Civil Service Reform, Self-Selection, and Bureaucratic Performance

Daniel Gibbs*

October 6, 2019

Abstract

I use a formal model to analyze the effect of civil service protections on bureaucratic performance. In a two-period model, a public manager observes a bureaucrat’s actions for a period and decides whether to retain or attempt to remove the bureaucrat. Bureaucrats vary in terms of their intrinsic motivation and choose between careers in government or the private sector. I show that bureaucratic performance is greater in any separating equilibrium in which motivated bureaucrats choose government than in all equilibria in which they do not. Stronger civil service protections reduce the amount of effort that motivated bureaucrats must exert to distinguish themselves from their unmotivated peers in order to ensure retention. This strengthens incentives for motivated bureaucrats to choose careers in government. Stronger civil service protections, however, also reduce the ability of public managers to remove unmotivated bureaucrats. These competing effects yield a non-monotonic and discontinuous relationship between civil service protections and bureaucratic performance. I use the model to analyze recent reforms to U.S. state and federal personnel management that have significantly rolled back traditional job protections.

Keywords: bureaucracy, civil service, public service motivation

JEL Classification: D73, H11, M59

*Ph.D. Candidate, Department of Politics, Princeton University. Email dagibbs@princeton.edu. I thank Charles Cameron, Brendan Cooley, John de Figueiredo, Matias Iaryczower, Janet Kohlhase, Nolan McCarty, seminar participants at the Princeton University Political Economy Colloquium, conference participants at the Emory Conference on Institutions and Lawmaking 2019, Annual Meetings of the Public Choice Society 2019, and Annual Conference of the Society for Institutional and Organizational Economics 2019 for helpful comments and suggestions.
1 Introduction

Civil service protections typically provide public employees with considerable protection from job dismissal. Public managers must often navigate extensive legal procedures in order to terminate the employment of a bureaucrat. Traditionally justified as a means to prevent politically motivated personnel decisions and patronage, these complex personnel rules present a potential source of government inefficiency. When too rigid, civil service protections impede the ability of public managers to remove bureaucrats who underperform (Osborne and Gaebler 1992; Johnson and Libecap 1994).

Concerns over government efficiency have driven a bipartisan civil service reform movement in the United States that began in the late 1980s inspired by public choice theory and the New Public Management approach to public administration (Suleiman 2003; Kearney and Hays 1998). Broadly, reformers have sought to implement private-sector management techniques in public agencies. At the center of this movement is an effort to drastically weaken traditional civil service protections for government workers. A major reform strategy is the reclassification of traditionally protected bureaucrats as at-will employees who may be dismissed for any reason without warning (Williams and Bowman 2007).

Reforms have been most extensively embraced at the state level. Civil service reform bills passed in Texas (1985), Georgia (1996), Colorado (2003), Kansas (2005), Florida (2001), Indiana (2011), and Arizona (2012) significantly expanded the proportion of at-will state employees and reduced the number of positions classified under traditional civil service protections. Over the course of the 1990s and early 2000s, traditional job protections were curtailed in 28 U.S. states as reform programs were implemented (Hays and Sowa 2007). At the federal level, one of the primary agendas of President Clinton’s “National Performance Review” initiative was to give managers greater discretion to hire and fire government employees (Pfifter 1997). Enthusiasm for more flexible personnel policy in the federal bureaucracy carried over into subsequent administrations. When created in 2002 the Department of Homeland Security adopted a personnel system with weaker civil service protections than
the traditional civil service system. In May of 2018, President Trump issued an executive
order that reduced the amount of time employees are granted to demonstrate acceptable per-
formance and removed the requirement that managers use progressive discipline. In addition
to this formal reorganization of state and federal personnel systems, governments have in-
creasingly relied on subcontracts over the last three decades to sidestep rigid personnel rules
that apply to public employees (Verkuil 2017).

Although this radical transformation of public sector personnel management has received
considerable empirical and normative attention, civil service protection reform has largely
escaped formal theoretical analysis. In this paper I develop a model to identify how civil ser-
vice protections influence bureaucratic performance. I use the results of this model to analyze
the implications of radical civil service reform in the United States. Two ideas are central to
my approach. First, a significant body of empirical literature in public administration and
political science acknowledges that bureaucrats vary in terms of the degree to which they are
intrinsically motivated by public service (Perry 1997; Perry and Hondeghem 2008; Gailmard
2010). In the language of public administration, bureaucrats have heterogeneous levels of
public service motivation (PSM). Left to their own devices, bureaucrats with high PSM vol-
untarily exert effort and advance the policy goals of the agency. Unmotivated bureaucrats, on
the other hand, prefer to exert as little effort as possible. Second, a large literature in public
administration (Paarlberg, Perry and Hondeghem 2008; Perry, Hondeghem and Wise 2010;
Perry and Hondeghem 2008; Moynihan and Pandey 2007; Vandenabeele 2008; Wright and
Grant 2010) and a smaller literature in political science and economics (Gailmard and Patty
2007; Gailmard 2010; Banuri and Keefer 2016; Dal Bo, Finan and Rossi 2013; Valasek 2018;
Delfgaauw and Dur 2010) considers both how PSM can be developed within bureaucracies
and how citizens with high levels of PSM can be effectively recruited into government. This
research is motivated by the idea that a straightforward way to foster quality bureaucratic
performance is to staff agencies with intrinsically motivated bureaucrats.

In the results of the model, the career decisions of motivated bureaucrats are a critical
determinant of bureaucratic performance. Bureaucracies perform better when intrinsically motivated bureaucrats self-select into government careers than when they do not. Civil service protections play a crucial role in the career decisions of motivated bureaucrats. In the model, a bureaucrat serves for a period followed by a performance review. At the performance review, a public manager decides whether to retain the bureaucrat or attempt to remove him from the agency.\(^1\) In order to ensure retention, motivated bureaucrats must distinguish themselves from their unmotivated peers. If the amount of effort that the motivated bureaucrats must exert to set themselves apart is too significant, they choose not to pursue careers in government.

Civil service protections influence the amount of effort that motivated bureaucrats must exert prior to review in order to distinguish themselves. Strong civil service protections make it unlikely that a bureaucrat is successfully terminated in the event that his manager tries to remove him. As civil service protections are made more robust, unmotivated bureaucrats become more willing to be identified as unmotivated and thus less willing to exert effort to imitate a motivated bureaucrat. This reduces the amount of additional effort that motivated bureaucrats must expend to distinguish themselves. Civil service protections take this pressure off of motivated bureaucrats which makes a career in government more appealing to these candidates. Motivated agents who initially self-selected out of government under weaker civil service protections can therefore be brought into government under a stronger protection regime.

The overall result is a non-monotonic relationship between the robustness of civil service protections and bureaucratic performance. At low levels of civil service protections, motivated bureaucrats do not self-select into government. At some greater level of protections, motivated types switch into government which results in a discontinuous rise in bureaucratic performance. On either side of the discontinuity, bureaucratic performance declines continuously as unmotivated bureaucrats become increasingly difficult to remove.

\(^1\)Consistent with convention in the principal-agent literature, I refer to the bureaucrat (agent) with male pronouns and the manager (principal) with female pronouns.
These results imply that the success of radical reform and the optimal reform strategy depend on the conditions that prevail prior to reform. Locally, small reductions in civil service protections yield overall gains to bureaucratic performance as unmotivated bureaucrats become easier to fire without affecting the career decisions of motivated bureaucrats. Major reforms, on the other hand, can affect whether or not motivated bureaucrats participate in government. If motivated bureaucrats initially choose careers in government, radical reform can force these bureaucrats out of government. This drags down government performance overall. If initially these bureaucrats do not choose government, the overall effect of radical reductions in protections is positive. Such a reform, however, is suboptimal. When motivated bureaucrats do not choose government prior to reform, the optimal reform strategy is to increase the robustness of protections in order to induce motivated types to switch into government careers. Moreover, because bureaucratic performance is always greater when motivated types choose government, a suboptimal policy that grants public employees too much protection outperforms any suboptimal policy that provides too little protection. These results urge caution upon radical reformers who implement organizational change in an uncertain environment.

2 Related Literature

The model is closely related to a literature on bureaucracies that examines the recruitment and retention of bureaucrats with desirable preferences (Besley and McClaren 1993; Prendergast 2007; Cameron, de Figueiredo and Lewis 2016; Cameron and de Figueiredo 2019; Lewis 2008; Valasek 2018). A small strand of this literature considers how civil service protections may induce motivated bureaucrats to self-select into government. One argument in this literature maintains that strong civil service protections raise the expected long-term earnings of desirable candidates relative to less protected private sector jobs (Bowman and West 2006; Battaglio 2010; Moe 2011). By offering relative certainty over long-term employment,
government agencies can offer competitive compensation at a lower nominal wage.

Gailmard and Patty (2007, 2013) provide a different account of how job stability encourages motivated candidates to choose government. In their model, civil service protections provide job security which strengthens the incentives of motivated agents to invest in relationship-specific expertise necessary to effectively implement policy. Under robust civil service protections, motivated types are likely to capture future productivity gains from specific investment in expertise, as it is likely that they will be employed in government in the future. In both the accounts, the job security of motivated bureaucrats is key to inducing their participation. I demonstrate the existence of an additional mechanism through which job security induces the participation of motivated bureaucrats in government. In my model, civil service protections relieve motivated bureaucrats of the pressure to exert significant effort to distinguish themselves from their unmotivated peers. This is accomplished because of the job security offered to unmotivated bureaucrats.

A larger segment of the bureaucratic recruitment literature considers how compensation affects the pool of citizens who choose careers in government. Several theoretical studies have observed that high wages may attract citizens with low public service motivation to careers in government (Francois 2000; Delfgaauw and Dur 2010; Dal Bo, Finan and Rossi 2013). If a job candidate’s intrinsic utility from public service is independent of wage, then any candidate who is sufficiently motivated to choose government at one wage also chooses government at a higher wage. Candidates who are insufficiently motivated to participate at the original wage, on the other hand, find government sufficiently attractive at a higher wage and enter the bureaucracy. In my model, non-pecuniary and pecuniary utility are dependent in equilibrium. This is an endogenous result of screening: motivated types must demonstrate their high motivation in order to be retained. Motivated bureaucrats may have to exert more effort than they would like in order to secure a government career, cutting into the intrinsic utility they receive from public service. A higher government wage can compensate for this. As a result, a higher wage may induce the participation of more motivated agents. This
result can be obtained in existing models if a candidate’s motivation and the value of his outside option are positively correlated (Dal Bo, Finan and Rossi 2013).

My model obtains this result for uncorrelated motivation and reservation wages. It is not true, however, that higher wages always make a government career more appealing to motivated types. As wages rise, unmotivated bureaucrats become more willing to exert effort to imitate motivated types in order to secure long-term employment. The motivated type must therefore exert more effort to separate which further erodes his intrinsic utility. While the model does not admit intuitive conditions for when one effect dominates the other, the competing effects that it characterizes make recent conflicting experimental papers that show a positive (Dal Bo, Finan and Rossi 2013) and negative (Deserranno 2019) association between wages and the attractiveness of government work to motivated candidates less surprising.

This paper also contributes to a literature on the welfare effects of civil service protections that is primarily comprised of informal theoretical and empirical studies. Notable recent papers by Folke, Hirano and Snyder (2011) and Ujhelyi (2014) empirically examine the effect of civil service on policy and political outcomes. Rauch (1995) and Rauch and Evans (2000) are more closely related to my outcome of interest, bureaucratic performance. They examine the effect of civil service rules on bureaucratic performance and uncover a positive association between the two. A more recent empirical literature in public administration studies the effects of contemporary radical civil service reform in U.S. states (Bowman and West 2006; Coggburn 2006; Goodman and Mann 2010; Bowman et al. 2003; Coggburn et al. 2010; Condrey and Battaglio 2007). The findings of this literature have been mixed, yielding few consistently robust findings linking civil service reform to objective measures of agency performance (Jordan and Battaglio 2014). My model supplements this empirical literature by providing a clear predicted theoretical relationship between civil service protections and

\footnote{Valasek (2018) shows a non-monotonic result in a different setup where workers also value the collective reputation of an organization. The effect of higher wages on motivation is negative for high reputation organizations and negative for those of low-reputation.}

\footnote{A related literature examines why civil service reforms are implemented (Geddes 1995; Johnson and Libecap 1994; Mueller 2015; Ting et al. 2013; Cruz and Keefer 2015).}
bureaucratic performance.

In the informal theoretical literature there are two major competing arguments concerning the effect of civil service protections on bureaucratic performance. The traditional argument for civil service protections views these rules as a means of insulating a bureaucracy from patronage and politicization. In addition to the potential normative undesirability of such practices in a democracy, patronage and politicization may undermine the performance of an agency by dampening the morale of talented career civil servants and inserting under-qualified personnel into the ranks of the bureaucracy (Lewis 2008). While undoubtedly an important aspect of the role that civil service protections perform, I bracket off such concerns in my analysis. Accordingly, my model serves as a best case for the second argument in the informal theoretical literature which identifies civil service protections as a drag on government performance by making undesirable bureaucrats difficult to remove (Osborne and Gaebler 1992). This bracketing off of politicization facilitates identification of a novel mechanism linking civil service protections to bureaucratic performance.

Ujhelyi (2014a) and Forand (2019) are important exceptions to the informality of the theoretical literature on the welfare effects of civil service reform. Ujhelyi (2014a) presents a formal model that considers the strategic relationship between a politician who sets policy, a bureaucrat who implements policy, and a voter. Job security may weaken bureaucratic performance by allowing bureaucrats to strategically implement bad policies so that voters will replace a politician with one that the bureaucrat prefers to the incumbent. Overall voter welfare, however, may rise with a bureaucrat’s job security. This is because politicians lose the ability to remove good bureaucrats which in turn can incentivize them to select better policies. Forand (2019) also embeds a strategic bureaucrat within a political agency model. Without job protections, politicians can pressure bureaucrats to use public resources in the service of their reelection. With robust job protections, bureaucrats have no stake in the electoral outcome and therefore provide public goods. Voters understand this partisan inefficiency and consent to higher taxes for public goods when bureaucrats’ jobs are protected.
Larger government budgets under tenure, however, provide politicians with greater resources to co-opt bureaucrats. Public sector wages must rise to an inefficient level to offset this.

The welfare analysis of civil service reform in this paper takes a different approach than Ujhelyi (2014b) and Forand (2019), focusing on the career decisions of bureaucrats rather than their political concerns. Integration of these two approaches is an obvious avenue for future research as this small theoretical literature grows. Despite differences in approach, all three models highlight tradeoffs that civil service protections present to public welfare. Ujhelyi (2014a) shows that civil service protections prevent bad politicians from forcing good bureaucrats to implement bad policies but also allow bad bureaucrats to undermine good politicians. Forand (2019) uncovers a tradeoff between efficient public good provision and efficient government wages. My model reveals that strong protections make unmotivated bureaucrats difficult to remove but induce motivated bureaucrats to choose careers in government.

3 Model

The actors in the model are a bureaucrat and a manager. There are two time periods. In each period, a bureaucrat chooses a level of effort $a_t \in A = \mathbb{R}_+$. This action is observable to the manager and generates a policy payoff for her. After a bureaucrat’s first period of employment, the manager chooses whether to retain or fire the bureaucrat. Denote the manager’s decision with $r \in R = \{0, 1\}$ where $r = 1$ connotes retention. If the manager chooses to fire the bureaucrat, he is successfully removed with probability $(1 - \rho)$. The parameter $\rho$ formalizes the strength of civil service protections. High values of $\rho$ represent robust civil service protections where there is a low probability that a bureaucrat is removed from government if his manager attempts to fire him. For $\rho = 0$, a bureaucrat is employed at will. In this case the manager has full discretion to fire or retain the bureaucrat.

Prior to the manager’s decision, the bureaucrat chooses either to stay in government
or exit. Denote the bureaucrat’s choice \( h \in H = \{0, 1\} \) where \( h = 1 \) indicates that the bureaucrat stays and \( h = 0 \) indicates that he exits. A bureaucrat cannot be rehired after he exits or is fired.

In each period the manager’s payoff is given by

\[
u_t^m = v(a_t)\]

The function \( v(\cdot) \) is continuously differentiable, increasing, and strictly concave with \( v(0) = 0 \). This represents the payoff the manager earns from the everyday execution of policy tasks that the bureaucrat carries out. The manager’s utility from policy represents that of the general public or the political principal who selects the policy that bureaucrats pursue in their everyday work.

There are two types of bureaucrats, \textit{good} and \textit{bad}. Formally, a bureaucrat’s type is denoted \( \gamma \in \Gamma = \{g, b\} \). With probability \( \lambda \in (0, 1) \), a randomly selected bureaucrat is a good type. A bureaucrat’s type is his private information. A good type earns a payoff in employment-period \( t \) of

\[
u_t^g = v(a_t) - c(a_t) + R_G
\]

unless the manager has attempted to fire him. Good types suffer a cost \( k \) when employed in government after an unsuccessful termination. This cost reflects policy and psychological loss that good types bear when their manager believes that they are a bad type. In public agencies, unsuccessfully terminated bureaucrats are often reassigned to low-impact positions or not assigned a position at all. Former patrol officers are assigned to administrative desk jobs for example. In the New York City school district, unsuccessfully terminated teachers are frequently not assigned a class. The bad type does not mind such an arrangement and does not suffer \( k \) from remaining in an organization that attempted to remove him. The cost function \( c(\cdot) \) is continuously differentiable, increasing, and strictly convex with \( c(0) = 0 \). The term \( R_G \) is the constant difference between the non-policy benefits of office for the
good type and an outside option normalized to zero. The good type’s policy preferences are
congruent with those of the manager. Both prefer greater policy output. The bad type on
the other hand receives no payoff from executing the policy tasks of the agency. His payoff
in employment-period $t$ is

$$u^b_t = -c(a_t) + R_B$$

There is no discounting. The game is played with the following sequence of moves:

1. Nature randomly selects a bureaucrat for employment.

2. The bureaucrat selects a level of effort $a_1$ and chooses to stay in government or exit.
   If he stays, the game proceeds to (3). If he exits, the game proceeds to (4).

3. The manager observes $a_1$ and fires or retains the bureaucrat. If she fires the bureau-
crat, the game proceeds to (4) with probability $1 - \rho$ and to (5) with probability $\rho$. If she
   retains the bureaucrat, the game proceeds to (5).


5. The bureaucrat selects of level of effort $a_2$.

6. Payoffs are realized and the game ends.

To economize on notation in the main text, definitions of each player’s beliefs and strate-
gies as mappings from type and all possible information sets are formalized in the Appendix.
The analysis in the main text only requires notation for the manager’s beliefs. Let $\bar{\lambda}(a_1)$
denote the manager’s belief that a bureaucrat is a good type given that the bureaucrat re-
mains in the bureaucracy and exerts $a_1$ in the first period. I denote the unique level of effort
that the good type prefers to exert while employed in government with

$$\bar{a} \equiv \arg \max_a \quad v(a) - c(a)$$

and let

$$W \equiv v(\bar{a}) - c(\bar{a})$$

denote the policy utility that the good type receives in a period if he exerts his preferred
amount of effort. I focus on the case in which $R_G + W \geq 0$. When this inequality does not hold, non-policy benefits of office are so low that good types exit after one period in all equilibria. I also assume that $k \geq W + R_G$. This assumption ensures that policy-motivated bureaucrats prefer to exit rather than remain with an organization where their supervisors believe they are unmotivated and try to expel them. Finally, to rule out cases in which bad types do not seek long-term government employment, I restrict attention to $R_B \geq 0$.

Assumption 1

i. $W + R_G \geq 0$

ii. $k \geq W + R_G$

iii. $R_B \geq 0$

4 Equilibrium

I restrict attention to pure strategies and define equilibrium to be a perfect Bayesian equilibrium that satisfies the intuitive criterion. Because $R_B \geq 0$, exit after one period is a weakly dominated strategy for the bad type. To reduce the number of equilibria, I focus on equilibria in which the bad types always stay.

Second-period bureaucrat effort strategies and the manager’s retention strategy are straightforward. In any equilibrium, in the final period of employment bad types exert no effort. Good types choose $\tilde{\alpha}$. This is true regardless of whether a bureaucrat is newly hired or a hold over from the first period. Moreover, because the previous actions of all players are irrelevant to the payoff a bureaucrat receives when choosing a level of effort in the second period, these actions are uniquely optimal for all possible histories. Given these second-period bureaucrat strategies, the manager’s equilibrium strategy selects a bureaucrat for retention if and only if she believes that the bureaucrat is at least as likely to be a good type as his replacement.
I define two types of separating equilibria, differentiated by bureaucrat actions in the first period. In an exit equilibrium, good types exert their preferred level of effort and exit. Bad types exert no effort, stay, and are fired. In a stay equilibrium, good types exert a positive level of effort, stay, and are retained. Bad types exert no effort, stay, and are fired.\footnote{I provide complete formal definitions of each type of equilibria in the Appendix.}

To identify conditions under which each type of equilibrium exits, it will prove useful to define a unique level of effort, $s$, in terms of exogenous parameters such that the bad type is indifferent between exerting $s$ and being retained and exerting no effort and being fired. If a bad type exerts no effort and is fired, his expected payoff is $\rho R_B$. If he exerts a positive level of effort, $a_1 > 0$, and is retained, his expected payoff is $-c(a_1) + R_B$. The level of effort

$$s \equiv c^{-1}(1 - \rho)R_B$$

is the unique solution in terms of $a_1$ to

$$\rho R_B = -c(a_1) + R_B$$

In a stay equilibrium, the good type must exert at least $s$ in the first period in order to prevent the bad type from imitating him in order to be retained. If $s \leq \tilde{a}$, the good type’s preferred level of effort is sufficiently high that the bad type has no incentive to imitate the good type if he exerts his preferred level in equilibrium. If $s > \tilde{a}$, then the good type must exert more effort than he prefers to in a stay equilibrium in order to distinguish himself from the bad type. In particular, he exerts just enough effort to separate, $s$. Let $\hat{a} \equiv \max\{\tilde{a}, s\}$ denote the level of effort that the good type exerts in the first period in a stay equilibrium.

In order for a stay equilibrium to exist, it must be incentive compatible for the good type to exert $\hat{a}$ in the first period. The good type’s expected payoff in a stay equilibrium is

$$v(\hat{a}) - c(\hat{a}) + W + R_B$$
If \( \hat{a} = \tilde{a} \), the good type has no incentive to choose any other level of effort, as doing so reduces his policy utility. If \( \hat{a} = s \), the good type would like to choose a level of effort closer to \( \tilde{a} \) if the manager were to still retain him after observing this alternative action. Because \( \hat{a} \leq s \), however, the bad type also would strictly benefit from exerting \( a_1 < \hat{a} \). The manager understands this and under the intuitive criterion is free to believe anything after observing an unexpectedly low level of effort. For sufficiently unfavorable beliefs about a deviant bureaucrat’s type, the manager fires any bureaucrat who chooses \( a_1 < \hat{a} \). Under Assumption 1, this makes the good type strictly worse off than equilibrium, even if he exerts his preferred level of effort in the first period. The good type’s payoff-maximizing deviation from equilibrium is therefore to exert \( \tilde{a} \) and exit. This yields a payoff of \( W \). A stay equilibrium then requires that

\[
v(\hat{a}) - c(\hat{a}) + R_G \geq 0
\]

Note that if \( \hat{a} = \tilde{a} \), this participation constraint is satisfied under Assumption 1. For the constraint to fail, the bad type must be willing to exert more that \( \tilde{a} \) to ensure retention. In particular, this level of effort, \( s \), must exceed that which the good type is willing to exert to ensure retention. Let \( z > \tilde{a} \) be defined such that

\[
v(z) - c(z) + R_G = 0
\]

The good type is indifferent between choosing \( z \) and being retained and exerting \( \tilde{a} \) and exiting. Because \( s \leq \hat{a} \), if \( s \leq z \), the good type’s strategy in a stay equilibrium is incentive compatible and a stay equilibrium exists. In other words, if the good type is willing to exert more effort to ensure retention than the bad type is, a stay equilibrium exists. On the other hand, if \( z > s \), then the amount of effort required to distinguish himself from the bad type is more than the good type is willing to bear. In this case he strictly prefers to exit the bureaucracy after one period and a stay equilibrium does not exist.
Proposition 1 A stay equilibrium exists if and only if $s \leq z$.

All proofs are in the Appendix. In an exit equilibrium, the bad type can do no better than to choose no effort in the first period because the manager fires any bureaucrat who stays (good types exit in equilibrium). If the good type deviates from equilibrium and remains within the bureaucracy, the manager attempts to fire him, as only bad types stay in equilibrium. Under Assumption 1, this makes the good type worse off than if he plays his equilibrium strategy of leaving after exerting his preferred level of effort for one period. Given that staying in the bureaucracy only reduces his payoff relative to equilibrium, the only other alternatives to consider are to levels of effort other than $\tilde{a}$ and exit. Because $\tilde{a}$ maximizes his policy payoff, no action other than that prescribed by his equilibrium strategy can make him better off.

For an exit equilibrium to exist, the manager must believe that a bureaucrat has a low probability of being a good type whenever she observes a level of effort that the good type would strictly prefer to exert and be retained. Formally, this requires that $\bar{\lambda}(a_1) < \lambda$ for all $a_1 < z$. To sustain these beliefs, the intuitive criterion requires that there is no $a'_1$ such that

$$-c(a'_1) + R_B \leq \rho R_B$$

and

$$v(a'_1) - c(a'_1) + W + R_G > W$$

The first condition states that at $a'_1$, the bad type is (weakly) better off playing his equilibrium strategy than exerting $a'_1$ and being retained. The second condition states that the good type is strictly better off exerting $a'_1$ and being retained than playing his equilibrium strategy. If such an $a'_1$ exists, then all beliefs other than $\bar{\lambda}(a'_1) = 1$ are ruled out under the intuitive criterion. The manager must believe that it was the good type who exerted $a'_1$ as only he stands to strictly benefit from such a deviation from equilibrium. If there is no such $a'_1$, then any deviation that strictly benefits the good type also strictly benefits the bad type. The
beliefs required to sustain the manager’s strategy of firing bureaucrats who stay survive the
intuitive criterion and an exit equilibrium exits. Note that these conditions can be rewritten
using $s$ and $z$ as

\[ s \leq a_1' \]

and

\[ a_1' < z \]

Rearranging terms reveals that if $s < z$, an exit equilibrium does not exist. If on the other
hand $s \geq z$, an exit equilibrium exists.

**Proposition 2** An exit equilibrium exists if and only if $s \geq z$.

It is possible to rule out the existence of all other classes of separating equilibria, where
classes of equilibria are distinguished by first-period bureaucrat strategies. First note that in
all separating equilibria, the manager correctly identifies the bureaucrat’s type. Equilibrium
requires that he fire the bad type and retain the good type. There is therefore no separating
equilibrium in which the bad type exerts a positive level of effort: he can always do better
by exerting no effort and being fired. There is also no equilibrium in which the good type
exits and exerts a level of effort other than $\tilde{a}$. He can always do better by exerting $\tilde{a}$ and
exiting.

It remains only to consider equilibria in which the good type exerts some effort other
than $\hat{a}$ and is retained. First consider the case in which $s \leq \tilde{a}$. In this case $\hat{a} = \tilde{a}$. If the good
type’s equilibrium choice of effort is not $\hat{a}$, then he is strictly better off than in equilibrium
if he is retained after a deviation to $\hat{a}$. Because $s \leq \hat{a}$, the good type is the only type who
strictly benefits from this deviation. Therefore under the intuitive criterion the manager
must believe he is a good type after observing $\hat{a}$. The manager must retain the bureaucrat
in this case. Accordingly, no level of effort other than $\hat{a}$ is an equilibrium strategy for the
good type if $s \leq \tilde{a}$. If $s > \tilde{a}$, then $\hat{a} = s$. By the definition of $s$, the good type cannot exert
less than $\hat{a}$ in equilibrium because the bad type can profitably imitate the good type and
be retained. If the good type exerts more than \( \hat{a} = s \) in equilibrium, then the good type is strictly better off exerting slightly less effort than \( s \) if he is retained. The same level of effort strictly harms the bad type because it is greater than \( s \). Under the intuitive criterion, the manager believes that it is the good type who exerts this unexpected level of effort and retains him. There is therefore no separating equilibrium in which the good type stays and does not exert \( \hat{a} \).

**Remark 1**

If \( s > z \), the unique separating equilibrium is exit.

If \( s < z \), the unique separating equilibrium is stay.

If \( s = z \), there exists an exit equilibrium and a stay equilibrium.

An implication of Propositions 1 and 2 is that a separating equilibrium always exists. Whether that equilibrium is a stay equilibrium or an exit equilibrium depends on whether the good or bad type is willing to exert more effort in order to ensure retention. That is, the type of equilibrium that exists is determined by whether \( s \) or \( z \) is higher.

In addition to these separating equilibria, pooling equilibria exist for some regions of the parameter space. In a pooling equilibrium, both types exert the same amount of effort and are retained. This level of effort must be below both \( s \) and \( z \). Moreover, both types must exert at least \( \tilde{a} \). This second condition is a consequence of the intuitive criterion. If the pooling equilibrium prescribes that both types exert \( a_1 < \tilde{a} \), the good type strictly gains relative to equilibrium from exerting \( \tilde{a} \) if doing so results in retention. This is strictly harmful for the bad type. Therefore the manager must believe that a bureaucrat who unexpectedly exerts \( \tilde{a} \) is a good type and retain him. Accordingly, the good type can gain from exerting \( \tilde{a} \) instead of \( a_1 < \tilde{a} \) as required by equilibrium. Together these conditions imply that a pooling equilibrium exists if and only if \( s \geq \tilde{a} \).

**Remark 2** There exists a pooling equilibrium in which both types exert a positive level of effort in the first period and are retained if and only if \( s \geq \tilde{a} \).
In my analysis of how civil service protections influence bureaucratic performance, I focus on separating equilibria. The separating equilibria are more robust than the pooling equilibria in the following sense. In the separating equilibria, bad types behave the same in their first period of employment regardless of the period in which they are hired. In a pooling equilibrium, newly hired bad types act in a drastically different manner depending on the period in which they are hired. New bad types exert at least $\tilde{a}$ if they are hired in the first period but no effort if they are hired in the second period. If a manager understands that bad types exert effort imitating good types in order to be retained in the first period, it is plausible—but inconsistent with equilibrium—that she believes a newly hired bureaucrat will face the same incentives to work for a period before slacking off after retention. In such a scenario it is better to bring in a new bureaucrat knowing that he will either be a good type or a bad type who works hard to mimic the good type. In a Supplemental Appendix, I show that this logic undermines the existence of a pooling equilibrium in a repeated two-period version of the model.\footnote{In this version the finite sequence of moves is repeated with an infinitely lived manager and a sequence of agents who live at most two periods.}

In the extended model, there exist separating equilibria in which bad types and good types choose the same equilibrium actions as in the finite model under analogous conditions to those in the finite model.

5 Civil Service Reform

Civil service protections affect whether the separating equilibrium is exit or stay by determining the level of effort, $s$, that the bad type is willing to exert to ensure retention. Civil service reform therefore has the capacity to change the type of equilibrium that prevails and thus discontinuously affect bureaucratic performance.

Lemma 1 A unique $\hat{\rho}$ exists such that for all $\rho < \hat{\rho}$, the unique separating equilibrium is exit and for all $\rho > \hat{\rho}$ the unique separating equilibrium is stay.
The effect of civil service protections on the cost that the good type must pay to separate drives Lemma 1. With weak civil service protections, a bad type can be retained only if he imitates the good type to make the manager believe he is a good type. As civil service protections become more robust, the connection between the manager’s beliefs and the employment status of the bad type becomes weaker. A bureaucrat who is found to be a bad type may keep his job despite the wishes of the manager. Therefore the willingness of the bad type to imitate the good type is decreasing in the robustness of civil service protections. The amount of effort the good type must exert in the first period to distinguish himself from the bad type therefore also decreases as civil service protections are expanded. Lemma 1 identifies a unique level of protections such that for any weaker civil service regime, the cost of separation is higher than that which the good type is willing to pay. Whether or not the good type stays or exists is a significant determinant of bureaucratic performance.

**Proposition 3** The manager’s welfare is always greater in a stay equilibrium than in an exit equilibrium. Welfare is maximized in an exit equilibrium at \( \rho = 0 \) and maximized in a stay equilibrium at \( \rho = \hat{\rho} \). The manager’s welfare is strictly decreasing in \( \rho \) in each type of equilibrium. The manager’s welfare is the same in an exit equilibrium for \( \rho = 0 \) as it is in a stay equilibrium for \( \rho = 1 \).

Figure 1 illustrates Proposition 3. For all civil service protection regimes weaker than \( \hat{\rho} \), the unique equilibrium is exit. For all regimes stronger than \( \hat{\rho} \), the unique equilibrium is stay. Bureaucratic performance is always greater for \( \rho > \hat{\rho} \) than \( \rho < \hat{\rho} \). On either side of the discontinuity in performance, however, stronger civil service protections lower bureaucratic performance. In either type of equilibrium, a reduction in civil service protections raises the probability that bad types can be successfully removed. This leads to an improvement in bureaucratic performance.

In a stay equilibrium, a reduction in civil service protections has an additional positive effect on agency performance. With weaker protections, good types must work harder to distinguish themselves from bad types. Therefore as long as the amount of additional effort
that he must exert to separate does not exceed \( z \), weaker protections yield higher expected agency output than stronger protections.

Figure 1: Civil service protections and bureaucratic performance

![Figure 1: Civil service protections and bureaucratic performance](image)

The effect of civil service reform, therefore, depends on the type of equilibrium that initially prevails and the extent of the reform. If the bureaucracy is initially unable to retain good types, a reduction in civil service protections does nothing to encourage good types to stay. The reform does, however, make it easier for bad types to be removed from the agency which enhances agency performance continuously and monotonically. An increase in civil service protections, on the other hand, has a continuous and negative effect on agency output for precisely the same reason. A sufficiently large increase in civil service protections, however, induces a transition to stay equilibrium which results in a discontinuous and strictly positive overall rise in agency performance.
If the equilibrium is initially stay, a reduction in civil service protections makes bad types easier to fire and good types work harder to separate. Both effects result in a continuous rise in bureaucratic performance. However, if the reduction in protections is too severe, too much pressure is placed on good types to distinguish themselves from the bad types. In this case they self-select out of the bureaucracy. This results in an overall decline in bureaucratic performance.

Reformers are typically aware of and cite the direct welfare effects associated with the capacity of managers to remove unmotivated bureaucrats from an agency. Calls for reform generally follow from the conclusion that a particular agency is occupied by too many bad agents. The appeals of reformers are typically accompanied by reports that surround the difficulties associated with the removal of bad public servants. Although reduced civil service protections make it easier to fire unmotivated bureaucrats and therefore remove a source of inefficiency, the results of the model reveal that it is not obvious that such a policy is optimal. The optimal reform depends on the nature of the actual problem that reformers see manifest as an agency filled with slackers who seemingly cannot be removed. Does the agency appear to be filled with protected bad types because good types self-select out of government service? Or do good types self-select into government but happen be overshadowed by a focus on bad types? If the latter case prevails, small reductions in civil service protections allow bad types to be removed without placing too much additional pressure on good types to stand out. Major reforms, however, can push good types out of the bureaucracy as the pressure to distinguish oneself from one’s unmotivated peers in an environment with few protections becomes too burdensome for good types to bear. If the former problem prevails then reformers are better off providing greater protections to bureaucrats. This reform allows good types to carry out work they care about in government free of the intense pressure to perform that prevailed under a weaker civil service regime. Good types who were initially turned off of government service because of excessive demands to distinguish themselves are attracted to government by stronger civil service protections.
It is not necessarily always the case that radical reductions in civil service protections drive good types out of the agency. If \( \hat{\rho} = 0 \), good types are willing to participate in government even without civil service protections. When this happens, at-will employment is optimal and bureaucratic performance is continuously decreasing in \( \rho \). This is only a special case however. Whenever \( \hat{\rho} > 0 \), at-will employment is a suboptimal civil service arrangement.

Figure 1 illustrates that for \( \hat{\rho} > 0 \), bureaucratic performance is equivalent under at-will employment (\( \rho = 0 \)) and a system of lifetime appointment in which bureaucrats can never be fired (\( \rho = 1 \)). Under an at-will employment policy, the unique equilibrium is exit. Each type serves only one period after being hired. Good types choose \( \tilde{a} \) and exit after one period. Bad types are identified after one period and successfully fired after they produce no policy output. Therefore in each period the good type’s expected payoff is \( \lambda v(\tilde{a}) \). Under lifetime appointment, the unique equilibrium is separating. Both types serve two periods. Because they cannot be fired, both types choose their preferred level of effort in both periods. Ex ante, the manager’s expected payoff in each period is \( \lambda v(\tilde{a}) \). Neither policy extreme, therefore, outperforms the other.

**Proposition 4** \( \hat{\rho} \) is weakly increasing in \( R_B \) and weakly decreasing in \( R_G \).

Figure 1 illustrates the comparative static results of Proposition 4. The negative effect of \( R_G \) on \( \hat{\rho} \) is driven by the good type’s willingness to participate in government. Relative government wage acts on the good type’s participation constraint. As the wedge between the non-policy rewards to office for the good type and his reservation wage increases, government employment becomes more attractive to the good type. Therefore for some \( \rho < \hat{\rho} \) where the good type initially chooses to not participate, a higher relative government wage induces the good type to participate. For any of these \( \rho \), this rise in rewards to office and subsequent transition from exit to separating equilibrium leads to an increase in bureaucratic performance.
Conversely, a rise in $R_B$ makes the good type less willing to choose government. As $R_B$ rises, the bad type becomes willing to exert a greater amount of effort to imitate the good type and secure the higher $R_B$ in the second period. The good type must therefore exert greater effort to distinguish himself from the bad type as $R_B$ rises. This lowers his payoff in a separating equilibrium. Stronger civil service protections are necessary to offset this increased competition for second-period employment with bad types.

Proposition 5 rationalizes a robust finding in the empirical literature that bureaucrats employed at-will display lower morale and job satisfaction than bureaucrats employed in a traditional civil service system. While both types of bureaucrats are negatively affected by cuts to civil service protections, the mechanism by which their payoff declines is different. In equilibrium, the manager always tries to fire the bad type and never tries to fire the good type. Therefore only the bad type’s utility is directly affected by job protections. His expected payoff rises strictly and continuously as his probability of keeping his job after the manager attempts to terminate him rises.\footnote{Formally, his equilibrium expected utility as a function of $\rho$ is $EU^b(\rho) = \rho R_B$.}

**Proposition 5** *Equilibrium expected utility for both types of bureaucrats is increasing in $\rho$.*

The good type, on the other hand, is only indirectly affected by civil service protections through the bad type’s willingness to imitate him. Figure 2 illustrates the relationship between the good type’s payoff and $\rho$. For $\rho < \hat{\rho}$, the amount of effort the bad type is willing to exert to imitate the good type is sufficiently high such that the good type is unwilling to exert a high enough level of effort to distinguish himself from the bad type. He therefore exits government and earns a constant payoff of $W$ for all $\rho \in [0, \hat{\rho}]$. As $\rho$ rises, the amount of effort required to separate falls as the bad type becomes less willing to imitate him. For $\rho > \hat{\rho}$, the cost of separation to the good type is low enough to make the payoff from participation in government greater than $W$. As $\rho$ continues to rise, the amount of effort the good type exerts in a stay equilibrium falls until $\bar{\rho} \equiv 1 - \frac{c(a)}{R_B}$ after which the bad type
becomes unwilling to exert any effort greater than \( \tilde{a} \) to imitate the good type. For \( \rho \geq \tilde{\rho} \), the good type earns a constant payoff of \( 2W + RG \).

Figure 2: Good type’s equilibrium payoff

6 Discussion

The indirect effect of civil service protections on the good type’s utility from participation is the key mechanism that drives the results of the model. His decision to participate and the effort he exerts when he participates depends on the bad type’s willingness to imitate him. Informal theories of civil service protections in the literature similarly recognize the ability of robust civil service protections to induce motivated bureaucrats to seek government employment. These theories, however, focus on a direct mechanism that does not operate
in the model. In these alternative accounts, civil service protections offer desirable job candidates more stable employment which raises the expected wage of government employees relative to the private sector (Bowman and West 2006; Battaglio 2010; Moe 2011). Civil service protections can therefore make government jobs with modest salaries more attractive than at-will jobs in the private sector that pay a higher nominal wage. In the model, however, good types are never fired in equilibrium. Whenever a good type chooses to participate, a rise in \( \rho \) has no direct impact on his earnings. The good type’s probability of being retained is always one. It is only through the indirect mechanism that government employment becomes more attractive to the good type as \( \rho \) rises. Somewhat counterintuitively, the good type benefits from greater civil protections because the bad type’s job becomes more secure. As stronger civil service protections extend greater job security to the bad type, the bad type becomes less willing to imitate the good type which reduces the amount of additional effort the good type must exert to distinguish himself.

A consideration of the effects of an alternative reform designed to entice good types to choose careers in government—an increase in government wages—emphasizes the significance of this indirect mechanism. Intuitively, greater pecuniary income should monotonically close the gap between the good type’s reservation wage and his payoff from government employment. A simple rise in government wages such that the good type’s participation constraint is satisfied should therefore be sufficient to induce the good type to self-select into government.

This intuition correctly identifies the direct effect of wages on the good type’s relative payoff to government employment. However, this intuition does not take into account the indirect effect that higher wages have on his participation utility via the bad type’s willingness to imitate him. Because good and bad types cannot be differentiated ex ante, both benefit from the higher wage. The higher wage makes the bad type willing to exert greater effort in the first period to imitate the good type than he is willing to exert at a lower wage. This raises the level of effort that the good type must exert to distinguish himself from the bad type. An increase in wages therefore produces simultaneous positive and negative effects on
the good type’s payoff from government employment.

The results of the model broadly urge caution upon proposals to radically reform civil service protections. At-will employment is only optimal in a special case where $\hat{\rho} = 0$. This arises only when the reservation wages of good types are sufficiently low relative to those of bad types. In all other cases, at-will employment is a sub-optimal policy. The suboptimality of at-will employment is concerning given the direction of most major civil service reform bills introduced and enacted into law over the course of the past 30 years in the United States.

The efficiency arguments that often motivate radical civil service reforms correctly identify a source of efficiency gains in expediting the removal of ineffective bureaucrats. The model shows, however, that in most cases this positive effect translates to an overall improvement in bureaucratic performance only locally. Reform arguments do not anticipate the potentially serious discontinuous negative effect that a move to at-will employment can bring about by inducing motivated bureaucrats to leave government. Similarly, reform arguments rarely consider the possibility that greater civil service protections may improve agency performance. A more prudent reform strategy should take into account the non-monotonic and discontinuous relationship between civil service protections and bureaucratic performance. This strategy involves acknowledging that a tradeoff exists between flexible personnel policy and the attractiveness of government employment to desirable candidates. The optimal civil service regime is likely to be somewhere between the extremes of at-will employment and lifetime appointment.

The results of the model also imply that it is better for reformers to err on the side of making civil service protections too robust rather than too weak. This follows from Proposition 3 illustrated in Figure 1. Any civil service protection regime stronger than the optimal regime results in greater bureaucratic performance than any regime weaker than the optimal. This result suggests that caution be applied when radical reductions in civil service protections are considered. In an uncertain environment where the optimal level is unknown
to policymakers, reforms that extend rather than retract job protections may be preferable.

It should be noted that the model provides something of a best case for the efficiency argument of reformers. The literature on civil service protection reform emphasizes the potential for at-will employment to foster patronage or political abuse (Bowman 2002; Kearney and Hays 1998). Concerns over patronage echo the rationale for instituting civil service protections in the first place to curtail the spoils system and foster the development of a professional Weberian bureaucracy. Skeptics of reform argue that at-will employment makes government agencies vulnerable to political abuse and have found mixed evidence to support this claim in U.S. states (Bowman and West 2006; Battaglio and Condrey 2009; Bowman et al. 2003). Such politicization of the bureaucracy may undermine the capacity of public sector agencies to effectively carry out their administrative tasks (Lewis 2008; Suleiman 2003; Carpenter 2001). My model brackets off issues of patronage and politicization. I assume that bureaucrats are hired by the same random process under all civil service protection regimes. Insofar as political abuse occurs more often at lower levels of protections, my model overestimates the level of bureaucratic performance expected where at-will employment policies have been enacted.

Finally, while the substantive focus of this paper is restricted to government bureaucracies, the model may apply to other settings as well. One of the key features of the model is a restricted contract space: the principal and agent cannot contract on output and all agents are offered a standard wage. This rigid personnel system is common in government agencies. Another key feature is intrinsic motivation: only one type of agent continues working after retention. Public service motivation is only one possible variety of intrinsic motivation. Employees in other industries vary in terms of their intrinsic satisfaction from their work as well. The model is broadly applicable to any other environment in which a restricted contract space makes selection on intrinsic motivation important to the performance of an organization. One potential additional application is to academics. Tenure-track assistant professors vary in terms of their intrinsic motivation for carrying out research and teaching
students. A committee observes their research and teaching production for several years then decides whether to grant tenure or not. Only the motivated types continue to produce a high level of research after receiving tenure. In this environment, civil service protections are analogous to any university or departmental institution that limits the discretion of the department to deny tenure. A higher probability of receiving tenure can attract the best candidates by easing the burden on them to demonstrate their type to their senior colleagues.

7 Conclusion

I develop this model in order to understand the effect of civil service reform on bureaucratic performance. It is premised on a selection problem that arises when bureaucrats vary in terms of the degree to which they are intrinsically motivated by their work. In order to be retained by their managers, motivated bureaucrats must distinguish themselves from unmotivated bureaucrats. If separation is too burdensome for motivated bureaucrats, they choose not to pursue careers in government. Strong civil service protections reduce the willingness of unmotivated bureaucrats to imitate motivated bureaucrats and therefore make separation less costly to motivated bureaucrats. In this way civil service protections affect the career decisions of motivated bureaucrats. Any civil service reform that affects this decision discontinuously affects the performance of a bureaucracy. At the same time, strong civil service protections negatively affect bureaucratic performance by making unmotivated bureaucrats more difficult to remove. The overall relationship between civil service protections and agency performance is therefore non-monotonic. This paper’s main contribution to scholarly and policy debates on civil service reform is to highlight both this non-monotonic relationship and the importance of intrinsically motivated bureaucrats’ career decisions for understanding and predicting the effects of specific reform policies.
8 Appendix

8.1 Strategies and Beliefs

Retained or unsuccessfully fired bureaucrats observe the manager’s decision and, if applicable, observe that their manager’s attempted firing was unsuccessful. Because only the bureaucrat’s type is private information, all of his information sets are singletons. Let \( \phi \) denote an information set for a bureaucrat hired in the first period at which he chooses a level of effort for the second period. Let \( \Phi \) denote the set of all possible \( \phi \). A pure strategy for a bureaucrat of type \( \gamma \) hired in period one is a triple that prescribes first-period effort, a career decision, and second-period effort

\[
\{a_1(\gamma, 1), h(\gamma), a_2(\gamma, 1)\} : \Gamma \times \Phi \rightarrow A \times H \times A
\]

Bureaucrats hired in the second period after a bureaucrat exits or is successfully fired observe every action taken earlier in the game except Nature’s draw of the first-period bureaucrat’s type. Each of his information sets therefore contains two histories, one in which the original bureaucrat is bad and one in which he is good. Let \( \Psi \) denote the set of all information sets at which the newly hired bureaucrat moves. For any information set, \( \psi \), at which he moves, let \( \mu(a_1, h, r) \) denote the newly hired bureaucrat’s belief that the original bureaucrat is a good type given the first bureaucrat and manager’s actions. If \( h = 0 \) and the manager does not move, denote this as \( \mu(a_1, 0, \emptyset) \). A pure strategy for a type \( \gamma \) bureaucrat hired in period two is a prescription of a level of effort

\[
a_2(\gamma, 2) : \Gamma \times \Psi \rightarrow A
\]

The manager observes \( (a_1, h) \) but not the bureaucrat’s type at any information set at which she makes her retention decision. Note that for any such information set to be reached, the first-period bureaucrat must have played \( h = 1 \). Any information set at which she moves
contains two histories, one in which he is good and one in which he is bad. An information set for the manager can be represented by \((a_1, h) \in A \times H\). Note that for any such information set to be reached, the first-period bureaucrat must have played \(h = 1\). Accordingly, the manager’s strategy and beliefs can be represented as a mapping from \(A\). A pure strategy for the manager is a retention decision

\[ r : A \to R \]

For any \(a_1\), \(\bar{\lambda}(a_1)\) denotes the manager’s belief that the bureaucrat is a good type.

### 8.2 Definitions of Equilibria

Let \(\{a_1^*(\gamma, 1), h_1^*(\gamma), a_2^*(\gamma, 1)\}\) denote an equilibrium strategy for a bureaucrat of type \(\gamma\) hired in period one, \(a_2^*(\gamma, 2)\) an equilibrium strategy for a bureaucrat of type \(\gamma\) hired in period two, and \(r^*\) an equilibrium strategy for the manager.

**Definition 1 (Exit Equilibrium)** In an exit equilibrium,

1. \(\{a_1^*(g, 1), h_1^*(g), a_2^*(g, 1)\} = \{\bar{a}, 0, \bar{a}\} \text{ for all } \phi\)

2. \(\{a_1^*(b, 1), h_1^*(b), a_2^*(b, 1)\} = \{0, 1, 0\} \text{ for all } \phi\)

3. \(\hat{a}_2^b = \bar{a} \text{ for all } \psi\)

4. \(\hat{a}_2^b = 0 \text{ for all } \psi\)

5. \(r^* = 0 \text{ for all } a_1\)

6. \(\bar{\lambda}(0) = 0\)

7. \(\bar{\lambda}(a_1) = 0 \text{ for all } a_1 \neq 0\)

8. \(\mu(0, 1, 0) = 0\)

9. \(\mu(\bar{a}, 0, \emptyset) = 1\)
10. $\mu(a_1, h, r) = 0$ for all $(a_1, h, r) \notin \{(\tilde{a}, 0, 0), (0, 1, 0)\}$

Definition 2 (Stay Equilibrium) In a stay equilibrium,

1. $\{a_1^*(g, 1), h^*(g), a_2^*(g, 1)\} = \{\tilde{a}, 1, \tilde{a}\}$ for all $\phi$
2. $\{a_1^*(b, 1), h^*(b), a_2^*(b, 1)\} = \{0, 1, 0\}$ for all $\phi$
3. $\hat{a}_2^b = \tilde{a}$ for all $\psi$
4. $\hat{a}_2^b = 0$ for all $\psi$
5. $r^* = \begin{cases} 0 & \text{for all } a_1 < \tilde{a} \\ 1 & \text{for all } a_1 \geq \tilde{a} \end{cases}$
6. $\tilde{\lambda}(a_1) = 0$ for all $a_1 < \tilde{a}$
7. $\tilde{\lambda}(a_1) = 1$ for all $a_1 \geq \tilde{a}$
8. $\mu(\tilde{a}, 1, 1) = 1$
9. $\mu(0, 1, 0) = 0$
10. $\mu(a_1, h, r) = \lambda(a_1)$ for all $(a_1, h, r) \notin \{(\tilde{a}, 1, 1), (0, 1, 0)\}$

Definition 3 (Pooling Equilibrium) In a pooling equilibrium,

1. $\{a_1^*(g, 1), h^*(g), a_2^*(g, 1)\} = \{\tilde{a}, 1, \tilde{a}\}$ for all $\phi$
2. $\{a_1^*(b, 1), h^*(b), a_2^*(b, 1)\} = \{\tilde{a}, 1, 0\}$ for all $\phi$
3. $\hat{a}_2^b = \tilde{a}$ for all $\psi$
4. $\hat{a}_2^b = 0$ for all $\psi$
5. \( r^* = \begin{cases} 
1 & \text{for } a_1 = \bar{a} \\
0 & \text{otherwise} 
\end{cases} \)

6. \( \bar{\lambda}(\bar{a}) = \lambda \)

7. \( \bar{\lambda}(a_1) = 0 \) for all \( a_1 \neq \bar{a} \)

8. \( \mu(\bar{a}, 1, 1) = \lambda \)

9. \( \mu(a_1, h, r) = \lambda(a_1) \) for all \( (a_1, h, r) \neq (\bar{a}, 1, 1) \)

### 8.3 Proofs

**Proof of Proposition 1**

I first show that \( z \leq s \) is a sufficient for the existence of an exit equilibrium. Assume \( z \leq s \).

It is straightforward to check that \( \hat{a}^g_2 \) and \( \hat{a}^b_2 \) are optimal for all \( \psi \). A newly hired good type’s payoff at any information set \( \psi \) is \( v(a_2) - c(a_2) \) where \( a_2 \) is his choice of effort. This payoff independent is his belief about the first-period bureaucrat’s type. The unique level of effort that optimizes his payoff is \( \bar{a} \). Similarly, a newly hired bad type’s payoff at all \( \psi \) is \( -c(a_2) \) which is uniquely maximized at \( a_2 = 0 \). Because payoffs are independent of beliefs for both types, this strategy is optimal given the beliefs specified in equilibrium. I show below that the off-path beliefs specified are consistent with the intuitive criterion.

The optimality of an incumbent bureaucrat’s second-period strategy is established similarly. The mapping from effort to payoffs is identical at all information sets \( \phi \). Therefore it is optimal for the good type to exert \( \bar{a} \) for all \( \phi \) and the bad type to exert no effort for all \( \phi \).

Given the manager’s strategy, \( a_1^*(b, 1) = 0 \) and \( h^*(b) = 1 \) are optimal. If the bad type exits, he is worse off by at least \( \rho R_B \) compared to equilibrium. If the bad type stays and chooses a positive level of effort he is fired. This additional effort reduces his expected payoff compared to equilibrium in which he exerts no effort and is fired.
To establish that $a_1^*(g, 1) = \tilde{a}$ and $h^*(g) = 0$ are optimal given the manager’s strategy, first note that if the good type exerts some level of effort other than $\tilde{a}$ and exits, this strictly reduces his payoff because $\tilde{a}$ uniquely maximizes his policy payoff in a period. If instead the good type stays, he is fired regardless of the amount of effort he exerts. The highest payoff he can earn from staying is $W + \rho(W + R_G) - k$ by exerting $\tilde{a}$ and staying. By Assumption 1, this is strictly less than his equilibrium payoff, $W$.

If $\bar{\lambda}(a_1) < \lambda$, it is optimal for the manager to fire the bureaucrat given bureaucrats’ second period strategies. The manager’s strategy is therefore optimal given his beliefs. On the equilibrium path, $\bar{\lambda}(0) = 1$ by Bayes’ rule. To show that her off-path beliefs survive the intuitive criterion, note that the equilibrium payoff to a good type hired in the first period is $W$. A bad type hired in the first period earns an expected payoff of $\rho R_B$. From the definition of $s$, if the bad type exerts $a_1 \geq s$, his payoff cannot exceed 0. Similarly, from the definition of $z$, if the bad type exerts $a_1 \geq z$, his payoff cannot exceed $W$. Because $s \geq z$, neither type can strictly benefit from exerting $a_1 \geq s$. Therefore all beliefs $\lambda(a_1)$ survive the intuitive criterion for $a_1 \geq s$.

Recall that $z > \tilde{a}$ is defined such that $v(z) - c(z) + R_G = 0$. Because of the concavity of $v(z) - c(z)$, there may exist another positive level of effort $z' < z$ such that $v(z') - c(z') + R_G = 0$. This is a consequence of the intrinsic motivation of the good type. Good types lose policy payoff from exerting less effort as well as more effort than they would prefer. At $z' < \tilde{a}$, the good type is indifferent between exerting $\tilde{a}$ and exiting and exerting a low level of effort and being retained.

First consider the case in which $z'$ exists. Any $a_1 \in (z', z)$ and $h = 1$ equilibrium dominates the good type’s equilibrium actions. This follows immediately from the definition of $z'$ and $z$. Because $s \geq z$, any $a_1 < z$ and $h = 1$ also equilibrium dominates the bad type’s equilibrium actions from the definition of $s$. Because both types strictly benefit from exerting $a_1 \in (z', z)$ and staying, $\lambda(a_1)$ is free at all such $a_1$ under the intuitive criterion. For $a_1 \leq z'$, the good type is at least as worse off than he is in equilibrium if he exerts $a_1$ and
stays in the bureaucracy and is retained. The bad type is strictly better off. Therefore the
intuitive criterion rules out all beliefs other than $\lambda(a_1) = 0$. Now consider the case in which
$z'$ does not exist. That is, the good type prefers to exert $a_1 = 0$ and be retained rather than
exert $\tilde{a}$ and exit. In this case the same argument just applied to $a_1 \in (z', z)$ applies to the
interval $(0, z)$. The intuitive criterion does not rule out any beliefs on this interval.

Finally, consider the interval $[z, s)$. By the definitions of $s$, any level of effort in this
interval strictly improves the bad type’s payoff relative to equilibrium if he is retained. By
the definition of $z$, no level of effort greater than or equal to $z$ can improve the good type’s
payoff relative to equilibrium if he is retained. The intuitive criterion therefore rules out all
beliefs other than $\lambda(a_1) = 0$ on this interval.

A bureaucrat hired in the second period’s beliefs are consistent with Bayes’ rule on the
equilibrium path: $\mu(0,1,0) = 0$ and $\mu(\tilde{a},0,\emptyset) = 1$. Off path, if the first-period bureaucrat
plays an unexpected action, the new bureaucrat’s beliefs are consistent with the manager’s:
$\mu(a_1,h,r) = 0$. This belief survives the intuitive criterion for the same reason the manager’s
belief does. If the manager deviates from equilibrium and makes an unexpected retention
decision given first-period bureaucrat actions, the new bureaucrat’s beliefs are consistent with
the manager’s: $\mu(a_1,h,r) = \lambda(a_1)$. Because the first-period bureaucrat plays an equilibrium
action, this belief survives the intuitive criterion.

I now show that $s \geq z$ is necessary for the existence of an exit equilibrium. Assume
$s < z$. In equilibrium the good type’s payoff is

$$W = v(z) - c(z) + W + R_G < v(z - \epsilon) + c(z - \epsilon) + W + R_G$$

where the equality follows from the definition of $z$ and $\epsilon$ is arbitrarily small. The inequality
implies that if the good type exerts $z - \epsilon$, stays, and is retained, he is strictly better off than
in equilibrium. The bad type’s equilibrium payoff is

$$\rho R_B = -c(s) + R_B$$
Any level of effort $a_1 > s$ makes the bad type strictly worse off than equilibrium even if the bad type is retained. Because $z - \epsilon > s$, the intuitive criterion therefore rules out all beliefs other than $\bar{\lambda}(z - \epsilon) = 1$. The exit equilibrium therefore does not survive the intuitive criterion if $s < z$. □

Proof of Proposition 2

I first show that $s \leq z$ is a necessary condition for a stay equilibrium to exist. Let $s > z$. Note that because $\tilde{a} \leq z$, $s > z$ implies that $\hat{a} = s$. By the definition of $z$, the good type is strictly better off exerting $\tilde{a}$ and exiting than playing his equilibrium strategy. Therefore $s > z$ implies that a stay equilibrium does not exist.

I now show that $s \leq z$ is a sufficient condition for a stay equilibrium to exist.

An identical argument to that in the proof of Proposition 2 establishes that strategies for bureaucrats in the final period are optimal. To establish that the bad type’s strategy in the first period is optimal given the strategies of the other players, note first that exiting the bureaucracy makes the bad type strictly worse off by $\rho R_B$. If he stays and chooses a positive level of effort $a_1 < \hat{a}$, the manager fires him. This makes him strictly worse off than in equilibrium where he is fired but does not suffer any cost from exerting effort. If he exerts $a_1 \geq \hat{a}$, the manager retains him. However, because $\hat{a} \geq s$, this cannot strictly benefit him relative to equilibrium from the definition of $s$.

Because $s \leq z$, the amount of effort the good type exerts in equilibrium, $\tilde{a}$, does not exceed $z$. His equilibrium payoff therefore weakly exceeds $W$. The maximum payoff he can earn by exiting is $W$ so he cannot strictly gain by exiting. Now consider a deviation from equilibrium in which the good type stays but exerts $a_1 \neq \tilde{a}$. If he exerts $a_1 > \hat{a}$, the manager retains him but he is strictly worse off compared to equilibrium because $v(a_1) - c(a_1) < v(a_1') - c(a_1')$ for all $a_1' \geq \tilde{a}$ and $a_1 > a_1'$. If he exerts $a_1 < \hat{a}$, he is fired. Under Assumption 1, this yields a payoff of less than $W$. The good type’s strategy is therefore optimal given the strategies of the other players.

If $\bar{\lambda}(a_1) < \lambda$, it is optimal for the manager to fire the bureaucrat given bureaucrats’
second period strategies. The manager’s strategy is therefore optimal given his beliefs. On
the equilibrium path, \( \bar{\lambda}(\hat{a}) = 1 \) and \( \bar{\lambda}(0) = 0 \) by Bayes’ rule. To establish that off-path
beliefs are consistent with the intuitive criterion, consider an off-path level of effort \( a_1 \geq \hat{a} \).
Because \( \hat{a} \geq s \), the bad type is strictly worse off exerting any level of effort greater than \( \hat{a} \).
Because \( \hat{a} \leq \hat{a} \), the good type is also strictly worse off exerting more than \( \hat{a} \). Therefore the
intuitive criterion imposes no restriction on \( \bar{\lambda}(a_1) \) for \( a_1 \geq \hat{a} \). Now consider \( a_1 < \hat{a} \). If \( \hat{a} = \hat{a} \),
then \( a_1 < \hat{a} \) makes the good type strictly worse off. The bad type is no better off for any
\( a_1 \in [s, \hat{a}) \). Therefore the intuitive criterion imposes no restriction on \( \bar{\lambda}(a_1) \) for \( a_1 \in [s, \hat{a}) \)
if \( \hat{a} = \hat{a} \). For \( a_1 < s \), the bad type is strictly better off than in equilibrium. The intuitive
criterion therefore requires that \( \bar{\lambda}(a_1) = 0 \) for \( a_1 < s \) if \( \hat{a} = \hat{a} \). If \( \hat{a} = s \), then the bad type
is strictly better off for all \( a_1 < s \). Define \( s' < \hat{a} < s \) such that \( v(s') - c(s') = v(s) - c(s) \).
That is, \( s' \) is the level of effort less than \( \hat{a} \) at which the good type receives the same policy
utility that he does if he exerts \( s \). The good type is strictly better off if he exerts \( a_1 \in (s', s) \).
The intuitive criterion therefore places on restriction on \( \bar{\lambda}(a_1) \) for \( a_1 \in (s', \hat{a}) \) if \( \hat{a} = s \). For
\( a_1 \leq s' \), the good type in no better off compared to his equilibrium payoff. Therefore under
the intuitive criterion \( \bar{\lambda}(a_1) = 0 \) for \( a_1 \leq s' \) if \( \hat{a} = s \).

A bureaucrat hired in the second period’s beliefs are consistent with Bayes’ rule on the
equilibrium path: \( \mu(0, 1, 0) = 0 \) and \( \mu(\hat{a}, 0, 1) = 1 \). Off path, if the first-period bureaucrat
plays an unexpected action, the new bureaucrat’s beliefs are consistent with the manager’s:
\( \mu(a_1, h, r) = \lambda(a_1) \). This belief survives the intuitive criterion for the same reason the
manager’s belief does. If the manager deviates from equilibrium and makes an unexpected
retention decision given first-period bureaucrat actions, the new bureaucrat’s beliefs are
consistent with the manager’s: \( \mu(a_1, h, r) = \lambda(a_1) \). Because the first-period bureaucrat plays
an equilibrium action, this belief survives the intuitive criterion. □

**Proof of Remark 1**

Recall that I restrict attention equilibria in which the bad type stays. To rule out all
other types of separating equilibrium distinguished by first-period strategies, I show that all
combinations of strategies \((a^*_1(g,1), h^*(g), a^*_1(b,1)) \in A \times H \times A\) other than \((\hat{a}, 1, 0)\) (stay) and \((\tilde{a}, 0, 0)\) (exit) are inconsistent with equilibrium.

I first show that there is no separating equilibrium in which the bad type exerts a positive level of effort, \(a^*_1(b,1) > 0\). By Bayes’ rule, on the equilibrium path \(\bar{\lambda}(a^*_1(b,1)) = 0\). In order for the manager’s strategy to be optimal, she must fire the bureaucrat. The bad type’s payoff in equilibrium is therefore \(-c(a^*_1(b,1)) + \rho R_B\). If the bad type instead exerts no effort, the worst payoff he can do off-path is \(\rho R_B\). Therefore there is no separating equilibrium in which the bad type chooses a positive level of effort.

It is straightforward to check that there are no equilibria in which the good type exits and chooses a level of effort other than \(\tilde{a}\). If he exerts \(\tilde{a}\) and leaves, his payoff is \(W\). If he exerts \(a_1 \neq \tilde{a}\), his payoff is \(v(a_1) - c(a_1) < W\) by the optimality of \(\tilde{a}\). Therefore in no equilibrium can he choose \(a_1 \neq \tilde{a}\) and exit.

Now I show that there is no separating equilibrium in which he stays and exerts \(a^*_1(g,1) \neq \hat{a}\). First consider the case in which \(s < \tilde{a}\) and \(a_1(g,1)^* \neq \hat{a} = \tilde{a}\). If the good type chooses to exert \(\tilde{a}\) instead of \(a^*_1\) and is still retained, he earns a payoff of \(2W + RG\). This exceeds his equilibrium payoff of \(v(a_1(g,1)^*) - c(a_1(g,1)^*) + W + RG\). Because \(s < \tilde{a}\), only the good type is strictly better off from such a deviation from equilibrium. Therefore \(\bar{\tilde{a}} = 1\). This makes deviation to \(\tilde{a}\) strictly profitable for the good type. Now consider \(s \geq \tilde{a}\) and \(a^*_1(g,1) \neq \hat{a} = s\). If \(a^*_1(g,1) < s\), then the bad type can strictly benefit from imitating the good type rather than exerting no effort and being fired from the definition of \(s\). If \(a^*_1(g,1) > s\), then any level of effort \(a_1 \in (s, a^*_1(g,1))\) makes the bad type worse better off even if he retained following such a deviation. Because \(a^*_1(g,1) > \hat{a}\), the good type is strictly better off if he is retained following such a deviation. Under the intuitive criterion, sufficiently punishing beliefs to deter this deviation by the good type cannot be sustained. Therefore in any equilibrium in which the good type stays, he must exert \(\hat{a}\) in the first period. □

**Proof of Remark 2**

I first show that \(s \geq \hat{a}\) is a necessary condition for a pooling equilibrium. Let \(s < \hat{a}\).
Consider first $\bar{a} > s$. From the definition of $s$, the bad type is strictly better off exerting no effort and being fired than playing the equilibrium strategy. Now consider $\bar{a} \leq s$. Equilibrium requires that the manager fire a bureaucrat if he exerts $a_1 > \bar{a}$. This strategy is not optimal if $\bar{\lambda}(a_1) = 1$. The bad type is strictly worse off exerting more than $\bar{a}$. Because $\bar{a} < \tilde{a}$, the good type is strictly better off exerting $\tilde{a}$ if doing so still results in retention. Therefore under the intuitive criterion, all beliefs other than $\bar{\lambda}(\tilde{a}) = 1$ are ruled out. A pooling equilibrium therefore does not survive the intuitive criterion if $s < \tilde{a}$.

I now show that $s \geq \bar{a}$ is a sufficient condition for a pooling equilibrium to exist. I establish that bureaucrat strategies in the second period are optimal in the proof of Proposition 1. To establish that first-period actions are optimal, consider $\bar{a} \in [\tilde{a}, \min\{s, z\}]$. Because $z \geq \bar{a}$ by the definition of $z$ and $\bar{a} \leq s$ by assumption, this interval is non-empty. Given the manager’s strategy, if a bureaucrat exerts some level of effort less than $\bar{a}$, he is fired. By the definition of $s$, the bad type can do no better than to play his equilibrium strategy given the manager’s strategy. Under Assumption 1 and the definition of $z$, the good type can do no better than to play his equilibrium strategy.

If $\bar{\lambda}(a_1) < \lambda$, it is optimal for the manager to fire the bureaucrat given bureaucrats’ second period strategies. If $\bar{\lambda}(a_1) \geq \lambda$, it is optimal for the manager to retain the bureaucrat. The manager’s strategy is therefore optimal given and his beliefs. On the equilibrium path, $\bar{\lambda}(\tilde{a}) = \lambda$ by Bayes’ rule. Off path, the bad type strictly gains from any $a_1 < \bar{a}$ if he is retained. Therefore $\bar{\lambda}(a_1) = 0$ is not ruled out for all $a_1 < \bar{a}$. For $a_1 > \bar{a}$, the good type can be made no better off because $\bar{a} \geq \tilde{a}$. Therefore $\bar{\lambda}(a_1)$ is consistent with the intuitive criterion for all $a_1 > \bar{a}$.

A bureaucrat hired in the second period’s beliefs are consistent with Bayes’ rule on the equilibrium path: $\mu(\bar{a}, 1, 1) = \lambda$. Off path, if the first-period bureaucrat plays an unexpected action, the new bureaucrat’s beliefs are consistent with the manager’s: $\mu(a_1, h, r) = \lambda(a_1)$. This belief survives the intuitive criterion for the same reason the manager’s belief does. If the manager deviates from equilibrium and makes an unexpected retention decision given first-
period bureaucrat actions, the new bureaucrat’s beliefs are consistent with the manager’s: 
\[ \mu(a_1, h, r) = \lambda. \]
Because the first-period bureaucrat plays an equilibrium action, this belief survives the intuitive criterion. \(\square\)

**Proof of Lemma 1**

From Remark 1, the unique separating equilibrium is exit if \( z < s \) and stay if \( z > s \). From the definition of \( z \) as the largest solution to \( v(a_1) - c(a_1) + R_G = 0 \), \( z \) is constant in \( \rho \). Recall that \( s \equiv c^{-1}((1 - \rho)R_B) \). Because \( c(\cdot) \) is strictly increasing and continuous, its inverse is also strictly increasing and continuous. Moreover, \( s \) is unique. Let \( s(\rho) \) denote the value of \( s \) as a function of \( \rho \). The properties of \( c(\cdot) \), the continuity of the term \((1 - \rho)R_B\) in \( \rho \), and Assumption 1 part iii \((R_B > 0)\) imply that \( s(\rho) \) is continuously and strictly decreasing in \( \rho \). Additionally, because \( c(0) = 0 \), \( s(1) = 0 \).

There are now two cases to consider. If \( s(0) > z \), then at \( \rho = 0 \) the unique separating equilibrium is exit. Because \( s(1) = 0 \) and \( z > 0 \), the continuity and strict monotonicity of \( s(\rho) \) implies that there exists a unique \( \hat{\rho} \) such that \( s(\hat{\rho}) = z \). For all \( \rho > \hat{\rho} \), \( s(\rho) > z \) and the unique separating equilibrium is exit. For all \( \rho < \hat{\rho} \), \( s(\rho) < z \) and the unique separating equilibrium is stay. Now consider the case in which \( s(0) \leq z \). In this case \( s(\rho) < z \) for all \( \rho > 0 \). The unique separating equilibrium is stay for all \( \rho > 0 \). In this case \( \hat{\rho} = 0 \). \(\square\)

**Proof of Proposition 3**

The manager’s ex ante expected utility in an exit equilibrium is

\[ \lambda(v(\tilde{a}) + \lambda v(\tilde{a})) + (1 - \lambda)(1 - \rho)\lambda v(\tilde{a}) \]

Her ex ante expected utility in a separating equilibrium is

\[ \lambda(v(\hat{a}) + v(\tilde{a})) + (1 - \lambda)(1 - \rho)\lambda v(\tilde{a}) \]

Because \((1 - \lambda)\lambda v(\tilde{a}) > 0\), both expressions are strictly decreasing in \( \rho \). The manager’s
utility is greater in a separating equilibrium if and only if

\[ v(\hat{a}) + v(\tilde{a}) \geq v(\hat{a} + \lambda v(\tilde{a})) \]

or

\[ v(\hat{a}) \geq \lambda v(\tilde{a}) \]

Because \( \hat{a} \geq \tilde{a} \) and \( \lambda \in (0, 1) \), the inequality is strictly satisfied. Note that for \( \rho = 0 \), the manager’s utility in an exit equilibrium is \( 2\lambda v(\tilde{a}) \). If \( \rho = 1, s = 0 < \tilde{a} \). Therefore if \( \rho = 1, \hat{a} = \tilde{a} \). It follows that for \( \rho = 1 \), the manager’s utility in a stay equilibrium is \( 2\lambda v(\tilde{a}) \). □

**Proof of Proposition 4**

Recall from the proof of Lemma 1 that for \( s(0) \geq z \), \( \hat{\rho} \) solves \( s(\rho) = z \) where \( z \) solves \( v(a_1) - c(a_1) + R_G = 0 \) for \( a_1 \) and \( s(\rho) = -c^{-1}((1 - \rho)R_B) \). A rise in \( R_G \) yields a rise in \( z \) but no change in \( s(\rho) \). Because \( s(\cdot) \) is increasing, \( \hat{\rho} \) must fall to satisfy \( s(\hat{\rho}) = z \). A rise in \( R_B \) yields a rise in \( s(\rho) \) but does not change \( z \). To satisfy \( s(\hat{\rho}) = z \), \( \hat{\rho} \) must rise.

For \( s(0) < z \), \( \hat{\rho} = 0 \). Note that \( s(0) = c^{-1}(R_B) \) is increasing in \( R_B \). For all \( R_B \) such that \( s(0) \) remains less than \( z \), there is no change in \( \hat{\rho} \). If \( R_B \) rises such that \( s(0) \geq z \), \( \hat{\rho} \) solves \( s(\rho) = z \) and the proof for this case applies. Any further rise in \( R_B \) yields a rise in \( \hat{\rho} \).

If \( R_G \) rises, \( s(0) \) remains less than \( z \) and there is no change in \( \hat{\rho} \). If \( R_G \) falls such that \( s(0) \) remains less than \( z \), \( \hat{\rho} \) remains zero. If \( R_G \) falls such that \( s(0) \geq z \), \( \hat{\rho} \) solves \( s(\rho) = z \) and the proof for this case applies. Any further fall in \( R_G \) yields a rise in \( \hat{\rho} \). □

**Proof of Proposition 5**

From Propositions 1 and 2, the bad type’s expected payoff in any separating equilibrium is \( \rho R_B \) which is strictly increasing by Assumption 1 \((R_B > 0)\). The good type’s payoff in an exit equilibrium is \( W \) from Proposition 2. From Lemma 1, the unique separating equilibrium for all \( \rho < \hat{\rho} \) is exit. For \( \rho = \hat{\rho} \), both types of equilibria exist. In the stay equilibrium, \( \hat{a} = s \). From the definition of \( \hat{\rho} \), at \( \hat{\rho} \), \( z = s \) which implies that the good type’s payoff in a stay equilibrium is \( v(z) - c(z) + W + R_G = W \). For \( \rho > \hat{\rho} \), the unique equilibrium is stay. In a stay
equilibrium, the good type’s expected payoff is $v(\hat{a}) - c(\hat{a}) + W + R_G$. From the definition of $\hat{\rho}$, $\hat{a} = s$ for all $\rho \leq \hat{\rho}$. It is established in the Proof of Lemma 1 that $s$ is strictly decreasing in $\rho$. Therefore his payoff is increasing in $\rho$ on $[\hat{\rho}, \tilde{\rho}]$. At $\tilde{\rho}$, $s = \tilde{a}$. In a stay equilibrium, $\hat{a} = \tilde{a}$ for all $\rho \geq \tilde{\rho}$. Therefore his payoff is constant in $\rho$ for $\rho > \tilde{\rho}$. □

9 Supplemental Appendix

This Supplemental Appendix describes a repeated two-period version of the finite model presented in the main text. I show that equilibria analogous to exit and separating equilibria exist under analogous conditions to those in the finite model. I then show that pooling equilibria analogous to those that exist in the finite version do not exist.

9.1 Setup

The actors in the model are a bureaucrat and a manager. In each of an infinite number of periods, a bureaucrat chooses a level of effort $a_t \in A = \mathbb{R}_+$. This action is observable to the manager and generates a policy payoff for her. After a bureaucrat’s first period of employment, the manager chooses whether to retain or fire the bureaucrat. Let $r \in R = \{0, 1\}$ denote the manager’s action space where $r = 1$ denotes retain. If the manager chooses to fire the bureaucrat, he is successfully removed with probability $1 - \rho$. Prior to the manager’s decision, the bureaucrat chooses either to stay in government or exit. Denote this action space $h \in H = \{0, 1\}$ where $h = 1$ denotes that the bureaucrat stays. A bureaucrat serves at most two consecutive periods and cannot be rehired after he exits or is fired. The manager lives forever. In each period the manager’s payoff is given by

$$u^m_t = v(a_t)$$

as in the baseline model. The manager discounts the future by $\delta$.  

40
There are two types of bureaucrats, good and bad. Let $\gamma \in \Gamma = \{g, b\}$ denote a bureaucrat’s type. With probability $\lambda \in (0, 1)$, a randomly selected bureaucrat is a good type. A bureaucrat’s type is his private information. Good types earn a payoff in employment-period $t$ of

$$u^g_t = v(a_t) - c(a_t) + R_G$$

The cost function $c(\cdot)$ is strictly increasing and convex with $c(0) = 0$. The term $R_G$ is the constant difference between the non-policy benefits of office for the good type and an outside option normalized to zero. If a good type’s manager attempts to fire him, he suffers a cost $k \geq 0$. The bad type’s payoff in employment-period $t$ is

$$u^b_t = -c(a_t) + R_B$$

The game is played with the following sequence of moves

1) Nature selects a bureaucrat for employment

2) The bureaucrat selects a level of effort and chooses to stay in government or exit. If the bureaucrat exits the game returns to (1). If the bureaucrat stays, the game proceeds to (3).

3) The manager fires or retains the bureaucrat. If she retains the bureaucrat or unsuccessfully fires him, the game proceeds to (4). If she successfully fires him, the game returns to (1).

4) The bureaucrat selects a level of effort and the game returns to (1).

9.2 Strategies, Beliefs, and Equilibrium

I define equilibrium to be a perfect Bayesian equilibrium that satisfies the intuitive criterion. The equilibria I consider are symmetric and stationary in the following sense:

1) A type $\gamma$ bureaucrat plays the same strategy as all other type $\gamma$ bureaucrats independently of the period in which he is hired.
2) The manager’s strategy is a mapping from $A$ that does not depend on the time period. Formally, a stationary strategy for a bureaucrat is a mapping

$$\{a_1^\gamma, h^\gamma, a_2^\gamma\} : \Gamma \to A \times H \times A$$

where subscripts denote employment period and $h = 1$ denotes that a bureaucrat stays and $h = 0$ denotes that he leaves. A stationary strategy for the manager is a mapping

$$r : A \to R$$

where $r = 1$ denotes that she retains the bureaucrat and $r = 0$ denotes that she attempts to fire him.

The manager’s belief about a bureaucrat’s type at any node at which she chooses to fire or retain him does not depend on the time period due to the stationarity of the bureaucrats’ strategies. In PBE, the history of the game prior to the hiring of the particular bureaucrat is irrelevant to the principal’s belief about the bureaucrat’s type. The manager’s belief about any bureaucrat who she chooses to fire or retain can therefore be represented simply as a function of his action in his first period of employment. Let $\lambda(a_1)$ denote this belief. Let $\hat{\alpha}$, $W$, $s$, $z$, and $\hat{\alpha}$ be defined the same as in the main text.

**Definition 4** In a stay equilibrium,

1. $\{a_1^q, h^q, a_2^q\} = \{\hat{\alpha}, 1, \hat{\alpha}\}$

2. $\{a_1^b, h^b, a_2^b\} = \{0, 1, 0\}$

3. $r(a_1) = \begin{cases} 1 & \text{if } a_1 \geq \hat{\alpha} \\ 0 & \text{if } a_1 < \hat{\alpha} \end{cases}$

---

7 On the equilibrium path this is a consequence of Bayes’ rule. Off path, PBE requires that beliefs at nodes further than one away from a deviation be updated according to Bayes rule. In this context, aberrant behavior from a bureaucrat who has left or been fired does not influence the principal’s belief about the current bureaucrat.
4. \( \tilde{\lambda}(a_1) = \begin{cases} 
1 & \text{if } a_1 \geq \hat{a} \\
0 & \text{if } a_1 < \hat{a} 
\end{cases} \)

In a stay equilibrium, the manager’s ex ante expected payoff when a new bureaucrat is hired is

\[
S(\rho) = \frac{\lambda v(\hat{a}) + \delta \lambda v(\tilde{a})}{1 - \lambda \delta^2 - (1 - \lambda)(1 - \rho)\delta - (1 - \lambda)\rho \delta^2}
\]

It is incentive compatible for the manager to retain the good bureaucrat if and only if

\[
\tilde{\lambda}(\hat{a})v(\tilde{a}) \geq (1 - \delta)S(\rho)
\]

For \( \hat{a} \) sufficiently large, it is not incentive compatible for the manager to retain the good type. In this case newly hired good types are so much more productive than retained good types that the principal prefers to hire a random new bureaucrat rather than retain a bureaucrat she knows is a good type. To prevent deviation by the manager from equilibrium when she observes \( \hat{a} \) and \( \tilde{\lambda}(\hat{a}) = 1, \hat{a} \) must be sufficiently low. In particular, \( \hat{a} \) must satisfy

\[
\frac{v(\tilde{a})}{v(\hat{a})} \geq \frac{\lambda}{1 + \delta \rho(1 - \lambda)}
\]

Define \( r \) such that

\[
r \equiv v^{-1}\left(\frac{v(\tilde{a})[1 + \delta \rho(1 - \lambda)]}{\lambda}\right)
\]

The manager’s incentive compatibility constraint can now be rewritten simply as

\[
\hat{a} \leq r
\]

I assume that \( r \) is sufficiently high such that \( r \geq s \) when \( \rho = 0 \). This assumption brackets off intuitively unappealing cases in which the principal fires a bureaucrat after learning that he is a good type. I also assume that Assumption 1 from the main text continues to hold.

**Definition 5** *In an exit equilibrium,*
1. \( \{a^a_1, h^g, a^g_2\} = \{\bar{a}, 0, \bar{a}\} \)

2. \( \{a^b_1, h^b, a^b_2\} = \{0, 1, 0\} \)

3. \( r(a_1) = 0 \) for all \( a_1 \)

4. \( \bar{\lambda}(0) = 0 \) for all \( a_1 \)

**Proposition 6**  
A stay equilibrium exists if and only if \( s \leq z \). An exit equilibrium exists if and only if \( s \geq z \).

**Proof:** If \( s > z \), then the good type’s strategy in a stay equilibrium is not incentive compatible. He is better off exerting \( \bar{a} \) and exiting by the definition of \( \hat{a} \) and \( z \). Therefore \( s \leq z \) is a necessary condition for the existence of a stay equilibrium.

I now show that \( s \leq z \) is a sufficient condition for a stay equilibrium to exist. In their second employment period, a bad type’s payoff is uniquely maximized by \( a_2 = 0 \). Similarly the good type’s payoff is uniquely maximized by \( a_2 = \bar{a} \) in his final period of employment. Given these second-period strategies and the assumption that \( s \leq r \), the manager’s strategy is optimal under the belief specified at each information set in the definition of equilibrium. To show that these beliefs survive the intuitive criterion, first note that because \( \hat{a} \geq \bar{a} \), the good type can never benefit from exerting more than \( \hat{a} \). Because \( s \leq \hat{a} \), the bad type can also never benefit from exerting more than \( \hat{a} \). Therefore \( \bar{\lambda}(a_1) = 1 \) for \( a_1 > \hat{a} \) survives the intuitive criterion. From the definition of \( s \), the bad type is strictly better off compared to equilibrium if he can exert \( a_1 < s \) and be retained. Therefore \( \bar{\lambda}(a_1) = 0 \) for \( a_1 < s \) survives the intuitive criterion. If \( s \geq \hat{a}, \) then \( \hat{a} = s \) and \( \bar{\lambda}(a_1) \) as specified in the definition survives the intuitive criterion. If \( s < \hat{a} \), then \( \hat{a} = \bar{a} \) and beliefs on \( (s, \bar{a}) \) still need to be shown to survive the intuitive criterion. From the definition of \( \bar{a} \), the good type is strictly better off exerting \( \bar{a} \) and being retained than exerting any other \( a_1 \) and being retained. Therefore she is strictly worse off exerting any \( a_1 \in (s, \bar{a}) \), regardless of the manager’s response. From the definition of \( s \), the bad type is also strictly worse off exerting any \( a_1 \) on this interval. Therefore \( \bar{\lambda}(a_1) = 0 \) survives the intuitive criterion on the interval.
Finally, it remains to be shown that first-period bureaucrat strategies are compatible with equilibrium. The bad type can never do better by exiting the bureaucracy as this yields a maximum payoff of zero in the second period which is less than his equilibrium payoff, $\rho R_B$. If he exerts $a_1 \in (0, \hat{a})$, he suffers the cost of exerting effort and is fired, making him strictly worse off. If he exerts $a_1 \geq \hat{a}$, he is retained. Because $\hat{a} \geq s$, this makes him weakly worse off compared to equilibrium.

The good type achieves his globally maximal payoff if $s \leq \bar{a}$. If $s > \bar{a}$, the good type earns an equilibrium payoff of at least $W$ because $s \leq z$. If the good type exerts $a_1 < \hat{a} = s$, he is fired which makes him worse off. Any $a_1 > s$ makes him worse off even if he is retained. Therefore $s \leq z$ is a sufficient condition for a stay equilibrium to exist.

I now show that $s \geq z$ is a sufficient condition for an exit equilibrium to exist. Optimality of second-period actions for the bureaucrats is established above. It is straightforward to check that given the manager’s belief, his strategy is optimal. Let $E$ denote the manager’s ex ante expected payoff when a new bureaucrat is hired in an exit equilibrium. If the manager retains the bureaucrat, her payoff is $\delta E$. If she fires him, her payoff is $(1 - \rho)E + \rho \delta E$. Because $E \geq \delta E$, the manager’s strategy is optimal given her beliefs.

These beliefs survive the intuitive criterion. If $s = z$, neither type can strictly benefit relative to equilibrium by exerting $a_1 = s = z$ and staying. Therefore $\lambda(s) = 0$ survives the intuitive criterion. From the definition of $s$, $\lambda(a_1) = 0$ survives the intuitive criterion for any $a_1 < s$. Similarly, $\lambda(a_1)$ survives the intuitive criterion for any $a_1 > z$ by the definition of $z$. Therefore the off-path beliefs specified survive.

The bad type’s strategy in the first period is optimal. He can never do better by exiting. If he stays and exerts $a_1 \neq 0$, he is fired which makes him worse off. The good type’s strategy is also optimal. If he exits, he can do no better than exerting $\bar{a}$. If he stays, he is fired regardless of the amount of effort he exerts. Under Assumption 1, this yields a smaller payoff than his equilibrium payoff, $W$. Therefore $s \geq z$ is a sufficient condition for the existence of an exit equilibrium.
Finally I show that $s \geq z$ is a necessary condition for the existence of an exit equilibrium. Consider the manager’s beliefs if $s < z$ and $a_1 = s + \epsilon < z$ for $\epsilon$ arbitrarily small. If the deviating bureaucrat is a good type, his payoff if the manager retains him is

$$v(s + \epsilon) - c(s + \epsilon) + W + R_G > W$$

where the inequality follows from the definition of $z$. If the bureaucrat is a bad type, his payoff if the manager retains him is

$$-c(s + \epsilon) + R_B < \rho R_B$$

where the inequality follows from the definition of $s$. Only the good type can strictly benefit from this deviation. The bad type is made strictly worse off. Therefore the only belief that survives the intuitive criterion is $\bar{\lambda}(s + \epsilon) = 1$. The manager must retain the bureaucrat after such a deviation which therefore is profitable for the good type. Therefore $s < z$ implies that an exit equilibrium does not exist. □

**Definition 6** In a pooling equilibrium,

1. $\{a_1^q, h^q, a_2^q\} = \{\bar{a}, 1, \tilde{a}\}$
2. $\{a_1^b, h^b, a_2^b\} = \{\bar{a}, 1, 0\}$
3. $r(\bar{a}) = 1$
4. $\bar{\lambda}(\bar{a}) = \lambda$

**Proposition 7** A pooling equilibrium does not exist.

**Proof:** A pooling equilibrium requires that the manager retains the bureaucrat if $a_1 = \bar{a}$. On the equilibrium path, $\lambda(\bar{a}) = \lambda$ by Bayes’ rule. Therefore in a pooling equilibrium the manager’s expected payoff in the second period of a bureaucrat’s employment is $\lambda v(\bar{a})$. 46
Because both types choose the same level of effort in the first period, the manager earns a payoff of $v(\bar{a})$ in the first period of a bureaucrat’s employment. In order for the manager to be willing to retain the first-period bureaucrat, $v(\bar{a})$ must be less than $\lambda v(\bar{a})$. That is, good types must exert less than their preferred amount of effort in their first period of employment in order to be retained in equilibrium. The good type’s equilibrium payoff is therefore

$$v(\bar{a}) - c(\bar{a}) + W + R_G < 2W + R_G$$

If the good type chooses $\bar{a}$ instead of $\bar{a}$ in the first period and is retained, his payoff is $2W + R_G$. The bad type’s payoff in equilibrium is

$$v(\bar{a}) - c(\bar{a}) + R_B > v(\bar{a}) - c(\bar{a}) + R_B$$

If the bad type chooses $\bar{a}$ instead of $\bar{a}$ and is retained, his payoff is $v(\bar{a}) - c(\bar{a}) + R_B$. Therefore under the intuitive criterion, $\lambda(\bar{a}) = 1$ in a pooling equilibrium. But then there exists a profitable deviation for the good type from his equilibrium strategy. Therefore a pooling equilibrium does not exist. □

Note that the bureaucrats must exert less than $\bar{a}$ in the first period of a pooling equilibrium due to the logic described in the main text. If bureaucrats pool on a level of effort near $\bar{a}$, the manager is better off hiring a new bureaucrat to ensure bureaucratic output in the next period rather than gamble that with probability $\lambda$ that the old bureaucrat is good and exerts effort. If new bureaucrats are sufficiently less productive than second-period bureaucrats, this temptation is mitigated. However, under the intuitive criterion bureaucrats cannot pool on any level of effort less than $\bar{a}$.  

47
References


48


