Inducing Leaders to Take Risky Decisions: Dismissal, Tenure, and Term Limits

Philippe Aghion and Matthew O. Jackson *

October 2014

Abstract

In this paper we analyze the problem of how a principal can induce an agent or 'leader' to make the right decision when the leader's competence rather than effort is being evaluated. If the value of foregone projects are not observed, and the leader's competency is only indirectly inferrable through the success or failure of projects that she undertakes, then the leader's incentives to act and make the right decision depend on the principal's dynamic dismissal (or replacement) strategy. If the principal can commit to a replacement strategy in advance, then (approximately) optimal mechanisms either involve a probationary period and then indefinite tenure, or else a random dismissal strategy. If instead commitment is impossible, and for instance voters (the principals) regularly choose whether to replace the leader, then the leader's incentives are poor and the principal's equilibrium payoffs are inefficiently low, even below that of simply replacing the leader in every period. However, incentives can be somewhat improved in this no-commitment case via term limits.

Keywords: Information, tenure, dismissal, replacement, elections, term limits, leadership, discretion

JEL Classification Codes: D72, D82, D86, M12, C72

---

*Aghion is at the Department of Economics, Harvard University, Cambridge, Massachusetts USA, and is a member of CIFAR. Jackson is at the Department of Economics, Stanford University, Stanford, California 94305-6072 USA, and is an external faculty member at the Santa Fe Institute and a member of CIFAR. Emails: paghion@fas.harvard.edu and jacksonm@stanford.edu. We gratefully acknowledge financial support from ARO MURI award No. W911NF-12-1-0509. We thank Daron Acemoglu, Tim Besley, Kirill Borusyak, Renee Bowen, Andres Drenik, Bengt Holmstrom, Nadezhda Kotova, Stephen Nei, Alessandro Pavan, Torsten Persson, Larry Samuelson, Ken Schotts, Bruno Strulovici, Guido Tabellini, Jean Tirole, Lin Zhang, Yulia Zhestkova, and seminar participants at CIFAR, MIT, Northwestern, Stanford, INSEAD and the IIES at Stockholm University, for helpful comments and suggestions.

If General McClellan isn't going to use his army, I'd like to borrow it for a time. *Abraham Lincoln, Jan 10, 1862, before relieving George B. McClellan of command (for the first time).*

1 Introduction

How can one incentivize agents who have discretionary power to make the right decision in situations where: (i) decisions can reveal the competence of the agent; (ii) effort provision is not an issue, and (iii) monetary incentives are fixed and only dismissal can be used for incentives. We refer to such agents as 'leaders' because they are self-motivated by the time they enjoy discretionary power. Thus, this is really a model of agents such as generals, politicians, academics, and other people in many senior roles. When dealing with a general of an army, a high level bureaucrat or administrator, a politician, or an academic researcher, what matters is that they make the right decisions, not so much the worry that would not put enough effort: a bad decision at such high level of discretion can have lasting consequences which is rarely the result of low effort.

Indeed, Abraham Lincoln's main concern with General McClellan was not that McClellan was not putting enough effort, but rather that McClellan was missing opportunities to act and to act efficiently, meanwhile the Confederate Army was dangerously close to Washington. Eventually Lincoln decided to dismiss McClellan in order to prompt action. What Lincoln was seeking from McClellan was not more effort - generals work long hours, and in fact, McClellan had a well-organized army - but rather that he act when appropriate and pick the right battles. Monetary incentives were basically fixed by a military pay scale, McClellan was not being lazy. It was really indecisiveness and a fear of failure that led to his inaction and eventual replacement.[1]

Similarly, the primary concern in motivating a high-ranking politician is not getting them to work long hours, as they tend to already be driven to do so, but instead to motivate them to make the 'right' decisions. The main lever to induce politicians to act is the threat of voting them out, as generally there is little discretion over monetary incentives. The interesting incentive issue, however, is that dismissal threats may induce politicians to 'play it safe' and avoid decisions that might reveal (or be mistaken for) a lack of competence. The question then is: how can dismissal threats be designed to induce politicians, generals, high-level bureaucrats, and administrators, and others in similar situations to act optimally?

Like generals or high-ranking politicians, academic researchers are also self-motivated: that is, the main issue with them is not so much effort provision (academics typically enjoy doing research), but rather their ability to pick the right research questions and to significantly push the knowledge frontier when addressing them. In contrast to the previous cases (generals, politicians, and bureaucrats), the many universities that employ academics can

[1] Some background on the interesting saga of finding a competent general who was willing to lead can be found in McPherson (1988).

commit ex ante to a (contingent) dismissal/no-dismissal policy.[2]

Our model is one where a principal can hire a new agent to make decisions - henceforth the agent is called the 'leader' - in any period at some fixed cost. The leader can be either competent or incompetent, and whether or not she is competent is initially unknown to the principal (and it does not matter whether it is known to the leader in our model). Moreover, the leader has discretion over a choice of actions. In each period, she can choose to take a 'conservative' action which yields a sure payoff (normalized to zero) but does not reveal anything about her level of competence. We can think of this as 'inaction.' Or, she can take a 'risky' action, which could lead to a positive or negative payoff. In each period the leader receives a signal about which is the better action in that period. An incompetent leader receives uninformative signals about whether the risky action is likely to lead to positive or negative payoffs, whereas a competent leader receives informative signals about whether the risky action is more likely to lead to positive or negative payoffs. Signals are privately observed by the leader- and therefore the state is only indirectly observable through the leader's actions.[3]

Each day there is a new state of nature, new information, and a new choice of action to be made, thus over time a leader's competence can be learned by tracking the payoffs on the days on which she took the risky action. Based on that information, the principal decides each period whether to keep or replace the current leader. If the leader takes the conservative action, nothing is learned, while if the leader takes the risky action some information is revealed. In this setting, the dismissal threat can lead to inaction, as the leader is reluctant to choose the risky action and thereby risk increasing the belief of incompetence and risking future dismissal, whereas inaction generally (in equilibrium) leaves beliefs unchanged .

We focus attention on the many important settings in which monetary incentives are not

[2]For some academics there are also monetary incentives dependent upon their success. Although this is true of some fields (e.g., economics), many fields (e.g., many humanities) operate on a fairly rigid pay scale, especially in public universities and in countries outside of the United States.

[3]In the Appendix we also consider the benchmark case where the signal and the state realization are observed by the principal and the agent. So even if the conservative decision is taken, the principal sees what would have happened if the risky decision had been taken. That benchmark case delivers three main predictions. First, the probability that the principal replaces the current leader follows either an increasing, or a decreasing, or bell-shaped and saw-toothed pattern. The intuition for the bell-shape is that early on it is worth keeping the current leader on the job even if she takes the risky action unsuccessfully, as there is an option value of acquiring more information about her level of competence. Then, as time goes on and the number of failed attempts increases, so does the probability of dismissing the current leader as incompetent leaders are identified. Eventually, leaders who have survived longer are increasingly more likely to be competent, and therefore the dismissal probability eventually decreases over time. The bell-shaped pattern is more likely to be observed when actions are not too revealing, so that learning takes at least a few periods. The bell-shaped pattern is in line with evidence on the dynamics of job separation, for example as shown by Farber (1999) for the US labor market, and or more recently by Jia et al (2014) when looking at the patterns of dismissals and promotions of provincial leaders in China. Finally in this benchmark case, we also get the surprising finding that the cumulative probabilities of replacing incompetent leaders, as well as mistakenly replacing competent leaders, are non-monotonic in the leader's actual degree of competence.

contingent upon performance and therefore the leader only draws marginal benefits from being kept on the job (so that there is a discrete benefit relative to the outside option). In this model, firing/replacement becomes the main incentive instrument, but it involves subtle trade-offs. On the one hand, the principal really needs to replace incompetent leaders. On the other hand, this threat of replacing based on observed competence may induce leaders to avoid risky decisions that reveal competence. The question is then to design the sequence of replacement decisions over time in a way that incentivizes the leader to take appropriate decisions, but still results in learning and the replacement of incompetent leaders.

We first consider the situation where principals can commit in advance to a dismissal/no dismissal dynamic strategy. The university-academics relationship is an illustrative example of this case, as are other bureaucratic positions and some other appointed (non-elected) administrative positions. We argue that in this case there are essentially two main mechanisms that emerge as optimal ways to motivate the agent, namely the 'carrot' and the 'stick'. The 'carrot' consists in granting tenure to the agent after a finite number of periods (sometimes immediately). The 'stick' consists in replacing the leader if there are enough failures, but *also* sometimes when she chooses the safe action in order to incentivize risky choices. The cost of choosing the carrot, i.e. the tenure strategy, is that the principal might get stuck forever with an incompetent leader. The cost of choosing the stick, i.e., dismissing sometimes after safe decisions, is that firing a competent leader for taking the safe action, even when it was the right thing to do. We show that the stick strategy dominates when the replacement cost is sufficiently low, and otherwise the carrot strategy dominates. More generally, we show if the principal is sufficiently patient, then it is always (approximately) optimal to use a tenure mechanism.

Before describing our results on no-commitment, it is interesting to contrast our model with the classical incentive models where effort is a main issue. Despite the simplicity and relevance of our model, it has surprisingly not been analyzed before, and is not a special case of previous moral hazard models, and it leads to very different insights. In standard models where effort is a variable along with a competence dimension, what generally happens is that competent agents work harder to signal their type and to generate good impressions - at least initially depending on the timing. Thus, tenure is not needed to sort out competent from the incompetent leaders, and in fact in standard models tenure actually discourages effort and worsens the outcomes once agents get tenure.

We then turn to the situation in which the principal cannot commit in advance to keeping or replacing the leader in future periods. A leading example of such a situation is that of an elective democracy, where voters (the principal) choose whether or not to replace the leader after each period, without any precommitment. We abstract from voter bias issues to concentrate on voters' decisions of whether to replace the leader as a function of the information voters receive over time about the leader's level of competence. One of our main results in that case is that every (Markov Perfect) equilibrium results a payoff that is strictly lower than the (highly) inefficient benchmark of simply replacing the leader in every period. The reason is that an incumbent leader's best response in the no-commitment case involves taking the safe action when they are faced with reelection probabilities: indeed, doing so maintains voters' belief about the leader's competence at its initial level; this, together with the fact that replacing the current leader by a new leader is costly but leads to the same probability of competence, implies that the current leader will not be fired. On the other hand, if current leaders are sure to be replaced then they are willing take risky decisions, which leads to a net positive payoff for the principal. So, even though constantly replacing the leader is highly costly and inefficient, it provides incentives to act that cannot be achieved via (Markovian) replacement decisions. We also prove that something similar is true of a fairly general class of non-Markov equilibria: *no matter how successful the leader has been in the past the leader eventually stops taking risky actions and is replaced.* Thus, voting settings can provide particularly poor incentives. We then show that an optimal term limit can substantially improve the leader's incentives to take the risky action. The optimal term limit in turn depends on the speed of learning and on the cost of replacement.

Again these results stand in contrast to settings in which effort is a major issue: there, setting a term limit would be suboptimal as it would discourage effort by the agent in her later periods. So, again the incentive results are quite contrasting, depending on whether risk-taking rather than effort is the substantial choice variable for the agent.

The paper relates to several strands of literature. Most obviously, there is the literature on incentives and moral hazard that we have already alluded to. Here, the important reference point for us is Holmstrom (1982, 1999)'s model of risk-taking incentives and career concerns.[4] In that model a manager proposes a project to the firm. The project may succeed or fail, and the probability of success is higher if the manager is talented than if she is untalented. The manager draws utility from career concerns in the sense that her (future) wage is proportional to the market's (or firm's) posterior probability that she is talented. In that model, a risk-averse manager refrains from proposing risky projects to the firm in order to avoid facing a risky (future) wage profile. Our set-up also has risk-taking as a major incentive issue in organizations and perception about talents as the (sole) source of risk-divergence between principals and agents. But to this we add the principal who can decide whether or not to replace the agent at a positive cost. This in turn allows us to analyze the role of dynamic dismissal mechanisms, including tenure and term limits, in providing risk-taking incentives to the agent.

The paper also contributes to a political economy literature on term limits and the role of elections. In particular, Banks and Sundaram (1990) look at how repeated elections can help select competent politicians. Besley and Case (1995)[5] build on Holmstrom (1982)'s *first* career concerns model with effort (developed in the first part of his paper) to argue and verify empirically that allowing for reelection improves political leaders' incentives.[6] In

[4]See also Holmstrom and Ricart i Costa (1986).

[5]See also Alt, Bueno de Mesquita and Rose (2011).

[6]For other related signalling-based career concerns models, see Dewatripont, Jewitt, and Tirole (1999 a and b) and Alesina and Tabellini (2007, 2008). More generally, a literature on reputations (Scharfstein and Stein (1990), Allen and Gorton (1993), Tirole (1996), Tadelis (1999), Taylor (2000), Mailath and Samuelson

their model agents have costs of effort that are decreasing in their type (competency). The equilibria that they focus upon (there can be many) are such that more competent agents have incentives to work harder given the greater marginal payoff to their effort. In that context, offering longer term limits (more chances for reelection) increases incentives to put in effort - especially in early periods.[7] Fearon (1999) analyzes a two-period model where election occurring at the end of the first period can play the double role of sanctioning bad policies and of selecting more competent politicians. Canes-Wrone et al (2001) also analyze a two-period model with election taking place at the end of the first period. In their model, executives have private information about both the state of nature (and therefore the effect of policies on outcomes) and their level of expertise. Moreover, executives care both about efficient decision making and reelection. A main insight from that model is that executives may choose the policy which (less informed) voters consider to be the best initially if this maximizes the chances of a positive updating of voters' beliefs about their ability. In other words, the fear of being found out as being incompetent may lead incumbent leaders to sometimes 'pander' to their current voters' beliefs.[8]

We emphasize that our model provides very different conclusions than this literature. In particular, in our analysis which does not have effort as a separation device, but does have replacement costs, we find that very regular retention decisions (without commitment), e.g., elections, actually worsen incentives as they induce excessively safe behavior. This contrasts with what happens when effort is available, or there is no safe option, or replacement is costless. In those settings the threat of dismissal can help incentives. So, a main implication of our work, in light of the previous literature, is that the incentive properties of regular evaluations is context dependent. Although regular evaluations can help in eliciting effort, or inducing certain actions when replacement is costless, they can backfire in settings where replacement is costly and where one is trying to induce correct decision-making. Repeated elections/retention decisions biases leaders towards sitting on their laurels and taking safe actions: much as the fear of failure that might have led to McClellan's inaction and Lincoln's frustration in finding a general who would act in the early parts of the U.S. Civil War.

Moreover, we show that tenure contracts or term limits can alleviate fear and provide leaders with incentives to make choices that they would not make under the repeated mi-

(2001,2006)) couples actions and some hidden type with outcomes.

[7]More recently, Smart and Sturm (2013) developed a model where incumbent politicians may be either 'public-spirited' (i.e. with payoffs that coincide with voters' payoff) or 'biased' towards a particular choice of action. By reducing her expected payoff from reelection, term limits reduces a 'public-spirited' politician's incentives to otherwise deviate from efficient decision making in order to signal her type so as to increase her probability of reelection. Like in our model, seeking reelection induces inefficient decision making by current leaders, and term limits help overcome this problem. However, once again, our analysis does not rely on any standard kind of effort problem (be it moral hazard or signaling) and moreover in our model politicians are not a priori biased against voters' preferences.

[8]See also Rodriguez-Barraquer and Tan (2014), who examine students' choices of fields to work on, which then can reveal their abilities. Their model is quite different from ours and focuses on herding effects in competitive settings, and not on dynamic incentives.

croscope of retention. Again, when contrasted with the previous literature, our results show that the effects of such mechanisms is context dependent, as they help in incentivizing appropriate decision making, while they generally fail to provide appropriate effort incentives in later periods.

Our paper also relates to the literature on tenure and 'up-or-out' contracts (e.g. see Kahn and Huberman (1988), Carmichael (1988), Waldman (1990), or Burdett and Coles (2003)). A main argument in this literature is that tenure promotion serves as a commitment device for the principal not to underreport the agent's value ex post (if the agent's value was truly low then the up-or-out contract allows the principal to simply fire the worker), which in turn preserves the worker's ex ante incentives to provide effort. In Carmichael (1988), granting academic tenure to current faculty helps ensure that good potential candidates to become new faculty are not dismissed by current faculty because the former would represent a threat for the latter. Thus, the tenure literature focuses on incentives of principals to evaluate agents, while our analysis is driven entirely by the agents' incentives, which provides new insights. Indeed, we provide a completely different perspective on tenure based on three important features that such contracts provide in a dynamic model with risk-taking as the main incentive issue, namely: incentives to take risky actions during the probationary period, ability to sort competent from incompetent, and incentives to take appropriate actions during the tenure phase.

The remaining part of the paper is organized as follows. In Section 2 we lay out the model. In Section 3 we analyze the case where the principal can commit in advance on a (contingent) dynamic dismissal/retention strategy and we discuss the role of tenure mechanisms. In Section 4 we analyze the case where the principal cannot commit in advance on a dismissal/no dismissal strategy and discusses the role of term limits. In Section 5 we conclude by mentioning various directions in which the model and analysis can be extended. Finally, the Appendix contains the proofs and also the analysis of the benchmark case where states and signals are both public information.

2 The model

2.1 The players

An organization is operated by an 'agent', whom we often refer to as a 'leader', since the person is an important decision maker. A 'principal' decides on keeping or replacing the leader.

We focus on settings in which the leader is paid a fixed wage and so the only relevant decision is whether to keep or replace the leader.

A leader is either 'competent' or 'incompetent', denoted by *Comp* and *Incomp*. The prior probability that the leader is competent is $\lambda_0 \in (0, 1)$.

A given leader's type does not change over time. If a leader is replaced, then the new leader is competent with probability λ_0. This is also the prior about the initial leader's competence with which the principal begins at the start of period 1: in other words, there is no asymmetric information ex ante about the leader's level of competence.

2.2 Time, states, and signals

Time proceeds in discrete periods $t \in \{1, 2, ..., T\}$. On occasion, we consider the case in which $T = 2$ or $T = 3$ for illustrative purposes, but then look at the infinite period case $T = \infty$ for general results.

In each period a state of nature, $\omega_t \in \{X, Y\}$, is realized. States X and Y occur with equal probability, independently across periods.

In the beginning of each period t, the leader sees a signal $s_t \in \{X, Y\}$ that may be informative about the state of nature. If the leader is *competent* then the probability that the signal is equal to the state is $p > 1/2$, and thus the signal is informative. If the leader is *incompetent* then the two signals $s_t = X$ and $s_t = Y$ are equally likely in both states X and Y, and so an incompetent leader's signals are completely uninformative.

Thus, for each ω_t,

$$
\begin{aligned}
\Pr(s_t &= \omega_t | \omega_t, Comp) = p > 1/2 \\
\Pr(s_t &= \omega_t | \omega_t, Incomp) = 1/2.
\end{aligned}
$$

2.3 Actions and replacement costs

We can think of the state being whether there is some sort of 'opportunity' (e.g., profitable investment, or beneficial action) available: in state Y there is such an opportunity and in state X there is not. If the state is Y then it is best to 'invest' or 'act': choosing action y then pays 1 to the principal; in contrast, in state X it is better not to act or not to invest as in that state choosing y results in a loss of value, $-v$, to the principal. Not investing or not acting (choosing action x) always leads to a payoff of 0 to the principal. Thus, we can think of y as 'taking action', while x can be interpreted as not acting or sticking with a status quo. Foregoing actions is a safe alternative that might provide no information of the value of the foregone opportunities.

In summary, the principal's payoff from the action as a function of the state in any period is:

	X	Y
x	0	0
y	$-v$	1

Given that $v > 1$, if there is no information about the state, then action x offers a higher expected payoff than action y, and so it is only competent leaders whom the principal would ever want to take the action y.

Throughout what follows, we focus on the nontrivial case in which $p - v(1-p) > 0$. If this were violated, then even a competent leader could never lead to a positive expected payoff from action y, and so action x is the only action that should ever be taken.

2.4 The order of moves

The sequence of moves within each period is as follows. At the beginning of the period the current leader sees her signal. Next, the leader takes an action, x or y. Subsequently, the principal sees the payoff and updates his beliefs about the leader's competence. At the end of the period the principal decides whether to keep the current leader or replace her with a new leader.

The system repeats itself with a new draw of signal and state in each period, but the type of any given leader remains fixed over time. Once a leader is replaced that leader never returns to the game. Replacing the leader leads to a cost of $c \geq 0$ in that period for the principal.

2.5 Payoffs

All players are expected payoff maximizers and discount time with the same discount factor, δ, such that $0 \leq \delta \leq 1$. Payoffs are as follows.

A leader gets private benefits b per period that she is on the job.[9] This is the benefit relative to her outside option. Note that this admits situations where the value of the outside option might change with beliefs about competence, as long as the marginal benefit from the current job is b-better than the outside option for any beliefs.

The principal gets the per-period payoffs from the matrix above as a function of the action taken and the state, less any costs of replacing leaders.

2.6 Histories

If the leader chooses x then the principal's payoff is 0 and the principal does not learn the true state nor anything about the leader's signal.[10] If the leader chooses y, then the principal's payoff is either 1 or $-v$ and therefore he can infer the state from observing his payoff. Neither the leader nor the principal sees the state except via inference from the realized outcome of the action.[11]

So, histories differ for the leader and the principal, as they observe different things.

[9] This isolates the incentive problem. We could also allow the leader to prefer to make successful decisions without changing the main content of the results, as long as the payoff from being in office was large enough.

[10] The principal could hypothetically make inferences about competence based on the frequency of x actions as a function of equilibrium behavior. However, as we will see, inactive behavior will not reveal information about competence in equilibrium.

[11] The leader can make inferences about the state via the signal, but learns nothing more than that about the state without taking action y.

A t-period history for the principal is a sequence

$$h_P^t = (u_1, d_1 \ ; \ \dots \ ; \ u_t, d_t),$$

where $u_t \in \{0, -v, 1\}$ is the payoff and $d_t \in \{K, R\}$ indicates whether the leader was Kept or Replaced at the end of the period by the principal.

Here the history for the leader also includes the signals she observed, but only since that leader was in office. So, for a leader that was in place since time τ, a history is then

$$h_L^t = (u_1, d_1; \dots, u_{\tau-1}, R \ ; \ s_\tau, u_\tau, d_\tau \ ; \dots \ ; \ s_t, u_t, d_t).$$

So, the latest leader has the same information as the principal about past leaders and then additional information about her own signals during her own reign from time τ through t. [12]

So, a history in this second scenario is a pair $h^t = (h_P^t, h_L^t)$ of related histories, and again we let H^t denote the set of possible histories. The state, signal, and payoff are observed by both parties at the end of each period regardless of which action was taken in that period.[13]

Let H be the set of all finite histories, and H_L and H_P be the sets of all finite histories for the leader and principal, respectively.

2.7 Priors and beliefs

The principal always starts with a prior of $\lambda_0 \in (0, 1)$ whenever there is a new leader.

The principal's belief at *the end of period* t is denoted λ_{tP} and is obtained via Bayes' Rule from the history h_P^t and beginning with a prior of λ_0 at the beginning of any period that begins with a new leader.

The leader may have a *different* prior from the principal, and we denote it by $\lambda_{\tau L}$, where τ is the first period the leader is employed. We only assume that $\lambda_{\tau L} > 0$. So, we rule out the degenerate case in which the leader begins being absolutely sure that she is incompetent. In that case the leader knows that signals are completely irrelevant and so the optimal strategies become numerous and uninteresting. So, we focus on the case in which she holds some positive probability of being competent. When we examine the leader's beliefs λ_{tL}, it is defined *at the beginning of period* t.[14]

[12]Note that u_t fully reveals the action a_t, so we do not need to add that to the histories.

[13]In the Appendix we analyze the case in which the state, signal, and payoff are observed by both parties at the end of each period regardless of which action was taken in that period.

In that case, a t-period history is a sequence

$$h^t = (\omega_1, s_1, a_1, d_1 \ ; \dots \ ; \ \omega_t, s_t, a_t, d_t),$$

where ω_t is the state, s_t is the leader's signal, $a_t \in \{x, y\}$ is the action that the leader took, and $d_t \in \{K, R\}$ indicates whether the leader was Kept or Replaced at the end of the period by the principal.

[14]The reason for defining the principal's belief at the end of period t and the leader's beliefs at the beginning of the period, is that the leader takes an action at the beginning of the period, while the principal takes an action at the end of the period, and this saves on having to define two different beliefs for each period, conserving on notation.

Each leader may have a different prior and it could be correlated with her competence. As it turns out, the leader's prior will not actually matter at all in equilibrium, and so how it is correlated with competence turns out to be completely irrelevant.

2.8 Strategies

A strategy for the principal is a function $\sigma_P : H_P \to \Delta\{K, R\}$, which is a function of the possible histories of the principal.

Strategies are similarly defined for the leader, $\sigma_L : [0,1] \times H_L \to \Delta\{x, y\}$, but can also depend on the prior of the leader (at the time of her hiring[15], which we have allowed to differ across leaders.

2.9 Equilibria

Unless otherwise stated, we consider perfect Bayesian equilibria of the games.

Let $V_\sigma(h_P^t, \lambda_{tP})$ denote the value function to the principal in the continuation of a perfect Bayesian equilibrium σ if the leader is kept after a history h_P^t and a corresponding posterior λ_{tP}.

2.10 Some useful preliminaries

2.10.1 Posteriors and updating

A useful expression is the posterior that a leader is competent if the leader is known to have had m correct signals out of n total signals given a prior of λ, denoted $\Lambda(m, n, \lambda)$. This is:

$$\Lambda(m, n, \lambda) = \frac{\lambda p^m (1-p)^{n-m}}{\lambda p^m (1-p)^{n-m} + (1-\lambda)/2^n} \tag{1}$$

Note that $\Lambda(m, n, \lambda)$ is increasing in m and λ, and decreasing in n.

A relevant expression is the probability that the leader's signal is correct when the leader sees a Y (or equivalently an X), conditional on a current posterior probability λ :

$$f(\lambda) = \lambda p + (1-\lambda)/2. \tag{2}$$

2.10.2 The principal's expected payoff

The following lemma is direct but useful.

LEMMA **1** *The principal's expected payoff when a leader takes an action that matches the signal and the leader is believed to be competent with probability λ is:*

$$u(\lambda) = \frac{1}{2}[f(\lambda) - (1 - f(\lambda))v].$$

[15]Note that the posterior can then always be deduced from the prior and the history.

The lemma follows from noting that with probability 1/2 the signal is Y and that playing y when the signal is Y yields expected payoff $[f(\lambda)-(1-f(\lambda))v]$ to the principal.[16]

3 Full commitment and tenure mechanisms

We examine two cases in order. In the first, the principal can commit to specific evaluation times and decisions conditional on histories. Effectively, this becomes a mechanism design problem. In the second case, the principal cannot commit to specific strategies, but instead can replace the leader at any time. The two different scenarios have different applications and their contrast provides some of our central insights.

To induce the leader to take the risk of choosing y when the true state of nature is Y, there are a variety of mechanisms that can be used. In particular, the principal can randomly dismiss the leader if she chooses action x : this we refer to as the stick. Or the principal can provide incentives by simply guaranteeing to keep the leader, we refer to as the carrot.

The commitment case, is one in which the principal chooses a strategy σ_P at the beginning of the game and it is fixed, it is observed by all prospective leaders. This mechanism design problem has some obvious relation to Stackelberg problems.

So, here we need only consider what the optimal responses of a leader are to any strategy of the principal, and then taking those as given, optimize for the principal. Thus, we are using a form of Stackelberg equilibrium rather than just perfect Bayesian equilibrium in this section.

To gain intuition on the commitment case, we begin by looking at a couple of examples with two and three periods, in which we can fully characterize the optimal mechanism. For these examples, we presume that leaders have priors of their competence of λ_0. In cases where the leader's prior can vary, we end up with more subcases without substantially more insight.[17] The general case where leaders' priors can vary is fully addressed in the infinite case.

3.1 The two period case

In this scenario, immediate tenure dominates if the replacement cost is sufficiently large as immediate tenure avoids replacement altogether; while the stick of random replacement after safe actions dominates if the replacement cost is small, as it avoids being stuck with a leader who turns out to be incompetent. Let us examine this in more detail.

We presume that $u(\Lambda(0,1,\lambda_0))>0$, so even after a failure in the first period, there is no question of whether or not the principal would like the leader to follow the signal, so the only issue is how to provide proper incentives for her to do so.

[16]The two states are equally probable and the probability of a leader getting the correct signal is the same no matter the state.

[17]The incentive compatibility conditions become more complicated, depending on the leaders' beliefs, while the relevant payoff maximization remains relative to the principal's beliefs.

More formally, a mechanism is characterized by two parameters: π, the probability of retention if y is chosen and the leader fails in period 1; and q, the probability of retention if x is chosen in period 1. It is clear that any optimal mechanism (i.e., which maximizes the principal's expected payoff subject to incentive compatibility of the leader) involves keeping the leader if the leader chooses y and is successful.

To induce the leader to choose the risky action y requires that the following incentive compatibility constraint be satisfied:

$$q \leq f(\lambda_0) + \pi(1 - f(\lambda_0)),$$

where the right-hand side is the overall probability that the leader is retained at the end of period 1 if she chooses action y in that period (this is equal to the success probability $f(\lambda_0)$ times 1 plus the failure probability $1 - f(\lambda_0)$ times π). Given the cost of replacement, it is straightforward to see that this constraint is binding in equilibrium, and thus

$$q = f(\lambda_0) + \pi(1 - f(\lambda_0)). \tag{3}$$

The overall ex ante expected profit of the principal is then

$$\begin{aligned} U(\lambda_0, \pi, q) &= \frac{1}{2}\begin{bmatrix} f(\lambda_0)\left[1 + \delta u(\Lambda(1,1,\lambda_0))\right] + \\ (1 - f(\lambda_0))\left[-v + \pi\delta u(\Lambda(0,1,\lambda_0)) + (1-\pi)(-c + \delta u(\lambda_0))\right] \end{bmatrix} \\ &\quad + \frac{1}{2}\left[-(1-q)c + \delta u(\lambda_0)\right], \end{aligned}$$

Substituting from (3), maximizing the principal's expected payoff with respect to π and q is equivalent to maximizing:

$$(1 - f(\lambda_0))\left[\pi\delta u(\Lambda(0,1,\lambda_0)) + (1-\pi)(-c + \delta u(\lambda_0))\right] + \left[f(\lambda_0) + \pi(1 - f(\lambda_0))\right] c.$$

Noting the linearity of the payoff in π, if

$$c > \frac{\delta}{2}[u(\lambda_0) - u(\Lambda(0,1,\lambda_0))] = \bar{c}(\lambda_0, p, v), \tag{4}$$

then it is better to set $\pi = q = 1$, and with the reverse inequality it is better to set $\pi = 0$ and $q = f(\lambda_0)$.

The above inequality (4) compares between the cost of replacing the leader and the potential gain of having a more competent leader, and leads to the following propositions.

PROPOSITION 1 *There exists a cut-off value $\bar{c}(\lambda_0, p, v)$ such that the optimal mechanism for the principal is: (i) if $c > \bar{c}(\lambda_0, p, v)$, then 'use the carrot' - grant immediate tenure and retain the leader regardless of the outcome; (ii) if $c < \bar{c}(\lambda_0, p, v)$, then 'use the stick' - fire the leader if she takes the risky action and fails in the first period, keep the leader if she takes the risky action and succeeds in the first period, and keep the leader with probability $q = f(\lambda_0)$ if she chooses x in the first period.*

Higher costs clearly favor non-replacement and so the tenure mechanism instead of the random dismissal mechanism. We also note how the optimal mechanism varies with other parameters.

PROPOSITION **2** *The cutoff value is given by*

$$\overline{c}(\lambda_0, p, v) = \frac{\delta(1+v)}{4}\left(\frac{\lambda_0(1-\lambda_0)(p-\frac{1}{2})^2}{\lambda_0(1-p)+(1-\lambda_0)/2}\right).$$

Thus, the random dismissal mechanism is optimal for a wider set of costs as p increases and as v increases. The set of cost values for which it is optimal is initially increasing in λ_0 and eventually decreasing in λ_0 and so non-monotone in the prior probability of competence.

A higher accuracy of signals of competent leaders, and a higher v both increase the relative value of having a competent leader compared to an incompetent one – this increases the relative payoff from the dismissal mechanism compared to the instant tenure mechanism. In the case of signals, higher accuracy also increases the probability of making correct dismissal decisions. The comparative statics in λ_0 are not monotone. If λ_0 is near 0 there is no value in replacing the leader as the replacement is likely to be incompetent. If λ_0 is close to 1, then the leader is likely to be competent and there is no reason to replace her regardless of the first period outcome. It is in intermediate cases that it becomes worthwhile to replace the leader.

Proposition 2 follows directly from the fact that

$$\frac{\delta}{2}[u(\lambda_0) - u(\Lambda(0,1,\lambda_0))] = \frac{\delta(1+v)}{4}\left(\frac{\lambda_0(1-\lambda_0)(p-\frac{1}{2})^2}{\lambda_0(1-p)+(1-\lambda_0)/2}\right).$$

Note that this two-period case does not capture all of the aspects of a tenure contract – as it is essentially a guaranteed contract - there is no decision made after seeing some output from the agent. The comparison between contracts is that of a carrot (guaranteed employment) with a stick (firing for either decision) in terms of motivating the leader. In order to get richer tenure possibilities, we move to three periods.

3.2 The three period case

Moving from two to three periods introduces the possibility that non-immediate tenure dominates immediate tenure or no tenure for suitable parameter values. More specifically, if costs of replacement are sufficiently high, then it makes sense to keep the leader in place forever. If costs are low, then it makes sense to constantly evaluate the leader, with some threat of random replacement for a choice of x in order to maintain incentives to choose action according to signal. For intermediate costs, it can become optimal to conditionally evaluate the leader: if the leader performs well in early periods then the leader is tenured and kept in later periods without any fear of replacement.

More formally, we now compare the following three mechanisms.

(i) A *Random Retention Mechanism* whereby at the end of each period the leader is replaced if she takes action y and fails, and is randomly replaced with positive probability if she takes the safe action x.

(ii) A *Probationary Tenure Mechanism* whereby the leader is kept for sure after the first period, and then is kept after the second period if: either she took action y and was successful in the first period, or if she took action x in the first period and y in the second period and was successful; and she is kept with probability $f(\lambda_0)$ if she took action x in both the first and second periods. The leader is fired after the second period in all other cases. [18]

(iii) An *Immediate Tenure Mechanism* whereby the leader is never replaced regardless of her actions and the outcomes.

These three mechanisms do not comprise the full space of mechanisms, as there are some hybrids. But the full exploration of the space yields little additional insights, and so for the purpose of exposition we compare on these three, which can each be optimal for some parameter values.[19]

We presume that $u(\Lambda(0, 2, \lambda_0)) > 0$, so that even after two failures it is better to have the leader follow signals than to stop taking the risky action, as that simplifies some of the calculations.[20] We also consider the case in which $\delta = 1$ to simplify calculations, as with the finite horizon the discounting case adds little insight.

The following proposition follows from comparison of payoffs of the principal.

[18] Thus, the leader is tenured immediately after being successful on the first attempt; and also with some random probability if both choices in the first two periods were x, and is replaced otherwise, but only replaced after the second period. She cannot be replaced immediately after failure in the first period, or that will distort incentives. The random probability of tenuring after two x's keeps the leader from being forced to take action y regardless of signal in the second period.

[19] More formally, let x denote an x choice, 1 denote a successful y attempt and 0 denotes a failed y attempt and then p_1, p_0, p_x be the corresponding retention probabilities after the first period. So $p_x = .7$ indicates that if x was chosen in the first period then the leader is fired with probability .3 and retained with probability .7. Let p_{x0} (resp. p_{x1}) denote the probability of retention in the second period if x was played in the first period and then y was played and failed (resp. succeeded) in the second period. Let p_{00} (resp. p_{01}) denote the probability of retention in the second period if y was chosen in the first period and failed and then y was played and failed (resp. succeeded) in the second period. And let p_{10} (resp. p_{11}) denote the probability of retention in the second period if y was successfully played in the first period and then y was played and failed (resp. succeeded) in the second period. So, a contract is a specification of

$$p_1, p_0, p_x, p_{11}, p_{10}, p_{1x}, p_{01}, p_{00}, p_{0x}, p_{x1}, p_{x0}, p_{xx}.$$

There are various incentive constraints tying these together, but the full variations of potential mechanisms extend beyond the three considered for our illustration.

[20] More generally, under the immediate tenure mechanism, it could be optimal to incentivize the leader to only take action x in some circumstances, but that case adds little insight to the analysis.

PROPOSITION 3 *There exist cut-off values $c^{IP} > 0$, $c^{PR} > 0$, and $c^{IR} > 0$ of the replacement cost, such that:*

- *the immediate tenure mechanism leads to higher payoffs for the principal than the probationary tenure mechanism if $c > c^{IP}$, with the reverse if $c < c^{IP}$,*
- *the probationary tenure mechanism leads to higher payoffs for the principal than the random retention mechanism if $c > c^{PR}$, with the reverse if $c < c^{PR}$, and*
- *the immediate tenure mechanism leads to higher payoffs for the principal than the random retention mechanism if $c > c^{IR}$, with the reverse if $c < c^{IR}$.*

Thus, for high costs, the immediate tenure mechanism is optimal and for low costs the random retention mechanism is optimal. For some parameter values (combinations of p, v, λ_0), $c^{IP} > c^{PR}$, in which case the probationary tenure mechanism is optimal for intermediate costs.

Intuitively, when the replacement cost c is very large, then it is optimal to never replace the leader and then immediate tenure is optimal. When the replacement cost is very small, it is optimal to re-evaluate the leader in each period, but then random dismissal following action x must be used as otherwise the leader would never take action y since it could lead to failure and dismissal. Tenuring the leader after a success in the first two periods, but firing after a failure (after the second period), emerges as an optimal solution for intermediate values of the replacement cost - as then it is worthwhile to replace a leader who has failed, but to provide incentives by guaranteed employment to a leader who has demonstrated sufficient competence via success.

3.3 The infinite horizon case

We now allow the leader's beliefs to differ arbitrarily from the principal's, and be correlated with her type, as long as she begins with a prior that is positive.

Moving to the infinite-period case, we say that a principal uses a *tenure mechanism* if:

- The principal sets a date τ and a nonnegative integer M and a fraction f.
- The principal commits to:

  - keep the leader forever if the leader chooses y *exactly* M times[21] and is successful at least a fraction f of the time by date τ, and

  - replace the leader otherwise.

[21] If the leader allows a choice of *at least* M choices of y, this motivates an incompetent leader to try y more than M times which is costly for the principal.

Note that we could also define a class of mechanisms which require the principal's posterior be at least some λ by time τ. These are not exactly equivalent, but would provide a result similar to the one that we now state.[22]

PROPOSITION 4 *For any $\varepsilon > 0$ there exists $\overline{\delta} < 1$ and a tenure mechanism such that if $\delta \geq \overline{\delta}$ and principal employs that tenure mechanism, then in all equilibria of the mechanism (in which a leader follows signals once tenured)*[23] *the principal's discounted stream of expected utilities is at least $(1-\varepsilon)$ times the utility the principal would have from having a competent leader in all periods who always chooses based on the signal.*

The implication is that there exists a tenure mechanism that does two things:

- ends up firing all incompetent leaders and eventually keeping a competent leader with an arbitrarily high probability, and
- gives the leader the correct incentives to choose actions consistent with signals in subsequent periods.

Intuitively, with long horizons and sufficient patience, the principal can prolong the test period during which the leader is assessed before deciding whether to tenure her or not, without damaging the overall long run expected utility. This allows the principal to make sure that incompetent leaders are fired while competent leaders are kept with arbitrarily large probability. The advantage of probationary tenure over the random replacement mechanism is that the former saves on replacement costs, while still providing good incentives for risk-taking both during the test period and then during the post-tenure period.

3.4 Excessive risk taking

There is one important aspect of the tenure mechanism design that is worth emphasizing. The tenure mechanism not only requires a given number of successes, it also limits the number of times y can be chosen in total - it requires that y be chosen *exactly* M times.

[22]Working only off of the posterior can lead an incompetent leader to try action y excessively often following failures in order to try to boost the posterior, which is both costly for the principal and can lead to higher chances of incorrectly passing incompetent leaders. This is related to the literature on "calibration"and the literature that followed (e.g., Dawid (1982), Foster and Vohra (1998), Sandroni, Smorodinsky, and Vohra (2003), Dekel and Feinberg (2006)). Our more general replacement strategy avoids the negative results from that literature (that any checking rule in some class can be passed) by allowing restrictions on the strategies and enriching the way in which outcomes can be checked. Here, competent and incompetent agents can be statistically distinguished with high accuracy.

[23]It would be enough to have the leader have lexicographic preferences for following signals, after maximizing total benefits. The proposition is also true if a tenured leader only follows signals as long as that leads to positive expected payoffs, but stops if it is eventually revealed that she is likely enough to be incompetent enough to lead to negative expected continuation values. So, one could have the leader value the principal's payoff lexicographically.

Without this, the leader's strategy would become distorted, as then leaders who end up with too many failures would have incentives to try y more often in order to 'catch up', even without even seeing a Y signal. This would generate excessive risk taking. Generally, without careful design a mechanism could induce too much risk taking, just as it can induce too little risk taking. Thus, (nearly) optimal mechanisms need to be set carefully, with implicit or explicit limits on *both* minimum and maximum numbers of times that the risky action y should be taken.[24]

3.5 Lessons from the commitment case

The take-away from our analysis in this section is that when the case in which the state and signal are privately observed by the leader but yet the principal can commit in advance to a dynamic replacement strategy contingent upon past payoff history: (i) when the replacement cost is sufficiently low, the principal will use the stick: fire the leader with positive probability at the end of some periods if the leader either chose the safe action or chose the risky action but failed in that period;[25] (ii) when the replacement cost is very high, he will grant immediate or early tenure to the leader; (iii) when the replacement cost is intermediate, and there is sufficient patience, it is (always at least approximately) optimal to resort to a tenure mechanism: during a trial period the leader's performance is evaluated and the principal decides whether to keep or replace the leader, and then the leader is kept forever if thought to be competent with sufficient probability.

4 No commitment: voting

The mechanisms we have considered up until now presume that the principal can fully commit to firing a leader in specific situations. In this section we turn our attention to the case where the principal cannot commit to a mechanism ex ante but instead decides in each period whether or not to keep the current leader. The principal makes the choice that maximizes his expected future discounted stream of payoffs in each period.

In this case, we refer to the principal as a "voter," as a central application of this scenarios is to voting settings. Here, we need only one voter since we focus on the leader's competence and abstract from partisan policies.[26]

[24] Fully optimal mechanisms may not involve hard limits, but instead probabilities of retention that approach 0 or 1 in complicated ways as a function of history, but are well-approximated by the simple and prominent tenure mechanisms.

[25] In the fully optimal mechanism, these firing probabilities, both for failures and safe actions, can depend in complex ways on the history and current beliefs.

[26] See Canes-Wrone, Herron, and Shotts (2001) and Kartik and McAfee (2007) analyses of how a quality/competence dimension can interact with a policy dimension.

4.1 A benchmark and some preliminary results

A useful benchmark is that of the one-period term limit institution whereby the current leader is replaced every period no matter her past achievement. The payoff to the principal of this benchmark is

$$V^1 = \frac{u(\lambda_0) - \delta c}{1-\delta}. \tag{5}$$

Note that this is a highly inefficient benchmark as successful leaders are replaced immediately and so there is no opportunity to take advantage of any learning about a leader's competence. Beyond the lack of learning and optimal retention, the cost of replacing a leader is incurred in every period. The only advantage to such a system is that the leader has an incentive to choose according to signal since the leader will be replaced regardless of the outcome.[27]

For some, but not all, results we consider Markov strategies that condition only upon the beliefs. In this case, the principal can only condition upon his beliefs, while the leader knows both beliefs, and so can condition upon both her beliefs and the principal's.

So, Markov strategies for the principal are functions $\sigma_P : [0,1] \to \Delta(\{K, R\})$; while Markov strategies for the leader are functions $\sigma_L : [0,1]^2 \times \{X, Y\} \to \Delta(x, y)$, with $\sigma = (\sigma_P(\lambda_{tP}), \sigma_L(\lambda_{tP}, \lambda_{tL}, s_t))$ representing a generic strategy.[28]

We first note that there always exists an equilibrium and in fact there always exists a Markov perfect equilibrium.

Lemma **2** *There exists a (Markov perfect) equilibrium. In particular, there always exists an equilibrium in which a leader is never voted out of office and all leaders choose x in all situations if elected.*

The proof of Lemma 2 is obvious and so omitted.

4.2 Non-Markovian equilibria

There are many equilibria of the game, but they can be sorted into two broad classes as we now show.

4.2.1 Commitment-like Equilibria

First, there is a class of equilibria in which, through threats of future leaders always choosing x, some sort of current behavior is enforced. Effectively, this allows the principal an arbitrary

[27]For the benchmark, the leader is assumed to choose according to signal (as she is indifferent given that she will be replaced regardless). So this would be the best-case payoff for this institution for the principal.

[28]The principal could be inferring something from the time period about the leader's beliefs. So, the Markov assumption is actually that the strategy conditions only upon the posterior belief, but we allow that belief to be derived from the full history. So, to be careful, a Markov strategy σ is actually a function of h^t, but depends on h^t only through the corresponding induced $\lambda_{tP}, \lambda_{tL}$, as generated via the strategy.

degree of commitment. The basic idea is that all leaders must act as they would under some mechanism as if that mechanism was committed to, and if a leader deviates from how he should act under that mechanism then all leaders resort to playing the equilibrium in which x is chosen forever by all leaders regardless of history. This can enforce many (but not all) mechanisms as if the leader could commit to the mechanism. The reason that it cannot enforce arbitrary mechanisms is that it can only motivate actions by current leaders that provide nonnegative continuation values to each leader at each point on the equilibrium path.

PROPOSITION **5** *For any $\varepsilon > 0$ there exists $\overline{\delta} < 1$ and a tenure mechanism (τ, M, f) as described above, such that if $\delta \geq \overline{\delta}$, then there is an equilibrium in which the principal and leaders act as if the tenure mechanism was in place with probability at least $1 - \varepsilon$ and the principal earns ex ante utility of at least $(1-\varepsilon)u(1)/(1-\delta)$, and if the leader ever deviates from behavior that would be consistent with the tenure mechanism then all players in the future resort to the equilibrium described in Lemma 2.*

The proof of Proposition 5 is straightforward and therefore omitted.[29] Essentially, it points out that arbitrarily efficient equilibria exist, via threats of reversion to a bad equilibrium. This is a way of enforcing commitment by the principal.

On the one hand, such equilibria appear somewhat natural: the principal states that he will follow a certain pattern over time, but if the principal deviates from that then he is no longer trusted and leaders revert to the equilibrium of never taking action y, which is self-enforcing. On the other hand, this sort of construction relies heavily upon the ability of future leaders to see the full play of the game, and thus would be ruled out for example if the model were such that leaders do not see or pay attention to the history of payoffs before their appointment.[30] Also, if there are new voters over time, and current voters are not punished for the actions of past generations, then such constructions would be precluded.

[29] We offer a sketch: The play is as follows, which can be directly checked to be an equilibrium. On the equilibrium path, the principal's actions are as in a tenure mechanism. The leader's actions are such that any leader who is tenured then takes action y if and only if there is a Y signal *and* the current belief λ_{Pt} is such that $u(\lambda_{Pt}) > 0$ (there is an arbitrarily small chance (via choice of mechanism) that the leader is tenured and turns out to be incompetent, which then eventually leads to the leader taking action x thereafter). Leaders who are not tenured try y the first M times where they get a Y signal and otherwise choose action x, unless $\tau - t = M - m$ where t is the current period and m is the cumulative number of y actions taken by the leader to date, in which case they take action y for each remaining period. If the principal ever deviates from prescribed play, then all players play as described in Lemma 2. If a tenured leader ever deviates from prescribed play, then all players play as described in Lemma 2 in the continuation. If a non-tenured leader deviates from prescribed play, then the play continues as in the tenure mechanism (note than the only detectable deviations would be such that the leader takes action y too many times, or too few times, to reach exactly M after τ periods, in which case the leader will already be fired after the τ periods). In any other situation, all players play as described in Lemma 2.

[30] Refinements that impose some sort of renegotiation-proofness can also rule out such constructions, but can be difficult to work with in such settings.

4.2.2 Impossibility of Non-Trivial and Weakly Efficient equilibria

The 'commitment' like equilibria were forced via a very special off-path behavior, requiring all future leaders to resort to never choosing risky decisions. Let us now turn to see how vital that is to establishing an approximate efficiency result.

We show that if we look at a class of equilibria that simply require that new leaders lead to a positive expected payoff for the leader, then not even a weak form of efficiency is possible. That is, the commitment induced in the previous section can only come from a very particular form of equilibria, and a natural class with any form of positive payoffs in the future must lead to very inefficient outcomes. In particular in cases where there is no reversion to a 0 equilibrium, then the approximate efficiency achieved under Proposition 5 is precluded, as we now show.

Let us say that an equilibrium σ is *has non-trivial continuations* if a new leader always provides a net positive expected continuation payoff of at least some $\gamma > 0$ to the principal.[31]

Theorem 1 *Consider any perfect Bayesian equilibrium, σ that has nontrivial continuations. For any λ that is reached on the equilibrium path (no matter how high), there is some (h_P^t, λ_{tP}) reached with positive probability such that $\lambda_{tP} = \lambda$, $V_\sigma(h_P^t, \lambda_{tP}) \leq V_\sigma(\emptyset, \lambda_{0P}) - c$, and the leader is replaced with positive probability conditional upon that history.*

Theorem 1 implies that regardless of how high the posterior is that the leader is competent, the leader cannot avoid being eventually replaced with positive probability - as long as future leaders provide some positive expected payoffs when hired. This underscores the fundamental difficulty in motivating any leader with *any* history to take the risky action y in the absence of term limits.

The intuition is as follows. If conditional upon some history, the leader was never replaced for any continuation that had the same belief, then by taking action x indefinitely, the leader would be guaranteed to never be fired. Thus, the leader cannot expect to be fired in any continuation that lies on the equilibrium path. Then in the (nonzero probability) case in which she is incompetent, the leader must eventually stop taking the risky action in order to be kept and not replaced. However, this leads to a continuation value of 0 for the principal, and since this is less than the value of a new leader she must be replaced at that time, which is a contradiction. Thus, there must be some histories with the original belief under which the leader is replaced with positive probability.

We point out an important corollary of Theorem 1.

Let us say that an equilibrium is weakly efficient if there exists some $\overline{\lambda}$ such that for any $(h^t, \lambda_{tL}, \lambda_{tP})$, if $\lambda_{tP} > \overline{\lambda}$, then σ is such that the leader follows signals.

Thus, weak efficiency is a minimal sort of requirement in terms of having the leader take appropriate actions. It merely requires that there is some high enough belief, such that whenever beliefs of the leader's competency are above that high enough level, the leader

[31]That is, $V_\sigma(h_P^t, \lambda_{0P}) - c \geq \gamma$ for some $\gamma > 0$ whenever h_P^t has a new leader at the end of period t ($h_P^t = (\ldots, R)$).

follows signals. It does not require that the leader always feel compelled or secure in following signals, only that this happen at least while beliefs are close enough to 1 that the leader is competent. As the following corollary states, there is no weakly efficient equilibrium.

Corollary **1** *There do not exist any weakly efficient perfect Bayesian equilibria that have non-trivial continuations.*

The corollary is derived from Theorem 1 as follows. If an equilibrium is weakly efficient, then there are some high enough levels of λ_{tP}, for which the leader always follows the signal. This then implies that the leader will not be fired by the principal in the next period, conditional upon such histories.[32] Thus, the leader will not be replaced if she takes x an arbitrary number of times in a row since beliefs do not change and she is known to be following signals. This contradicts the Theorem.

4.3 Markov perfect equilibria

Theorem 1 shows a weak form of efficiency is impossible. The set of equilibria is still quite complex, and the theorem does not really give us a full feeling for how inefficient equilibria really are. To get a very full characterization of the possible payoffs, we turn to a more restricted, but still natural and interesting class, of equilibria: Markov Perfect equilibria. Here we show that equilibria are extremely inefficient.

Let $V_\sigma(\lambda_{tP})$ denote the expected discounted payoff of the principal/voter conditional upon starting a period with a belief that the leader is competent with probability λ_{tP} and given Markov strategies σ.

Theorem **2** *The entire present discounted value to the principal of* any *Markov perfect equilibrium, σ, is no more than the one-period cost of replacement: $V_\sigma(\lambda_{0P}) \leq c$. Thus, if the value of constant replacement exceeds the cost of replacement ($V^1 > c$), then the value of any Markov perfect equilibrium is worse than simply replacing the leader in every period.*

Theorem 2 makes a strong statement about the inefficiency of *all* Markov perfect equilibria in a democracy. In particular, if $V^1 > c$, then the ex ante discounted expected value value $V_\sigma(\lambda_0)$ of unlimited democracy (as well as the value in any possible continuation, as

[32]The careful argument is more subtle that it appears, as the principal can expect the posterior to drop in the next period with some probability, and if he knows that he will replace the leader in the next period with some probability, he may also be willing to replace her today. The argument that for high enough λ_{tP} he will keep her for at least one period is as follows. Consider any number of periods $T > 0$. If the leader follows signals, then there must then exist some level $\lambda'_{tP} > \overline{\lambda}$ which is reached with positive probability on the equilibrium path (simply by hitting a sequence of successes given that the leader follows the signals), and such that starting from that level of at least λ'_{tP}, beliefs are guaranteed to stay above $\overline{\lambda}$ for at least T periods. Thus, the principal is guaranteed at least $u(\overline{\lambda})$ for at least T periods. By choosing T large enough, this exceeds the value of replacing the leader, and so the leader will be kept.

shown in the proof) is strictly less than the value of the benchmark one-term institution whereby the current leader is replaced in every period.

A basic intuition or heuristic proof for this result (see the Appendix for the full proof) is as follows. First, given the cost of replacement, any leader who has the same probability of being competent as a replacement must be kept with probability one. This then implies that a new leader can keep taking action x and never be replaced. Thus, for the leader to continue to take action y, it must be that she has the assurance of never being replaced thereafter no matter the sequence of successes and failures that ensues. However, with positive probability the leader is incompetent, in which case if she continues to take action y, by the law of large numbers the principal's belief about her ability will end up dropping to a point where she will have to stop taking action y in order to avoid being replaced (as otherwise she yields a negative utility for the principal). However, at that point the continuation value of keeping her on the job is simply zero (as she must take action x forever after) whereas the continuation value of hiring a new leader is $V_\sigma(\lambda_{0P}) - c$. For the current leader not to be fired in equilibrium even at that point, it must be the case that $V_\sigma(\lambda_{0P}) \leq c$.

4.4 The costs and benefits of longer term limits

Given that unlimited democracy may be incapable of providing proper incentives for choosing risky actions indefinitely, we now investigate whether term limits can provide better incentives. This is a limited form of commitment, where the principal/voter cannot commit to any particular decision in each period, but can commit to an absolute limit on re-elections.

For this analysis, we consider the case in which $V^1 > c$, so that there is some benefit in having the leader make appropriate choices, even if that involves frequent replacement. Under this condition, we know that at a minimum, setting a term limit of one period, already improves over the unlimited equilibrium. We compare short versus longer term limits.

4.4.1 Choice between one-period and two-period term limits

On the one hand longer term limits lower the leader's incentives since the leader has more to lose in terms of foregone future private benefits if she chooses the risky action early and is consequently voted out of office in case she failed: in other words, the longer the term limit, the more one has to wait before the leader takes the risky action. On the other hand, a longer term limits lowers the aggregate inter-temporal replacement cost. The higher the replacement cost, the longer the optimal term limit should be.

To get some first intuition of how this trade-off plays out, suppose first that only one-period and two-period term limits are available and also suppose that $V^1 > c$. Theorem **2** implies that some term limit dominates unlimited terms, and in particular dominated by one-period term limits. Let V^2 denote the expected utility to the principal of replacing a leader every two periods and having leaders follow signals in all periods.

The following cases may occur:

Case 1: The replacement cost c is sufficiently large so that:

$$V^2 - c \leq u(\Lambda(0, 1, \lambda_0)) + \delta(V^2 - c). \tag{6}$$

In this case, under a two-period term limit it is always optimal to keep the current leader no matter her performance on action y in her first period of employment. In this case, the leader has proper incentives to take action y in both periods, as she does not fear replacement.

This lead to an equilibrium payoff to the principal under a two-period term limit equal to

$$V^2 = \frac{u(\lambda_0)}{1 - \delta} - \frac{\delta^2 c}{1 - \delta^2}. \tag{7}$$

This equation allows us to re-express condition (6) as

$$c > \widetilde{c}(\lambda_0, p, v) = (1 + \delta)[u(\lambda_0) - u(\Lambda(0, 1, \lambda_0))]$$

or, using our analysis in Section 3.1 above:

$$c > \widetilde{c}(\lambda_0, p, v) = \frac{(1 + \delta)(1 + v)}{2} \left(\frac{\lambda_0(1 - \lambda_0)(p - \frac{1}{2})^2}{\lambda_0(1 - p) + (1 - \lambda_0)/2} \right).$$

Equation (7) also implies that

$$V^2 > V^1 = \frac{u_(\lambda_0) - \delta c}{1 - \delta}.$$

Hence in this case a two-period term limit dominates a one-period term limit whenever condition (6) is satisfied.

Case 2: The replacement cost is sufficiently low that under a two-period term limit it would be optimal to dismiss the current leader if she took action y and failed in her first period of employment. Without commitment, the principal cannot commit to randomly firing the leader if x is chosen in the first period, and then that leads the leader to no longer want to choose y in the first period for fear of failure. In this case a one-period term limit becomes optimal.

4.4.2 Longer term limits

The above logic extends to term limits of greater lengths.

Let V^T be the present expected discounted value starting at λ_0 of having a leader follow the signal in all T periods and replacing the leader every T-th period, but not sooner. This is given by

$$V^T = \frac{u(\lambda_0)}{1 - \delta} - \frac{\delta^T c}{1 - \delta^T}.$$

Let T^* be the largest number of periods T for which it is better to keep the current leader after $T^* - 1$ periods even if she failed in all periods rather than replace her, but not after T^*. That is,

$$V^{T^*-1} - c \leq u(\Lambda(0, T^* - 1, \lambda_0)) + \delta(V^{T^*-1} - c),$$

but

$$V^{T^*} - c > u(\Lambda(0, T^*, \lambda_0)) + \delta(V^{T^*} - c),$$

where

Proposition 6 *The term limit that maximizes the principal's ex ante expected discounted utility is at least T^*.*

The proof is straightforward, and so we just point out the obvious extension from the case of $T = 2$. Just as in the two-period example, the T^*-period term limit provides the leader with correct incentives, and leads to higher expected utility ex ante than any shorter term limit, as the per period cost is lower but the expected utility in every period is the same irrespective of the term limit (provided it is not more than T^*).

Once one moves beyond the T^* expressed above, the equilibria become more complicated, as they can involve mixing or non-signal following in some periods by leaders with some probability. Whether the payoffs to such equilibria actually exceed the T^*-period term limit depends on the circumstances, but it could be optimal to have a term limit longer than T^*.

Note that with a term limit not exceeding T^*, the leader is never replaced except when reaching the constraint. Under the optimal term limit, which could exceed the T^* discussed above, there would still be a strong incumbent advantage.

4.5 Wrapping-up

In this section we have shown that if the signal of the leader is private information and the principal only observes payoff realizations *and* dismissals are decided at the end of each period, then the equilibria that have nontrivial continuations lead to poor incentives for risk-taking by leaders. This contrasts with the commitment case, in which tenure contracts achieve (approximate) efficiency.

5 Summary and discussion

We have analyzed the problem of how to motivate individuals who have discretion over decision making in situations in which firing/replacement threats are the main motivation. In contrast to models where effort rather than discretion interacts with competence, we find that contracts involving commitment to specific periods of evaluation (tenure contracts) or without commitment but then with limited horizons (term limits) are optimal. We have seen that in this world there is a marked contrast between the case in which the principal can commit to keep the leader from cases in which he cannot.

Undoubtedly, there are settings in which both discretion and effort are both at play, and given that we are finding different results from the effort-based literature, it makes sense to examine how the two interact when effort does not satisfy the usual signaling assumptions.

The analysis in this paper can be extended in several other potentially interesting directions. Two such extensions are straightforward. First, what happens if job separation also occurs exogenously with positive probability? Second, what about if a new leader inherits bad outcomes from previous leaders, which impacts on the probability of failure this period?

On the former (exogenous job separation): suppose that no matter the past history of success and failures, in each period there is some probability the current leader leaves the job for some exogenous reason. Effectively, this does acts like a decrease in the discount factor for the leader, and a random cost and restart for the principal. In the case of the tenure result, it means that the patience of the principal must now also include a low probability of exogenous separation, otherwise a random retention mechanism will now dominate and a long tenure evaluation period becomes obsolete since leaders are never expected to stay long enough to be properly evaluated.

Now moving to the case where past failures affect current failure: in our model so far we have assumed that the occurrence of successes and failures is uncorrelated across successive leaders. But suppose instead that the occurrence of a failure last period increases the likelihood of a failure this period (even) when action y is chosen in state Y. This naturally occurs in various settings such as coaching or managing (where a past coach or manager's athletes still comprise the team, and it may take some time to evaluate the new coach/manager's ability to choose good athletes), as well as when politicians inherit various economic or political crises (as well as successes) from previous administrations. This has an interesting effect as it can provide a lag before any real updating can take place. Thus, with complete or incomplete information, it can lead to some time periods where the leader is kept for sure, before we return to a setting where the decisions lead to informative signals, and then to the analysis as we have conducted it here.

Other extensions are equally interesting to investigate. One would be to introduce the possibility of promotions rather than just hiring and firing. A conjecture is that as long as there is a limited set of possible promotions, many of the underlying insights would still remain, but that requires a careful analysis.

Another extension would be to allow for multiple leaders and/or multiple principals. Having several leaders and one principal would allow the principal to observe the outcome from the challenger's actions over time, not just the outcome from the incumbent's action. A conjecture is that this should enhance the leaders' motivation to take risky actions (even to an extreme), as it results in a sort of contest. On the other hand, having several principals compete for a limited number of leaders should increase the likelihood of early tenure, with probationary periods that would be shorter than what would be the optimal mechanism from any single principal's perspective. This emerges from the fact that once a leader begins to have success, a principal can entice that leader away from other principals by offering early tenure.

For example, in the two-period case considered in Section 3, we know that it is optimal for a single principal not to grant immediate tenure to the agent whenever

$$c < \frac{\delta}{2}[u(\lambda_0) - u(\Lambda(0,1,\lambda_0))] = \overline{c}(\lambda_0, p, v).$$

Yet, with multiple principals competing for the same leader, having all principals propose immediate tenure to the leader will be the unique Nash-equilibrium: first given that all other principals propose immediate tenure, a principal's payoff will drop down to zero if he proposes anything less than immediate tenure to the leader; but

$$0 < U(\lambda_0, 1, 1)$$

where[33]

$$\begin{aligned} U(\lambda_0,1,1) &= \frac{1}{2}\left[\begin{array}{c} f(\lambda_0)\left[1+\delta u(\Lambda(1,1,\lambda_0))\right]+ \\ (1-f(\lambda_0))\left[-v+\delta u(\Lambda(0,1,\lambda_0))\right] \end{array}\right] \\ &\quad +\frac{1}{2}\delta u(\lambda_0). \end{aligned}$$

More generally, we conjecture that having more leaders than principals should lead to excessive risk-taking by current leaders, whereas having more principals than leaders should lead to earlier tenure - in either case leading to increased inefficiencies. This provides a rich agenda for future research.

One could analyze the implications, in the no-commitment case, of having elections being held every X periods instead of being held every period, where $X > 1$. However, the basic inefficiency issues underlying Theorems 1 and 2 would still hold in that case.

Finally, the model delivers empirical predictions, in particular on the efficiency of tenure or term limit arrangements, which should be confronted with the evidence on the performance of firms and countries with differing contractual practices, corporate charters or political constitutions. These and other extensions await further research.

References

[1] Alesina, A., and Tabellini, G. (2007) "Bureaucrats or politicians? Part I: a single policy task," *American Economic Review,* 97, 169–179.

[2] Alesina, A., and Tabellini, G. (2008) "Bureaucrats or politicians? Part II: Multiple policy tasks," *Journal of Public Economics*, 92, 426–447.

[3] Allen, F. and Gorton, G. (1993) "Churning Bubbles" *Review of Economic Studies,* 60:4, 813-836.

[33] Here we are making implicit use of the assumption that $u(\Lambda(0,1,\lambda_0)) > 0$.

[4] Alt, J, Bueno de Mesquita, E, and Rose, S (2011) "Disentangling Accountability and Competence in Elections: Evidence from US Term Limits. *Journal of Politics,* 73, 171-186.

[5] Banks, J, and Sundaram, R (1998), "Optimal Retention in Agency Problems", *Journal of Economic Theory,* 82, 293-323.

[6] Besley, T (2006), *Principled Agents? The Political Economy of Good Government,* Oxford: Oxford University Press.

[7] Besley, T, and Case, A (1995a), "Does Electoral Accountability Affect Economic Policy Choices? Evidence from Gubernatorial Term Limits", *Quarterly Journal of Economics,* 110, 769-98.

[8] Besley, T, and Case, A (1995b), "Incumbent Behavior: Vote-Seeking, Tax-Setting, and Yardstick Competition." *American Economic Review*, 85, 25-45.

[9] Besley, T, and Case, A (2003), "Political Institutions and Policy Choices: Evidence from the United States", *Journal of Economic Literature,* 41, 7-73.

[10] Burdett, K, and Coles, M (2003), "Equilibrium Wage-Tenure Contracts", *Econometrica,* 71, 1377-1404.

[11] Canes-Wrone, B, and Shotts, K (2007), "When Do Elections Encourage Ideological Rigidity?", *American Political Science Review,* 101, 273-288.

[12] Canes-Wrone, B, Herron, M, and Shotts, K (2001), "Leadership and Pandering: A Theory of Executive Policymaking", *American Journal of Political Science,* 45, 532-50.

[13] Carmichael, L (1988), "Incentives in Academics: Why Is There Tenure?", *Journal of Political Economy,* 96, 453-472.

[14] Dawid, A. P. (1982), "The Well Calibrated Bayesian," *Journal of the American Statististical Association*, 77, 605-613.

[15] Dekel, E. and Feinberg, Y. (2006) "Non-Bayesian testing of a stochastic prediction," *Review of Economic Studies* 73:4, 893 - 906.

[16] Dewatripont, M, Jewitt, I, and Tirole, J (1999a), "The Economics of Career Concerns, Part I: Comparing Information Structures", *Review of Economic Studies,* 66, 183-198

[17] Dewatripont, M, Jewitt, I, and Tirole, J (1999a), "The Economics of Career Concerns, Part II: Application to Missions and Accountability of Government Agencies", *Review of Economic Studies,* 66, 199-217

[18] Farber, H (1999), "Mobility and Stability: The Dynamics of Job Changes In Labor Markets", in *Handbook of Labor Economics,* Ashenfelter, O, and Card, D (Eds) , Elsevier: North-Holland

[19] Fearon, J (1999), "Electoral Accountability and the Control of Politicians" In *Democracy, Accountability and Representation*, Przeworski, A, Stokes, S, and Manin, B (Eds) New York: Cambridge University Press, 55-97.

[20] Ferejohn, J (1986), "Incumbent Performance and Electoral Control", *Public Choice* 50, 5-26.

[21] Ferraz, C, and Finan, F (2009), "Electoral Accountability and Corruption: Evidence from the Audits of Local Governments", *NBER Working Paper No 14937.*

[22] Foster, D. P., and Vohra, R. (1998) "Asymptotic Calibration," *Biometrika,* 85:2, 379 - 390.

[23] Gordon, S, Huber, G, and Landa, D (2007), "Challenger Entry and Voter Learning", *American Political Science Review*,101, 303-20.

[24] Harrington, J (1993), "Economic Policy, Economic Performance and Elections," *American Economic Review*, LXXXIII, 27-42.

[25] Holmstrom, B (1982) "Managerial Incentive Problems: A Dynamic Perspective," in *Essays in Economics and Management in Honor of Lars Wahlbeck,* Helsinki: Swedish School of Economics.

[26] Holmstrom, B, and Ricart i Costa, J (1986), "Managerial Incentives and Capital Management", *Quarterly Journal of Economics,* 101, 835-860.

[27] Jackson, M, and Lazear, E (1991) "Stock, Options, and Deferred Compensation," *Research in Labor Economics*, 2, 41–62.

[28] Jia, R, Kudamatsu, M, and Seim, D (2014), "Political Selection in China: Complementary Role of Connections and Performance", Working Paper, IIES, Stockholm University.

[29] Jovanovic, B (1979), "Job Matching and the Theory of Turnover", *Journal of Political economy,* 87, 972-990.

[30] Kahn, C, and Huberman, G (1988), "Two-Sided Uncertainty and Up-or-Out Contracts, *Journal of Labor Economics,* 6, 423-444.

[31] Kartik, N. and McAfee, R.P. (2007) "Signaling Character in Electoral Competition," *American Economic Review*, 97(3), 852-870.

[32] Mailath, G. and L. Samuelson (2001), "Who Wants a Good Reputation?" *Review of Economic Studies* 68(2): 415-441.

[33] Mailath, G.J., and Samuelson, L. (2006), *Repeated Games and Reputations*. Oxford University Press.

[34] Mattozzi, A, and Merlo, A (2008), "Political Careers or Career Politicians?" *Journal of Public Economics,* 92 , 597-608.

[35] McPherson, J.M. (1988) *Battle Cry of Freedom,* Oxford University Press.

[36] Padro i Miquel, G, and Snyder, J (2006), "Legislative Effectiveness and Legislative Careers", *Legislative Studies Quarterly,* 31, 347-81.

[37] Rodriguez-Barraquer, T, and Tan, X (2014), "A Model of Competitive Signaling," Unpublished Working Paper, Universitat Autonoma de Barcelona.

[38] Sandroni, A., Smorodinsky, R. and Vohra, R. (2003) "Calibration with Many Checking Rules," *Mathematics of Operations Research*, 28:1, 141 - 153.

[39] Scharfstein, D, and Stein, J (1990) ",Herd Behavior and Investment", *American Economic Review,* 80, 465-479

[40] Smart, M, and Sturm, D. (2013), "Term Limits and Electoral Accountability", mimeo London School of Economics.

[41] Tadelis, S. (1999), "What's in a Name? Reputation as a Tradeable Asset," *American Economic Review*, 89(3), 548-563

[42] Taylor, C. (2000), "The Old-Boy Network and the Young-Gun Effect," *International Economic Review*, 41(4), 871-891

[43] Tirole, J. (1996), "A Theory of Collective Reputations (With Applications to the Persistence of Corruption in the Firm Quality)," *Review of Economic Studies*, 63, 1, 1-22.

[44] Waldman, M (1990), "Up-Or-Out Contracts: A Signaling Perspective", *Journal of Labor Economics,* 8, 230-250.

Appendix A: Proofs

Proof of Proposition 3: We begin by characterizing the principal's payoff from each of the mechanisms.

Recall that $\delta = 1$ and that $u(\Lambda(0, 2, \lambda_0)) > 0$, so that it is optimal for a leader to follow signal in any period regardless of history.

The principal's payoff from the immediate tenure mechanism It follows directly that the ex ante expected utility from the immediate tenure mechanism is

$$U^{IT} = 3u(\lambda_0).$$

The principal's payoff from the probationary tenure mechanism Recall that tenure is offered with probability $f(\lambda_0)$ in the case of two x actions in order to satisfy the incentive compatibility condition to induce the leader to follow the signal in the first two periods. Then, careful calculation (considering each possible realization of signals and actions) shows that the ex ante expected utility from the probationary tenure mechanism is

$$\begin{aligned} U^{PT} &= \frac{1}{2}\left[\begin{array}{c} f(\lambda_0)(1+2u(\Lambda(1,1,\lambda_0))) \\ +(1-f(\lambda_0))(-v+u(\Lambda(0,1,\lambda_0))-c+u(\lambda_0)) \end{array}\right] \\ &+\frac{1}{2}\left[\frac{f(\lambda_0)}{2}(1+u(\Lambda(1,1,\lambda_0)))+\frac{(1-f(\lambda_0))}{2}(-v-c+u(\lambda_0))+\frac{1}{2}(u(\lambda_0)-c(1-f(\lambda_0)))\right], \end{aligned}$$

which can be rewritten as

$$\begin{aligned} U^{PT} &= \frac{f(\lambda_0)}{2}\left[1+2u(\Lambda(1,1,\lambda_0))\right] \\ &+\frac{1-f(\lambda_0)}{2}\left[-v-c+u(\lambda_0)+u(\Lambda(0,1,\lambda_0))\right] \\ &+\frac{1}{2}\left[u(\lambda_0)-c(1-f(\lambda_0))+\frac{f(\lambda_0)}{2}u(\Lambda(1,1,\lambda_0))+\frac{(1-f(\lambda_0))}{2}u(\lambda_0)+\frac{1}{2}u(\lambda_0)\right]. \end{aligned}$$

The principal's payoff from the random retention mechanism The leader is kept after a success and fired after a failure, so to characterize the payoffs from this mechanism, we first need to solve for the various retention probabilities following x actions in both the first and second period.

First, note that the second period incentive compatibility constraints imply that

$$f(\lambda_0) \geq p_{xx} \geq 1-f(\lambda_0) \tag{8}$$

and

$$f(\Lambda(1,1,\lambda_0)) \geq p_{1x} \geq 1-f(\Lambda(1,1,\lambda_0)).$$

Next there is a first period incentive constraint. If the leader takes action y conditional on a Y signal, then the expected payoff is

$$b\left[1+f(\lambda_0)\left(1+\frac{1}{2}(f(\Lambda(1,1,\lambda_0))+p_{1x})\right)\right].$$

If the leader instead chooses x in the first period (and acts according to signal in the second period) then the payoff is

$$b\left[1+p_x\left(1+\frac{1}{2}(f(\lambda_0)+p_{xx})\right)\right].$$

The first period constraint is thus[34]

$$f(\lambda_0)\left(1+p_{1x}\right)) \geq p_x\left(1+\frac{1}{2}(f(\lambda_0)+p_{xx})\right).$$

Setting $p_{1x} = f(\Lambda(1,1,\lambda_0))$ (the maximum possible under the second period constraint) is optimal as it lowers costs, keeps high productivity leaders, and also eases the first period constraint.

Thus, the first period constraint becomes

$$f(\lambda_0)\left(1+f(\Lambda(1,1,\lambda_0))\right) \geq p_x\left(1+\frac{1}{2}(f(\lambda_0)+p_{xx})\right). \tag{9}$$

There are many combinations of p_x and p_{xx} that satisfy this, subject to still satisfying the other second period incentive constraint (8). To see which combination maximizes the principal's expected payoffs, note that the following expression is the principal's expected payoff from the random retention mechanism, U^{RR} (the first part is what the principal gets if action y is chosen and is successful in the first period, the second part is what he gets if action y is chosen and fails in the first period; the third part is what the principal gets if x is chosen in the first period and the leader is retained, and the fourth part is what happens if x is chosen and the leader is replaced; all assuming that the leader follows signals in all periods): [35]

$$\begin{aligned}
U^{RR} = & \frac{f(\lambda_0)}{2}\Bigg[1+u(\Lambda(1,1,\lambda_0))+\frac{f(\Lambda(1,1,\lambda_0))}{2}\left(u(\Lambda(2,2,\lambda_0))+u(\Lambda(1,1,\lambda_0))\right) \\
& \qquad +\frac{1-f(\Lambda(1,1,\lambda_0))}{2}\left(2u(\lambda_0)-2c\right)\Bigg] \\
& +\frac{1-f(\lambda_0)}{2}\left[-v-c+u(\lambda_0)+\frac{f(\lambda_0)}{2}\left(u(\Lambda(1,1,\lambda_0))+u(\lambda_0)\right)+\frac{1-f(\lambda_0)}{2}\left(2u(\lambda_0)-2c\right)\right] \\
& +\frac{p_x}{2}\left[\begin{array}{c} 0+u(\lambda_0)+\frac{f(\lambda_0)}{2}u(\Lambda(1,1,\lambda_0))+ \\ \frac{p_{xx}}{2}u(\lambda_0)+\frac{1-f(\lambda_0)}{2}\left(u(\lambda_0)-c\right)+\frac{1-p_{xx}}{2}\left(u(\lambda_0)-c\right) \end{array}\right] \\
& +\frac{1-p_x}{2}\left[-c+u(\lambda_0)+\frac{f(\lambda_0)}{2}\left(u(\Lambda(1,1,\lambda_0))+u(\lambda_0)\right)+\frac{1-f(\lambda_0)}{2}\left(2u(\lambda_0)-2c\right)\right].
\end{aligned}$$

[34] There is a further constraint that the leader not wish to take action y conditional on an x signal, but that is automatically satisfied in the first period given that this equation will bind with equality (as we argue next).

[35] Note that, for instance, in the first line the expressions $u(\Lambda(2,2,\lambda_0))+u(\Lambda(1,1,\lambda_0))$ represent the two different continuation utilities for the third period, one a second period y action and success where the leader is kept, and the other for an x action and the leader is randomly retained, and each happens with probability $\frac{f(\Lambda(1,1,\lambda_0))}{2}$. Similarly, the $(2u(\lambda_0)-2c)$ represent the cost of a new leader and the third period expected utility for a third leader if the leader either chooses y and fails in the second period, or is randomly replaced after a second period x action; both of which occur with probability $1-\frac{f(\Lambda(1,1,\lambda_0))}{2}$. Other expressions are similarly derived.

Note that this expression can be written as

$$U^{RR} = A + \frac{cp_x}{2}[p_{xx} + 1 - f(\lambda_0)], \tag{10}$$

where A is an expression that is independent of p_x and p_{xx}. Given that this is increasing in both p_x and p_{xx}, it follows that (9) must bind (or else that p_x and p_{xx} are set to their highest possible values, subject to the second period constraint). From solving (9) with equality for p_x and p_{xx} and then substituting it into (10), we find that

$$U^{RR} = A' - \frac{cp_x}{2}, \tag{11}$$

where A' is another expression that is independent of p_x and p_{xx}. Thus, the solution is to set p_{xx} to its highest value that still satisfies (8) of $f(\lambda_0)$, and then to set p_x to the corresponding minimum subject to the constraint (9) binding which is then $p_x = f(\lambda_0)\frac{1+f(\Lambda(1,1,\lambda_0))}{1+f(\lambda_0)}$.

Given these three probabilities, we then have the following expression for the random retention mechanism (the first part is what the principal gets if there is failure in the first period, the second part is what he gets if there is success in the first period; the third part is what the principal gets if x is chosen in the first period):

$$\begin{aligned} U^{RR} \;=\; & \frac{f(\lambda_0)}{2}\Bigg[1 + u(\Lambda(1,1,\lambda_0)) + \frac{f(\Lambda(1,1,\lambda_0))}{2}\left(u(\Lambda(2,2,\lambda_0)) + u(\Lambda(1,1,\lambda_0))\right) \\ & \qquad + \frac{1 - f(\Lambda(1,1,\lambda_0))}{2}\left(2u(\lambda_0) - 2c\right)\Bigg] \\ & + \frac{1 - f(\lambda_0)}{2}\left[-v - c + u(\lambda_0) + \frac{f(\lambda_0)}{2}\left(u(\Lambda(1,1,\lambda_0)) + u(\lambda_0)\right) + \frac{1 - f(\lambda_0)}{2}\left(2u(\lambda_0) - 2c\right)\right] \\ & + \frac{p_x}{2}\left[0 + u(\lambda_0) + \frac{f(\lambda_0)}{2}\left(u(\Lambda(1,1,\lambda_0)) + u(\lambda_0)\right) + \frac{1 - f(\lambda_0)}{2}\left(2u(\lambda_0) - 2c\right)\right] \\ & + \frac{1 - p_x}{2}\left[-c + u(\lambda_0) + \frac{f(\lambda_0)}{2}\left(u(\Lambda(1,1,\lambda_0)) + u(\lambda_0)\right) + \frac{1 - f(\lambda_0)}{2}\left(2u(\lambda_0) - 2c\right)\right], \end{aligned}$$

or, further simplifying,

$$\begin{aligned} U^{RR} \;=\; & \frac{f(\lambda_0)}{2}\Bigg[1 + u(\Lambda(1,1,\lambda_0)) + \frac{f(\Lambda(1,1,\lambda_0))}{2}\left(u(\Lambda(2,2,\lambda_0)) + u(\Lambda(1,1,\lambda_0))\right) \\ & \qquad + \frac{1 - f(\Lambda(1,1,\lambda_0))}{2}\left(2u(\lambda_0) - 2c\right)\Bigg] \\ & + \frac{1 - f(\lambda_0)}{2}\left[-v - c + u(\lambda_0) + \frac{f(\lambda_0)}{2}\left(u(\Lambda(1,1,\lambda_0)) + u(\lambda_0)\right) + \frac{1 - f(\lambda_0)}{2}\left(2u(\lambda_0) - 2c\right)\right] \\ & + \frac{1}{2}\left[0 + u(\lambda_0) - c(1 - f(\lambda_0)) + \frac{f(\lambda_0)}{2}u(\Lambda(1,1,\lambda_0)) + \frac{1 - f(\lambda_0)}{2}u(\lambda_0) + \frac{1}{2}u(\lambda_0)\right] \\ & - c\frac{1 - p_x}{2}. \end{aligned}$$

where

$$p_x = f(\lambda_0)\frac{1 + f(\Lambda(1,1,\lambda_0))}{1 + f(\lambda_0)}.$$

We now prove the result by direct comparison of the expressions.

A) *Comparison between Immediate Tenure (IT) and Probationary Tenure (PT):*

The difference between Immediate Tenure (IT) and Probationary Tenure (PT) only occurs in the third period. There are four possible cases that could have occurred in terms of the signals observed in the first two periods (in order), namely: XX, XY, YX and YY. And among the last two cases only the first Y matters for the tenure decision: so they collapse into $Y\cdot$. Now using the facts that: (i) XX and XY occur with probability $1/4$ each and $Y\cdot$ happens with probability $1/2$; (ii) the difference between IT and PT occurs only under firing; (iii) that in every case (including XX by the IC constraint), firing happens with probability $1 - f(\lambda_0)$, IT dominates PT whenever:[36]

$$\begin{aligned}
& \frac{1}{2}(1 - f(\lambda_0))[-c + u(\lambda_0)] \\
& +\frac{1}{4}(1 - f(\lambda_0)[-c + u(\lambda_0)] \\
& +\frac{1}{4}\{(1 - f(\lambda_0))[-c + u(\lambda_0)] \\
< \; & \frac{1}{2}(1 - f(\lambda_0))u(\Lambda(0, 1, \lambda_0)) \\
& +\frac{1}{4}(1 - f(\lambda_0)u(\Lambda(0, 1, \lambda_0)) \\
& +\frac{1}{4}(1 - f(\lambda_0))u(\lambda_0)
\end{aligned}$$

or equivalently

$$c > \frac{3}{4}\left[u(\lambda_0) - u(\Lambda(0, 1, \lambda_0))\right] = c^{IP}. \tag{12}$$

B) *Comparison between Probationary Tenure (PT) and Random Retention (RR):*

PT will dominate RR if and only if $U^{PT} - U^{RR} > 0$ where:

$$\begin{aligned}
U^{PT} - U^{RR} \;=\; & \frac{f(\lambda_0)}{2}\left[u(\Lambda(1, 1, \lambda_0)) - \frac{f(\Lambda(1, 1, \lambda_0))}{2}\left(u(\Lambda(2, 2, \lambda_0)) + u(\Lambda(1, 1, \lambda_0))\right)\right. \\
& \left. - \frac{(1 - f(\Lambda(1, 1, \lambda_0)))}{2}\left(2u(\lambda_0) - 2c\right)\right] \\
& + \frac{1 - f(\lambda_0)}{2}\left[u(\Lambda(0, 1, \lambda_0)) - \frac{f(\lambda_0)}{2}\left(u(\Lambda(1, 1, \lambda_0)) + u(\lambda_0)\right) - (1 - f(\lambda_0))(u(\lambda_0) - c)\right] \\
& - c\frac{1 - p_x}{2}.
\end{aligned}$$

Hence, $U^{PT} - U^{RR} > 0$ if and only if

$$c > c^{PR} = N/D,$$

[36] The first terms are the continuation utilities under PT for $y\cdot$, xy and xx conditional upon firing (failures in the first two and random firing in the third) and the next three are what the continuations would be under IT in exactly the same cases.

where

$$N = f(\lambda_0)\left[\frac{f(\Lambda(1,1,\lambda_0))}{2}(u(\Lambda(2,2,\lambda_0)) + u(\Lambda(1,1,\lambda_0))) + (1 - f(\Lambda(1,1,\lambda_0)))u(\lambda_0) - u(\Lambda(1,1,\lambda_0))\right]$$
$$+(1 - f(\lambda_0))\left[\frac{f(\lambda_0)}{2}(u(\Lambda(1,1,\lambda_0)) + u(\lambda_0)) + (1 - f(\lambda_0))u(\lambda_0) - u(\Lambda(0,1,\lambda_0))\right]$$

and

$$D = f(\lambda_0)(1 - f(\Lambda(1,1,\lambda_0))) + (1 - f(\lambda_0))^2 + 1 - p_x,$$

where

$$p_x = f(\lambda_0)\frac{1 + f(\Lambda(1,1,\lambda_0))}{1 + f(\lambda_0)}.$$

C) *Comparison between Immediate Tenure (IT) and Random Retention (RR):*

The existence of the cost threshold c^{IR} follows along similar lines. Here is sufficient to note that U^{RR} is decreasing in c and U^{IT} is independent of c, and that $U^{RR} > U^{IT}$ for $c = 0$.

To complete the proof we need only verify that there exist parameter values for which $c^{IP} > c^{PR}$.

Here it is useful to consider the limit case where $p = v = 1$, as the calculations simplify dramatically and it is then easy to see that the results hold in a neighborhood of that case.[37] In particular, when $p = v = 1$ (and $\lambda_0 \in (0,1)$), it follows that, for any $\lambda \in (0,1)$:

$$f(\lambda) = \frac{1+\lambda}{2},$$

$$u(\lambda) = f(\lambda) - \frac{1}{2} = \frac{\lambda}{2},$$

and

$$\Lambda(0,1,\lambda) = 0, \qquad \Lambda(1,1,\lambda) = \frac{2\lambda}{\lambda+1}, \qquad \Lambda(2,2,\lambda) = \frac{4\lambda}{3\lambda+1},$$

so that:

$$f(\Lambda(0,1,\lambda)) = \frac{1}{2}, \qquad f(\Lambda(1,1,\lambda)) = \frac{\lambda}{\lambda+1} + \frac{1}{2}, \qquad f(\Lambda(2,2,\lambda)) = \frac{2\lambda}{3\lambda+1} + \frac{1}{2}.$$

Then, if we use $\lambda_0 = .2$, it follows that from direct computation with the expressions above that:

$$c^{IP} = \frac{3}{8}\lambda_0 = .075$$

and

$$c^{PR} = \frac{N}{D}$$

[37] We did not formally admit $v = 1$ in our definition of the model, but the model works without difficulties in that case, and the payoffs are all continuous in v and so any strict inequalities in the limit remain strict for v close to 1.

is approximately .07.

This completes the proof of the proposition. ▮

Proof of Proposition 4:

Let f be halfway between 1/2 and p, so $f = \frac{1}{4} + \frac{p}{2}$.

Let $M = \tau/2$.

Let us show that, for such M, f, if τ is chosen to be large enough, then the conclusions of the proposition hold.

The proposition is established via the following claims.

(1) For any $\varepsilon' \in (0, 1/2)$ there exists a large enough τ such that

- with probability at least $1 - \varepsilon'$ a competent leader will have at least a fraction of f successes out of any number of more than $\tau/2$ choices of y made conditional on Y signals, and
- with probability less than ε' an incompetent leader will have a fraction of f successes under y choices from $\tau/2$ choices of y (regardless of signals).

(2) For any λ_0 and $\varepsilon'' \in (0, 1)$ there exists some number K such that the probability of at having at least one competent leader out of K tries is at least $1 - \varepsilon''$.

Item (2) follows directly from setting K such that $(1 - \lambda_0)^K < \varepsilon''$.

(1) follows from the law of large numbers. Choose τ large enough so that the probability of a fraction of f successes out of $\tau/2$ tries is no more than $1 - \varepsilon'$ when the coin has a probability of 1/2 but is more than $1 - \varepsilon'$ when the coin has a probability p.

By following a strategy of choosing according to signals, a competent leader will successfully pass the tenure with a probability that is above $1 - 2\varepsilon'$ for large enough τ. *Regardless* of the strategy followed, an incompetent leader will pass with probability less than $1 - \varepsilon'$.

Let us now argue that in *all* equilibria (regardless of leader's beliefs) competent leaders will pass the test with probability at least $1 - 2\varepsilon'$ and the incompetent leaders will fail the test with probability at least $1 - \varepsilon'$.

Let us say that a leader is *forced to make choices* for remaining periods she has already chosen y at least M times, or has made no more than $M - (\tau - t)$ choices of y. *Regardless of a leader's beliefs*, given that they are positive, the leader has a better chance of being successful when choosing y conditional upon a Y signal, rather than upon an X signal. It is then easy to check that it is weakly dominant for a leader (for any beliefs) to choose y if and only if a Y signal is chosen up until: either M choices have been made (and then choose x after), or a period t is reached when exactly $M - (\tau - t)$ choices of y have been chosen and then choose y thereafter. Any equilibrium cannot lead to higher payoffs than these weakly dominant strategies, and so cannot lead to higher probabilities of retention than under these strategies.

By (2) applied to $\varepsilon'' = \varepsilon/4$, we can find K such that within K draws of leaders the probability of having a competent leader is at least $1 - \varepsilon/4$.

Find ε' such that $(1-\varepsilon'^K > 1-\varepsilon/4$. It then follows from 1, using 2 to set K as above, that there exists τ such that within K periods there is a probability of at least $(1-\varepsilon/4)^2 > (1-\varepsilon/2)$ of having a competent leader pass the test and not having any incompetent leader pass the test. This follows since there is a probability of at least $(1 - \varepsilon/4$ of having at least one competent leader in the first K draws. Then under 1, the probability that all incompetent leader in the first K draws fail and any competent leader in the first K draws (if that many are needed) passes the test is at least $(1 - \varepsilon'^K$. By the selection of ε', this later probability is at least $1 - \varepsilon/4$.

Thus, for any ε we can find τ and K such that there is a probability of at least $1 - \varepsilon/2$ that a competent leader will pass the before time $K\tau$ and such that the probability of having an incompetent leader pass the test is less than $\varepsilon/2$.

Set $\overline{\delta}$ such that $\overline{\delta}^{K\tau} > 1 - \varepsilon/2$.

Thus, for any $\varepsilon \in (0, 1/2)$ we can find a tenure mechanism such that there is a probability of at least $1 - \varepsilon/2$ that a competent leader will pass the test (an no incompetent leader will) within $K\tau$ periods, and then lead to the truthful and competent payoffs for all time after $K\tau$. For $\delta > \overline{\delta}$ this leads to a total payoff of at least $1 - \varepsilon/2$ of the payoff as if truthful and competent payoffs were obtained in all periods. Thus, the ex ante expected discounted sum of payoffs from this mechanism are at least $(1 - \varepsilon/2)^2$ of the fully competent and truthful payoffs, which establishes the proposition. ∎

Proof of Theorem 1:

Suppose the contrary of the proposition. Let h_P^t, λ_{tP} be reached with positive probability and suppose that a leader is never fired in any continuation $h_P^{t\prime}, \lambda_{tP}$ that has the same posterior λ_{tP} that is reached with positive probability. This implies that the leader is never fired in any continuation that is reached with positive probability on the equilibrium path, since the leader can guarantee a payoff of b in perpetuity by simply choosing x forever (as any finite sequences of x choices occurs with positive probability and leaves the principal's belief unchanged). Next, note that with some positive probability the leader is incompetent, conditional on the principal's observations (or, actually, under any other observations of history to date). Let us then consider the case in which the leader is actually incompetent.

Consider any $\varepsilon > 0$, and let A be the set of all sequences of possible payoffs (for which all finite truncations have positive probability in equilibrium) for the principal that have the limsup of the fraction of successes divided by failures not exceed $1 + \varepsilon$. Then, if the leader is incompetent, the probability that the continuation sequence lies in A is 1 regardless of the strategy and the leader's beliefs. It follows from Kolmogorov's 0-1 law, the beliefs of the principal on the event A (which is a tail event) converge to 1. For small enough $\varepsilon > 0$ this implies that the continuation value of the principal must lie below γ with positive probability (where γ is as defined the in the statement of the theorem). This is a contradiction, since the principal will then replace the leader. ∎

Proof of Theorem 2: First note that $V_\sigma^D(\lambda_0) > V_\sigma^D(\lambda_0) - c$. This implies that $P(\lambda_0) = 1$, as keeping a leader with λ_0 is better than paying the cost of replacement and getting a leader with the same probability of competence who would follow the same strategy. This implies if a leader takes action x in perpetuity the principal's belief about her competence will remain equal to λ_0 and the leader will be kept in perpetuity. This implies that since the leader has a strategy that guarantees being kept forever response to a Markovian strategy of the voter, the leader must be kept with probability 1 under the leader's best response since the leader values only being in office. Note that since $\lambda_0 < 1$, there is a positive probability that the leader is incompetent (namely, $1 - \lambda_0$). For an incompetent leader (regardless of the leader's beliefs), either there is a positive probability that the leader hits some λ_t after which the leader no longer takes action Y, and then $V(\lambda_t) = 0$, or else the there probability one that the leader continues to take action Y an infinite number of times, in which case λ_t converges to 0 by the law of large numbers. In that second case, it must be that $V(\lambda_{tP}) \leq 0$ for some λ_{tP} reached with positive probability on the path, since an incompetent leader has a negative expected payoff from taking action Y. Thus, in either case, there is a positive probability of hitting some λ_{tP} such that $V(\lambda_{tP}) \leq 0$. Yet, as we argued above the leader is never replaced. This implies that $0 \geq V(\lambda_{tP}) \geq V_\sigma^D(\lambda_{0P}) - c$, which implies that $V_\sigma^D(\lambda_{0P}) \leq c$. ∎

Online Materials
Appendix B: The benchmark case with public information

We now discuss the case where the state and signal are publicly observed after each period.

Given this symmetric information, we examine a situation in which the leaders simply follow the signal in each period. This is easily enforced simply by using a mechanism in which a leader who does not follow signals is fired immediately. Thus, we simply examine the principal's optimal choice of how long to keep any given leader. This is a variation on a standard "bandit problem", and the optimal strategy for the principal can be expressed via a simple cut-off belief such that the leader is replaced at the end of period t whenever the posterior belief on her competence is lower than the cutoff.

Here, the "typical" replacement probability is bell-shaped over time, i.e., first increasing and then eventually decreasing. The intuition is follows. Given that replacing the leader is costly, unless signals are extremely accurate it will not be optimal for the principal to replace the leader immediately. In other words, there is an initial "honeymoon" period where the leader is not replaced as there is some initial number and fraction of failures before the leader's competence could begin to be revealed. This is for the short run. As for the very long run, if the leader survives for a long enough time, by the law of large numbers she is very likely to be competent, in which case she is unlikely to be replaced. Hence, in the very long run, the replacement probability also becomes small. It is in the middle range where the substantial replacement probability falls, as enough information to identify competence with some confidence has accumulated. Overall, we thus expect a bell-shaped replacement probability of the incumbent leader over time.

We suppose that $u(\lambda_0) - c > 0$. This guarantees that it is better to get a new leader than to keep a leader who is thought sufficiently incompetent that the principal would rather have them not even try to take action y even with a good signal.

Let $P(t)$ be the probability that a principal replaces a leader at time t (and kept the leader in all periods before t). An optimal strategy for the principal is to retain the leader as long as $\lambda_t \geq \lambda$ for some $0 < \lambda \leq \lambda_0$. We now show that the replacement probability follows a sawtooth pattern for any threshold strategy (optimal or not).

Claim: Sawtooth Patterns Suppose that the principal starts with some prior λ_0 and retains the leader at the end of period t if and only if $\lambda_t \geq \lambda$ for some $0 < \lambda \leq \lambda_0$. There exists $t > 1$ such that $P(t) > 0$ while $P(t+1) = 0$ and $P(t+k) > 0$ for some $k \geq 1$.

The sawtooth pattern suggested in the proposition is illustrated in Figure 1. That $P(t)$ is strictly positive in period 7 but zero in period 8 can be explained as follows. A leader who is fired in period 7 but not before necessarily had 7 failures in a row over the first seven periods. This follows since a leader is not fired in any previous period regardless of the

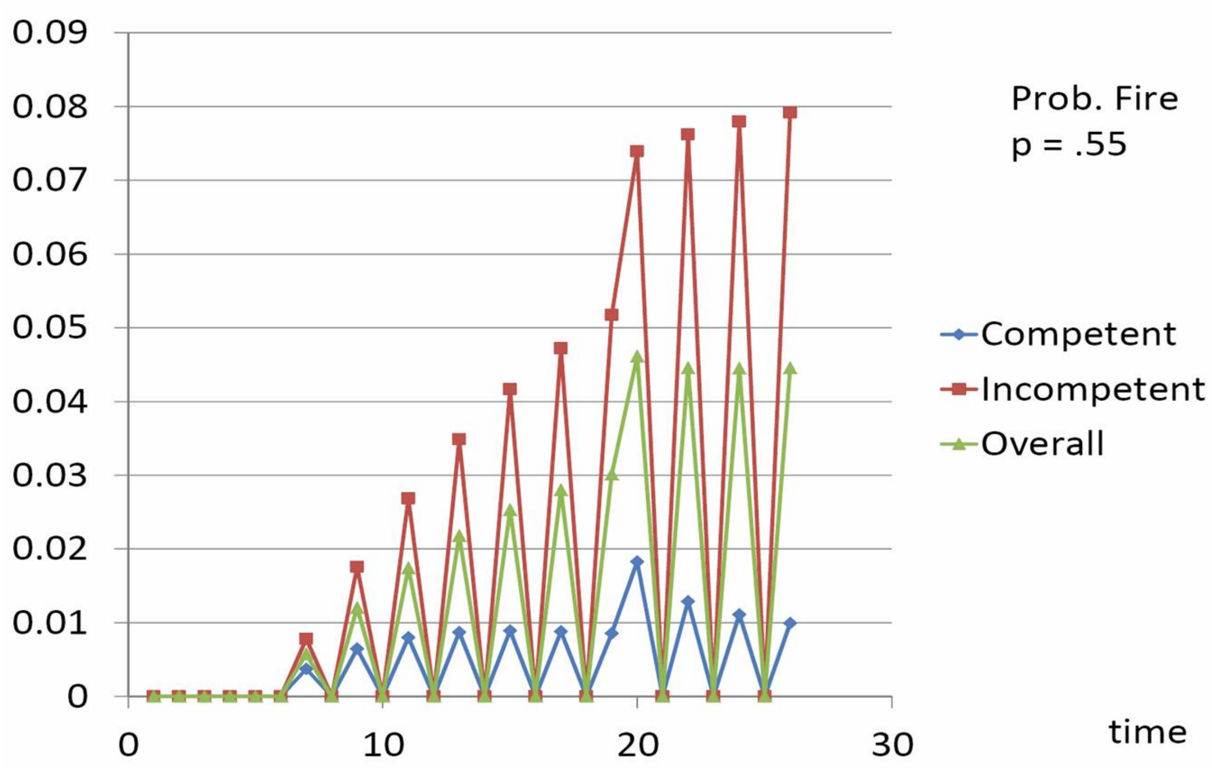


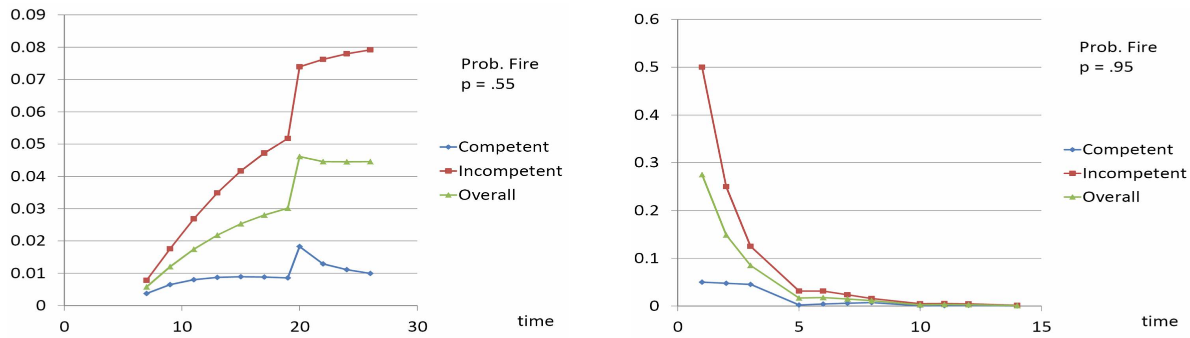
Figure 1: The probability of replacing a leader (for the first time) in various periods for $p = .55$, $\lambda_0 = 1/2$ and $\lambda(c) = 1/3$.

number of failures, including six straight failures. So consider a leader who survived until period 8. Such a leader could not have had eight failures and no success as in that case she would already been fired at the end of period 7. Nor could she have had failures in all first seven periods followed by one success in period 8 because once again she would have been fired at the end of period 7. Thus the only possibility is that she had at most six failures and at least one success over the first seven periods. But then even if she had another failure in period 8 this is better than having had seven failures in a row over the first seven periods, which was the threshold for replacing the leader, and having seven failures and one success is closer to the posterior of having six failures and no successes than seven failures and no success.[38] This in turn implies that a leader who survived until period 8 will not be fired in period 8. In a nutshell: *a success* over the first periods *buys the current leader some additional "grace period" where she is not fired.*

We can also examine just the positive probability dates as pictured in Figure 2. The fact that successes and failures come in integers leads the curves to be non-monotonic even when

[38]Although our reasoning is particular to our discrete time setting, the sawtooth pattern will not fully disappear if we move to continuous time. In the continuous time case, after any success there will still be periods of time during which the leader is kept for sure. Again, a leader who makes it past some particular time must have had some success, and so there can be periods in which the leader is fired with positive probability followed by ones in which the leader is not fired at all.

we look only at dates with positive probabilities.



(a) $p = .55$, $\lambda_0 = 1/2$ and $\lambda(c) = 1/3$ (b) $p = .95$, $\lambda_0 = 1/2$ and $\lambda(c) = 1/3$

Figure 2: The probability of replacing a leader (for the first time) in various periods, *just the dates positive probabilities.*

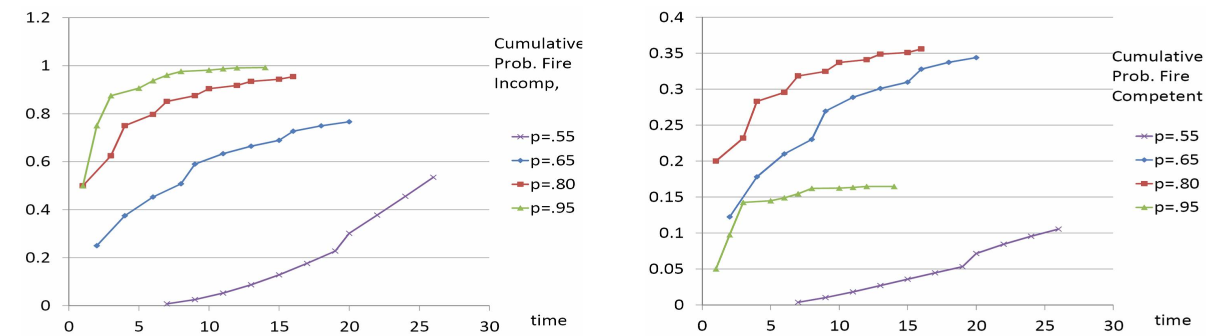
We also see some interesting comparisons between situations where competent leaders are "barely competent" so that $p = .55$ and so hard to tell apart from incompetent leaders, compared to situations where competent leaders are "highly competent" so that $p = .95$ and very different from incompetent leaders. In the left-hand panel where competent leaders are barely competent, it takes time to sort out leaders, and so the probability of replacement is growing over time. Also, the probability of making a "false-positive" or type I error (replacing a competent leader) conditional on making a replacement starts out at roughly 1/2 and then drops over time reaching about 1/8 by period 25. Notice also that the probability of replacing a leader in any given period is quite small, less than .1 in all of the first 25 periods. Also, the first period where any replacement occurs is not even until period 7. In contrast, when competent leaders are "highly competent" then they are much easier to distinguish from incompetent ones, and replacements begin in period 1, and have a much higher probability (.5 for incompetent leaders in the first period). Moreover, the probabilities in that case are decreasing over time. The relative probability a type I error conditional on making a replacement starts out at 1/10 and then actually increases for a few periods.[39]

Cumulative Firing Probabilities in the Complete Information Case

In Figure 3 we see that the cumulative probabilities are not ordered with respect to how competent a competent leader is. Although the replacement curves in the case of incompetent leader are ordered by the p's (see Panel 3a), those for competent leaders are not ordered monotonically (see Panel 3b). The highest probability of replacing a good leader starts out highest for the case of $p = .8$ and lowest for $p = .55$, with $p = .95$ and $p = .65$ in between. The two cases of .65 and .55 eventually overtake the others, but start out lower because

[39]The cumulative probabilities of replacing incompetent leaders, as well as mistakenly replacing competent leaders, also exhibit some interesting patterns as pictured in Figure 3.

information is slow to accumulate in those cases and so replacement probabilities are low initially.



(a) $\lambda_0 = 1/2$ and $\lambda(c) = 1/3$

(b) $\lambda_0 = 1/2$ and $\lambda(c) = 1/3$

Figure 3: The cumulative probability of replacing a leader (for the first time).

By the law of large numbers, incompetent leaders will be recognized with a probability 1 (so $\lambda_t \to 0$ with probability 1 in the case of an incompetent leader), and so eventually any incompetent leader will be replaced as long as the cost of replacement is not so high that one would not replace a incompetent leader even if known to be incompetent. There are two main differences for what happens as a function of the cost of replacement: how quickly incompetent leaders are replaced, and with what probability competent leaders are replaced. These present a trade-off.

Proof of the Sawtooth Claim:

In the benchmark case in which states are observed as well as signals, we can let $o_t = 1$ if the leader is correct (a "success") and $o_t = 0$ if the leader is incorrect (a "failure"). A sufficient statistic for λ_t given λ_0 is the cumulative number of successes through time t:

$$O_t = \sum_{t' \leq t} o_t.$$

As noted above, the posterior probability of having a competent leader $\lambda_t = \Lambda(O_t, t, \lambda_0)$ is increasing in the number of successes through a given time, O_t.

Let

$$O_t(\lambda_0, \lambda) = \{O \in \mathbb{N} : \Lambda(O, t, \lambda_0) < \lambda\}$$

denote the set of possible cumulative successes through time t for which the posterior falls below the threshold λ. Given that $\Lambda(O, t, \lambda_0)$ is increasing in O, there exists

$$\underline{O}_t(\lambda_0, \lambda) = \max O_t(\lambda_0, \lambda).$$

Thus, $O \in O_t(\lambda_0, \lambda)$ if and only if $O \leq \underline{O}_t(\lambda_0, \lambda)$ and $O \in \mathbb{N}$.

The proof of the claim makes use of the following two lemmas.

Lemma B1: Suppose that the principal starts with some prior λ_0 and continues to retain the leader as long as $\lambda_t \geq \lambda$ for some $0 < \lambda \leq \lambda_0$. Then $\underline{O}_t(\lambda_0, \lambda)$ is the *unique* value of O_t for which the principal fires the leader at period t and not before t (if $P(t) > 0$).

Proof of Lemma B1:

If $\underline{O}_t(\lambda_0, \lambda) = 0$, then this holds since this is the maximum value out of $O_t(\lambda_0, \lambda)$ and since there are no lower values, it must be unique. So, suppose $\underline{O}_t(\lambda_0, \lambda) > 0$, and also that to the contrary $o = \underline{O}_t(\lambda_0, \lambda) - k$, with $k > 0$, is another value for which the follower stops at period t and not before t. Note that since $k > 0$

$$\Lambda(o, t-1, \lambda_0) < \Lambda(o+k, t, \lambda_0).$$

Since $\Lambda(o+k, t, \lambda_0) < \lambda$ (noting that $o + k = \underline{O}_t(\lambda_0, \lambda)$), it follows that $\Lambda(o, t-1, \lambda_0) < \lambda$. This implies that $O_{t-1} \leq o$ would lead to the follower stopping at time $t-1$ (or possibly before), and so o could not lead to a first stopping at time t since $O_t = o$ implies $O_{t-1} \leq o$ which would lead to stopping by $t-1$.

Lemma B2: Suppose that the follower starts with some prior λ_0 and continues to retain the leader as long as $\lambda_t \geq \lambda$ for some $0 < \lambda \leq \lambda_0$. Let $P(t') > 0$ and $P(t) > 0$. Then $\underline{O}_{t'}(\lambda_0, \lambda)) > \underline{O}_t(\lambda_0, \lambda))$ whenever $t' > t$, and $\underline{O}_{t'}(\lambda_0, \lambda)) = 1 + \underline{O}_t(\lambda_0, \lambda))$ if $t' = t + 1$.

Proof of Lemma B2:

First, note that $\underline{O}_{t'}(\lambda_0, \lambda)) \geq \underline{O}_t(\lambda_0, \lambda))$. To see this, suppose to the contrary that $o = \underline{O}_{t'}(\lambda_0, \lambda)) < \underline{O}_t(\lambda_0, \lambda))$. However, if $O_{t'} = o$ then $O_t \leq o < \underline{O}_t(\lambda_0, \lambda))$ which implies that the follower would have stopped by t.

Next let us show that $\underline{O}_{t'}(\lambda_0, \lambda)) \neq \underline{O}_t(\lambda_0, \lambda))$. This follows since $O_t \leq O_{t'}$ and so then the principal would have fired the leader by t if the two were the same. Next, let us show that $\underline{O}_{t+1}(\lambda_0, \lambda)) = \underline{O}_t(\lambda_0, \lambda)) + 1$. Note that $\underline{O}_t(\lambda_0, \lambda)) \geq \underline{O}_{t+1}(\lambda_0, \lambda)) - 1$ since

$$\Lambda(o, t, \lambda_0) < \Lambda(o+1, t+1, \lambda_0),$$

for any o. So if $\Lambda(o+1, t+1, \lambda_0) < \lambda$), it follows that $\Lambda(o, t, \lambda_0) < \lambda$. ▮

To show the Sawtooth Claim, let $\underline{t}$ be the first t for which $P(t) > 0$.[40]

We now invoke the two lemmas above. They imply that if $P(t) > 0$ for all $t > \underline{t}$, then it would have to be that

$$\underline{O}_t = \underline{O}_{\underline{t}} + t - \underline{t}.$$

This would imply that $\underline{O}_t/t \to 1$, and so $\lambda_t \to 1$, which contradicts the fact that the leader would be followed. Thus, there is some finite t for which $P(t) = 0$. Note also, that $P(t) > 0$ infinitely often: it is possible that the leader is correct for any arbitrary initial number of periods and then incorrect thereafter for any given number of periods. Thus, for any τ and $\varepsilon > 0$ it is possible to have $\lambda_t > \lambda_0$ for all $t < \tau$, and then $\lambda_t < \varepsilon$ for large enough $t > \tau$. This implies that $P(t) > 0$ for some $t > \tau$. Since this is true for any τ, it is true infinitely often. ▮

[40]There exists such a t since a long enough string 0's will lead λ_t into any given neighborhood of 0. Any long enough string of 0's has positive probability.