressive Republicans under Hiram Johnson had swept into office. This group had gained control of the Republicans after the introduction of direct primaries in 1910, and passed the initiative and referendum by huge margins – unanimously in the Assembly. In North Dakota, direct legislation in its final form was passed in 1918 by the Nonpartisan League of small farmers – a kind of throwback to the populists which collapsed shortly thereafter. This fits the "short-term power" story. But reform-oriented legislators, having gained the upper hand over the party machines, passed no less than four bills introducing direct democracy in 1911. In the end a relatively weak form, with high signature requirements to put a measure on the ballot, won out; nevertheless this multitude of alternatives, some quite radical, does not fit the picture of legislators being pushed unwillingly to act.

Why were these reformist members of the main parties keen on binding their future hands? Surely, one reason was the short-term reputational advantage to be gained by introducing the measures. But there were plenty of other reforms that voters wanted, and that would not have tied their hands as legislators in future. It is plausible that, as in the model, they traded off a direct loss in freedom of action, for an increase in freedom of action and reelectability due to future voters' trusting politicians more.

In other cases, traditional parties introduced direct legislation without strong pressure from outside. In Missouri, at a time when the Missouri Direct Legislation League had been weakened by the voters' rejection of direct legislation in 1903, a Democratic governor, Senate and House reintroduced the idea in 1907. Far from embracing the new institutions, Missourians then approved precisely none of the fourteen initiative measures brought before them between 1910 and 1918 (although they also rejected several referred statutes and legislative constitutional amendments). Again, it seems that at least some legislators foresaw they could work better within the new system. Similarly, in Arkansas, after the original direct democracy measure of 1910 had been weakened by the courts, a new law was brought in by the legislature in 1916, but failed at the polls after opposition by prohibitionists. In Michigan, where the Republicans were dominant and were to remain so until the 1930s, a constitu-

tional convention passed a weak measure in 1908; in 1912, prodded by the state governor, the legislature strengthened it.

In fact, state Governors were often key players in calling for direct democracy, including the already-mentioned Hiram Johnson in California and Joseph Folk in Missouri. In Colorado, the popular Progressive Democrat John Shafroth was elected Governor in 1908 after 5 terms in the state Congress. He campaigned for the initiative and referendum throughout the legislative session, brought in a special session to adopt it, and supported the subsequent measure when it came before the voters, at the time of his successful reelection bid. In Arizona, the president of the 1916 constitutional convention was one George Hunt, a state Senator who had first introduced a bill supporting direct democracy in 1899, and who later became Governor.

Governors' support of direct democracy has continued to this day, as witness Arnold Schwarzenegger's use of ballot initiatives in 2005. Perhaps one reason is that it allows them to get round an obdurate legislature. But representative politicians introducing their own initiatives is a relatively modern trend (Smith and Tolbert 2001). In the Progressive era the trade-off modeled above seems a more likely cause of their support.

To sum up, outside pressure on legislators was not enough to introduce direct democracy. Nor, except in the cases of North and South Dakota, did Populist radicals achieve this aim. As Piott puts it (p. 254): "progressive control of state government was key to winning legislative approval of these measures." Often, these progressives were within the main parties and had long-term goals to keep their political power, and often, they introduced reforms not when electorally weak or divided but after winning elections with large majorities. They must therefore have gained something beyond short-term advantage from doing so. The model shows there were two long-term advantages: direct democracy could restrain their opponents who might regain office through the remaining corruption in the political system; and it would win back the electorate's trust, allowing them to pursue their goals more effectively when in power.[7]

One question is harder to answer. How predictable was the political environment of the Progressive era – in terms of the model, how high was σ_ε^2? Clearly, the 1890s were a time of economic change and dislocation. At the same time, by the 1900s, a coherent Progressive program had been formed on a number of issues, and certain states and cities had undertaken reforms and thus acted as pathfinders for others. It seems probable that by the 1910s, when most of the direct democratic reforms were brought in, many voters saw politicians less as stewards to face an unknown and uncertain environment, and more as agents who ought to be implementing a thoroughly worked-out program of reform, whose outline at least was clear.

----

[7]. A point not discussed here is that in the era described, the median voter under direct democracy could be very far from the median voter for the legislature, due to malapportionment and other forms of corruption. If Progressives believed that they were close to the real median voter, while their future opponents might be far on the other side, this would strengthen their motives to support direct democracy.

5   Conclusion

Top-down institutional reform, in which those in power voluntarily give it up, is the hardest kind of institutional change to explain. The introduction of direct democracy is one such tough case. Existing theories fail to appreciate how hard it is: in general, to change the constitution, a supermajority of legislators must be won over. This paper has suggested that even without internal divisions between those in power, it may still be possible for democratic representative to do better under direct democracy. The power handed over is compensated by increased trust at the time of representative elections. Historical data and voting data supports my case.

Two broader conclusions come out of this research. For research on direct democracy, there are complex links between representative and direct elections. We already know how legislators may be threatened by the initiative (Gerber 1996), learn from initiatives in nearby states (Boehmke 2005), or seek to weaken their implementation (Gerber et al. 2001). Scholars have observed that politicians are getting directly involved in initiative campaigns (Smith and Tolbert 2001). For example, Governor Arnold Schwarzenegger called a special election for five initiatives in 2005. Modelling this process, in particular of politicians' and initiative proponents' role as trusted agents, could enhance our understanding further (see also Garrett 2004). In particular, the interaction between House, Senate and Governor seems likely to help explain direct democracy's appeal to the political class: Governors can use ballot initiatives to bypass the legislature as a whole and the House majority may wish to override recalcitrant Senators the same way.

More generally, our present models of institutional change tend to see the world as a bargaining process in which people know what they are getting. There is room for a middle ground between this rational-choice approach and the historicists who emphasize uncertainty and habit. Incomplete information, signalling and trust ought to be brought in to our understanding of institutional change.

Appendix: proofs

Throughout the proofs I examine the situation of politicians with $k > 0$ and/or voters when $w > 0$. Symmetric proofs will hold when $k < 0$ or $w < 0$. For ease of exposition I prove the Lemmas before Proposition 1.

Lemma 1. *Under representative democracy, any equilibrium satisfying the Intuitive Criterion can be characterized by a pair of points $w_\star < 0$ and $w^\star > 0$.*

Equilibrium behaviour is as follows: in period 2 incumbents and challengers implement their ideal point $p_2 = q_2$.

In period 1, the voters reelect the incumbent if and only if the outcome $w_1 \in [w_\star, w^\star]$. Incumbents with $w_\star \leqslant k \leqslant w^\star$

implement their ideal point $p_1 = q_1$. So do incumbents with $k \geqslant k^\star$, where $k^\star > w^\star$ and $k^\star$ is uniquely determined

by $w^\star$; and incumbents with $k \leqslant k_\star < w_\star$ where $k_\star$ is uniquely determined by $w_\star$. Incumbents with $k \in (w^\star, k^\star)$

(respectively $k \in (k_\star, w_\star)$) set $p_1 = w^\star + \varepsilon_1$ (respectively $p_1 = -w_\star + \varepsilon_1$) so that the outcome is $w_1 = w^\star$ ($w_1 = w^\star$).

The set of possible equilibria is as follows: $w^\star$ may be any point satisfying $\underline{w} \leqslant w^\star \leqslant \bar{w}$; similarly $w_\star$ may be any

point satisfying $-\underline{w} \geqslant w_\star \geqslant -\bar{w}$. The bounds $\underline{w}$ and $\bar{w}$ are non-negative and fixed by the parameters. When $\delta \geqslant 1/$

2, $\underline{w} = 0$, i.e. there are equilibria in which politicians must implement the median voter's ideal point, $w_1 = 0$ to be

reelected.

Proof. In period 2, as politicians face no election, they maximize their utility by setting $p_2 = q_2$.

Thus, the politician's expected period 2 loss is 0. The median voter's expected loss is $E(k_2^2) = (Ek_2)^2 + \text{Var}(k_2)$.

If the voter has elected a challenger, this is just σ_k^2. For a voter or politician with ideal point l, the expected loss

is $(E(k_2 - l))^2 + \text{Var}(k_2)$, and when a challenger has been elected this is just $l^2 + \sigma_k^2$.

The remainder of the proof has 4 parts. Voter strategy in period 1 can be defined by a *reelection region R* such

that voters reelect iff $w_1 \in R$. In part 1, I show that given voter behaviour as described in the Lemma, incumbent

behaviour will be as described. In part 2, I show that given incumbent behaviour as described, if R is a closed

interval $[w_\star, w^\star]$, then $w^\star$ must be within bounds $[\underline{w}, \bar{w}]$ which are fixed from the parameters (and similarly for

$w_\star$). In part 3, I show that if an equilibrium satisfies the Intuitive Criterion, then R must be a closed interval. In

part 4 I show that if R is a closed interval with $w^\star$ and $w_\star$ within the appropriate bounds, then the equilibrium

satisfies the Intuitive Criterion. From now on I drop period subscripts.

1. Suppose that in period 1, $R = [w_\star, w^\star]$ with $0 \in R$. We then show that incumbent behaviour is as stated in the

Lemma. Proofs are for $k \geqslant 0$; the case for $k \leqslant 0$ is analogous.

If the incumbent's $k \in R$ then he chooses $p = q$, where $q = k + \varepsilon$ is his induced policy ideal point. The outcome is

then k and the incumbent is reelected, thus achieving his best possible outcome.

An incumbent with $k \notin R$ can either choose $p = q$ or $p = p^\star \equiv w^\star + \varepsilon$. Any other choice will be strictly dominated

by one of these: if $p \neq q$ and $p \notin R$ then q is preferable as neither p nor q get the incumbent reelected, and p

uniquely minimizes the period 1 loss; If $p < p^\star$ and $p \in R$ then $p^\star$ is preferable since both get the incumbent

reelected and $p^\star$ is nearer to q than p. If the incumbent chooses $p = q$ we say he or she is "honest"; choosing $p =$

$p^\star$ is called being "disciplined".

With a discount rate of δ the loss from choosing $p = q$ is

$$0 + \delta(k^2 + \sigma_k^2) \tag{1}$$

and the loss from discipline, i.e. choosing $p = p^\star$, is

$$(k - w^\star)^2 + \delta\, 0. \tag{2}$$

Thus, discipline is chosen iff both $k \notin R$ and the loss from discipline is less than the loss from honesty:

$$(k - w^\star)^2 - \delta(k^2 + \sigma_k^2) \;\leqslant\; 0 \tag{3}$$

This expression does not depend on ε. I show that for $k > w^\star$, if a k-type is honest then every k'-type is honest for $k' > k$. Hence, only types within a fixed interval $k \in [w^\star, \bar{k}]$ are disciplined.

First, differentiating the left hand side of (3) with respect to k gives

$$2(k - w^\star) - 2\delta k = 2((1 - \delta)k - w^\star) \tag{4}$$

which is non-positive if and only if $k \leqslant w^\star/(1 - \delta)$. Thus, when $w^\star < k \leqslant w^\star/(1 - \delta)$, the LHS of (1) is decreasing in k. Since (3) is satisfied at $k = w^\star$, it is satisfied on $w^\star < k \leqslant w^\star/(1 - \delta)$.

Taking $k = w^\star/(1 - \delta)$ gives

$$(k - w^\star)^2 \;=\; \left(\frac{w^\star - (1 - \delta)w^\star}{1 - \delta}\right)^2 = \delta^2\left(\frac{w^\star}{1 - \delta}\right)^2 \tag{5}$$

while

$$\delta(k^2 + \sigma_k^2) \;>\; \delta k^2 = \delta\left(\frac{w^\star}{1 - \delta}\right)^2 > \delta^2\left(\frac{w^\star}{1 - \delta}\right)^2. \tag{6}$$

Thus (1) is strictly satisfied at $k = w^\star/(1 - \delta)$. For $k > w^\star/(1 - \delta)$ the left hand side of (3) is increasing and therefore hits zero only once. The crossover point solves (3) with equality. Rewriting this as

$$(1 - \delta)k^2 - 2w^\star k + w^{\star 2} - \delta\sigma_k^2 = 0, \tag{7}$$

the relevant solution, from the quadratic formula and some algebra, is

$$\bar{k}(w^\star) = \frac{w^\star + \sqrt{\delta(w^{\star 2} + (1 - \delta)\sigma_k^2)}}{(1 - \delta)}. \tag{8}$$

Thus, for $k > \bar{k}$ the incumbent prefers to choose her own ideal point q. $\bar{k}$ increases in δ, $w^\star$ and σ_k^2 – thus more patient incumbents are more likely to be disciplined, as are those facing more ideologically diverse challengers. Setting $k^\star = \bar{k}(w^\star)$ gives us the incumbent behaviour specified in the Lemma.

2. Given incumbent behaviour as described in the Lemma, I continue to assume that $R = [w_\star, w^\star]$, and derive bounds on $w^\star$ and $w_\star$. Write $R^\circ = (w_\star, w^\star)$ for the interior of the reelection region. We examine best responses to $w \geqslant w_\star$; the other case is analogous. The median voter's optimal cutpoint $w^\star$ must satisfy three conditions:

$$E(k^2|w \in R^\circ) \;\leqslant\; \sigma_k^2 \tag{9}$$
$$E(k^2|w = w^\star) \;\leqslant\; \sigma_k^2 \tag{10}$$
$$E(k^2|w > w^\star) \;\geqslant\; \sigma_k^2 \tag{11}$$

That is, the voter reelects the incumbent if the expected loss from reelection is less than the expected loss from electing a challenger (conditions 9-10) and throws the incumbent out of the expected loss is greater (condition 11). In all cases, the expected loss is just the variance of the politician's ideal point around 0, and for a challenger this is just the variance of the distribution of politicians, σ_k^2.

The link between w and the expectation of k^2 is provided by the voters' beliefs, a probability distribution over k. Along the equilibrium path, i.e. for $|k| \leqslant w^\star$ and $|k| \geqslant k^\star$, these must be updated by Bayes' rule. Thus, for $w < w^\star$ and $w \geqslant k^\star$, the voter must believe that $k = w$ for sure; $w = w^\star$ lets the voter infer that $k \in [w^\star, \bar{k}(w^\star)]$. $w \in (w^\star, \bar{k}(w^\star))$ is not observed in equilibrium – instead, players with $k \in (w^\star, \bar{k}(w^\star))$ choose $p = p^\star$ and thus $w = w^\star$. Thus the posterior over k after observing such an outcome is unrestricted by standard equilibrium concepts.

Conditions (9,11) are slack when (10) is satisfied with equality. For, if

$$E(k^2|w = w^\star) \equiv E(k^2|k \in [w^\star, \bar{k}(w^\star)]) = \sigma_k^2 \tag{12}$$

then for $w \in (-w^\star, w^\star)$, as $E(k^2|w) = E(k^2|k = w)$, we have that $E(k^2|w) < \sigma_k^2$. Similarly, for $w > \bar{k}(w^\star)$, $E(k^2|w) = E(k^2|k = w) > \sigma_k^2$. These facts follow from well-known properties of truncated distributions.

Suppose (10) is satisfied strictly, $E(k^2|w = w^\star) < \sigma_k^2$. Then (9) is again slack. And as $\bar{k}(w^\star) > w^\star$ (11) could also hold even as $w \searrow \bar{k}$. So there are multiple equilibria. The largest possible one, in terms of R, has $w^\star = \bar{w}$ satisfying (12) with equality; the smallest has $w^\star = \underline{w}$ satisfying (11) with equality on the equilibrium path, i.e. with

$$\bar{k}(\underline{w}) = \sigma_k; \tag{13}$$

then clearly (10) holds and hence (9) also. Note that (12) has a unique solution, since $\bar{k}(w^\star)$ is increasing in $w^\star$, and $E(k^2|k \in [a, b])$ is increasing in a and b when both are positive.

When $\delta \geqslant \frac{1}{2}$, then (13) has no positive solution. For, from (8), $\bar{k}(0) = \sqrt{\delta/(1-\delta)}\,\sigma_k$, and thus $\bar{k}(0) \geqslant \sigma_k$ for $\delta \geqslant \frac{1}{2}$. We then set $\bar{w} = 0$ – thus, $R = \{0\}$ is possible in these cases.

3. Points 1-2 have proved that there is a set of Perfect Bayesian Equilibria with $R = [w_\star, w^\star]$. However, there are other possible PBEs. For example, suppose $R = [w_\star, w^\star] \setminus (-e, e)$ for some small $e > 0$ and some $w^\star$ in the range $[\underline{w}, \bar{w}]$. Then for e small enough all incumbents with $k \in (-e, e)$ will choose $w = \pm e$. Since $w_1 \in (-e, e)$ is thus not observed in equilibrium, the voter's beliefs on that interval are unrestricted and can be set so that $E(k^2) > \sigma_k^2$, justifying the voters in rejecting the incumbent. This is then a PBE. To limit the set of equilibria I use Cho and Kreps' Intuitive Criterion (see e.g. McCarty and Meirowitz 2007).

First, in any equilibrium, $R \neq \varnothing$, i.e. some outcomes result in reelection. For otherwise, every type k would choose $w = k$ and so not reelecting after w would not be sequentially rational for $w^2 < \sigma_k^2$. Also, R must be closed, since otherwise an incumbent with ideal outcome $k \notin R$ but k a limit point of R could achieve a net expected loss approaching zero from above by choosing $w \in R$ arbitrarily close to k, but could not actually achieve a loss of zero and would thus have no equilibrium strategy.

Next, fix R and let $\mathrm{BR}(k)$ be the set of best response outcomes w to implement for a politician with ideal point k. I prove the following.

> **Lemma.** *Say $k' > k$. Then $\mathrm{BR}(k') \geqslant \mathrm{BR}(k)$, where the inequality holds for every element of the best response sets.*

> **Proof.** If $k \in R$ then $\mathrm{BR}(k) = \{k\}$ and $\mathrm{BR}(k') \geqslant k$ for every element, as the best response to k' is always either k' or the closest element of R to k'.

> If $k > \max R$, then $\mathrm{BR}(k) \subset \{\max R, k\}$ and if $\max R \notin \mathrm{BR}(k)$ then $\max R \notin \mathrm{BR}(k')$ since the loss from choosing $\max R$ increases when k increases, while the loss from choosing $w = k$ is constant. Thus the Lemma holds in this case. Similarly when $k < \min R$.

> In the remaining case $k \notin R$ but there are points of R below and above k. As R^C is open, define $(\underline{w}, \bar{w})$ to be the largest open interval in R^C around k, and define $\hat{w} = \frac{\underline{w} + \bar{w}}{2}$, the halfway point of this interval. Thus $\bar{w} \in R$. If $k' \geqslant \bar{w}$ then $\mathrm{BR}(k') \geqslant \bar{w}$ and $\mathrm{BR}(k) \leqslant \bar{w}$, by the same argument as the previous paragraph. If $k' < \bar{w}$ then

$$\mathrm{BR}(k) \;\subset\; \{\underline{w}, k, \bar{w}\} \tag{14}$$
$$\mathrm{BR}(k') \;\subset\; \{\underline{w}, k', \bar{w}\}. \tag{15}$$

> Also, as $k \searrow \underline{w}$, $\mathrm{BR}(k) \to \{\underline{w}\}$ since the loss from $\underline{w}$ goes to zero while the loss from k or $\bar{w}$ remains positive; similarly as $k' \nearrow \bar{w}$, $\mathrm{BR}(k') \to \{\bar{w}\}$. Each pair of the 3 functions $u(\underline{w}, k) = (k - \underline{w})^2$, $u(k, k) = \delta(k^2 + \sigma_k^2)$ and $u(\bar{w}, k) = (\bar{w} - k)^2$ has at most one crossing point in $(\underline{w}, \bar{w})$: the two endpoint utilities cross at the halfway point $\hat{w} = \frac{\underline{w} + \bar{w}}{2}$; for the other two crossing points, the uniqueness can be shown from the quadratic formula.

Clearly if $\mathrm{BR}(\hat{k})$ never includes $\hat{k}$ for $\hat{k} \in (\underline{w}, \bar{w})$, then we are done, for the best response sets are simply $\{\underline{w}\}$ then $\{\bar{w}\}$, with $\{\underline{w}, \bar{w}\}$ at the halfway point. Also, if there is only a single connected interval I on which $k \in \mathrm{BR}(k)$ for $k \in I$, we are done as the best response sets are $\{\underline{w}\}$ then $\{k\}$ then $\{\bar{w}\}$ with one or two crossover points.

I show that if $\hat{k} \in \mathrm{BR}(\hat{k})$ anywhere on $(\underline{w}, \bar{w})$, then the halfway point $\hat{w} \in \mathrm{BR}(\hat{w})$. Take the contrapositive: suppose $u(\hat{w}, \hat{w}) < u(\underline{w}, \hat{w}) = u(w, \hat{w})$. Then, on $\hat{k} \in (\underline{w}, \hat{w})$, $u(\hat{k}, \hat{k}) < u(\underline{w}, \hat{k})$, since $\lim_{k \searrow \underline{w}} u(\underline{w}, \hat{k}) - u(\hat{k}, \hat{k}) > 0$ and there is only one crossing point in the interval; similarly on $\hat{k} \in (\hat{w}, \bar{w})$ $u(\hat{k}, \hat{k}) < u(\bar{w}, \hat{k})$. Thus $k \notin \mathrm{BR}(k)$ for any $k \in (\underline{w}, \bar{w})$.

Now if $\hat{w} \in \mathrm{BR}(\hat{w})$ we must have only a single connected interval I including $\hat{w}$ on which $k \in \mathrm{BR}(k)$ for $k \in I$, and this completes the proof. For otherwise, $u(k, k)$ would have to cross either $u(\underline{w}, k)$ twice to the left of $\hat{w}$ where $u(\underline{w}, k) > u(\bar{w}, k)$, or $u(w, k)$ twice to the right of $\hat{w}$. And this cannot be true because each pair of functions has only one crossing point. $\qquad\square$

Furthermore, except at isolated points of indifference, $\mathrm{BR}(k)$ is a singleton set. Therefore, any possible best response function will be increasing in k. Given this, along the equilibrium path $E(k^2|w)$ is strictly increasing in $|w|$: if $w' > w$ then the lowest k-type playing w' is higher than the highest type playing w.

Next, if $w, w' \in R$ and $w < w'$ then $w'' \in R$ for any $w'' \in (w, w')$. This establishes that R is a single connected interval. The reasoning is as follows. Suppose false. Then it must be that

$$E(k^2|w'') \geqslant E(k^2|w') \text{ and } E(k^2|w'') \geqslant E(k^2|w). \tag{16}$$

This can only be the case if w'' is never chosen in equilibrium. Otherwise, (16) would be contradicted by $E(k^2|w)$'s being strictly increasing in $|w|$ along the equilibrium path, given that w and w' are both chosen in equilibrium (by types $k = w$ and $k = w'$).

If w'' is never chosen in equilibrium, we invoke the Intuitive Criterion to bound $E(k^2|w'')$. The Criterion states that off equilibrium at w'', the support of the voters' beliefs must be composed only of k-types who would improve on their equilibrium utility by playing w'', given some best response to w'' (and some belief). So, this support can at most contain k-types who would improve on their equilibrium utility by playing w'' and getting reelected. (If reelecting after w'' is not a best response to any possible belief, the support is even more limited, as utility from playing w'' and not getting reelected is lower.) Clearly, all types $k \leqslant w$ (resp. $k \geqslant w'$) could do better than playing w'' and getting reelected by playing w (resp. w') and getting reelected; and so they must be doing at least as well as this in equilibrium. Thus the support of voters' beliefs is limited to types between w and w'. But now, repeating the argument of the last paragraph, (16) must be violated.

Thus R is a single connected interval. It remains to show that $0 \in R$. In fact this is trivial. Suppose for example that $w^\star = \max R < 0$. Then types with $k > \bar{k}(w^\star)$, as given in (8), will prefer honesty to discipline. As (8) is increasing in w, it is then clear that such types will violate condition (11).

We have proved that an equilibrium cannot satisfy the Intuitive Criterion unless R is a single interval $[w_\star, w^\star]$ containing 0. Part 1 and 2 show that such intervals define BNE, so long as $w_\star$ and $w^\star$ satisfy certain bounds.

4. We also need to show that if R is a single interval satisfying these bounds, the resulting BNE does indeed fulfil the Intuitive Criterion. To do this, we examine beliefs after an out-of-equilibrium $w \in (w^\star, \bar{k}(w^\star))$. Consider an incumbent with $k = \bar{k}(w^\star)$. This incumbent would prefer to deviate to $w > w^\star$ and get accepted than to implement $w^\star$ and be accepted, and is indifferent between the latter and his equilibrium strategy of playing $w = k$ and being rejected. Thus beliefs after w that put 100% weight on k will satisfy the Criterion. If $\bar{k}(w^\star) \geqslant \sigma_k$, given these beliefs the voters prefer not to reelect. As $\bar{k}(\underline{w}) = \sigma_k$ and $\bar{k}(\,\cdot\,)$ is increasing, this condition is satisfied in every equilibrium. $\qquad\square$

Lemma 2. *In any equilibrium satisfying the Intuitive Criterion, in period 2 under direct democracy, write A for the acceptance region such that $p_2 \in A$ is accepted, and write Φ_k for the voters' probability distribution over the politician's k at the start of period 2.*

1. Politicians with ideal policy $q_2 = k_2 + \varepsilon_2$ choose $p_2 = q_2$ if $q_2 \in A$. Otherwise they choose the closest point in A to q_2. (If q_2 lies closer to 0 than to any other point in A, then any point in A^C may be chosen; we assume they choose 0.)

2. $\mathbb{R}_- \subset A$ when the politician is known by the voters to have $k \geqslant 0$. When $k \leqslant 0$ for sure, $\mathbb{R}_+ \subset A$.

3. If the voters always know the sign of k and have bounds on k (e.g. $\Phi(k) = 0$ for $k < k'$, $\Phi(k) = 1$ for $k > k''$, $0 < k'$), at the start of period 2, then the following is a possible equilibrium: $A = \mathbb{R}_-$ when k is known positive, $A = \mathbb{R}$ when $k = 0$, $A = \mathbb{R}_+$ when k is known negative.

Proof.

1. First, consider the period 2 incumbent's behaviour. The incumbent observes the shock ε and his own ideal point k, and has a resulting policy ideal point $q = k + \varepsilon$. He must propose a policy p. Suppose this policy is accepted iff $p \in A$, the *acceptance region*. Otherwise the default policy of 0 is chosen. We will assume $0 \in A$, since the politician can always get 0 by proposing it. The following describes his best response: if $q \in A$, propose $p = q$ for a loss of 0, the best possible outcome; otherwise, propose the closest point in A to p. (If 0 is the closest point in A to p then we assume he proposes 0 rather than proposing any point not in A.) Write A^C for the complement of A, the *rejection region*.

2. Given this politician behaviour, the voters must figure out whether to accept p or not. If p is in the interior of A, then by single-peakedness of preferences, voters can infer $q = p$. If p is a boundary point of A, they will have a truncated probability distribution over q, updated from their prior. (Their prior belief over q, with cdf Φ_q, is the convolution of Φ_k with the distribution of ε. For example, if they know k for sure, then $\Phi_q(q) = \Phi_\varepsilon(q - k)$.) From our description of the politician's behaviour, it follows that for $p \in A$, the voters' posterior $\Phi_{q|p}(q|p)$ has a support $[\underline{q}(p), \bar{q}(p)]$ and for $p' > p$ with both in A, $\underline{q}(p') \geqslant \bar{q}(p)$, i.e. the voters' belief is increasing in p.

Beliefs after observing $p \notin A$ are off the equilibrium path and unconstrained. Because of this, many equilibria are possible in period 2. For example, the voters could reject all proposals, $A = \varnothing$, and could hold beliefs such that any proposal was worse for them than the default policy. We reduce this proliferation using the Intuitive Criterion. In this case, this reduces to the following: suppose $p \notin A$ so that p is not observed in equilibrium. We require that the voters' belief $\varphi(q|p)$ satisfies $\varphi(q|p) > 0$ only if p would be strictly preferable for the q-type to their equilibrium action, under some rationalizable action by the voters. Call the q-type's equilibrium action $\bar{p} \in A$. Since $0 \in A$ we have that $(\bar{p} - q)^2 \leqslant (\bar{p} - 0)^2$, so if the voters reject p this would not be strictly preferred to $\bar{p}$. Thus we require that the q-type strictly prefers that the voters accept p: $(q - p)^2 < (q - \bar{p})^2$. For instance, if the interval $(p', p'') \subset A^C$ with $\{p', p''\} \subset A$, then the voters' out-of-equilibrium belief for $p \in (p', \frac{p' + p''}{2})$ must have $q > \frac{p + p'}{2}$ whenever $\varphi(q|p) > 0$, since only these types would prefer acceptance at p to their equilibrium payoff of acceptance at p'.

Suppose $p \leqslant 0$, $p \in A^C$ and the voters believe that $k \geqslant 0$ for sure. Then by the Intuitive Criterion $q < p/2$ for sure, since otherwise the q-type would prefer 0 to p. Therefore $\varepsilon = q - k \leqslant q < p/2 \leqslant 0$. Since $\varepsilon < p/2$, $p - \varepsilon > p/2$. Say $p - \varepsilon > 0$. Then $0 < p - \varepsilon < -\varepsilon$, hence $(p - \varepsilon)^2 < \varepsilon^2$ so the voters prefer to accept p, contradicting $p \in A^C$. Say $p - \varepsilon \leqslant 0$. Then $\varepsilon < p/2 < p - \varepsilon \leqslant 0$ so again the voters prefer to accept p, contradicting $p \in A^C$. This shows that $\mathbb{R}_- \subset A$ when k is known to be non-negative. A similar proof shows $\mathbb{R}_+ \subset A$ when k is known to be positive.

3. To show the equilibrium described is possible, I need to show that there are some beliefs satisfying the Intuitive Criterion such that $p > 0$ is always rejected when k is known to be positive and bounded. The Criterion requires only that $q > p/2$ so that the politician proposing p would strictly prefer this to the equilibrium policy of 0. Suppose the voters know $k < k''$. Say $p \leqslant 2k''$. Then $k = k'' = q$ is a possible belief, so that $\varepsilon = 0$ and the voters prefer to reject p.

Say $p > 2k''$. Set the voters' belief to $q = p/2$. Then $\varepsilon = q - k \in [p/2 - k'', p/2 - k']$ and by $p > 2k''$, $\varepsilon > 0$. The median voter's payoff from accepting p is $p - \varepsilon \in [p - p/2 + k', p - p/2 + k''] = [p/2 + k', p/2 + k'']$. As $k' > 0$ we have $0 < \varepsilon \leqslant p/2 - k' < p/2 + k' \leqslant p - \varepsilon$ and hence the voters prefer to reject p.

It has not yet been shown that this equilibrium is incentive-compatible with period 1; this will become clear in Proposition 2. □

Larger equilibria are possible, with $A = \mathbb{R}_- \cup [\underline{p}, \infty)$ for some large enough $\underline{p}$. Focusing on this equilibrium has two advantages. First, it simplifies the calculation of period 1 equilibria. Second, it is the smallest possible equilibrium in terms of A, in other words, it is the unkindest equilibrium for the politician, in which he is most constrained in period 2. If some politicians prefer direct democracy under this equilibrium, the result will hold when A is larger.

Lemma 3. *Under direct democracy, in period 2, if a challenger has been elected, and if $\sigma_k^2 > \sigma_\varepsilon^2$, the voters always reject the challenger's proposal and choose $p_2 = 0$ instead. If $\sigma_k^2 \leqslant \sigma_\varepsilon^2$ then the challenger proposes $p_2 = q_2$ and the voters always accept.*

Proof. If the challenger has been elected after period 1, then k is distributed normally, with mean 0 and variance σ_k^2. The necessary condition for a policy $p \geqslant 0$ to be in A° is that $p \leqslant 2E(\varepsilon|p) = 2E(\varepsilon|q = p)$. The conditional distribution of the shock, $\varepsilon|q = p$, is itself normal with mean $\rho\, q\, \sigma_\varepsilon / \sqrt{\sigma_\varepsilon^2 + \sigma_k^2}$, where ρ is the correlation between q and ε: $\rho = \mathrm{cor}(q, \varepsilon) = \mathrm{cov}(q, \varepsilon)/\sqrt{\sigma_\varepsilon^2 + \sigma_k^2}\, \sigma_\varepsilon = \sigma_\varepsilon^2/\sqrt{\sigma_\varepsilon^2 + \sigma_k^2}\, \sigma_\varepsilon$. (These are well-known facts about normal distributions.) Thus $E(\varepsilon|q) = \lambda q$ where $\lambda = \sigma_\varepsilon^2/(\sigma_\varepsilon^2 + \sigma_k^2)$. Thus $p \leqslant 2E(\varepsilon|q = p)$, for any $p \geqslant 0$, if and only if

$$\sigma_\varepsilon^2 \geqslant \sigma_k^2. \tag{17}$$

In other words, when the challenger is unknown, there is a straight choice between trusting to the challenger's ideal point and accepting the shock. A similar argument shows that $2E(\varepsilon|q = p) \leqslant p < 0$ iff $\sigma_\varepsilon^2 \geqslant \sigma_k^2$.

Thus if $\lambda \geqslant 1/2$, the largest possible acceptance region $A = A^\circ = \mathbb{R}$, i.e. the politician's proposal p_2 is accepted. There are no boundary conditions to check. If $\lambda < 1/2$, $A = \varnothing$ so that the politician's proposal is always rejected. Recalling that $\lambda = \sigma_\varepsilon^2/(\sigma_\varepsilon^2 + \sigma_k^2)$, this completes the proof. $\square$

Proposition 1. *In period 2, any politician weakly prefers representative democracy to direct democracy, and an incumbent reelected from period 1 strictly prefers it.*

Proof. The first statement is trivial: a period 2 politician implements her ideal point under representative democracy, thus getting maximum utility of 0. The second statement follows from $A \neq \mathbb{R}$ and the fact that a reelected incumbent has ideal policy $q = k + \varepsilon$ distributed with support $\mathbb{R}$. Therefore, sometimes the incumbent does not get her ideal point under direct democracy. $\square$

Proposition 2. *Under direct democracy, if $\sigma_\varepsilon^2 < \sigma_k^2$, then all politicians are reelected in period 1. Furthermore, at least some politicians who would be reelected under representative democracy nevertheless strictly prefer direct democracy.*

Proof.

1. We show that $R = \mathbb{R}$ is an equilibrium. Since $\sigma_\varepsilon^2 < \sigma_k^2$, electing a challenger is pointless, as by part 1, his period 2 proposal will always be rejected: the median voter's expected period 2 loss from electing a challenger is just σ_ε^2. On the other hand, reelecting an incumbent cannot give a loss greater than σ_ε^2, since the median voter could choose to reject any proposal that the incumbent makes and hence again get σ_ε^2. In fact, since even an extreme incumbent sometimes makes an informative proposal $p \in \mathbb{R}_-$, the median voter will do better.

Given that all period 1 proposals are accepted, a period 1 k-type's loss is bounded above by the loss he would get from playing $w_1 = k$. (It is not necessarily always true that the k-type does play k. Under direct democracy, there may be a gain from pooling with another type, since this alters the period 2 acceptance region A.) This loss is

$$\delta L_{\text{pol}}(k, A(k)) \leqslant \delta \int_{A(k)} q^2 d\Phi_{q|k}(q|k) = \delta \int_{A(k)} (k+\varepsilon)^2 d\Phi_\varepsilon(\varepsilon) \leqslant \delta \int_{-\infty}^{\infty} (k+\varepsilon)^2 d\Phi_\varepsilon(\varepsilon) = \delta(k^2 + \sigma_\varepsilon^2). \tag{18}$$

That is, if the politician's q-type is within the acceptance region, he proposes and gets his ideal point in period 2, with zero loss; otherwise his loss is at most that from a policy of zero. Since the loss from stepping down is $\delta(k^2 + \sigma_\varepsilon^2)$, no politician steps down after period 1.

Take any equilibrium under representative democracy, and consider the politician with $k = \bar{k} \equiv \bar{k}(w^\star)$ – the most extreme politician who is reelected. Under representative democracy, this type is indifferent between reelection and choosing his own ideal point; thus his loss is $\delta(\bar{k}^2 + \sigma_k^2)$. His loss under direct democracy is $\delta L_{\text{pol}}(\bar{k}, A(\bar{k})) \leqslant \delta(\bar{k}^2 + \sigma_\varepsilon^2)$, as above, and since $\sigma_\varepsilon^2 < \sigma_k^2$ it is strictly less. Continuity of all the utility functions shows that there is an interval of reelected politicians who lose strictly less under direct than under representative democracy.

2. We show that no equilibrium with $R \neq \mathbb{R}$ satisfies the Intuitive Criterion. Suppose $\max R = \bar{w}$. Consider $w_1 > \bar{w}$. The Criterion requires that the resulting belief has a support only of types who might gain from playing w_1 instead of their equilibrium play. This requires that $k > (w_1 + \bar{w})/2$. But given any belief with support on $[(w_1 + \bar{w})/2, \infty)$, the voters know $k > 0$ for sure and hence $A = \mathbb{R}_-$ by Lemma 2 point 2. Therefore, given this belief the voters do strictly better by reelecting the incumbent than by electing a challenger: with an incumbent, when $q \in \mathbb{R}_-$ they improve their utility from the default policy of 0, and when $q \in \mathbb{R}_+$ they choose the default policy; with a challenger they always choose the default policy. Thus, it cannot be rational for the voters to reject after observing w_1. Hence R cannot be bounded above. Similarly, R cannot be bounded below. Finally, a proof similar to that in Lemma 1 point 3 shows that R cannot have "holes". $\square$

Proposition 3. *if voters learn k with certainty in each period, then only politicians with $|k_1| > \sigma_k$, who will not be reelected, strictly prefer direct democracy in period 2; all other politicians strictly prefer representative democracy.*

Proof. Omitted. The logic is obvious from the text in Section 3.2.   □

Bibliography

Acemoglu, Daron, and James A Robinson. 2006. *Economic Origins of Dictatorship and Democracy.* Cambridge University Press.

Acemoglu, Daron. 2003. Why not a Political Coase Theorem? Social Conflict, Commitment, and Politics. Journal of Comparative Economics 31, no. 4 (December): 620-652.

Allswang, John M. 2000. *The Initiative and Referendum in California, 1898-1998.* Stanford University Press.

Bali, Valentina A., and Belinda C. Davis. 2007. One More Piece to Make Us Puzzle: The Initiative Process and Legislators' Reelection Chances. Political Research Quarterly 60(2): 215.

Banks, Jeff S, and John Duggan. 2006. A Social Choice Lemma on Voting Over Lotteries with Applications to a Class of Dynamic Games. Social Choice and Welfare 26, no. 2: 285-304.

Bawn, Kathleen. 1993. The Logic of Institutional Preferences: German Electoral Law as a Social Choice Outcome. American Journal of Political Science 37, no. 4 (November): 965-989.

Benoit, Kenneth. 2004. Models of Electoral System Change. Electoral Studies 23, no. 3 (September): 363-389.

Besley, Timothy J., and Stephen Coate. 2001. Issue Unbundling via Citizens' Initiatives. SSRN eLibrary. http://ssrn.com/paper=275032.

Boehmke, Frederick J. 2005. The Indirect Effect of Direct Democracy: How Institutions Shape Interest Group Systems. Ohio State University Press, Columbus, OH.

Boix, Carles 2003. *Democracy and Redistribution.* Cambridge: Cambridge University Press.

Boix, Carles. 1999. Setting the Rules of the Game: The Choice of Electoral Systems in Advanced Democracies. American Political Science Review 93, no. 3 (September): 609-624.

Bowler, Shaun, and Todd Donovan. 2006. Direct Democracy and Political Parties in America. Party Politics 12, no. 5 (September 1): 649-669.

Bowler, Shaun, Todd Donovan, and Jeffrey A. Karp. 2002. When Might Institutions Change? Elite Support for Direct Democracy in Three Nations. Political Research Quarterly 55, no. 4 (December 1): 731-754.

Bowler, Shaun, Todd Donovan, and Jeffrey A. Karp. 2006. Why Politicians Like Electoral Institutions: Self-Interest, Values, or Ideology? Journal of Politics 68 (May): 434-446.

Broder, David S. 2000. *Democracy Derailed: Initiative Campaigns and the Power of Money.* Harcourt, New York.

Cho, In-Koo, and David M. Kreps. 1987. Signaling Games and Stable Equilibria. The Quarterly Journal of Economics 102, no. 2 (May): 179-221.

Cukierman, Alex, and Mariano Tommasi. 1998. When Does It Take a Nixon to Go to China? American Economic Review 88, no. 1 (March): 180-197.

David, Paul T., and Ralph Eisenberg. 1961. Devaluation of the Urban and Suburban Vote. Bureau of Public Administration, University of Virginia.

Diaz-Cayeros, Alberto, and Beatriz Magaloni. 2001. Party Dominance and the Logic of Electoral Design in Mexico's Transition to Democracy. Journal of Theoretical Politics 13, no. 3 (July 1): 271-293.

Dunleavy, Patrick, and Helen Margetts. 1995. Understanding the Dynamics of Electoral Reform. International Political Science Review 16, no. 1 (January 1): 9-29. .

Figueiredo, Rui J. P. de. 2002. Electoral Competition, Political Uncertainty, and Policy Insulation. American Political Science Review 96, no. 2 (June): 321-333.

Garrett, Elizabeth. 2004. Hybrid Democracy. George Washington Law Review 73: 1096.

Geddes, Barbara. 1991. A Game Theoretic Model of Reform in Latin American Democracies. American Political Science Review 85, no. 2 (June): 371-392.

Gerber, Elisabeth R. 1996. Legislative Response to the Threat of Popular Initiatives. American Journal of Political Science 40 (1):99-128.

Gerber, Elisabeth R., Arthur Lupia, and Matthew McCubbins. 2001. *Stealing the Initiative: How State Government Responds to Direct Democracy.* Upper Saddle River, NJ: Prentice Hall.

Goebel, Thomas. 2002. *A Government By the People: Direct Democracy in America, 1890-1940.* Chapel Hill: UNC Press.

Greif, Avner, and David D. Laitin. 2004. A Theory of Endogenous Institutional Change. American Political Science Review 98, no. 4: 633-652.

Hamm, Keith E., and Gary F. Moncrief. 1999. Legislative Politics in the States. In Virginia Gray and Russell L. Hanson, eds., *Politics in the American States: A Comparative Analysis* 7th edition. CQ Press.

Haynes, George H. 1919. How Massachusetts Adopted the Initiative and Referendum. Political Science Quarterly 34, no. 3 (September): 454-475.

Kelleher, Christine A., and Jennifer Wolak. 2007. Explaining Public Confidence in the Branches of State Government. Political Research Quarterly 60, no. 4 (December 1): 707-721.

Kessler, Anke S. 2005. Representative versus Direct Democracy: The Role of Informational Asymmetries. Public Choice 122, no. 1/2 (January): 9-38.

Maskin, Eric, and Jean Tirole. 2004. The Politician and the Judge: Accountability in Government. American Economic Review 94, no. 4: 1034-1054.

Matsusaka, John G., and Nolan M. McCarty. 2001. Political Resource Allocation: Benefits and Costs of Voter Initiatives. Journal of Law, Economics, and Organization 17, no. I2.

Matsusaka, John G. 2004. *For the Many Or the Few: The Initiative, Public Policy, and American Democracy.* University of Chicago Press.

McCarty, Nolan M., and Adam Meirowitz. 2007. *Political Game Theory: An Introduction.* 1st ed. Cambridge University Press.

McCubbins, Matthew D., Roger G. Noll, and Barry R. Weingast. 1987. Administrative Procedures as Instruments of Political Control. Journal of Law and Economic Organization 3, no. 2 (January 1): 243-277.

Milesi-Ferretti, Gian Maria. 1995. The Disadvantage of Tying Their Hands: On the Political Economy of Policy Commitments. Economic Journal 105, no. 433 (November): 1381-1402.

North, Douglass C., and Barry R. Weingast. 1989. Constitutions and Commitment: The Evolution of Institutional Governing Public Choice in Seventeenth-Century England. Journal of Economic History 49, no. 4 (December): 803-832.

O'Donnell, Guillermo, Philippe C. Schmitter, and Laurence Whitehead. 1986. *Transitions from Authoritarian Rule: Tentative Conclusions about Uncertain Democracies.* Johns Hopkins University Press.

Piott, Steven L. 2003. *Giving Voters a Voice: The Origins of the Initiative and Referendum in America.* University of Missouri Press.

Przeworski, Adam. 1991. D*emocracy and the Market: Political and Economic Reforms in Eastern Europe.* Cambridge University Press.

Riker, William H. 1988. *Liberalism Against Populism: A Confrontation Between the Theory of Democracy and the Theory of Social Choice.* Waveland Press.

Rogowski, Ronald. 1987. Trade and the Variety of Democratic Institutions. International Organization 41, no. 2 (Spring): 203-223.

Rokkan, Stein. 1970. Citizens, Elections, Parties: Approaches to the Comparative Study of the Processes of Development. Imprint unknown.

Romer, Thomas, and Howard Rosenthal. 1978. Political resource allocation, controlled agendas, and the status quo. Public Choice 33, no. 4 (Winter 1978): 27-43.

Scarrow, Susan E. 1997. Party Competition and Institutional Change: The Expansion of Direct Democracy in Germany. Party Politics 3, no. 4 (October 1): 451-472.

Schrag, Peter. 2004. *Paradise Lost: California's Experience, America's Future.* Berkeley: University of California Press.

Smith, Daniel A. and Dustin Fridkin. 2008. Delegating Direct Democracy: Interparty Legislative Competition and the Adoption of the Initiative in the American States. American Political Science Review 102, no. 3: 333-350.

Smith, Daniel A., and Joseph Lubinski. 2002. Direct Democracy During the Progressive Era: A Crack in the Populist Veneer. Journal of Policy History 14, no. 4: 349-83.

Smith, Daniel A., and Caroline J. Tolbert. 2001. The Initiative to Party: Partisanship and Ballot Initiatives in California. Party Politics 7, no. 6 (November 1): 739-757.

Tables and Figures

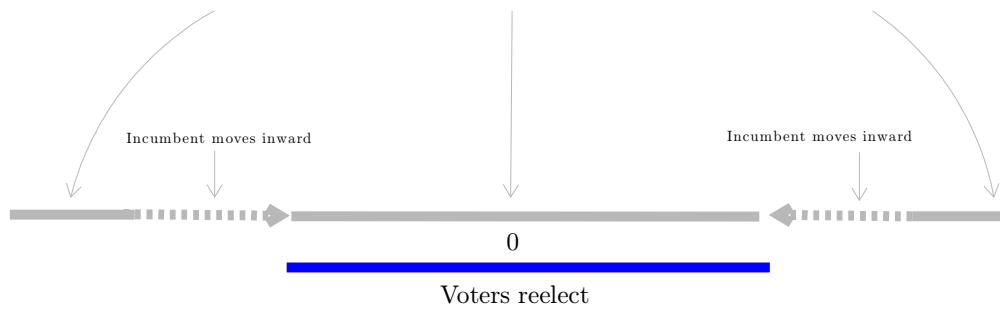| |
|---|
| Incomplete voter information |
| Imperfect electoral discipline |
| Direct and representative elections interact |
| Direct elections are *ex ante*, representative elections *ex post* |

Table 1. Key assumptions of the model

Figure 1. Form of equilibria in period 1.

$\sigma_\varepsilon^2 > \sigma_k^2$: all policies accepted

$\sigma_\varepsilon^2 < \sigma_k^2$: all policies rejected in favour of default

Reelected incumbent

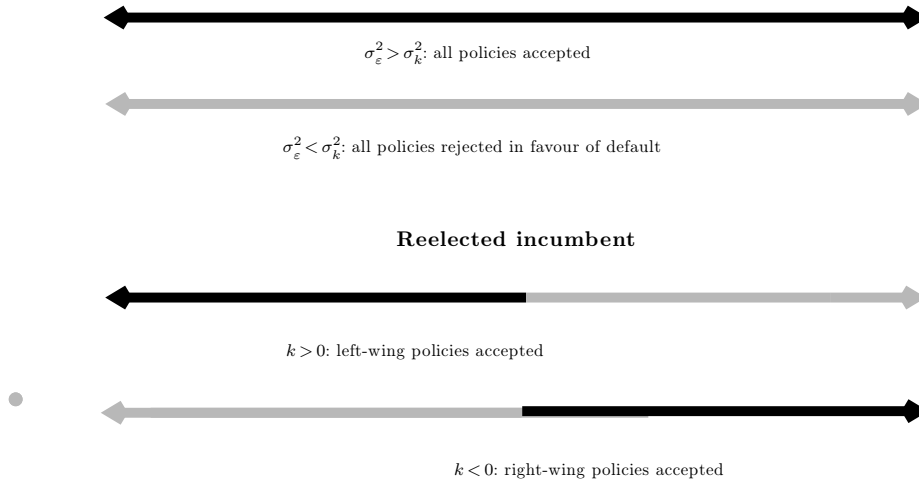$k > 0$: left-wing policies accepted

$k < 0$: right-wing policies accepted

Figure 2. Period 2 equilibrium under direct democracy