Career Concerns, Beijing Style*

Jiahua CHE†
Kim-Sau Chung‡
Xue Qiao§

July 5, 2014

Abstract

We study a model of overlapping principal-agent problems, where one of yesterday's agents would be selected/promoted as today's principal, who then wields absolute power free of checks and balances, and has discretion over how to select/promote one of today's agents as tomorrow's principal. We use this model to theorize how a political system building on career concerns instead of checks and balances may function. We call this a model of career concerns, Beijing style, which differs fundamentally from one of career concerns, Holmström style, in that the disciplinary effect of career concerns exhibits inherent indeterminacy. The proper functioning of such a political system also relies on strong enough state capacity and an intermediate level of decentralization. Small improvement in the rule of law or a mandatory merit-based promotion rule may inadvertently render such a political system unworkable.

JEL: D72, H11, P48

Keywords: career concerns, accountability, power abuse

1 Introduction

Why has China been able to achieve phenomenal growth despite a political system free of checks and balances? An emerging popular answer is that, thanks to regional tournament, the otherwise unaccountable Chinese government officials become enthusiastic in

*We thank Melody Lo, Chia-hui Lu, Vai-lam Mui, and seminar/conference audience at various seminars/conferences for helpful comments. Yunyun Lu provided excellent research assistance. All errors are ours.

†Department of Economics, Chinese University of Hong Kong.
‡Department of Economics, Chinese University of Hong Kong.
§School of Economics and Management, Tsinghua University.

promoting economic growth out of their career concerns (Li and Zhou, 2005; Xu, 2011). To the extent that career concerns can serve as substitutes of checks and balances in disciplining government officials, many believe that China has discovered a viable alternative to democracy. Indeed, *The Economists* went so far to claim that the success of China "poses a far more credible threat than communism ever did to the idea that democracy is inherently superior."[1]

However, while there are thousands of studies on how democracy functions—from how voters vote, how parties determine their campaign platforms, to how policies are affected by election cycles—there are very few theories on how a political system building on career concerns instead of checks and balances may function in different economies. Any country seriously choosing between democracy and such a political system would like to know much more about the latter other than the mere fact that it seems to work for China at one point of time. For example:

1. Would the very same career concerns generate occasional disasters as during the history of Communist China (Li and Yang, 2005)?

2. What kind of next-generation leaders would be selected/promoted through the regional tournament that generates these career concerns? Would they be competent civil servants aching to serve? Or would they be patient crooks craving to abuse the absolute power at the top of the government?

3. Would the next-generation leaders so selected/promoted be interested in carrying on such a political system? Or would they change the rules of tournament (and hence reshape the accompanying career concerns) once they get to the top of the government?

We propose a simple model of overlapping principal-agent problems to theorize how a political system building on career concerns instead of checks and balances may function. A model of overlapping principal-agent problems captures the following key feature of such a political system: one of yesterday's agents would be selected/promoted as today's principal, who then wields absolute power free of checks and balances, and has discretion over how to select/promote one of today's agents as tomorrow's principal.

The key question we ask is how such a political system may generate accountability at the top of the government—where the disciplinary effect of career concerns is absent. We believe this question is the mother of all concerns about a political system building on career concerns instead of checks and balances. In order not to simply assume away this concern, our model explicitly allows for the possibility of power abuse at the top of the government.

Our model offers one possible answer to this key question. Leaders at the top of the government may refrain from abusing their absolute power if doing so may decrease their ability to distinguish different types of bureaucrats, and hence diminish their prospects of selecting/promoting their favorite type of bureaucrats as the next-generation leaders. In other

[1] *The Economist*, "What's Gone Wrong with Democracy?" March 1, 2014.

words, while accountability-at-the-bottom is generated by career concerns, accountability-at-the-top can be generated by *selection concerns*.

The postulation that power abuse by the leaders would diminish the distinguishability of different types of bureaucrats seems plausible to us. A non-performing leader would waste good efforts of his bureaucrats, whereas a performing leader would create room for his bureaucrats to perform to their best potential.

Although our model presents only one possible theory of how a political system building on career concerns instead of checks and balances may manage to generate accountability-at-the-top, it is the first such theory as far as we are aware of, and hence it is interesting to push this model as far we can and investigate its various implications. Our model yields the following seven implications.

First, although career concerns and selection concerns seem to be exerting their disciplinary effects at different levels of the government (career concerns generate accountability-at-the-bottom, selection concerns generate accountability-at-the-top), the two kinds of concerns are complementary. If leaders have weak selection concerns—perhaps because bureaucrats are homogenous, or because the difference between different types of bureaucrats is small—leaders will abuse power more, which in turn decreases the distinguishability of different types of bureaucrats, and hence weakens the disciplinary effects of career concerns on the bureaucrats as well. Similarly, if bureaucrats have weak career concerns—perhaps because there is a lot more to gain from power abuse at the bottom than at the top of the government—different types of bureaucrats will behave too similarly for leaders to tell them apart, which weakens the disciplinary effect of selection concerns and hence also accountability-at-the-top. Therefore, any exogenous factor that weakens one kind of concerns will also weaken the other kind and results in the lack of accountability at all levels of the government.

Second, there is a fundamental difference between career concerns in overlapping principal-agent problems (*career concerns, Beijing style*) and career concerns in standard principal-agent problems where the identities of principal(s) and agent(s) are fixed (*career concerns, Holmström style*). The disciplinary effect of career concerns, Beijing style, is increasing in the gain an agent can enjoy when he is selected/promoted as the next-generation principal. However, the magnitude of this gain depends on whether (once he becomes the principal) his own agents will be sufficiently disciplined by their career concerns, which in turn depends on whether (once they become principals) their own bureaucrats will be sufficiently disciplined by their career concerns, which in turn depends on .... Anticipation of strong (weak) disciplinary effect of career concerns in the future results in strong (weak) disciplinary effect of career concerns now. But both kinds of anticipation are internally consistent, resulting in an inherent indeterminacy in the disciplinary effect of career concerns, Beijing style. This indeterminacy is in itself a reason to be cautious when transplanting a political system building on career concerns instead of checks and balances to other economies.

Third, there is no hope of achieving full accountability at every level of the government in such a political system. Some type of government officials must abuse power at some level of the government. The intuition is actually quite simple. If bureaucrats expect that they will not abuse power even after being selected/promoted to the top of the government, then there is not much to gain from being selected/promoted, and hence the disciplinary effect of career concerns will be weak, resulting in power abuse at the bottom of the government.

This, however, does not imply that there must be power abuse at the top of the government. The above intuition only rules out full accountability at every level of the government, but does not preclude the possibility of full accountability at the top of the government and some power abuse at the bottom of the government. This prompts the question of, if the society has to live with the fact that some type of bureaucrats will abuse power at the bottom of the government, which type of bureaucrats would the society rather let them abuse power? The answer depends on the distribution of types among government officials. Different promotion rules encourage different types of bureaucrats to work accountably and different types bureaucrats to abuse power, and hence can be desirable (from the society's point of view) given different distributions of types. This leads to our fourth implication: institutionalizing a merit-based promotion rule, instead of letting leaders have discretion over who to promote, may inadvertently render a political system building on career concerns instead of checks and balances unworkabe.

Fifth, a small decrease in the private gain of power abuse either at the top or at the bottom of the government—perhaps because of a small improvement in the rule of law—may inadvertently render a political system building on career concerns instead of checks and balances unworkable as well. This implication is surprising because a smaller private gain of power abuse, especially at the bottom of the government, ought to induce more accountability from bureaucrats at least. However, aside from this direct consequence, there is also an indirect consequence: as different types of bureaucrats behave more similarly, they are also less distinguishable. As a result, leaders become less eager to refrain from abusing power, thanks to the weakened disciplinary effect of selection concerns. The small decrease in the private gain of power abuse at the top of the government may not be enough to counter-balance this indirect consequence. The complementarity between career concerns and selection concerns then implies that bureaucrats' career concerns will also have a weaker disciplinary effect, resulting in a collapse of performance at the bottom of the government as well.

Sixth, the proper functioning of a political system building on career concerns instead of checks and balances relies on strong enough state capacity, which we define as the productivity of the government in producing public goods given a fixed budget. This implication can be understood as a corollary to the complementarity between career concerns and selection concerns. When state capacity is weak, it makes little difference for leaders to selecting/promoting their favorite type of bureaucrats as the next-generation leaders. As a result, leaders' selection concerns have a weak disciplinary effect. The complementarity between career concerns and selection concerns then implies that bureaucrat's career concerns will also have a weak disciplinary effect, resulting in a collapse of performance at all levels of the government.

Finally, the proper functioning of a political system building on career concerns instead of checks and balances relies on an intermediate level of decentralization. This can be understood as a corollary to the last implication above. Too much centralization (decentralization) increases the private gain of power abuse at the top (bottom) of the government at the expense of state capacity, making it difficult to contain power abuse at the top (bottom) of the government. By the complementarity between career concerns and selection concerns, the collapse of accountability-at-the-top can lead to the collapse of accountability at the bottom, and vice versa. Therefore, both too much centralization and too much decentralization can render a political system building on career concerns instead of checks and balances unworkable.

The rest of this paper is organized as follows. The rest of this section reviews the related literature. Section 2 introduces the model. Section 3 provides preliminary analysis of the model. Section 4 characterizes all the equilibria. Section 5 formally derives the various implications outlined earlier. Section 6 concludes.

1.1 Related Literature

This paper is obviously related to the literature of career concerns. The literature is pioneered by Holmström (1982, 1999). It has since been applied to study government agencies (Tirole, 1994; Dewatripont, Jewitt, and Tirole, 1999; and Acemoglu, Kremer, and Mian, 2007), bureaucrats and politicians (Alesina and Tabellini, 2007 and 2008), and performance comparisons (Meyer and Vickers, 1997), among others. As explained earlier, this paper differs from this literature in that the identities of the principal(s) and agent(s) are not fixed but are instead endogenous,

This paper is also related to the literature on leadership accountability. Myerson (2010) examines the leader's choice between rewarding his agents for their good efforts and pocketing the reward for himself. Rauch (2001) studies the leader's decision in spending time between monitoring his agents against their corruption and engaging in corruption himself. In both papers, the leader's accountability is what makes his agents accountable, while his agents' career concerns play no role. Since the goal of this paper is to study how a political system building on career concerns instead of checks and balances may function, our starting point is necessarily different: it is the agents' career concerns that make them accountable. As a result, in this paper, the leader's accountability problem takes an alternative form. Instead of resting in making his agents' accountable, the leader's accountability rests in providing public goods.

This paper is also related to the literature of political selection (Besley, 2005). For example, Acemoglu, Egorov, and Sonin (2010) study political selection in a broad range of political institutions. They paint a rather grimy picture for political selection in autocracies, where bad leaders will eventually remain in power forever after eliminating other, better candidates. This paper, instead, paints a more complicated picture for political selection in a political system building on career concerns instead of checks and balances. Proper functioning of such a political system is possible, even though it relies on many factors, which we painstakingly enumerate.

This paper also complements the literature on other sources of motivations for (governmental) agents. For example, Besley and Ghatak (2005) and Prendergast (2008) study intrinsic motivations; and Olson (1993), Maskin and Tirole (2004), and Besley and Kudamatsu (2008) study re-election concerns. In this paper, government officials are self-interested and free from intrinsic motivations, nor do they (and in particular the leaders) face any re-election.

The model of overlapping principal-agent problems employed in this paper resembles the one employed in Che, Chung, and Qiao (2013). The two papers differ in their treatment of checks and balances. While this paper studies a political system free of checks and balances, Che, Chung, and Qiao (2013) study a particular source of checks and balances, namely the civil society, whose strength in turn is endogenous.

2 The Model

This is an overlapping-generations model. Every generation lives for three periods: when they are young, middle-aged, and old, respectively. There are many citizens in each generation. Each citizen is born with a competence type $\theta \in \{h, l\}$, with h meaning a competent type and l an incompetent type. Each citizen's draw of his type is independent, with the probability of $\theta = h$ being ρ. Types are private information.

There is one government, consisting of a leader and two bureaucrats. Bureaucrats are randomly drawn from young citizens. When they turn middle-aged, one of the two bureaucrats will be promoted by the retiring leader to be the next leader. The one who does not get promoted will become one of the ordinary (middle-aged) citizens. A leader is hence necessarily middle-aged. He retires when he turns old. Right before he retires, he promotes one of the two (young-turning-middle-aged) bureaucrats as the next leader. A retired leader becomes an ordinary (old) citizen.

In any given period, the sequence of events is as follows. First, two young citizens are randomly chosen as the bureaucrats. The leader (who was promoted at the end of the previous period) chooses whether to work or to abuse power, which takes the form of embezzlement in this model. If the leader works, he may succeed in performing, or he may fail; his probability of performing is 1 if he is competent, and is $\gamma \in (0, 1)$ if he is incompetent. If the leader embezzles, he for sure will not perform, but he will gain a private benefit of a.

Second, the bureaucrats observe whether or not the leader has performed, and then decide simultaneously whether to work or to embezzle. If the leader has performed, and if a bureaucrat works, then this bureaucrat may succeed in performing, or he may fail; same as the leader, his probability of performing is 1 if he is competent, and is $\gamma \in (0, 1)$ if he is incompetent. If he embezzles, he for sure will not perform, but he will gain a private benefit of $\tilde{a}$. If the leader has not performed, then a bureaucrat for sure cannot perform regardless of his action and his type. This assumption captures the idea that a non-performing leader would waste good efforts of his bureaucrats, whereas a performing leader would create room for his bureaucrats to perform to their best potential.

Third, the leader observes whether or not each bureaucrat has performed, and then decides which of the two to promote as the succeeding leader for the next period. We assume that bureaucrats' actions are not observable. We also assume that if both bureaucrats have performed, or if both have not performed, then the leader will promote each of them with equal chance. A real promotion decision follows only the event where one bureaucrat has performed while the other has not. We say that the promotion is merit-based if the leader promotes the one who has performed; and that it is anti-merit-based if the leader promotes the one who has not performed.

Finally, each performing bureaucrat generates 1 unit of public good, whereas a non-performing bureaucrat generates nothing. Only old citizens consume public goods at the end of each period, while young and middle-aged citizens are assumed to not consume public goods. This assumption not only simplifies, but also sharpens, our analysis. With this assumption, each government official derives no static gain from his effort, in sync with the standard career concerns model à lá Holmström.

For expositional simplicity, we also assume zero discounting. Any agent's life-time payoff is hence the simple sum of (*i*) any private benefit from embezzling as a (young) bureaucrat, (*ii*) any private benefit from embezzling as a (middle-aged) leader, and (*iii*) any consumption of public good as an (old) citizen.

3 Preliminaries

Our solution concept is pure-strategy Markov perfect equilibrium, which we refer to simply as *equilibrium.*

An equilibrium is characterized by a vector $(\varepsilon_h^L, \varepsilon_l^L, \varepsilon_h^B, \varepsilon_l^B, r)$, where

- $\varepsilon_h^L = 1$ if a competent leader chooses to embezzle, and $\varepsilon_h^L = 0$ if he chooses to work accountably;
- similarly for ε_l^L;
- $\varepsilon_h^B = 1$ if a competent bureaucrat chooses to embezzle conditional on the leader having performed, and $\varepsilon_h^B = 0$ if he chooses to work accountably;[2]
- similarly for ε_l^B; and
- $r = mb$ if promotion is merit-based, and $r = amb$ if promotion is anti-merit-based.[3]

An equilibrium is a vector $(\varepsilon_h^L, \varepsilon_l^L, \varepsilon_h^B, \varepsilon_l^B, r)$ that satisfies:

[2]Note that the bureaucrats must embezzle conditional on the event that the leader has *not* performed, as they for sure cannot perform if the leader has not.

[3]A retiring leader's equilibrium promotion strategy is independent of his competence type because both types of retiring leader share the same payoff function which is the payoff function of a typical old citizen, rendering his competence type payoff irrelevant.

1. r is optimal for a retiring leader given his belief that all current and future government officials follow strategies $(\varepsilon_h^L, \varepsilon_l^L, \varepsilon_h^B, \varepsilon_l^B, r)$;

2. ε_θ^L is optimal for a type-θ leader given his belief that all current and future government officials follow strategies $(\varepsilon_h^L, \varepsilon_l^L, \varepsilon_h^B, \varepsilon_l^B, r)$; and

3. ε_θ^B is optimal for a type-θ bureaucrat given his belief that all current and future government officials follow strategies $(\varepsilon_h^L, \varepsilon_l^L, \varepsilon_h^B, \varepsilon_l^B, r)$.

We deal with these three conditions one by one below.

(I) *r is optimal for a retiring leader given his belief that all current and future government officials follow strategies $(\varepsilon_h^L, \varepsilon_l^L, \varepsilon_h^B, \varepsilon_l^B, r)$*

When choosing his successor, a retiring leader calculates the amount of public good he will consume when he is old (and retired). Let y_h and y_l denote the amount of public good he will consume when he is old conditional on the next leader being competent and incompetent, respectively. That is,

$$y_h \equiv 2(1-\varepsilon_h^L)\left[\rho(1-\varepsilon_h^B) + (1-\rho)(1-\varepsilon_l^B)\gamma\right], \tag{1}$$

$$\text{and} \quad y_l \equiv 2(1-\varepsilon_l^L)\gamma\left[\rho(1-\varepsilon_h^B) + (1-\rho)(1-\varepsilon_l^B)\gamma\right]. \tag{2}$$

Recall that a real promotion decision follows only the event where one bureaucrat has performed while the other has not. As a bureaucrat for sure cannot perform regardless of his action and his type when the leader fails to perform, the leader formulates his promotion strategy r only after he himself has performed. Conditional on the leader having performed, the probability that the next leader will be competent given promotion strategy r is

$$q_r = \rho^2 + 2\rho(1-\rho)A_r, \tag{3}$$

where A_r is the probability that a competent bureaucrat will be promoted under promotion strategy r, conditional on there being exactly one competent and one incompetent bureaucrats, and conditional on the leader having performed.[4] Accordingly, conditional on the leader having performed, his retirement payoff can be written as:

$$Z_r = q_r y_h + (1-q_r)y_l. \tag{4}$$

[4] In particular,

$$\begin{aligned} A_{mb} \equiv\ & (1-\varepsilon_h^B)\left[\varepsilon_l^B + (1-\varepsilon_l^B)(1-\gamma)\right] \\ & + \frac{1}{2}(1-\varepsilon_h^B)(1-\varepsilon_l^B)\gamma \\ & + \frac{1}{2}\varepsilon_h^B\left[\varepsilon_l^B + (1-\varepsilon_l^B)(1-\gamma)\right], \end{aligned}$$

where the first term refers to the event that the competent bureaucrat performs while the incompetent one does not; the second term the event that both perform; the third term the event that both do not; and the fraction $\frac{1}{2}$ reflects the fact that the leader will have to randomly promote one of the two bureaucrats

Condition **(I)** hence boils down to

$$r = \begin{cases} mb & \text{only if} \quad Z_{mb} - Z_{amb} = \Delta q \cdot \Delta y \geq 0 \\ amb & \text{only if} \quad Z_{mb} - Z_{amb} = \Delta q \cdot \Delta y \leq 0 \end{cases}, \tag{5}$$

where

$$\begin{aligned} & \Delta q \equiv q_{mb} - q_{amb} = 2\rho(1-\rho)\Delta A, \\ & \Delta A \equiv A_{mb} - A_{amb} = \varepsilon_l^B - \varepsilon_h^B + (1-\varepsilon_l^B)(1-\gamma), \\ \text{and} \quad & \Delta y \equiv y_h - y_l. \end{aligned}$$

(II) ε_θ^L *is optimal for a type-θ leader given his belief that all current and future government officials follow strategies* $(\varepsilon_h^L, \varepsilon_l^L, \varepsilon_h^B, \varepsilon_l^B, r)$

A leader's expected retirement payoff is Z_r if he has performed, and is $\rho y_h + (1-\rho)y_l$ if he has not (where Z_r is defined in (4), and y_h and y_l are defined in (1) and (2), respectively). The difference between the two is $Z_r - \rho y_h - (1-\rho)y_l = (q_r - \rho)\Delta y$.

Condition **(II)** hence boils down to:

$$\varepsilon_h^L = \begin{cases} 0 & \text{only if} \quad (q_r - \rho)\Delta y \geq a \\ 1 & \text{only if} \quad (q_r - \rho)\Delta y \leq a, \end{cases} \tag{6}$$

$$\text{and} \quad \varepsilon_l^L = \begin{cases} 0 & \text{only if} \quad \gamma(q_r - \rho)\Delta y \geq a \\ 1 & \text{only if} \quad \gamma(q_r - \rho)\Delta y \leq a. \end{cases} \tag{7}$$

(III) ε_θ^B *is optimal for a type-θ bureaucrat given his belief that all current and future government officials follow strategies* $(\varepsilon_h^L, \varepsilon_l^L, \varepsilon_h^B, \varepsilon_l^B, r)$

Recall that a bureaucrat can not perform if the leader fails to perform. Therefore, our discussion of ε_θ^B focuses on the event that the leader has performed. Conditional on the leader having performed, the type-θ bureaucrat's life-time payoffs following him working or embezzling are then

$$\begin{aligned} & B_\theta(0) = m_\theta(0)\varepsilon_\theta^L a + n_\theta(0)Y_h + [1 - n_\theta(0)]\, Y_l \\ \text{and} \quad & B_\theta(1) = \tilde{a} + m_\theta(1)\varepsilon_\theta^L a + n_\theta(1)Y_h + [1 - n_\theta(1)]\, Y_l, \end{aligned}$$

---

if their performances are the same. Similarly,

$$\begin{aligned} A_{amb} \equiv\ & \varepsilon_h^B(1-\varepsilon_l^B)\gamma \\ & + \frac{1}{2}(1-\varepsilon_h^B)(1-\varepsilon_l^B)\gamma \\ & + \frac{1}{2}\varepsilon_h^B\left[\varepsilon_l^B + (1-\varepsilon_l^B)(1-\gamma)\right], \end{aligned}$$

where the first term refers to the event that the competent bureaucrat does not perform while the incompetent one does; the second term the event that both perform; and the third term the event that both do not.

respectively, where $m_\theta(\epsilon_\theta^B)$ represents the probability that this type-θ bureaucrat gets promoted as the next leader given his strategy ϵ_θ^B, $n_\theta(\epsilon_\theta^B)$ represents the probability that the next leader is competent ($\theta' = h$) given his strategy ϵ_θ^B, and $Y_{\theta'}$ with $\theta' \in \{h, l\}$ represents the bureaucrat's expected level of public good consumption (when he is old), conditional on the next leader being type θ'.

Let $\Delta m_\theta \equiv m_\theta(1) - m_\theta(0), \Delta n_\theta \equiv n_\theta(1) - n_\theta(0)$ and $\Delta Y \equiv Y_h - Y_l$. Condition **(III)** hence boils down to:

$$\varepsilon_\theta^B = \begin{cases} 0 & \text{only if} \quad B_\theta(1) - B_\theta(0) = \tilde{a} + \Delta m_\theta \cdot \varepsilon_\theta^L \cdot a + \Delta n_\theta \cdot \Delta Y \leq 0 \\ 1 & \text{only if} \quad B_\theta(1) - B_\theta(0) = \tilde{a} + \Delta m_\theta \cdot \varepsilon_\theta^L \cdot a + \Delta n_\theta \cdot \Delta Y \geq 0 \end{cases}. \tag{8}$$

It is tedious, and yet straightforward, to calculate $m_\theta(\epsilon_\theta^B)$ and $n_\theta(\epsilon_\theta^B)$, as we show in the Appendix. However, it turns out that Δm_θ follows a simple formula:

$$\Delta m_h = \begin{cases} -1/2 & \text{if} \quad r = mb, \\ 1/2 & \text{if} \quad r = amb, \end{cases}$$

$$\text{and} \quad \Delta m_l = \begin{cases} -\gamma/2 & \text{if} \quad r = mb, \\ \gamma/2 & \text{if} \quad r = amb. \end{cases}$$

To see this, first suppose the leader's promotion strategy is $r = mb$. Then, regardless of whether the other bureaucrat performs or not, by performing a bureaucrat will increase his probability of getting promoted by 1/2.[5] Symmetrically, suppose the leader's promotion strategy is $r = amb$. Then, regardless of whether the other bureaucrat performs or not, by performing a bureaucrat will *decrease* his probability of getting promoted by 1/2.[6] Conditional on the leader having performed, by choosing to work, a competent bureaucrat will for sure perform, while an incompetent one will perform only with probability γ.

Likewise, Δn_θ follows a simple formula as well:

$$\Delta n_h = \begin{cases} -(1-\rho)/2 & \text{if} \quad r = mb, \\ (1-\rho)/2 & \text{if} \quad r = amb, \end{cases},$$

$$\text{and} \quad \Delta n_l = \begin{cases} \rho\gamma/2 & \text{if} \quad r = mb, \\ -\rho\gamma/2 & \text{if} \quad r = amb. \end{cases}$$

To see this, first suppose the leader's promotion strategy is $r = mb$. A competent bureaucrat can affect the next leader's type θ' only if the other bureaucrat is incompetent (which happens with probability $1 - \rho$); in which case, by performing he will increase the

[5]If the other bureaucrat performs, by performing a bureaucrat will increase his probability of getting promoted from 0 to 1/2. If the other bureaucrat does not perform, by performing a bureaucrat will increase his probability of getting promoted from 1/2 to 1.

[6]If the other bureaucrat performs, by performing a bureaucrat will decrease his probability of getting promoted from 1 to 1/2. If the other bureaucrat does not perform, by performing a bureaucrat will decrease his probability of getting promoted from 1/2 to 0.

probability that $\theta' = h$ by 1/2. Similarly, an incompetent bureaucrat can affect θ' only if the other bureaucrat is competent (which happens with probability ρ); in which case, by performing he will *decrease* the probability that $\theta' = h$ by 1/2. The case of $r = amb$ is symmetric to that of $r = mb$.

Finally, define

$$Pr(h|h) \equiv \varepsilon_h^L \rho + (1 - \varepsilon_h^L) q_r \tag{9}$$

$$\text{and} \quad Pr(h|l) \equiv \left[\varepsilon_l^L + (1 - \varepsilon_l^L)(1 - \gamma)\right] \rho + (1 - \varepsilon_l^L)\gamma q_r \tag{10}$$

as the probabilities that the succeeding leader is competent conditional on the sitting leader being competent and being incompetent, respectively. In equations (9) and (10), ε_h^L and $\varepsilon_l^L + (1 - \varepsilon_l^L)(1 - \gamma)$ are the probabilities that a competent and an incompetent sitting leaders, respectively, fail to perform; and q_r is defined in (3), and is the probability that the succeeding leader will be competent given the sitting leader's promotion strategy r and conditional on the sitting leader having performed. We have:

$$Y_{\theta'} = Pr(h|\theta')y_h + [1 - Pr(h|\theta')]\, y_l;$$

and ΔY can be readily calculated.

Definition 1 *An equilibrium is a vector* $(\varepsilon_h^L, \varepsilon_l^L, \varepsilon_h^B, \varepsilon_l^B, r)$ *that satisfies* (5)*,* (6)*,* (7)*, and* (8)*.*

To compare different equilibrium outcomes, we measure welfare by computing the expected level of public good provision, where expectation is taken with respect to the stationary distribution in a given equilibrium. More precisely, given any equilibrium $(\varepsilon_h^L, \varepsilon_l^L, \varepsilon_h^B, \varepsilon_l^B, r)$, we can compute the stationary probability that a leader is competent by solving:

$$p = p \cdot Pr(h|h) + (1 - p) \cdot Pr(h|l),$$

where $Pr(h|h)$ and $Pr(h|l)$ are defined in (9) and (10), respectively, and are the probabilities that a succeeding leader is competent conditional on the (sitting) leader being competent and being incompetent, respectively.

Our welfare measure is hence

$$W = py_h + (1 - p)y_l,$$

where y_h and y_l are defined in (1) and (2), respectively, and are the expected levels of public good consumption conditional on the sitting leader being competent and incompetent, respectively. We do not count government officials' private benefits from embezzlement into our welfare measure, as government officials are infinitesimal to the rest of the economy.

4 Equilibria

In this section, we analyze when and to what extent a political system building on career concerns instead of checks and balances can generate accountability. We do so by solving all possible equilibria of this model. We breakdown accountability by looking at accountability of the leaders and that of the bureaucrats. We say that a leader is *accountable* as long as he works, regardless of whether that leader succeeds in performing or not. Similarly, we say that a bureaucrat is *accountable* as long as he works conditional on the leader having performed, regardless of whether that bureaucrat succeeds in performing or not.

By identifying equilibria that are observationally indistinguishable,[7] we can prove that there are only six possible equilibria. In none of these equilibria can full accountability be achieved; i.e., some power abuse must appear in equilibrium. In one of these equilibria, the government is completely corrupt, with no accountability achieved at any level of the government. This equilibrium always exists regardless of the parameters $(a, \tilde{a}, \rho, \gamma)$. We shall therefore refer to it as the trivial, corruptive equilibrium. In the other five equilibria, certain types of government officials at some levels of the government behave accountably. We shall hence refer to them as equilibria with limited accountability. Unlike the trivial, corruptive equilibrium, these equilibria exist only in certain parameter ranges. In other words, while complete corruption is always possible, accountability (albeit only limited) arises only under certain conditions.

4.1 The Impossibility of Full Government Accountability

We begin by establishing that full accountability is impossible, that is, $(\varepsilon_h^L, \varepsilon_l^L, \varepsilon_h^B, \varepsilon_l^B) = (0, 0, 0, 0)$ cannot be part of any equilibrium. Power abuse must appear sometimes and somewhere in this model, where government officials are free from checks and balances.

To see why full accountability is impossible in any equilibrium, let's first assume that promotion is merit-based; i.e., let's assume that $(\varepsilon_h^L, \varepsilon_l^L, \varepsilon_h^B, \varepsilon_l^B, r) = (0, 0, 0, 0, mb)$ could be sustained as equilibrium, contrary to our claim. Note that an accountable leader collects no private benefit for himself from the office. Therefore, an incompetent bureaucrat, if he foresees that being the next leader brings no private benefit, has no desire to be the next leader. If anything, he has a desire *not* to be the next leader: he knows very well that he himself is incompetent, while there is some chance that the other bureaucrat is competent, so he would rather have the other bureaucrat being promoted as the next leader. What is the best way for an incompetent bureaucrat to avoid being promoted? Since promotion is expected to be merit-based, the best way to avoid being promoted is to embezzle. Therefore, an incompetent bureaucrat will deviate to $\varepsilon_l^B = 1$, upsetting the supposed equilibrium of $(\varepsilon_h^L, \varepsilon_l^L, \varepsilon_h^B, \varepsilon_l^B, r) = (0, 0, 0, 0, mb)$.

[7] In this paper, some equilibria are observationally indistinguishable. For example, any two equilibria with $(\varepsilon_h^L, \varepsilon_l^L) = (1, 1)$ are observationally indistinguishable. When the leader embezzles regardless of his type, he will never perform, and hence neither will his bureaucrats, again regardless of their types. Any difference in $(\varepsilon_h^B, \varepsilon_l^B, r)$ hence can only be observed off the equilibrium path.

The above argument relies on the assumption that $r = mb$. Can $(\varepsilon_h^L, \varepsilon_l^L, \varepsilon_h^B, \varepsilon_l^B) = (0,0,0,0)$ be part of an equilibrium where $r = amb$? The answer is no. If all government officials (leader and bureaucrats included) are accountable, then a retiring leader would rather promote a competent successor, because a competent successor generates a higher expected level of public goods. To increase the chance of promoting a competent successor, a retiring leader would use a merit-based promotion strategy, because he understands that both bureaucrats are working hard.

Proposition 1 *Full accountability is impossible in equilibrium; i.e.,* $(\varepsilon_h^L, \varepsilon_l^L, \varepsilon_h^B, \varepsilon_l^B) = (0,0,0,0)$ *cannot be part of any equilibrium.*

Proof: See the discussion above. ■

4.2 The Trivial, Corruptive Equilibrium

In contrast, the corruptive equilibrium, in which every type of every government official chooses to embezzle, always exists regardless of the parameters $(a, \tilde{a}, \rho, \gamma)$. Among the observationally indistinguishable equilibria of this class, one particular simple example is $(\varepsilon_h^L, \varepsilon_l^L, \varepsilon_h^B, \varepsilon_l^B, r) = (1,1,1,1, amb)$. To see that this is an equilibrium, note that the only reason why a sitting leader works instead of embezzles is to create an environment where he can distinguish different types of bureaucrat and make a more informed choice of the succeeding leader. If a sitting leader believes that every type of succeeding leader will embezzle, the whole point of distinguishing different types of bureaucrat becomes moot. The sitting leader will no longer have any incentive to work. In other words, $\varepsilon_h^L = \varepsilon_l^L = 1$ (for the sitting leader) is the best response to the belief that $\varepsilon_h^L = \varepsilon_l^L = 1$ (for the succeeding leader).

Also, when a sitting leader believes that every type of succeeding leader will embezzle, he is indifferent in who to promote, and hence indifferent in which promotion strategy to choose. In other words, every promotion strategy r, including $r = amb$ of course, is a best response to the belief that $\varepsilon_h^L = \varepsilon_l^L = 1$.

How different types of bureaucrat behave conditional on the leader having performed (i.e., ε_h^B and ε_l^B) depends on the leader's promotion strategy r. For example, $\varepsilon_h^B = \varepsilon_l^B = 1$ is a best response to $r = amb$. However, as the leader never performs on the equilibrium path, no bureaucratic performance can be observed on the equilibrium path either, regardless of how they intend to behave off the equilibrium path.

Proposition 2 *A trivial, corruptive equilibrium, where every government official embezzles on the equilibrium path, always exists regardless of the parameters* $(\rho, \gamma, a, \tilde{a})$*. The equilibrium welfare is* $W = 0$*.*

Proof: See the discussion above. ■

4.3 Equilibria with Limited Accountability

Other than the trivial, corruptive equilibrium which always exists, there are five equilibria with limited accountability. We label these equilibria alphabetically and discuss in details below.

The first equilibrium, Equilibrium A, is characterized by $(\varepsilon_h^L, \varepsilon_l^L, \varepsilon_h^B, \varepsilon_l^B, r) = (0, 0, 0, 1, mb)$. In this equilibrium, leaders of both types are accountable, but only competent bureaucrats are accountable, while incompetent ones embezzle. Accordingly, we refer to this equilibrium as a *top-accountable equilibrium*.

In this equilibrium, leaders of both types work instead of embezzle because they want to distinguish competent bureaucrats from incompetent ones (recall that if a leader embezzles, his bureaucrats will also embezzle, rendering different types of bureaucrats indistinguishable). A competent bureaucrat works instead of embezzles because he knows very well that he is competent, while there is some chance that the other bureaucrat is incompetent, so he would rather have himself promoted as the next leader. Since promotion is merit-based, a competent bureaucrat works to get promoted.

Evidently, in order for $(0, 0, 0, 1, mb)$ to be an equilibrium, the temptation for embezzlement must not be too large, as the next lemma suggests. In particular, the private gain from embezzlement at the top of the government, a, must be small enough in order for both types of leaders to choose work over embezzlement. Similarly, the private gain from embezzlement at the bottom of the government, $\tilde{a}$, must be small enough in order for the competent bureaucrats to choose work over embezzlement. Meanwhile, as we reason in Proposition 1, in this particular equilibrium with promotion being merit-based and all others working accountably, incompetent bureaucrats will have no incentive to fight for promotion, and as a result, they will embezzle no matter how small $\tilde{a}$ is.

Lemma 1 *A top-accountable equilibrium* $(\varepsilon_h^L, \varepsilon_l^L, \varepsilon_h^B, \varepsilon_l^B, r) = (0, 0, 0, 1, mb)$ *exists if and only if*

1. $a \leq 2\rho^2(1-\rho)\gamma(1-\gamma)$ *and*
2. $\tilde{a} \leq \rho^2(1-\rho)^2(1-\gamma)^2$.

The equilibrium welfare is $W = \dfrac{\rho + (1-\rho)\gamma}{1 - \rho(1-\rho)(1-\gamma)}\rho.$

Equilibrium B is characterized by $(\varepsilon_h^L, \varepsilon_l^L, \varepsilon_h^B, \varepsilon_l^B, r) = (0, 0, 1, 0, amb)$. This equilibrium resembles the top-accountable equilibrium in that both types of leaders are accountable. However, in this equilibrium, a competent bureaucrat embezzles instead of works, while an incompetent bureaucrat works instead of embezzles. The reason for these differences is that promotion is anti-merit-based in this equilibrium, and hence embezzling is the way to get promoted, while working is the way to avoid being promoted. Anti-merit-based promotion arises in equilibrium exactly because a retiring leader understands that competent bureaucrats (counter-)signal their competence by deliberately not working. We therefore refer to this equilibrium as a *top-accountable equilibrium with anti-merit-based promotion.*

As in the case of Equilibrium A, in order for $(0, 0, 1, 0, amb)$ to be an equilibrium, a must be small enough in order for both types of leaders to choose work over embezzlement. Compared to Equilibrium A, $\tilde{a}$ must be even smaller in order for the incompetent bureaucrats to choose work over embezzlement (as incompetent bureaucrats, being less productive, are more tempted to embezzle, *ceteris paribus*). No condition is needed for competent bureaucrats to embezzle because, with anti-merit-based promotion, they will be happy to embezzle while at the same time wining promotion to the top.

Lemma 2 *A top-accountable equilibrium with anti-merit-based promotion* $(\varepsilon_h^L, \varepsilon_l^L, \varepsilon_h^B, \varepsilon_l^B, r) = (0, 0, 1, 0, amb)$ *exists if and only if*

1. $a \leq 2\rho(1-\rho)^2\gamma^3(1-\gamma)$ *and*
2. $\tilde{a} \leq \rho^2(1-\rho)^2\gamma^3(1-\gamma)^2$.

The equilibrium welfare is $W = \dfrac{\rho + (1-\rho)\gamma}{1-\rho(1-\rho)\gamma(1-\gamma)}(1-\rho)\gamma.$

Equilibrium C is characterized by $(\varepsilon_h^L, \varepsilon_l^L, \varepsilon_h^B, \varepsilon_l^B, r) = (0, 1, 0, 1, mb)$, which we refer to as a *partially accountable equilibrium.*

In this equilibrium, no matter whether it is at the top or at the bottom of the government, competent government officials always work, and incompetent ones always embezzle. This stark contrast between competent and incompetent government officials in turn motivates competent leaders to work in order to select competent successors, and also motivates competent bureaucrats to work in order to save the economy from being led by incompetent leaders.

Some conditions on the parameters $(a, \tilde{a}, \rho, \gamma)$ need to be satisfied in order for $(\varepsilon_h^L, \varepsilon_l^L, \varepsilon_h^B, \varepsilon_l^B, r) = (0, 1, 0, 1, mb)$ to be an equilibrium. First, in order for a competent leader to be willing to work instead of embezzle, while an incompetent leader to embezzle instead of work, a must lie in some intermediate range. Second, in order for a competent bureaucrat to be willing to work instead of embezzle, $\tilde{a}$ cannot be too big. Third, in order for an incompetent bureaucrat to be willing to embezzle instead of work—which lowers his chance of being promoted to be a corruptive leader—a cannot be too tempting relative to $\tilde{a}$.

Lemma 3 *A partially accountable equilibrium* $(\varepsilon_h^L, \varepsilon_l^L, \varepsilon_h^B, \varepsilon_l^B, r) = (0, 1, 0, 1, mb)$ *exists if and only if*

1. $2\rho^2(1-\rho)\gamma \leq a \leq 2\rho^2(1-\rho)$,
2. $\tilde{a} \leq \rho^2(1-\rho)^2$, *and*
3. $\left[\dfrac{a}{2} - \dfrac{\tilde{a}}{\gamma}\right] \leq \rho^3(1-\rho)$.

The equilibrium welfare is $W = \dfrac{\rho^2}{1-\rho(1-\rho)}$.

Equilibrium D is characterized by $(\varepsilon_h^L, \varepsilon_l^L, \varepsilon_h^B, \varepsilon_l^B, r) = (0, 1, 0, 0, mb)$. We call this a *bottom-accouantable equilibrium* because bureaucrats of both types choose to work in this equilibrium. Meanwhile, only competent leaders are accountable, while incompetent leaders are not. Competent bureaucrats work hard in order to save the economy from being led by incompetent leaders. Incompetent bureaucrats also work hard, but they are instead after the private gain of power abuse at the top of the government.

Some conditions on the parameters$(a, \tilde{a}, \rho, \gamma)$ need to be satisfied in order for $(\varepsilon_h^L, \varepsilon_l^L, \varepsilon_h^B, \varepsilon_l^B, r) = (0, 1, 0, 0, mb)$ to be an equilibrium. First, in order for a competent leader to be willing to work instead of embezzle, while an incompetent leader to embezzle instead of work, a must lie in some intermediate range. Second, in order for a competent bureaucrat to be willing to work instead of embezzle, $\tilde{a}$ cannot be too big. Third, in order for an incompetent bureaucrat to be willing to work instead of embezzle, a has to be tempting enough relative to $\tilde{a}$.

Lemma 4 *A bottom-accountable equilibrium* $(\varepsilon_h^L, \varepsilon_l^L, \varepsilon_h^B, \varepsilon_l^B, r) = (0, 1, 0, 0, mb)$ *exists if and only if*

1. $2\rho(1-\rho)[\rho + (1-\rho)\gamma]\gamma(1-\gamma) \leq a \leq 2\rho(1-\rho)[\rho + (1-\rho)\gamma](1-\gamma)$,
2. $\tilde{a} \leq \rho(1-\rho)^2[\rho + (1-\rho)\gamma](1-\gamma)$, *and*
3. $\left[\dfrac{a}{2} - \dfrac{\tilde{a}}{\gamma}\right] \geq \rho^2(1-\rho)[\rho + (1-\rho)\gamma](1-\gamma)$.

The equilibrium welfare is $W = \dfrac{\rho + (1-\rho)\gamma}{1-\rho(1-\rho)(1-\gamma)}\rho$.

The last equilibrium with limited accountability, Equilibrium E, is characterized by $(\varepsilon_h^L, \varepsilon_l^L, \varepsilon_h^B, \varepsilon_l^B, r) = (0, 1, 1, 0, amb)$. It resembles Equilibrium C in that there is some power abuse at both levels of the government. The main difference is that, in this equilibrium, competent bureaucrats embezzle while incompetent bureaucrats work, and equilibrium promotion is anti-merit-based, reflecting the leaders' desire to promote a competent successor. We therefore refer to this equilibrium as a partially accountable equilibrium with anti-merit-based promotion.

Some conditions on the parameters$(a, \tilde{a}, \rho, \gamma)$ need to be satisfied in order for $(\varepsilon_h^L, \varepsilon_l^L, \varepsilon_h^B, \varepsilon_l^B, r) = (0, 1, 1, 0, amb)$ to be an equilibrium. First, in order for a competent leader to be willing to work instead of embezzle, while an incompetent leader to embezzle instead of work, a must lie in some intermediate range. Second, in order for an incompetent bureaucrat to be willing to work instead of embezzle—which lowers his chance of being promoted to be a corruptive leader under an anti-merit-based promotion strategy—the combined temptation from a and $\tilde{a}$ cannot be too big.

Lemma 5 *A partially accountable equilibrium with anti-merit-based promotion* $(\varepsilon_h^L, \varepsilon_l^L, \varepsilon_h^B, \varepsilon_l^B, r) = (0, 1, 1, 0, amb)$ *exists if and only if*

1. $2\rho(1-\rho)^2\gamma^3 \leq a \leq 2\rho(1-\rho)^2\gamma^2$, *and*
2. $\left[\frac{a}{2} + \frac{\tilde{a}}{\gamma}\right] \leq \rho^2(1-\rho)^2\gamma^2$.

The equilibrium welfare is $W = \dfrac{\rho(1-\rho)\gamma}{1-\rho(1-\rho)\gamma}$.

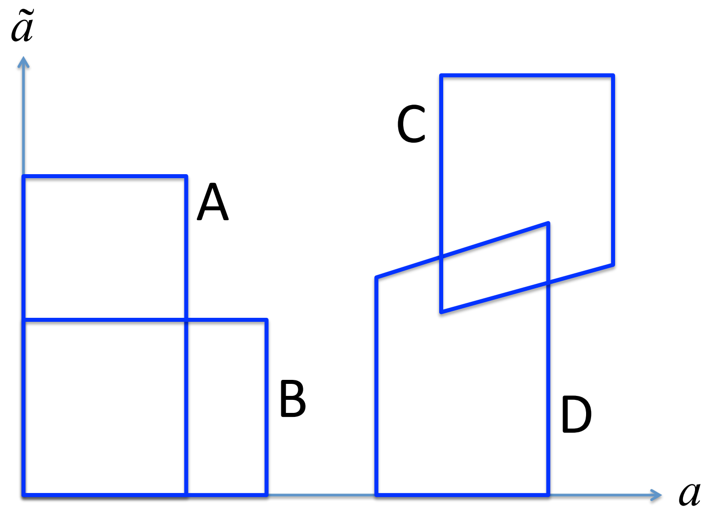

Figure 1: Equilibria with Limited Government Accountability

The four areas in Figure 1 offer a heuristic illustration of the possible parameter ranges for Equilibria A through D. The parameter range for Equilibrium E is not shown because it is empty whenever rectangle A does not completely contain rectangle B (as is the case in Figure 1).[8]

Rectangles A and B represents the parameter range for Equilibria A and B, respectively. Per Lemmas 1 and 2, rectangle B is always shorter than rectangle A. It is weakly narrower than rectangle A (and hence completely contained in the latter) if and only if $\rho \geq \gamma^2/(1+\gamma^2)$. We choose to depict the opposite case in Figure 1 to facilitate some of our discussion in Section 5.

[8]It can be shown using Lemma 5 that the set of $(a, \tilde{a})$ for Equilibrium E to exist is empty if and only if $\rho < \gamma$. Meanwhile, rectangle A does not completely contain rectangle B (as is the case in Figure 1) if and only if $\rho < \gamma^2/(1+\gamma^2)$. Since $\gamma^2/(1+\gamma^2) < \gamma$, we have the parameter range for Equilibrium E to exist being empty whenever rectangle A does not completely contain rectangle B. It can also be shown using Lemmas 1 and 5 that, when the parameter range for Equilibrium E to exist is not empty, it is always a subset of rectangle A.

Trapezoid C represents the parameter range for Equilibrium C. In Figure 1, trapezoid C is disjoint from rectangles A and B. In particular, the left side of trapezoid C is on the right of the right sides of both rectangles A and B. It can be shown from Lemmas 1 and 3 that trapezoid C is always disjoint from rectangle A, and from Lemmas 2 and 3 that trapezoid C is disjoint from rectangle B if $\rho > \gamma^2 / (1/(1-\gamma) + \gamma^2)$. We choose to depict this case in Figure 1 in order to facilitate some of our discussion in Section 5.

Trapezoid D represents the parameter range for Equilibrium D. It can be shown using Lemmas 1, 2, and 4 that trapezoid D is always to the right of, and always disjoint from, rectangles A and B. In Figure 1, trapezoid D is also "on the left" of trapezoid C in the sense that the right side of trapezoid D is on the left of the right side of trapezoid C. This, however, is not always the case. It can be shown using Lemmas 3 and 4 that trapezoid D is "on the left" of trapezoid C in the above sense if and only if $\rho > (1-\gamma)/(2-\gamma)$. We choose to depict this case in order to facilitate some of our discussion in Section 5.

4.4 Other Impossibilities

Aside from the equilibria we identified above, there are two other potential candidates. The first candidate features $(\varepsilon_h^L, \varepsilon_l^L) = (1, 0)$; that is, competent leaders embezzle while incompetent leaders work. The second candidate features $(\varepsilon_h^B, \varepsilon_l^B) = (1, 1)$ but $(\varepsilon_h^L, \varepsilon_l^L) \neq (1, 1)$; that is, both types of bureaucrat embezzle without the government being completely corrupt.

Neither candidate can be sustained as an equilibrium. To see why $(\varepsilon_h^L, \varepsilon_l^L) = (1, 0)$ cannot be part of an equilibrium, recall that both types of leaders will choose the same promotion strategy r that is independent of their types. Accordingly, anticipating others following the equilibrium strategies $(\varepsilon_h^L, \varepsilon_l^L, \varepsilon_h^B, \varepsilon_l^B, r)$, a competent leader will embezzle whereas an incompetent leader will work only if

$$Y_r - [\rho y_h + (1-\rho) y_l] \leq a \leq \gamma \left(Y_r - [\rho y_h + (1-\rho) y_l]\right)$$

by (6) and (7), which is impossible because $\gamma \in (0, 1)$.

In other words, competent leaders always have a (weakly) stronger incentive to work, because successful performance gives these leaders a chance to select their desired successors, and this selection value of performance is stronger for competent leaders. Even though competent leaders do lead to (weakly) higher expected level of public goods, it is worth being reminded that it is not due to this direct consequence that competent leaders have a (weakly) stronger incentive to work, because leaders do not consume any concurrent public goods in this model.

Proposition 3 *Competent leaders have a (weakly) stronger incentive to work accountably in equilibrium; i.e., $(\varepsilon_h^L, \varepsilon_l^L) = (1, 0)$ cannot be part of any equilibrium.*

PROOF: See the discussion above. ■

The reason why $(\varepsilon_h^B, \varepsilon_l^B) = (1, 1)$ cannot be part of any equilibrium with limited government accountability can be simply understood as well. With $(\varepsilon_h^B, \varepsilon_l^B) = (1, 1)$, a leader is not able to differentiate different types of bureaucrat regardless whether he succeeds in performing or not. As a result, he has no incentive to work regardless of his type.[9] Therefore, no accountability can be achieved at the top of the government either.

Proposition 4 *If career concerns fail to generate accountability at the bottom of the government, selection concerns will also fail to generate accountability at the top of the government; i.e.,* $(\varepsilon_h^B, \varepsilon_l^B) = (1, 1)$ *while* $(\varepsilon_h^L, \varepsilon_l^L) \neq (1, 1)$ *cannot be part of any equilibrium.*

PROOF: See the discussion above. ■

5 Implications

We started off with the task of theorizing how a political system building on career concerns instead of checks and balances may function. In any model of such a political system that explicitly allows for the possibility of power abuse at the top of the government, such as ours, accountability-at-the-top can only come from something other than career concerns. In our specific model, it comes from the leaders' selection concerns: abusing power would decrease their ability to distinguish different types of bureaucrats. In this section we explore the implications of such a source of accountability-at-the-top.

5.1 Indeterminacy

Our model reveals that it can be misleading to think about a political system building on career concerns instead of checks and balances using classical models of career concerns á là Holmström. In a political system building on career concerns instead of checks and balances, it follows almost immediately from definition that leaders are being selected/promoted from bureaucrats instead of being elected. These leaders then select/promote from their own bureaucrats the next-generation leaders. In other words, the overlapping-generations structure is part and parcel of such a political system. Such an overlapping-generations structure is susceptible to indeterminacy. Such indeterminacy takes the starkest form in our particular model: regardless of parameters, good governance can never be guaranteed, as the trivial, corruptive equilibrium always exists.

[9] Formally, if $(\varepsilon_h^B, \varepsilon_l^B) = (1, 1)$, then, following (6), we have:

$$(q_r - \rho)\Delta y \geq a,$$

which is impossible because $A_{mb} = A_{amb} = 1/2$ and hence $q_{mb} = q_{amb} = \rho^2 + \rho(1 - \rho) = \rho$.

5.2 Complementarity Between Career Concerns and Selection Concerns

If abusing power would decrease leaders' ability to distinguish different types of bureaucrats, it must then follow that any exogenous factor that make abusing power more tempting for leaders would indeed make bureaucrats less distinguishable, and hence weaken the disciplinary effect of the latter's career concerns. Similarly, any exogenous factor that make abusing power more tempting for bureaucrats would make them less distinguishable, and hence would reduce the costs for leaders to abuse power as well. In short, career concerns and selection concerns are complementary. Any exogenous factor that weakens the disciplinary effect of one kind of concerns will weaken that of the other kind and render the whole political system unworkable.

It is hence not surprising that such a political system will not "work" if government officials are too homogenous, because homogeneity makes selection pointless. It is a straightforward exercise to verify that the range of parameters for each non-trivial equilibrium vanishes when either $\rho \to 0$ (too few government officials are competent), or $\rho \to 1$ (too few government officials are incompetent), or $\gamma \to 1$ (different types of government officials are too similar in terms of their competence).

Proposition 5 *If there are too few competent officials ($\rho \to 0$), or too few incompetent officials ($\rho \to 1$), or if both types of officials are similarly competent ($\gamma \to 1$), the parameter ranges of all equilibria with limited accountability vanish, leaving the trivial, corruptive equilibrium as the only sustainable equilibrium.*

Proof: By direct inspection of Lemmas 1–5. ■

Similarly, either a sufficiently large a or a sufficiently large $\tilde{a}$ alone, not both, suffices to induce power abuse at all levels of the government.

Proposition 6 *Either a sufficiently large a or a sufficiently large $\tilde{a}$ alone suffices to make the trivial, corruptive equilibrium the unique equilibrium.*

Proof: By direct inspection of Lemmas 1–5. ■

5.3 Mandating a Merit-Based Promotion Rule Can Backfire (Reason 1)

One may think that, in a political system building on career concerns instead of checks and balances, government officials will be more accountable if a merit-based promotion rule is somehow committed. This logic turns out to be flawed, for two reasons.

To see the first reason, observe by our previous analysis that when leaders are free to choose their promotion strategies, there are always a and $\tilde{a}$ small enough so that both

Equilibria A and B can exist for any fixed (ρ, γ). Fixing any γ, as $\rho \to 0$, the welfare under Equilibrium A converges to 0, while the welfare under Equilibrium B converges to $\gamma^2 > 0$. In other words, for ρ sufficiently small, Equilibrium A generates lower welfare than Equilibrium B does. This observation is summarized in the next proposition.

Proposition 7 *There exists some range of parameters $(\rho, \gamma, a, \tilde{a})$ such that two equilibria with limited accountability coexist, one featuring merit-based promotion and the other featuring anti-merit based promotion, while the latter results in strictly higher welfare.*

The intuition behind this observation is as follows. When ρ is sufficiently small, there are many more incompetent bureaucrats than competent ones. It is hence better to motivate the incompetent ones, who are the majority, than the competent ones to work. Merit-based promotion increases the incentives of the competent ones, while decrease those of the incompetent ones, to work, resulting in lower welfare as compared to anti-merit-based promotion. This observation implies that, while either promotion strategies can be sustained as part of an equilibrium when leaders maintain their discretion over promotion strategies, taking away such a discretion and imposing on every leader a merit-based promotion rule instead will eliminate the better equilibrium.[10]

5.4 Mandating a Merit-Based Promotion Rule Can Backfire (Reason 2)

The previous section concerns the possibility that when two equilibria with limited accountability coexist without any promotion rule being imposed, mandating a merit-based promotion rule destroys the better equilibrium. Mandating a merit-based promotion rule can also backfire if the only equilibrium with limited accountability features anti-merit-based promotion, in which case such a mandate will do nothing but leading the economy to plunge into the trivial, corruptive equilibrium.

This second reason is illustrated in Figure 1, where we demonstrate the possibility of Equilibrium B being the only equilibrium with limited accountability under certain conditions. Such a possibility is highlighted by the area of rectangle B outside rectangle A and to the left of the two trapezoids. This area represents parameters under which the only equilibrium with limited accountability is Equilibrium B, a top-accountable equilibrium with anti-merit-based promotion.[11]

Fixing parameters in the aforementioned area where the only equilibrium with limited accountability to exist features anti-merit-based promotion, if we change the game into

[10] Technically, the game with a mandatory merit-based promotion rule is a different game, and hence we should redo the equilibrium analysis in Section 4, instead of directly appealing to those previous results. However, the equilibrium analysis of this new game is highly similar and repetitive, and hence are omitted.

[11] Following Lemma 1, 3, and 4, we can show that rectangle A is disjoint from trapezoids C and D. Furthermore, as explained in Section 4, the rectangle B is wider than rectangle A provided that $\rho < \gamma^2/\left(1+\gamma^2\right)$. Therefore, when $\rho < \gamma^2/\left(1+\gamma^2\right)$, there exists an area of rectangle B outside rectangle A and to the left of the two trapezoids.

one where a retiring leader's discretion of who to promote is taken away, and is replaced with a mandatory merit-based promotion rule, then the only sustainable equilibrium under the same parameters is the trivial, corruptive equilibrium $(\varepsilon_h^L, \varepsilon_l^L, \varepsilon_h^B, \varepsilon_l^B) = (1, 1, 1, 1)$. In other words, mandating a merit-based promotion rule can destroy the only sustainable equilibrium with limited government accountability and lead to lower equilibrium public good provision.

Proposition 8 *There exists some range of parameters $(\rho, \gamma, a, \tilde{a})$ such that,*

1. *if retiring leaders have discretion in choosing their promotion strategies, then some equilibrium with limited accountability can exist, where equilibrium welfare is positive; while*
2. *if a merit-based promotion rule is exogenously imposed, then only the trivial, corruptive equilibrium $(\varepsilon_h^L, \varepsilon_l^L, \varepsilon_h^B, \varepsilon_l^B) = (1, 1, 1, 1)$ can exist, where equilibrium welfare is zero.*

The intuition has something to do with the fact that career concerns alone are not enough for such a political system to "work"—some other supplementary factors need to exist. In our model, these supplementary factors are leaders' selection concerns. The benefit of selecting a competent successor, however, is proportional to the productivities of the next government as led by different types of succeeding leader. Scaling up/down these productivities by the same factor will strengthen/weaken the disciplinary effect of the leader's selection concerns. The productivity of the next government, in turn, depends on who among the succeeding leader's bureaucrats will work and who will embezzle. In Equilibrium A, competent bureaucrats work while incompetent bureaucrats embezzle; whereas in Equilibrium B, competent bureaucrats embezzle while incompetent bureaucrats work. Therefore, when there are relatively more incompetent bureaucrats than competent ones, the productivity of the next government can be lower in Equilibrium A than in Equilibrium B. When this happens, there can be a range of a against which the lower productivity of the next government is not sufficient to keep the sitting leader accountable, while the higher productivity of the next government is. In terms of Figure 1, we have rectangle B being wider than rectangle A when there are relatively more incompetent bureaucrats than competent ones.

5.5 "Pure Water Kill the Fishes"

The very reason why a government official may not be accountable is that embezzlement brings private benefits. One hence wonders whether welfare will increase if the country can somehow lower a and $\tilde{a}$. While this logic may work in a political system with checks and balance, it does not always work in a political system where government officials are disciplined by career concerns.

To see an example, let's refer to Figure 1 again. Suppose the country starts with a point near the upper right corner of trapezoid C. And let's also assume that the country manages to coordinate on Equilibrium C instead of the trivial, corruptive equilibrium initially.

Suppose the country now lowers a or $\tilde{a}$ through some policy that is outside of our model, such as by improving the rule of law. Lowering a, while keeping other parameters intact, moves the point leftward, and potentially moves it outside of trapezoid C. The country is now represented by a point that does not belong to any of the rectangles or trapezoids. As a result, the only equilibrium is the trivial, corruptive equilibrium. Welfare falls to zero consequently.

Similarly, lowering $\tilde{a}$, while keeping other parameters intact, moves the point downward, and potentially moves it outside of trapezoid C. The country is now represented by a point below trapezoid C and on the right of trapezoid D, which does not belong to any of the rectangles or trapezoids. As a result, the only equilibrium is the trivial, corruptive equilibrium. Welfare falls to zero consequently.

Why may a smaller a hurt? Note that when a leader decides whether to work or to embezzle, he is trading off selection against private benefit a. The disciplinary effect of his selection concerns depends on his belief about how different a competent and an incompetent successor will behave. In Equilibrium A or B, both types of successor behave in the same way ($\varepsilon_h^L = \varepsilon_l^L = 0$), and hence selection is less important, and a sufficiently low a is needed to sustain Equilibrium A or B. In Equilibrium C or D, different types of successor behave differently ($\varepsilon_h^L = 0 \neq 1 = \varepsilon_l^L$), and hence selection is more important, and Equilibrium C or D can survive on bigger a. However, a sufficiently high a is needed to lure an incompetent successor to behave differently from a competent one, which is what makes selection more important in the first place. Therefore, there can be a range of a that is not sufficiently low to sustain Equilibrium A or B, and not sufficiently high to sustain Equilibrium C or D. When a is lowered into this range, only the trivial, corruptive equilibrium can exist.

Why may a smaller $\tilde{a}$ hurt? Note that another determinant of the disciplinary effect of a leader's selection concerns is how different a competent and an incompetent bureaucrat behave. If they behave in the same way, it is more difficult to select a competent successor even if the leader works, and hence the disciplinary effect of his selection concerns is weaker. In Equilibrium D, both types of bureaucrats behave in the same way ($\varepsilon_h^B = \varepsilon_l^B = 0$), and hence selection is more difficult, and a sufficiently low a is needed to sustain Equilibrium D. In Equilibrium C, different types of bureaucrats behave differently ($\varepsilon_h^B = 0 \neq 1 = \varepsilon_l^B$), and hence selection is easier, and Equilibrium C can survive on bigger a. This results in trapezoid D being "on the left" of trapezoid C. However, to sustain Equilibrium C, a sufficiently high $\tilde{a}$ is also needed to lure an incompetent bureaucrat to behave differently from a competent one, which is what makes selection easier in the first place. When $\tilde{a}$ is lowered while a remains in the above mentioned range, neither Equilibrium C nor Equilibrium D can survive.

Proposition 9 *A decrease in the private benefits of embezzlement (a or $\tilde{a}$), while holding other parameters intact, may paradoxically lower welfare by leaving the trivial, corruptive* *equilibrium as the only sustainable equilibrium.*

Interestingly, it seems that at a gut level Mao Zedong recognized the potential danger of lowering a or $\tilde{a}$ in a political system building on career concerns instead of checks and balances, as demonstrated by his famous warning that "pure water kills the fishes" (*shui zhi qing ze wu yu*).

5.6 The Importance of State Capacity

In this and the next sections, we extend our model to investigate how the capacity and the structure of the state affect the quality of a political system building on career concerns instead of checks and balances.

We have previously assumed that the maximum level of public good provision is 1 per bureaucrat, and the private benefits of embezzlement are a and $\tilde{a}$ for leaders and bureaucrats, respectively. In this and the next sections, we extend our model by introducing government budgets as exogenous variables and using these variables to endogenize the maximum level of public good provision and the private benefits of embezzlement. To do so, let's assume that the public good production function in each of the local government is

$$F = \beta x^{\delta} \tilde{x}^{1-\delta},$$

where F is the level of public good produced, β is a productivity parameter, and $\delta \in (0,1)$ is a weight parameter. The variables x and $\tilde{x}$ are the outputs of the leader and of the corresponding bureaucrat, respectively.

We assume that the output of any government official equals to his budget if he succeeds in performing, and equals 0 otherwise. Let α and $\tilde{\alpha}$ be the budgets of the leader and of a bureaucrat, respectively. The level of public good provision, F, in any local government hence equals to $f := \beta \alpha^{\delta} \tilde{\alpha}^{1-\delta}$ if both the leader and the corresponding bureaucrat succeed in performing, and equals to 0 otherwise.

Embezzling the budget hence brings a leader and a bureaucrat private benefits α and $\tilde{\alpha}$, respectively. In this section, we treat these budget parameters as exogenous, and focus our attention on the productivity parameter β, which we take as measuring the state capacity, as a more effective government can deliver the same amount of public good with smaller budgets.[12]

In our model, a government official's payoffs are linear in public good consumption and private benefits of embezzlement, and hence his decisions are invariant to positive scalar multiplication of $(f, \alpha, \tilde{\alpha})$. Therefore, our preceding analysis remains intact if we normalize the maximum level of public good provision to 1 and the private benefits of embezzlement to the following ratios:

$$a = \frac{\alpha}{f} \qquad \text{and} \qquad \tilde{a} = \frac{\tilde{\alpha}}{f}. \tag{11}$$

[12]Besley and Persson (2009, 2010) discuss two kinds of state capacity: the capacity to support markets and the capacity to levy tax. Our notion of state capacity is closer to the first kind of theirs.

Given any central and local government budgets, $(\alpha, \tilde{\alpha})$, the resulting $(a, \tilde{a})$ must satisfy

$$a^{\delta}\tilde{a}^{1-\delta} = \frac{1}{\beta}. \tag{12}$$

Conversely, for any pair $(a, \tilde{a})$ that satisfy equation (12), we can construct central and local government budgets $(\alpha, \tilde{\alpha})$ such that (11) holds.[13] Therefore, given β, equation (12) represents the set of all $(a, \tilde{a})$-pairs that can be resulted from different choices of central and local government budgets.

In Figure 2, we super-impose the set of all $(a, \tilde{a})$-pairs satisfying equation (12) onto Figure 1. A bigger β moves the locus of equation (12) closer to the origin. When $\beta \to 0$, the locus of equation (12) becomes arbitrarily far away from the origin. From Figure 2, it is readily verified that equilibria with limited accountability can exist only if β is big enough.[14] When state capacity is low enough, the only sustainable equilibrium is the trivial, corruptive equilibrium, regardless of the central and local government budgets.

Proposition 10 *When state capacity (measured by β) is low enough, the only sustainable equilibrium is the trivial, corruptive equilibrium, regardless of the central and local government budgets.*

Why does state capacity have a natural role to play in a political system building on career concerns instead of checks and balances? The intuition is that career concerns alone is not enough to substitute for checks and balance as a source of discipline for government officials. In the specific model we study here, career concerns "work" only when they are supplemented by the leaders' selection concerns. However, when state capacity is low, selecting competent successors is not that important, which weakens the disciplinary effect of leaders' selection concerns, resulting in the failure of such a political system.

5.7 The Importance of Moderate Decentralization

The extension of our model outlined in the previous section also highlights the importance of moderate decentralization. We can think of the ratio of central and local government budgets, $\alpha/\tilde{\alpha}$, as a measure of how centralized or decentralized the government structure is. As evident from Figure 2, even when state capacity is strong enough (β big enough) so that the locus of equation (12) passes through the regions where non-trivial equilibria can exist, the existence of non-trivial equilibria still relies on $\alpha/\tilde{\alpha}$ taking intermediate values.

[13]Let

$$\sigma = \frac{(a\beta)^{1/2(1-\delta)}}{(\tilde{a}\beta)^{1/2\delta}}.$$

and pick any $(\alpha, \tilde{\alpha})$ such that $\alpha/\tilde{\alpha} = \sigma$. It can be readily verified that (11) will be satisfied.

[14]We have not drawn in Figure 1 the range of parameters that sustain Equilibrium E. From Lemma 5, it can be easily seen that the range, if non-empty, takes the shape of a triangle, and hence is also bounded in size.

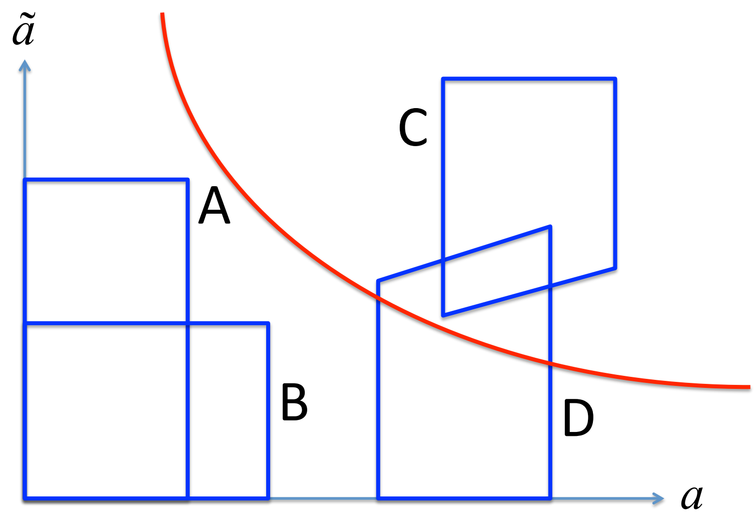

Figure 2: Locus of equation (12).

Proposition 11 *For any given level of state capacity (measured by β), if the government structure is very centralized ($\alpha/\tilde{\alpha} \to \infty$), or if it is very decentralized ($\alpha/\tilde{\alpha} \to 0$), the only sustainable equilibrium is the trivial, corruptive equilibrium.*

Why does moderate decentralization have a natural role to play in a political system building on career concerns instead of checks and balances? Again, the intuition lies in the supplementary role of selection concerns in association with career concerns. When the government is too centralized, most of the government budget concentrates in the hand of the leader. The temptation of embezzlement for a leader is thus too big compared to his selection concerns, resulting in the failure of such a political system. When the government is too decentralized, most of the government budget concentrates in the hand of the bureaucrats. The temptation of embezzlement for bureaucrats is thus too big compared to his career concerns, resulting in widespread corruption at the local government level. This, in turn, makes selecting a competent successor difficult for a retiring leader. Anticipating that, a leader also faces weak disciplinary effect from his selection concerns, resulting in the failure of such a political system.

6 Conclusion

We have proposed a stylized model of overlapping principal-agent problems to theorize how a political system building on career concerns instead of checks and balances may function. The key question we asked is how such a political system may generate accountability at the top of the government—where there are no more career concerns. Our model offered one possible answer to this important question. Leaders at the top of the government may refrain from abusing their absolute power if doing so may decrease their ability to distinguish different types of bureaucrats, and hence diminish their prospects of selecting/promoting their favorite type of bureaucrats as the next-generation leaders. In other words, while accountability-at-the-bottom is generated by career concerns, accountability-at-the-top is generated by selection concerns. Our model yields seven implications. First, any exogenous factor that weakens one of these concerns will also weaken the other and results in the lack of accountability at all levels of the government. Second, there is an inherent indeterminacy in the disciplinary effect of career concerns in such a political system. Third, achieving full accountability at every level of the government in such a political system is impossible. Fourth, institutionalizing a merit-based promotion rule may inadvertently render such a political system unworkable. Fifth, a small decrease in the private gain of power abuse—perhaps because of a small improvement in the rule of law—may inadvertently render such a political system unworkable. Sixth, the proper functioning of such a political system relies on strong enough state capacity. And, finally, the proper functioning of such a political system relies on a somwhat balanced state structure, neither over-centralized nor over-decentralized.

7 References

Acemoglu, Daron, Georgy Egorov, and Konstantin Sonin (2010), "Political Selection and Persistence of Bad Governments." *Quarterly Journal of Economics*, 125(4), 1511-1576.

Acemoglu, Daron, Michael Kremer, and Atif Mian (2008), "Incentives in Markets, Firms, and Governments." *Journal of Law, Economics, and Organization*, 24(2), 273-306.

Alesina, Alberto and Guido Tabellini (2007), "Bureaucrats or Politicians? Part I: A Single Policy Task." *American Economic Review*, 97(1), 169-179.

Alesina, Alberto and Guido Tabellini (2008), "Bureaucrats or Politicians? Part II: Multiple Policy Tasks." *Journal of Public Economics*, 92, 426-447.

Besley, Timothy (2005), "Political Selection." *Journal of Economic Perspectives*, 19(3), 43-60.

Besley, Timothy and Torsten Persson (2009), "The Origins of State Capacity: Property Rights, Taxation, and Politics." *American Economic Review*, 99(4), 1218-1244.

Besley, Timothy and Torsten Persson (2010), "State Capacity, Conflict, and Development." *Econometrica*, 78(1), 1-34.

Besley, Timothy and Masayuki Kudamatsu (2008), "Making Autocracy Work." in Elhanan Helpman (ed) *Institutions and Economic Performance*, Cambridge: Harvard University Press.

Besley, Timothy, and Maitreesh Ghatak (2005), "Competition and Incentives with Motivated Agents." American Economic Review, 616-636.

Che, Jiahua, Kim-Sau Chung, and Xue Qiao (2013), "The Good, the Bad, and the Civil Society", *Journal of Public Economics* Vol 106: 68-76.

Dewatripont, Mathias, Ian Jewitt, and Jean Tirole (1999). "The Economics of Career

Concerns, part II: Application to missions and accountability of government agencies." *The Review of Economic Studies*, 66(1), 199-217.

Holmström, Bengt (1982), "Managerial Incentive Problems: A Dynamic Perspective", in *Essays in Economics and Management in Honor o f Lars Wahlbeck* (Helsinki: Swedish School of Economics).

Holmström, Bengt (1999). "Managerial Incentive Problems: A Dynamic Perspective." *The Review of Economic Studies*, 66(1), 169-182.

Kung, James Kai-Sing and Shuo Chen (2011), "The Tragedy of the *Nomenklatura*: Career Incentives and Political Radicalism during China's Great Leap Famine." *American Political Science Review*, vol 105, no 1: 27-45.

Li, Hongbin and Li-An Zhou (2005), "Political Turnover and Economic Performance: The Incentive Role of Personnel Control in China." *Journal of Public Economics* vol 89, no 9/10: 1743-62.

Li, Wei, and Dennis Tao Yang (2005), "The Great Leap Forward: Anatomy of a Central Planning Disaster." *Journal of Political Economy,* 113(4), 840-877.

Maskin, Eric and Jean Tirole (2004), "The Politician and the Judge: Accountability in Government." *American Economic Review*, 94, 1034-1054.

Meyer, Margaret, and John Vickers (1997), "Performance Comparisons and Dynamic Incentives." *Journal of Political Economy*, 105(3), 547-581.

Myerson, Roger (2010), "Moral Hazard in High Office and the Dynamics of Aristocracy." Unpublished Manuscript, University of Chicago.

Olson, Mancur (1993), "Dictatorship, Democracy, and Development." *American Political Science Review*, 567-576.

Prendergast, Canice (2008), "Intrinsic Motivation and Incentives." *The American Economic Review*, 98(2), 201-205.

Rauch, James (2001), "Leadership Selection, Internal Promotion, and Bureaucratic Corruption in Less Developed Polities." *Canadian Journal of Economics*, v 34, n 1: 240-258.

Tirole, Jean (1994), "The Internal Organization of Government." *Oxford Economic Papers*, 1-29.

Xu, Chenggang (2011), "The fundamental institutions of China's reforms and development." *Journal of Economic Literature* 49(4), 1076-1151.

8 Appendix

8.1 Formulas for $m_\theta(\cdot)$ and $n_\theta(\cdot)$

The computation of $m_\theta(\cdot)$ and $n_\theta(\cdot)$ is straightforward though tedious, and their formulas are listed below:

$$m_h(0) = \begin{cases} \rho\left[\varepsilon_h^B + (1-\varepsilon_h^B)\frac{1}{2}\right] + (1-\rho)\left[\varepsilon_l^B + (1-\varepsilon_l^B)\left(\frac{\gamma}{2} + (1-\gamma)\right)\right] & \text{if} \quad r = mb \\ \rho\left[(1-\varepsilon_h^B)\frac{1}{2}\right] + (1-\rho)\left[(1-\varepsilon_l^B)\frac{\gamma}{2}\right] & \text{if} \quad r = amb \end{cases};$$

$$m_h(1) = \begin{cases} \rho\left[\varepsilon_h^B\frac{1}{2}\right] + (1-\rho)\left[\varepsilon_l^B\frac{1}{2} + (1-\varepsilon_l^B)\frac{1-\gamma}{2}\right] & \text{if} \quad r = mb \\ \rho\left[\varepsilon_h^B\frac{1}{2} + (1-\varepsilon_h^B)\right] + (1-\rho)\left[\varepsilon_l^B\frac{1}{2} + (1-\varepsilon_l^B)\left(\gamma + \frac{1-\gamma}{2}\right)\right] & \text{if} \quad r = amb \end{cases};$$

$$m_l(0) = \begin{cases} \rho\left[\varepsilon_h^B\left(\gamma + \frac{1-\gamma}{2}\right) + (1-\varepsilon_h^B)\frac{\gamma}{2}\right] + (1-\rho)\left[\varepsilon_l^B\left(\gamma + \frac{1-\gamma}{2}\right) + (1-\varepsilon_l^B)\frac{1}{2}\right] & \text{if} \quad r = mb \\ \rho\left[\varepsilon_h^B\frac{1-\gamma}{2} + (1-\varepsilon_h^B)\left(\frac{\gamma}{2} + (1-\gamma)\right)\right] + (1-\rho)\left[\varepsilon_l^B\frac{1-\gamma}{2} + (1-\varepsilon_l^B)\frac{1}{2}\right] & \text{if} \quad r = amb \end{cases};$$

$$m_l(1) = \begin{cases} \rho\left[\varepsilon_h^B\frac{1}{2}\right] + (1-\rho)\left[\varepsilon_l^B\frac{1}{2} + (1-\varepsilon_l^B)\frac{1-\gamma}{2}\right] & \text{if} \quad r = mb \\ \rho\left[\varepsilon_h^B\frac{1}{2} + (1-\varepsilon_h^B)\right] + (1-\rho)\left[\varepsilon_l^B\frac{1}{2} + (1-\varepsilon_l^B)\left(\gamma + \frac{1-\gamma}{2}\right)\right] & \text{if} \quad r = amb \end{cases};$$

$$n_h(0) = \begin{cases} \rho + (1-\rho)\left[\varepsilon_l^B + (1-\varepsilon_l^B)\left(\frac{\gamma}{2} + (1-\gamma)\right)\right] & \text{if} \quad r = mb \\ \rho + (1-\rho)\left[(1-\varepsilon_l^B)\frac{\gamma}{2}\right] & \text{if} \quad r = amb \end{cases};$$

$$n_h(1) = \begin{cases} \rho + (1-\rho)\left[\varepsilon_l^B\frac{1}{2} + (1-\varepsilon_l^B)\frac{1-\gamma}{2}\right] & \text{if} \quad r = mb \\ \rho + (1-\rho)\left[\varepsilon_l^B\frac{1}{2} + (1-\varepsilon_l^B)\left(\gamma + \frac{1-\gamma}{2}\right)\right] & \text{if} \quad r = amb \end{cases};$$

$$n_l(0) = \begin{cases} \rho\left[\varepsilon_h^B\left(\frac{1-\gamma}{2}\right) + (1-\varepsilon_h^B)\left(\frac{\gamma}{2} + (1-\gamma)\right)\right] & \text{if} \quad r = mb \\ \rho\left[\varepsilon_h^B\left(\gamma + \frac{1-\gamma}{2}\right) + (1-\varepsilon_h^B)\frac{\gamma}{2}\right] & \text{if} \quad r = amb \end{cases};$$

$$n_l(1) = \begin{cases} \rho\left[\varepsilon_h^B\frac{1}{2} + (1-\varepsilon_h^B)\right] & \text{if} \quad r = mb \\ \rho\left[\varepsilon_h^B\frac{1}{2}\right] & \text{if} \quad r = amb \end{cases}.$$

The reader can easily verify that Δm_θ and Δn_θ indeed have the simple expressions in the main text.

8.2 Omitted Proofs

Proof of Lemma 1:

We look for conditions on $(a, \tilde{a}, \rho, \gamma)$ such that $(\varepsilon_h^L, \varepsilon_l^L, \varepsilon_h^B, \varepsilon_l^B, r) = (0, 0, 0, 1, mb)$ satisfies (5), (6), (7), and (8). (5) is always satisfied because $\Delta q = 2\rho(1-\rho) > 0$ and $\Delta y = 2\rho(1-\gamma) > 0$. (6) is more stringent than (7), which in turn is satisfied if and only if $\gamma(q_{mb} - \rho)\Delta y \leq a$, which reduces to Condition 1 in the Lemma. (8) is satisfied if and only if $\tilde{a} + \Delta n_h \cdot \Delta Y \leq 0$ and $\tilde{a} + \Delta n_l \Delta Y \geq 0$. The former reduces to Condition 2 in the Lemma; while the latter is always satisfied because $\Delta n_l = \rho\gamma/2 > 0$ and $\Delta Y = (1-\gamma)\rho(1-\rho)\Delta y > 0$. ■

Proof of Lemma 2:

We look for conditions on $(a, \tilde{a}, \rho, \gamma)$ such that $(\varepsilon_h^L, \varepsilon_l^L, \varepsilon_h^B, \varepsilon_l^B, r) = (0, 0, 1, 0, amb)$ satisfies (5), (6), (7), and (8). (5) is always satisfied because $\Delta q = -2\rho(1-\rho)\gamma < 0$ and $\Delta y = 2(1-\rho)\gamma(1-\gamma) > 0$. (6) is more stringent than (7), which in turn is satisfied if and only if $\gamma(q_{amb} - \rho)\Delta y \leq a$, which reduces to Condition 1 in the Lemma. (8) is satisfied if and only if $\tilde{a} + \Delta n_h \cdot \Delta Y \geq 0$ and $\tilde{a} + \Delta n_l \Delta Y \leq 0$. The latter reduces to Condition 2 in the Lemma; while the former is always satisfied because $\Delta n_h = (1-\rho)/2 > 0$ and $\Delta Y = \rho(1-\rho)\gamma(1-\gamma)\Delta y > 0$. ■

Proof of Lemma 3:

We look for conditions on $(a, \tilde{a}, \rho, \gamma)$ such that $(\varepsilon_h^L, \varepsilon_l^L, \varepsilon_h^B, \varepsilon_l^B, r) = (0, 1, 0, 1, mb)$ satisfies (5), (6), (7), and (8). (5) is always satisfied because $\Delta q = 2\rho(1-\rho) > 0$ and $\Delta y = 2\rho > 0$. (6) and (7) are satisfied if and only if $(q_{mb} - \rho)\Delta y \geq a \geq \gamma(q_{mb} - \rho)\Delta y$, which reduces to Condition 1 in the Lemma. (8) is satisfied if and only if $\tilde{a} + \Delta n_h \cdot \Delta Y \leq 0$ and $\tilde{a} + \Delta m_l a + \Delta n_l \Delta Y \geq 0$. The former reduces to Condition 2 in the Lemma; while the latter reduces to Condition 3 in the Lemma. ■

Proof of Lemma 4:

We look for conditions on $(a, \tilde{a}, \rho, \gamma)$ such that $(\varepsilon_h^L, \varepsilon_l^L, \varepsilon_h^B, \varepsilon_l^B, r) = (0, 1, 0, 0, mb)$ satisfies (5), (6), (7), and (8). (5) is always satisfied because $\Delta q = 2\rho(1-\rho)(1-\gamma) > 0$ and $\Delta y = 2(\rho(1-\gamma) + \gamma) > 0$. (6) and (7) are satisfied if and only if $(q_{mb} - \rho)\Delta y \geq a \geq \gamma(q_{mb} - \rho)\Delta y$, which reduces to Condition 1 in the Lemma. (8) is satisfied if and only if $\tilde{a} + \Delta n_h \cdot \Delta Y \leq 0$ and $\tilde{a} + \Delta m_l a + \Delta n_l \Delta Y \leq 0$. The former reduces to Condition 2 in the Lemma; while the latter reduces to Condition 3 in the Lemma. ■

Proof of Lemma 5:

We look for conditions on $(a, \tilde{a}, \rho, \gamma)$ such that $(\varepsilon_h^L, \varepsilon_l^L, \varepsilon_h^B, \varepsilon_l^B, r) = (0, 1, 1, 0, amb)$ satisfies (5), (6), (7), and (8). (5) is always satisfied because $\Delta q = -2\rho(1-\rho)\gamma < 0$ and $\Delta y = 2(1-\rho)\gamma > 0$. (6) and (7) are satisfied if and only if $(q_{amb} - \rho)\Delta y \geq a \geq \gamma(q_{amb} - \rho)\Delta y$, which reduces to Condition 1 in the Lemma. (8) is satisfied if and only if $\tilde{a} + \Delta n_h \cdot \Delta Y \geq 0$ and $\tilde{a} + \Delta m_l a + \Delta n_l \Delta Y \leq 0$. The latter reduces to Condition 2 in the Lemma; while the former is always satisfied because $\Delta n_h = (1-\rho)/2 > 0$, and $\Delta Y = \rho(1-\rho)\gamma(1-\gamma) > 0$. ■

In the main text, we claim that there are only six possible equilibria (described respectively in Lemmas 2 and 1–5). We shall verify this claim here.

Lemma 6 *There is no equilibrium other than the six ones described in Lemmas 2 and 1–5.*

Proof: There are four possible combinations for $(\varepsilon_h^L, \varepsilon_l^L)$. Any equilibrium with $(\varepsilon_h^L, \varepsilon_l^L) = (1, 1)$ is observationally indistinguishable from the trivial, corruptive equilibrium as described in Lemma 2. Any equilibrium with $(\varepsilon_h^L, \varepsilon_l^L) = (1, 0)$ is precluded by Proposition

3.

Consider any equilibrium with $(\varepsilon_h^L, \varepsilon_l^L) = (0,0)$. The proof of Lemma 1 shows that, if $(\varepsilon_h^B, \varepsilon_l^B) = (0,1)$, then we must have $r = mb$, in which case we have Equilibrium A as described in Lemma 1. Similarly, the proof of Lemma 2 shows that, if $(\varepsilon_h^B, \varepsilon_l^B) = (1,0)$, then we must have $r = amb$, in which case we have Equilibrium B as described in Lemma 2. The case of $(\varepsilon_h^B, \varepsilon_l^B) = (1,1)$ is precluded by Proposition 4. Suppose $(\varepsilon_h^B, \varepsilon_l^B) = (0,0)$. Then we have an equilibrium with full accountability, which is impossible by Proposition 1.

Consider any equilibrium with $(\varepsilon_h^L, \varepsilon_l^L) = (0,1)$. The proof of Lemma 3 shows that, if $(\varepsilon_h^B, \varepsilon_l^B) = (0,1)$, then we must have $r = mb$, in which case we have Equilibrium C as described in Lemma 3. Similarly, the proof of Lemma 4 shows that, if $(\varepsilon_h^B, \varepsilon_l^B) = (0,0)$, then we must have $r = mb$, in which case we have Equilibrium D as described in Lemma 4. Finally, the proof of Lemma 5 shows that, if $(\varepsilon_h^B, \varepsilon_l^B) = (1,0)$, then we must have $r = amb$, in which case we have Equilibrium E as described in Lemma 5. The case of $(\varepsilon_h^B, \varepsilon_l^B) = (1,1)$ is precluded by Proposition 4. ■