

Rotation, Performance Rewards, and Property Rights

Weijia Li*

January 10, 2019

Abstract

Economic growth needs a strong and well-functioning government. But a government too strong can also dominate the private sector, leading to severe holdup problems. This paper studies how to constrain politicians through personnel rules, with a special focus on rotation and performance evaluation. Through a game theoretic model, I show that rotation or performance evaluation alone actually makes holdup problems even worse. But it is exactly their combination that covers each other's weakness and solves holdup problems together. Frequently rotated and carefully evaluated, politicians would also develop no entrenched interests in existing firms. This helps avoid crony capitalism and encourages Schumpeterian "creative destruction", solving another key problem with government-assisted development. Thus, rotation and performance rewards resolve the acute tradeoff between commitment and flexibility, a feature rarely satisfied by other commitment devices. Firm-level panel data from China are consistent with the main predictions of the model.

*Department of Economics, Monash University. Email: weijia.li@monash.edu. I am especially indebted to Gerard Roland, Ernesto Dal Bo, Peter Lorentzen, and Yang Xie for their invaluable advice, guidance, and encouragement. This work has also benefited particularly from the comments of Efthymios Athanasiou, Chong'en Bai, Arthur Campbell, Douglas Campbell, Kang Chen, Zeger Van Der Wal, Fred Finan, Gunes Gokmen, Yuriy Gorodnichenko, Sergei Izmalkov, Katerina Jensen, Ruixue Jia, Claudio Labanca, Matthew Leister, Lixing Li, Wei Li, Xi Lu, Yi Lu, Pushkar Maitra, Andrei Markevich, Marta Troya Martinez, Vai-Lam Mui, Martha Olney, Laura Puzello, Xue Qiao, Danny Quah, Ramkishan Rajan, Michelle Rendall, Michele Valsecchi, Brian Viard, Shaoda Wang, Yijiang Wang, Michael Ward, Binzhen Wu, Tianyang Xi, Guo Xu, Chenggang Xu, Yiqing Xu, Yang Yao, Leslie Young, and Hosny Zoabi. I am also grateful for comments from the participants of UCSD Young Scholar Conference, WPSA Mini-Conference on Non-democratic Politics, and seminars in UC Berkeley, Cheong Kung Graduate School of Business, Monash University, the National University of Singapore, New Economic School, Peking University, and Tsinghua University. I thank Shuo Chen's effort in making personnel data on city-level officials in China publicly available. This research is supported by the Helen Gan and Richard Aston Fellowship for China Studies and the John L. Simpson Memorial Research Fellowship for Comparative Studies. All errors are my own.

1. Introduction

The holdup problem is destructive to economic growth, especially in environments without well-defined property rights (Williamson, 1985; Grossman and Hart, 1986). When constitutional constraints on politicians are weak, entrepreneurs anticipate excessive *ex post* extraction from politicians (Persson and Tabellini, 2000; Acemoglu, 2003). Thus, authoritarian regimes have to rely on alternative constraints on politicians to solve the holdup problem against entrepreneurs. Many papers show that politicians can be disciplined by performance-based rewards that compensate (local) politicians for economic growth, especially promotion for growth (Maskin et al. 2000; Lazear and Oyer, 2012). However, models of performance-based rewards usually shy away from the holdup problem (Laffont, 2001), the very dilemma behind the urgency of strong property rights (Grossman and Hart, 1986).

Alternatively, politicians can enter a reciprocal relationship with the entrepreneurs so that reputation concerns can constrain the politicians. Among the many problems of reputation-based solutions (Hart, 1995), politicians inevitably develop entrenched interests in their cronies so that the politicians block the entrance of new firms and the associated “creative destruction”, endangering sustainable growth (Aghion and Howitt, 1995; Acemoglu et al. 2006). In other words, reputation-based solutions cannot extend credible protection to new firms.

I propose that political rotation is at the heart of credible constraints on politicians in authoritarian regimes, especially local politicians. Namely, political rotation provides *de facto* property rights for entrepreneurs that guard against politicians’ extraction. The starting point of this proposition is that a politician with prolonged tenure can become very knowledgeable about local conditions, which endows her with formidable power. To formalize the idea, I allow a politician to acquire detailed information about an entrepreneur’s project, especially the profitability of the project. Equipped with such information, the politician wields enormous *ex post* bargaining power against the entrepreneur when they bargain on the division of the surplus. Notice that power and information are interchangeable in my model. Political rotation plays a decisive role in alleviating the acute holdup problem. If a politician is rotated to a new jurisdiction, the information about the last jurisdiction becomes completely useless. This discourages the politician from power/information acquisition, establishing credible property rights to the entrepreneur.

The bulk of the paper investigates the interaction between rotation and performance-based rewards. I show that stronger performance-based rewards exacerbate the holdup problem, requiring more intense political rotation. Note that information acquisition helps the politician to avoid bargaining breakdown with the entrepreneur. This is because a

politician fully informed about the project can extract rent based on the realized profitability of the project so that the entrepreneur always accepts the extraction. Consequently, an informed/empowered politician can ensure that the project will always be finished, thus reaping the full benefit of performance-based rewards. By contrast, a politician uninformed about the project cannot condition her rent extraction on the realized profitability. An entrepreneur whose project yields fewer profits than the proposed extraction will reject cooperation with the politician, engendering substantial likelihood of bargaining breakdown. Thus, an uninformed/disempowered politician can capture only a partial benefit from performance-based rewards. To summarize, stronger rewards based on economic performance raise the return for a politician to empower herself, a temptation that has to be discouraged by more intense rotation. *When performance-based rewards are more intense, political rotation has to be more frequent.*

On the other hand, excessive political rotation itself can be disastrous. This is the traditional view that political rotation creates “roving bandits” (Olson, 1993). In my model, excessive rotation induces the politician to confiscate the capital invested by the entrepreneur into the project. The investment is necessary for the entrepreneur’s project to be productive, so capital confiscation denies future production. However, a politician with short tenure loses little by pursuing such shortsighted policy. The roving-bandit concern puts an upper bound on rotation frequency, while the prior analysis on information acquisition identifies the lower bound. Importantly, the upper bound also increases with performance-based rewards, which raises the stake of future production opportunity and the cost of capital confiscation. Stronger performance-based rewards make frequent rotation possible by counterbalancing the temptation to confiscate capital. *Thus, rotation and performance rewards are supporting each other. Rotation discourages information acquisition tempted by performance-based rewards. Performance rewards disincentivize shortsighted policy, which allows frequent rotation.*

Political rotation also eliminates entrenched interests and facilitates “creative destruction”. Thus, the solution to the holdup problem by rotation and performance-based rewards overcomes the standard trade-off between commitment and flexibility (Levin, 2003; Chas-sang, 2010). Suppose that a new project better than the old one may arrive. The politician has no time to learn about the new project; she can support either the old project or the new project, but not both. Without rotation, the informed politician can take all rents away from the old project, but as she is uninformed about the new project, she has to award significant information rents to the new entrepreneur. Thus, even if the new project is better than the old project, the politician will probably stick with the old one. Performance-based rewards induce politicians to focus on productivity rather than economic rents, so they can partially

reduce entrenched interests. With sufficient rotation, the politician is equally uninformed about the old and new projects. This situation completely eliminates entrenched interests, so the politician will endorse the new project that is on average “better” than the old one. Rotation complemented by performance-based rewards provides both flexibility and *de facto* property rights that render it possible to sustain long-term growth.

I confront predictions of the theory with Chinese firm-level data matched with politician characteristics. Specifically, the theory should apply especially well to investment in equipment and buildings that are difficult to move. Better political incentives should not boost firms’ holdings in liquid assets such as cash and intellectual property. A firm with large liquid assets can simply move to another city when predation is looming. The data shows that a firm indeed holds more physical capital when it anticipates rotation and promotion of the mayor of the city where the firm locates. The effect is especially strong for young mayors new in office, who face much stronger personnel incentives. For different types of ownership, private firms are especially responsive to anticipated rotation and promotion while state-owned enterprises (SOEs) are much less affected. This is consistent with the theory because SOEs already enjoy substantial bargaining power against local politicians. Taken together, the empirical evidence lends high credibility to the theory.

2. Literature

Starting from the influential work of Olson (1993), the political economy literature usually treats the “roving-bandits” created by political rotation as a formidable hurdle to reliable property rights (Rose-Ackerman and Palifka, 2016). My paper shows that the traditional view is appropriate when it abstracts political rotation as a stand-alone institution, but less so when rotation is one of the building blocks in an institutional cluster. My analysis demonstrates that political rotation and performance-based rewards interact in an intriguingly symbiotic way. The two institutions can be destructive as stand-alone arrangements, but their interaction helps each other restore desired disciplinary effects.

There is a relatively separate literature in personnel economics on rotation. Most of these studies take a human capital approach. They show that lateral moves of employees help 1) the employer learn the ability of the subordinates (Ortega, 2001; Eriksson and Ortega, 2006) and 2) the employees accumulate a diverse set of skills that improves their eligibility for promotion (Eriksson and Ortega, 2006; Friebel and Raith, 2014; Jin and Waldeman, 2017). My paper takes a distinct political approach to understanding rotation. The perspective of political economy proves to be very fruitful in generating fresh insights on the role of rotation.

In addition, Friebel and Raith (2014) and Jin and Waldeman (2017) focus on the interaction between rotation and narrowly-defined promotion, while performance-based rewards in my paper are easily applicable to many personnel phenomena.

Performance-based rewards are a key topic in political economy and economics of organizations (Roland, 2000; Lazear and Oyer, 2012). Researchers advocate and formalize many ideas why performance-based rewards may misfire and create unintended consequences (for example, see Gibbons, 1987 and Holmstrom and Milgrom, 1991). To the best of my knowledge, my paper is the first attempt to understand how performance rewards may exacerbate the temptation to accumulate information-based power. By focusing on the holdup problem, my theory explains why high-powered incentive, not carefully designed, can be especially destructive in autocracy. The insight is further applicable to generic organizations, where a paramount concern of headquarters is the potential loss of control over divisions (Qian, 1994). Apart from highlighting this dilemma, my paper also systematically explores the institutional solution that can limit the negative side of performance-based rewards.

My paper is closely related to the large literature on political connection and economic outcomes. Many influential papers demonstrate various ways that political connection can increase business value and economic efficiency (Khwaja and Mian, 2008; Ferguson and Voth, 2008; Bai et al. 2014). However, it is puzzling that private business can reap such huge benefit from political connection, as by definition politicians command formidable coercive power that should enable them to capture the bulk of the surplus (Weingast, North 1989). Moreover, political connection seems to matter most where constitutional constraints on politicians are weak. This makes it especially thought-provoking to investigate the alternative constraints on authoritarian politicians that justify the value of political connections for entrepreneurs.

The literature on state capacity (Besley and Persson, 2011) focuses on how fiscal capacity and legal capacity engender economic growth together. Less attention has been paid to how bureaucratic capacity is decisive for economic outcomes, which is the focus of the state capacity literature in its original contributions (such as Mann, 1986 and Evans and Rauch, 1999). My research offers a concrete micro-foundation on the importance of bureaucratic capacity, defined as impersonal rules and controls over bureaucrats and politicians (Evans and Rauch, 1999).

A huge literature shows that Schumpeterian “creative destruction” engenders sustainable growth (Aghion and Howitt, 1992; Aghion et al. 2013). It is well known that the key threat to “creative destruction” is the entrenched interests of powerful politicians who benefit from their cronies’ firms (Olson, 1983; Aghion and Howitt, 1996; Acemoglu et al. 2006).

This dilemma constitutes the foundational force behind the “middle-income trap” of many economies that rely on crony capitalism to initiate industrialization (Acemoglu et al. 2006; Acemoglu and Robinson, 2012). My analysis shows that the problem is solvable by institutionalized personnel control over politicians using strong rotation and performance-based rewards. This argument is consistent with the empirical literature showing that “creative destruction” can be active in emerging economies with little legal capacity but high state capacity (Brandt et al., 2012).

3. A Workhorse Model

3.1 Performance-based Rewards in the Benchmark Model

In a reduced-form manner, this section illustrates the basic insights of models on performance-based rewards. It serves as the key building block to formalize my own ideas in later sections. There are three players: a principal, a local politician, and an entrepreneur. The principal is the central authority who implements personnel policy. The game has two periods. The timeline and strategies of players are as follows:

At $t = 1$, the entrepreneur arrives and decides whether to invest in a project that costs him $k > 0$. The project’s output y is uncertain. It has a distribution $F(y)$ continuously distributed on $[y, \bar{y}]$ with the cumulative distribution function $F(\cdot)$, $F' \equiv f \cdot \frac{f(y)}{1-F(y)}$ increases in y ; in other words, the distribution satisfies the monotone hazard rate property. If the entrepreneur does not invest, the game ends and all players get 0. If he pays k , the project begins but remains unfinished.

After the entrepreneur pays k , the profitability of the project y is revealed to the entrepreneur. The principal and the local politician do not know y , but they both know $F(y)$.

At the end of $t = 1$, the principal rotates the politician with probability π and appoints another one, who will take office in $t = 2$. If the politician in $t = 1$ is rotated, she gets an exogenous payoff of \tilde{U} by serving in another jurisdiction.

At $t = 2$, the politician (either the same one as in $t = 1$ or a newly appointed politician) proposes to extract w from the entrepreneur. If the entrepreneur accepts, the politician lobbies the central government, and the value of the project y is fully realized. The entrepreneur gets $y - w - k$, as he cares about the net profit. The principal wants a large total output, so she gets y . For the politician, she gets $w + Ry$. The politician’s payoff has two parts: w is the economic rent, and Ry represents the performance-based rewards, where the exogenous parameter R measures the intensity of performance rewards. A large R means

that performance rewards are highly valued by the politician.

R is the key parameter in the model. A preferred interpretation is that R measures how much the politician values promotion opportunities. If R is large, higher political positions are prized by the local politician. In this case, the politician strongly dislikes bargaining breakdown, which will demolish the valuable opportunity of promotion. If R is small, higher political positions are not very lucrative. The politician cares less about her career path than the current opportunity of rent extraction. It is straightforward to provide a micro-foundation that validates the interpretation of R as the value of promotion opportunity, as shown in Appendix 3. I also demonstrate that R can be interpreted as the likelihood of promotion for a given output level. Thus, a polity enjoys high R if the personnel turnover is very active, which can occur with mandatory retirement rule that releases many political positions regularly (Svolik, 2012). In this case, the prospect of promotion is realistic and reasonable likely, so promotion opportunity constitutes strong incentive imposed the politician. If the polity is a gerontocracy with little hope of promotion for young local politicians, the politician's career concern cannot form effective constraints on the politician. The micro-foundation in Appendix 3 focuses on the interpretation of Ry as promotion opportunity, as promotion is far more important in motivating politicians than direct monetary compensation (Maskin et al. 2000; Svolik, 2012). In general, it is still very useful to organize the basic story of performance-based rewards in a reduced-form manner. The insights can be applied to many forms of performance-based rewards, and all theoretical results will be robust regardless of the specific way performance rewards are implemented. It is worth noting that the linearity of Ry is purely for expositional purposes. Appendix 3 shows that all key results are unchanged under very general functional forms¹.

If the entrepreneur does not accept the extraction, the politician will not help the entrepreneur and the entrepreneur has to abandon the project. Although the project's potential productivity is y , the realized output is 0. The entrepreneur gets $-k$. The principal gets 0. The politician gets 0. The solution concept is sequential equilibrium.

The necessity for the entrepreneur and the politician to strike a bargain is a key assumption in the model. What is the reason that the entrepreneur cannot finish his project by himself? The rationale is that markets in many economies are highly regulated and frag-

¹In Appendix 3, the politician gets the performance rewards with probability $V(y, \alpha)$, and she values the performance rewards at R . Thus, the expected value of performance rewards is $V(y, \alpha)R$. I assume that $\frac{\partial V}{\partial y} > 0$, $\frac{\partial V}{\partial \alpha} > 0$: the politician is more likely to reap the reward if the output is higher, and α is a parameter of the function V . We can see that R and α correspond to the two interpretations: R measures the value of promotion opportunities, and α affects the likelihood of promotion for any output level. Importantly, the only assumptions I need to impose on $V(y, \alpha)$ is that $\frac{\partial V}{\partial y} > 0$, $\frac{\partial V}{\partial \alpha} > 0$, and $V(0, \alpha) = 0$, and all but one key results in this paper are preserved.

mented. On the one hand, the central government installs many entry barriers to protect firms that directly benefit the central government (e.g., Gordon and Li, 2009). On the other hand, politicians have a strong interest in enforcing trade barriers so that non-local firms can be excluded from the local market (e.g., Young, 2000). The consequence is that an entrepreneur can invest in capacity to produce plenty of goods. However, it is very difficult to sell them and realize the profit. Local politicians are playing a decisive role here. They have the necessary political capital to lobby the central government for a permit so that the local firm can enter the national market (Wedeman, 2011; Bai et al. 2014). They can also try to break down trade barriers by lobbying the central government or negotiating with other politicians. The full value of the entrepreneur's investment can be realized only through the mobilization of a politician's capacity and connection.

Regulations can be even more strict in markets of production factors. For example, regulations over labor market are especially punishing in most countries. In many cases, it is more effective for the entrepreneur to mediate labor disputes and to address labor unions through politicians. Otherwise, the entrepreneur's firm can be paralyzed and produces nothing. This further validates the notion that politician's inputs are indispensable for a successful firm, especially in developing economies.

Let us analyze the simple game. At $t = 2$, the entrepreneur accepts any extraction $w \leq y$. The politician, either a newly appointed one or the same one as in $t = 1$, proposes to extract:

$$w^* = \operatorname{argmax}_w [1 - F(w)] \{w + E[Ry|w \leq y]\}. \quad (1)$$

$[1 - F(w)] \{w + E[Ry|w \leq y]\}$ is the expected payoff to a politician who proposes to extract w from the entrepreneur. The politician faces some risk of bargaining breakdown: an entrepreneur with $y < w$ will reject the rent extraction rather than collaborating with the politician. In this case, the politician gets 0 from rent extraction and also gets 0 from performance rewards as the project will not be finished. The probability that there will be an agreement is $1 - F(w)$. The politician gets w from the rent extraction, and $E[Ry|y \geq w]$ from performance rewards. Thus, the utility of an uninformed politician who proposes to extract w is $[1 - F(w)] \{w + E[Ry|w \leq y]\} + F(w) * 0 = [1 - F(w)] \{w + E[Ry|w \leq y]\}$. In this case, the expected surplus captured by the entrepreneur is $[1 - F(w)] \{E[y|y \geq w] - w\}$. This is the probability that the project will be finished (which is $1 - F(w)$) times the net surplus to the entrepreneur (which is $E[y|y \geq w] - w$).

The first order condition for (1) with respect to w characterizes w^* , the optimal rent extracted by an uninformed politician:

$$\max_w [1 - F(w)]\{w + E[Ry | w \leq y]\} = \max_w \left\{ w \int_w^{\bar{y}} f(z) dz + R \int_w^{\bar{y}} z f(z) dz \right\}$$

$$w^* \int_{w^*}^{\bar{y}} f(z) dz - w^* f(w^*) - R w^* f(w^*) = 0$$

$$\frac{w^* f(w^*)}{1 - F(w^*)} = \frac{1}{1 + R}. \quad (2)$$

The monotone hazard rate property guarantees that w^* is unique. It also ensures that w^* decreases with R , so the uninformed politician reduces rent extraction with stronger performance rewards. Basically, the politician faces a trade-off between more rents and a higher likelihood of bargaining breakdown. Performance-based rewards raise the cost of bargaining breakdown, which induces the politician to extract fewer rents from the entrepreneur.

In $t = 1$, the entrepreneur will invest if the expected surplus $[1 - F(w^*(R))]\{E[y|y \geq w^*(R)] - w^*(R)\}$ outweighs the investment cost k . I impose the following assumption:

Assumption 1:

$$[1 - F(w^*(0))]\{E[y|y \geq w^*(0)] - w^*(0)\} < k \text{ and } k < E[y].$$

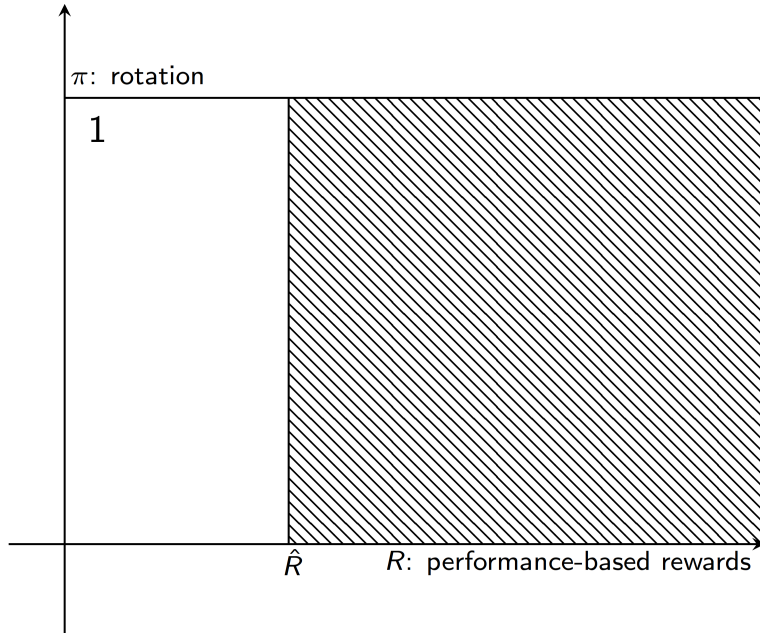


Figure 1

The assumption expresses that a politician who wants only to maximize economic rents (which is the case when $R = 0$) extracts too much from the entrepreneur. Thus, the entrepreneur refuses to initiate investment. Performance-based rewards can solve the problem, as a politician who obtains large rewards for economic growth dislikes bargaining breakdowns. She extracts less from the entrepreneur, which increases the probability of a successful project and enables the entrepreneur to cover his *ex ante* cost of investment. We have the following result.

Proposition 1: *In the benchmark model, the entrepreneur will invest under sufficiently strong performance-based rewards.*

Algebraically, $\exists! \hat{R} \in (0, \infty)$ such that $1 - F(w^(\hat{R}))\{E[y|y \geq w^*(\hat{R})] - w^*(\hat{R})\} = k$. The entrepreneur expects a non-negative profit if $R \geq \hat{R}$.*

Proofs of all propositions are in Appendix 1. Note that political rotation is irrelevant. Whether the politician is the same one as in $t = 1$ or newly appointed, she behaves exactly the same in $t = 2$ and extracts $w^*(R)$. In Figure 1, the horizontal axis is the intensity of performance-based rewards, while the vertical axis is the frequency of rotation. The shaded area is the parametric range in which the entrepreneur will invest. The entrepreneur will invest as long as the politician is sufficiently rewarded for economic growth.

3.2 Rotation and Performance-based Rewards when Politician can Accumulate Power

In this section, the key difference is that the politician can accumulate power. Specifically, after the entrepreneur has invested, the politician can pay a cost c such that the politician in $t = 2$ is fully informed about the project's value y . Everything else is the same as in the benchmark model. This formalizes the idea that a politician with a long tenure can become very knowledgeable about her jurisdiction, and consequently enormously powerful. As in classical models of information economics (Bolton and Dewatripont, 2005), power and information are interchangeable. A politician who knows everything about the entrepreneur's project wields formidable bargaining power against the entrepreneur, who will not be able to reap benefits from his own project. Rotation plays a key role here in discouraging the politician from information acquisition.

What are the practical ways to improve a politician's knowledge about local enterprises? She can rely on her office of assistants to implement inspection and networking for more

sensitive information about private enterprises. She can boost the capacity of government bureaus that register and review private enterprises (Besley and Persson, 2011). In the more extreme form, she can build her coercive power so that the local firms have to surrender such information. Given that unelected local politicians can be “petty dictators” in authoritarian regimes (Lieberthal, 2005), there are so many ways for her to become immensely knowledgeable about local economic conditions.

The setup implicitly assumes that the newly appointed politician automatically inherits the information acquired by the politician at $t = 2$. The rationale is that the politician initiates the learning process through mobilizing the bureaucracy. It is her bureaucrats and assistants who actually implement the investigation. When the politician is rotated, the bureaucrats stay and work for the newly appointed politician, who can readily harvest the knowledge embedded in the bureaucracy. This assumption helps highlight the key mechanism of the model. In Appendix 2, I show that my qualitative results do not rely on this assumption.

We need to derive restrictions on π to ensure that the politician in $t = 1$ does not pay c . It will guarantee some information rents to the entrepreneur, which is a necessary condition for the entrepreneur to gain some surplus from his project. Under the jurisdiction of an uninformed politician, the entrepreneur earns $[1 - F(w^*(R))]\{E[y|y \geq w^*(R) - w^*(R)]\}$. The entrepreneur will invest if the politician does not pay c and $[1 - F(w^*(R))]\{E[y|y \geq w^*(R) - w^*(R)]\} > k$ (the expected benefit from the project outweighs the cost).

At $t = 2$, the entrepreneur accepts any $w \leq y$. An informed politician extracts $w = y$, and the entrepreneur gets 0 surplus. An uninformed politician extracts w^* that satisfies equation (1), where $\frac{w^* f(w^*)}{1 - F(w^*)} = \frac{1}{1 + R}$. The extraction proposed by an uninformed politician is the same as the case in the benchmark model.

The politician in $t = 1$ does not pay c if:

$$-c + (1 - \pi)(1 + R)E[y] + \pi\tilde{U} \leq (1 - \pi)[1 - F(w^*)]\{w^* + E[Ry | w^* \leq y]\} + \pi\tilde{U}.$$

The left-hand side is the expected payoff to a politician who plans to pay c . If so, with probability $1 - \pi$ the politician continues her tenure. She sets $w^* = y$ and extracts all surplus y away, without worrying about bargaining breakdown. As the politician knows the productivity of the firm, she can always calibrate rent extraction based on the realized productivity, something that an uninformed politician cannot achieve. Hence, an agreement between the entrepreneur and the fully informed politician is guaranteed, and the politician also reaps the full benefit of performance-based rewards Ry . The expected payoff with

extended tenure is $(1 + R)E[y]$. With probability π , the informed politician is rotated, and she gets her exogenous payoff \tilde{U} .

The right-hand side is the expected payoff to a politician who plans to not pay c . With probability $1 - \pi$, she gets the same expected payoff as an uninformed politician, which is $[1 - F(w^*)]\{w^* + E[Ry | w^* \leq y]\}$, and with probability π she gets \tilde{U} . The inequality identifies the lower bound for rotation frequency π :

$$\pi \geq 1 - \frac{c}{(1 + R)E[y] - [1 - F(w^*)]\{w^* + E[Ry | w^* \leq y]\}} \equiv \underline{\pi}. \quad (3)$$

Thus, rotation has to be sufficiently frequent to forestall information acquisition. To guarantee that minimum frequency $\underline{\pi} \in (0, 1)$, I have the following assumption.

Assumption 2:

$$(1 + R)E[y] - [1 - F(w^*)]\{w^* + E[Ry | w^* \leq y]\} > c.$$

Recall that w^* is a function of R : $\frac{w^* f(w^*)}{1 - F(w^*)} = \frac{1}{1 + R}$. Assumption 2 is a restriction on exogenous parameters R and c . Suppose that Assumption 2 holds. Without rotation, the politician would pay c if the entrepreneur invests k , as the net benefit $(1 + R)E[y] - [1 - F(w^*)]\{w^* + E[Ry | w^* \leq y]\}$ outweighs the cost c . The entrepreneur refuses to initiate the investment since he predicts that he will get $-k$ by doing so. Thus, $\underline{\pi} > 0$ is guaranteed by Assumption 2. Without rotation, the politician in $t = 1$ will pay c and extract all rents.

In addition, it is clear that $(1 + R)E[y] - [1 - F(w^*)]\{w^* + E[Ry | w^* \leq y]\} > 0$:

$$\begin{aligned} & (1 + R)E[y] - [1 - F(w^*)]\{w^* + E[Ry | w^* \leq y]\} \\ &= (1 + R) \int_{\underline{y}}^{\bar{y}} z f(z) dz - w^* \int_{w^*}^{\bar{y}} f(z) dz - R \int_{w^*}^{\bar{y}} z f(z) dz \\ &= \int_{w^*}^{\bar{y}} (z - w^*) f(z) dz + (1 + R) \int_{\underline{y}}^{w^*} z f(z) dz > 0. \end{aligned}$$

Thus, $\underline{\pi} = 1 - \frac{c}{\int_{w^*}^{\bar{y}} (z - w^*) f(z) dz + (1 + R) \int_{\underline{y}}^{w^*} z f(z) dz} < 1$. We have $0 < \underline{\pi} < 1$.

Suppose that the principal chooses π that satisfies (3). As argued before, the entrepreneur pays k if the expected benefit from the project outweighs k :

$$k < [1 - F(w^*)]\{E(y|y \geq w^*) - w^*\}.$$

With Assumption 1 and Assumption 2, I can prove the main results of the paper.

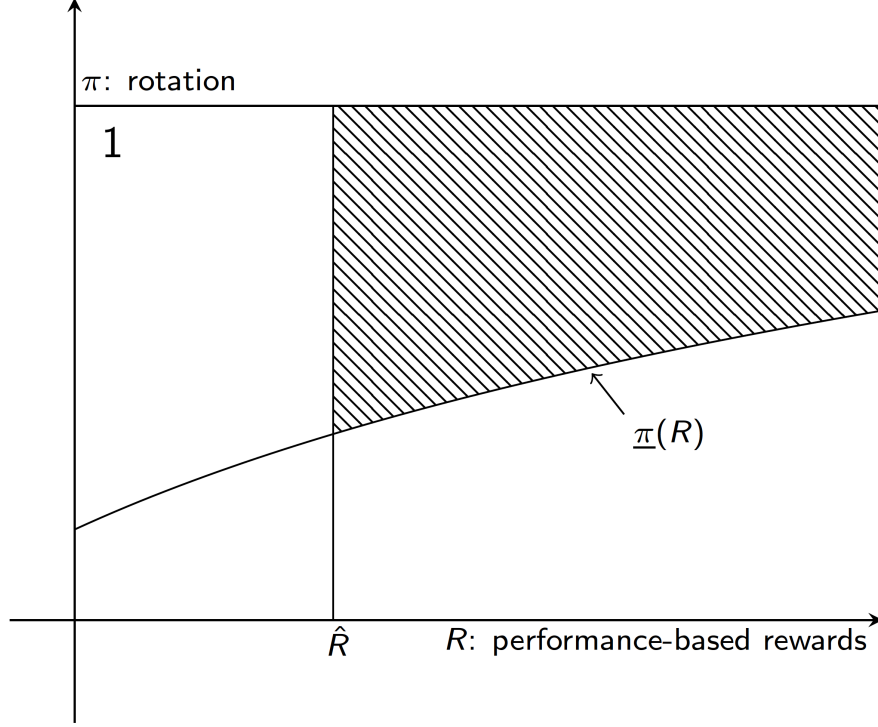


Figure 2

Proposition 2: 1. Sufficiently strong rotation and performance-based rewards incentivize the entrepreneur to invest.

In algebra, denote $\hat{R}(k)$ and $\underline{\pi}(R, c)$ such that $[1 - F(w^*(\hat{R}))]\{E(y|y \geq w^*(\hat{R})) - w^*(\hat{R})\} = k$ and $\underline{\pi}(R, c) = 1 - \frac{c}{(1+R)E[y] - [1 - F(w^*(R))]\{w^* + E[Ry | w^*(R) \leq y \}]}$. If $R > \hat{R}(k)$ and $\pi \geq \underline{\pi}(R, c)$, the politician does not pay c , and the entrepreneur will reap a non-negative return.

2. $\frac{\partial \underline{\pi}(R, c)}{\partial R} > 0$: the minimum rotation frequency increases when performance-based rewards are stronger. If the minimum rotation frequency does not change, more intense performance-based rewards incentivize the politician in $t = 1$ to pay c , and the politician in $t = 2$ will fully predate the entrepreneur.

Corollary 1: Suppose $\pi = 0$. $\forall R \geq 0$, the entrepreneur does not invest.

Propositions 1 and 2 and Corollary 1 can be summarized by Figure 1 (on page 7) and Figure 2. The shaded area is the parametric range such that the entrepreneur will invest. Figure 1 depicts Proposition 1, while Figure 2 illustrates Proposition 2 and Corollary 1. Figure 1 shows the effect of performance rewards when there is no severe holdup problem. The politician cannot invest c in Figure 1, so performance-based rewards alone can induce investment from the entrepreneur. Moreover, the probability of project completion $1 - F(w^*(R))$ also increases with R , as $w^*(R)$ decreases with R . Hence, economic performance will improve with a higher level of performance-based rewards. By contrast, Figure 2 shows the situation when the politician can acquire local information. In this case, we need a second constraint $\pi \geq \underline{\pi}(R, c)$ to ensure that the politician gives up the opportunity of information acquisition. Otherwise, an empowered politician would have extracted all surplus so that the entrepreneur never invests in the first place.

The key result of Proposition 2 is that $\frac{\partial \underline{\pi}(R, c)}{\partial R} > 0$. Notice that:

$$\underline{\pi}(R, c) = 1 - \frac{c}{(1 + R)E[y] - [1 - F(w^*(R))]\{w^* + E[Ry | w^*(R) \leq y]\}} \equiv 1 - \frac{c}{\Delta(R)}.$$

where $\Delta(R) \equiv (1 + R)E[y] - [1 - F(w)]\{w + E[Ry | y \geq w]\}$ is the difference between the expected payoffs to an informed politician and an uninformed one. In other words, $\Delta(R)$ measures the benefit of local knowledge, and the temptation for the politician to be informed. Hence, $\underline{\pi}(R, c)$ increases with R because $\Delta(R)$ increases with R :

$$\frac{d\Delta(R)}{dR} = \int_{\underline{y}}^{\bar{y}} z f(z) dz - [1 - F(w)] \frac{\int_{w^*}^{\bar{y}} z f(z) dz}{[1 - F(w)]} = \int_{\underline{y}}^{w^*} z f(z) dz > 0.$$

Performance-based rewards exacerbate the temptation to learn. This finding is a key insight of the paper that directly contributes to the complementarity between rotation and performance-based rewards. If unrotated, the informed politician can reap all the benefit from performance-based rewards, as she can calibrate the rent extraction based on the realized productivity. Doing so ensures that there will be no bargaining breakdown, so the project will always be finished, and the politician will always get her performance-based rewards.

By contrast, if the politician does not pay the cost of learning, she faces a substantial risk of bargaining breakdown. Apart from the lost economic rent, the rewards for economic growth also fail to materialize when the politician and the entrepreneur cannot settle upon an agreement. An uninformed politician can only reap a partial benefit from performance-based rewards, while an informed politician captures the full benefit. When rewards for economic

performance increase, it becomes even more tempting to invest in information acquisition because such information is more valuable under stronger performance rewards.

Such temptation to acquire local knowledge can be discouraged by rotation, as shown by the curve that represents $\underline{\pi}(R)$. With stronger performance rewards, the minimum rotation frequency has to increase to counterbalance the temptation to learn. Assumptions 1 and 2 guarantee that for any intensity of performance rewards $R \in [0, \infty)$, there is a minimum rotation frequency $\underline{\pi}(R, c) \in (0, 1)$. With the rotation prospect, the benefit from the detailed information realizes with such a small probability that it cannot justify the investment in local information.

$\pi \geq \underline{\pi}(R, c)$ such that the politician does not learn is a necessary condition for the entrepreneur to invest. As in Proposition 1, performance-based rewards also cannot be too low. In the extreme case where $R = 0$, the uninformed politician's problem reduces to $\max_w [1 - F(w)]w$; thus, the politician prefers to maximize expected rent. As $w^*(R)$ is a decreasing function of R , $w^*(R = 0)$ is large. Without performance-based rewards, the uninformed politician does not care too much about bargaining breakdown because she is not rewarded for economic achievements. Thus, she is willing to risk the high probability of bargaining breakdown in exchange for a higher rent. This results in small *ex ante* surplus for the entrepreneur because of the high rent extraction attempted by the politician and the associated high risk of bargaining breakdown. By Assumption 1, $[1 - F(w^*(0))] \{E(y|y \geq w^*(0)) - w^*(0)\} < k$, so the entrepreneur does not invest when an uninformed politician is not politically rewarded for growth. A moderate degree of performance-based rewards \hat{R} such that $[1 - F(w^*(\hat{R}))] \{E(y|y \geq w^*(\hat{R})) - w^*(\hat{R})\} = k$ coupled with $\pi \geq \underline{\pi}(\hat{R}, c)$ finally allows the entrepreneur to break even.

To summarize, if $\pi < \underline{\pi}(R, c)$, strong performance-based rewards cannot induce the entrepreneur to invest because the entrepreneur expects full extraction of his surplus in the future. A higher level of performance-based rewards, rather than alleviating the predicament, further justifies information acquisition and exacerbates the holdup problem. With frequent rotation such that $\pi \geq \underline{\pi}(R, c)$, the desired disciplinary effects of performance-based rewards are completely restored. The politician finds it unprofitable to invest in information acquisition. In this case, stronger rewards based on performance induce the uninformed politician to care a lot about striking a bargain with the entrepreneur. Frequent rotation limits the potentially destructive force of performance-based rewards and unleashes their role in aligning the incentive of the politician with the principal's objective. At the same time, if $R < \hat{R}$, frequent rotation achieves nothing, as the uninformed politician cares too little about the economy to behave benevolently. Thus, the effectiveness of rotation also re-

lies on sufficiently strong rewards for economic growth. Performance-based rewards support rotation in another important manner in the next section.

3.3 The Interdependence between Rotation and Performance Rewards

A key problem with rotation is that it can encourage shortsighted policy. Indeed, surrendering the opportunity of information acquisition is itself a shortsighted policy, but it is a particular category that improves welfare. However, there are many other shortsighted policies that can drastically reduce welfare. Specifically, the model in section 3.2 requires the politician to wait until period 2 when the entrepreneur can use the invested capital k to produce output. If rotation is too frequent, the politician may find it optimal to steal the invested capital in period 1. This is the famous “roving-bandit” problem articulated by Olson (1993). The policy instrument to discourage capital confiscation is performance-based rewards. Even if the politician anticipates a high likelihood of rotation, protection of private capital remains desirable under strong performance rewards. Basically, the reduced stake of future economic performance due to frequent rotation can be compensated by stronger performance rewards.

Assume $\pi \geq \bar{\pi}(R)$ so that the politician does not learn the value of the project. Suppose before that, the politician can decide whether to steal the capital k away. If the politician does, she can resell the capital and gain a value of ηk from it, $\eta \leq 1$. All capital depreciates away in $t = 2$.

The politician will not steal the capital if:

$$\pi \tilde{U} + (1 - \pi) * 0 + \eta k < \pi \tilde{U} + (1 - \pi)[1 - F(w^*)]\{w^* + E[Ry|y \geq w^*]\}. \quad (4)$$

The left-hand side is the expected payoff to a politician who steals capital. With probability π , she is rotated and gets \tilde{U} . With probability $1 - \pi$, she continues her tenure. If the capital has been stolen, the entrepreneur cannot produce anything, so the politician gets no rents and no performance-based rewards. In either case, the politician always gets ηk . The right-hand side is the expected payoff to a politician who does not steal capital or acquires local information. With probability $1 - \pi$, she gets the same payoff as an uninformed politician $[1 - F(w^*)]\{w^* + E[Ry|y \geq w^*]\}$. With probability π , she gets \tilde{U} from serving in another jurisdiction. Equation (4) gives us:

$$\pi < 1 - \frac{\eta k}{[1 - F(w^*)]\{w^* + E[Ry|y \geq w^*]\}} \equiv \bar{\pi}(R).$$

So the “roving-bandit” concern puts an upper bound on rotation frequency. If rotation is excessively frequent, the politician will steal the capital and resell it, and the project will yield no surplus in the future. The key observation is that the upper bound $\bar{\pi}(R)$ is also an increasing function of R :

$$\frac{\partial \bar{\pi}(R)}{\partial R} = \frac{\eta k}{\{[1 - F(w^*)]\{w^* + E[Ry | y \geq w^*]\}\}^2} [1 - F(w^*)] E[y | y \geq w^*] > 0.$$

Inequality (4) shows that the problem of rotation is that it reduces the desirability to protect private capital. Even if the politician steals the private capital today, she will not be affected by the destruction of growth opportunity if she gets rotated to another city. Thus, she does not care much about the consequence brought by her confiscation of private capital. Performance-based rewards discourage the temptation by restoring the stake in protecting private capital for the politician. Thus, a higher level of performance-based rewards allows more frequent rotation.

Now the question becomes: are there any parametric ranges so that

$$\underline{\pi} \equiv 1 - \frac{c}{(1 + R)E[y] - [1 - F(w^*)]\{w^* + E[Ry | y \geq w^*]\}} \leq 1 - \frac{\eta k}{[1 - F(w^*)]\{w^* + E[Ry | y \geq w^*]\}} \equiv \bar{\pi}?$$

I have an additional assumption:

Assumption 3:

$$\frac{c}{\eta k} > \max\left\{\frac{E[y]}{[1 - F(w^*(0))]w^*(0)} - 1, \frac{F(w^*(0))E[y | y \leq w^*(0)]}{[(1 - F(w^*(0)))]E[y | y \geq w^*(0)]}\right\}$$

Thus, we can prove:

Proposition 3: Denote $\underline{\pi}(R) \equiv 1 - \frac{c}{(1+R)E[y] - [1-F(w^*(R))]\{w^*(R) + E[Ry | y \geq w^*(R)]\}}$, $\bar{\pi}(R) \equiv 1 - \frac{\eta k}{[1-F(w^*(R))]\{w^*(R) + E[Ry | y \geq w^*(R)]\}}$. For $R \geq 0$, $\exists \pi \in [\underline{\pi}(R), \bar{\pi}(R)]$, such that the local politician neither acquires information nor steals capital.

Moreover, if $R' < R''$, then $\underline{\pi}(R') < \underline{\pi}(R'')$, and $\bar{\pi}(R') < \bar{\pi}(R'')$: Stronger performance-based rewards complement and are complemented by more frequent rotation.

A proper selection of π makes sure that the politician neither steals capital nor acquires information to capture all rent from the project. The key conclusion is that the complementarity between rotation and performance-based rewards persists, as the interval $[\underline{\pi}(R), \bar{\pi}(R)]$

is moving upward with a higher R . This is shown in the graph below. There is now an upper limit on the frequency of rotation, and the upper limit also increases with stronger performance-based rewards.

The two limits vividly characterize that rotation and performance-based rewards are supporting each other. Rotation discourages the temptation to dominate entrepreneurs, allowing higher-powered performance rewards. In turn, performance-based rewards make it less desirable to steal private capital, enabling more frequent rotation.

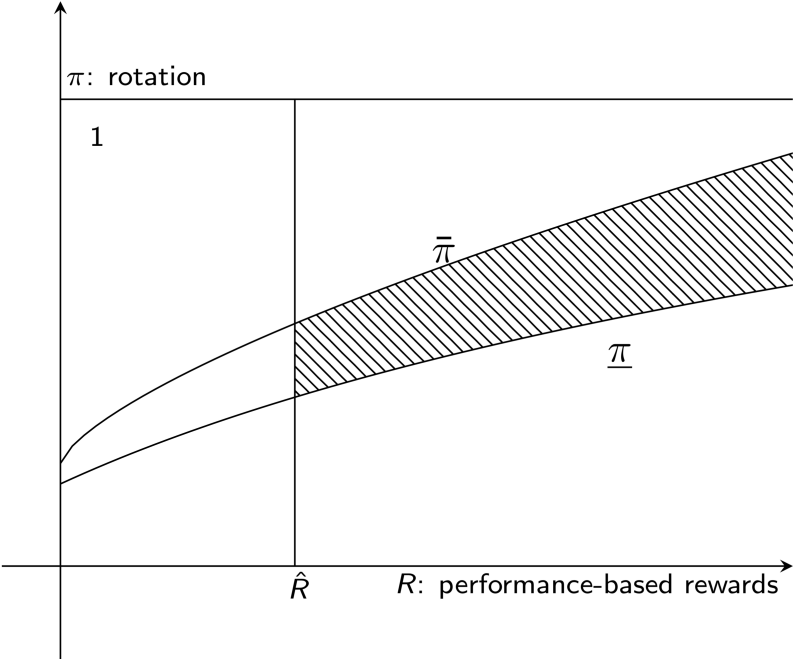


Figure 3

4. Rotation, Performance-based Rewards, and “Creative Destruction”

Rotation also creates active adaptation and “creative destruction”. As emphasized in the introduction, reputation-based solutions to the holdup problem rely on relational enforcement, which induces rigidity (Levin, 2003; Chassang 2010) and chocks “creative destruction” (Acemoglu et al. 2006). In my model, suppose that at the end of period 1 another entrepreneur arrives with a new project that is better than the old one. The question is whether the politician adapts to the new scenario and endorses the more productive project instead of the old one. Without rotation, the politician pays the cost of learning and can capture most surplus

from the old project. In other words, the politician has considerable entrenched interest in the old project. If instead, the politician endorses the new project, she will have to pay large information rent. On average, the new project produces more surplus than the old one; but most likely, the politician cannot extract more surplus or obtain more performance-based rewards from the new project. As a consequence, the politician will be very conservative and endorse new project only if the realization of old project's value is sufficiently low. With rotation, the politician has no entrenched interests with the old project. For either the new or old project, the politician needs to pay information rent to an entrepreneur. As the new project is more productive than the old one, the politician will always endorse the new project.

4.1 Setup

The timeline is similar to Section 3; but in $t = 2$, with probability p , a new project arrives that is “better” than the old one. There are still two periods. The old project's profitability still follows $F(\cdot)$ with support $[y, \bar{y}]$, $F' \equiv f$. At $t = 1$, we have exactly the same timeline except for the last stage. Specifically, the profitability of the project y is revealed to the entrepreneur but not the politician. The politician can pay a cost c to learn the realization of $y \sim F(\cdot)$, which will be revealed to the politician next period.

At the end of $t = 1$, with probability $1 - p$, there is no opportunity for a new project. The politician proposes to extract w from the entrepreneur. If the entrepreneur accepts, he will finish the project and payoffs are realized. If not, the politician gets 0 and the entrepreneur gets $-k$.

With probability p , another entrepreneur arrives with a new project such that the output $y \sim G(\cdot)$ with support $[y, \bar{y}]$, $G' \equiv g$. $G(\cdot)$ first order stochastically dominates $F(\cdot)$. The politician can choose to endorse the old project or the new one but not both. The rationale is that a local politician must devote all her political capital and connections to lobby for one firm. Dilution of the politician's resources means that firms in her jurisdiction cannot compete successfully with firms allotted exclusive support in other jurisdictions. The assumption can be relaxed, but I need at least some capacity constraint in the numbers of firms the politician can support.

Another difference with the workhorse model is that with probability $q \in (0, 1)$, the sunk cost investment $k = 0$. This is very important to generate the problem of entrenched interests: if the old entrepreneur never invests under an informed politician, the politician has no available project to form entrenched interests at all. Of course, we want to avoid q to be too small or too large. This ensures strong tensions regarding both the commitment

problem and entrenched interests. There are other ways to impose the tensions, but this is the most parsimonious one.

Notice that for the new project, there is too little time to do effective learning. Then, she proposes to extract w from the endorsed project (either the new one or the old one). If the entrepreneur accepts, he will finish the project, and payoffs are realized. If not, the politician gets 0.

Players' payoffs with a completed project are very similar to Section 3. Specifically, the informed politician gets $w + Ry - c$ if a project (either the old or the new one) is finished with productivity y and if the politician extracts w from it. The uninformed politician gets $w + Ry$.

4.2 Adaptability with or without Learning

Let us analyze the modified game using “backward induction”. At $t = 2$, suppose that a new project arrives and the politician did not pay c in $t = 1$. Then, she will get $U_1 \equiv \max_w [1 - F(w)]\{w + E_f[Ry|y \geq w]\}$ from endorsing the old project, and $U_2 \equiv \max_w [1 - G(w)]\{w + E_g[Ry|y \geq w]\}$ from the new one. Denote $w^* = \operatorname{argmax}_w [1 - F(w)]\{w + E_f[Ry|y \geq w]\}$ and $\tilde{w} = \operatorname{argmax}_w [1 - G(w)]\{w + E_g[Ry|y \geq w]\}$.

If c has been paid, she gets $(1 + R)y$ from supporting the old project in the case where the old project has been invested (the case where the old project has not been invested is trivial). We can prove the following result:

Proposition 4: *Suppose the politician paid c and is thus informed about the old project that has been invested.*

1. *She will endorse the new project with probability $F(\frac{U_2}{1+R}) = F(\frac{1}{1+R}[1 - G(\tilde{w})]\{\tilde{w} + E_g[Ry|y \geq \tilde{w}]\})$, which increases with R : performance-based rewards encourage adaptation.*

2. *$\forall R < \infty$, $F(\frac{1}{1+R}[1 - G(\tilde{w})]\{\tilde{w} + E_g[Ry|y \geq \tilde{w}]\}) < F(E_g(y))$, the first-best probability of adaptation. An informed politician holds entrenched interests in the old project, no matter how strong performance-based rewards are.*

The first part of Proposition 4 is simple to prove. The politician informed about the old project endorses the new project if:

$$(1 + R)y' < U_2 = [1 - G(\tilde{w})]\{\tilde{w} + E_g[Ry|y \geq \tilde{w}]\},$$

where y' is a realized draw from $F(\cdot)$. The left-hand side is the payoff to the informed

politician if she sticks with the old project. The right-hand side is the payoff to the politician if she endorses the new project instead. As $y' \sim F(\cdot)$, the probability that the politician does so is $F(\frac{1}{1+R}[1 - G(\tilde{w})]\{\tilde{w} + E_g[Ry | y \geq \tilde{w}]\})$. Notice that even if $G(y)$ first order stochastically dominates $F(y)$, the politician supports the old project in most cases. This is because the politician can extract all surplus from the old project and ensures an agreement with the entrepreneur at the same time. If the politician supports the new project instead, she suffers from a significant difficulty of rent extraction, as well as a substantial risk of bargaining breakdown. Thus, although the new project is “on average” much better than the old project, the politician will most likely support the old project. The politician has entrenched interests in the old project.

Take the derivative of $F(\frac{1}{1+R}[1 - G(\tilde{w})]\{\tilde{w} + E_g[Ry | y \geq \tilde{w}]\})$ with respect to R : we find that the derivative is positive. Moreover, $\lim_{R \rightarrow \infty} F(\frac{1}{1+R}[1 - G(\tilde{w})]\{\tilde{w} + E_g[Ry | y \geq \tilde{w}]\}) = F(E_g(y))$, which is the first-best probability of adaptation. Hence, $\forall R < \infty$, $F(\frac{1}{1+R}[1 - G(\tilde{w})]\{\tilde{w} + E_g[Ry | y \geq \tilde{w}]\}) < F(E_g(y))$: we confirm that the informed politician *must* have entrenched interests, no matter how strong performance-based rewards are.

Performance-based rewards can partially correct entrenched interests, as $F(\frac{1}{1+R}[1 - G(\tilde{w})]\{\tilde{w} + E_g[Ry | y \geq \tilde{w}]\})$ increases with R . However, entrenched interests always persist, as $U_2 = [1 - G(\tilde{w}(R))]\{\tilde{w}(R) + E_g[Ry | y \geq \tilde{w}(R)]\}$ always entails some risk of bargaining breakdown, while endorsement of the old project completely avoids it. Bargaining breakdown under the new project vanishes only when $R \rightarrow \infty$. In this case, the payoff from rent extraction is completely dwarfed by the payoff from performance rewards, so the politician wants to ensure the completion of the project. Thus, she extracts $\tilde{w}(R) \rightarrow \underline{y}$ as $R \rightarrow \infty$, so we have $U_2 = E_g(Ry)$. By contrast, endorsement of the old project gives a payoff of $Y' + Ry'$; y' is a realized draw from $F(\cdot)$. Again, as $R \rightarrow \infty$, the politician only cares about performance-based rewards. She simply compares $E_g(Ry)$ and Ry' , and endorses the new project if $E_g(y) \geq y'$. This utilizes all information possibly available to the politician in a manner perfectly aligned with the principal’s interests, thus achieving first best. However, this asymptotic ideal can never work unless R is unrealistically large.

What about the politician who did not learn? She has no entrenched interests with the old project, and thus will support the new project that is “on average” better than the new project:

Proposition 5: *Suppose the politician didn’t pay the cost of learning and a new project arrives with $G(y) \leq F(y)$, $\forall y$. The politician will always endorse the new project.*

For an uninformed politician, her payoffs from the new or the old project both entail a substantial risk of bargaining breakdown. In other words, it is equally challenging for her to extract from the old and the new projects. Given that the new project is “on average” more productive, the uninformed politician will support the new project that can offer more rents and more performance-based rewards. Indeed, the proof of Proposition 5 shows that both $[1 - G(\tilde{w})]\tilde{w} > [1 - F(w^*)]w^*$ (more rents from the new project) and $[1 - G(\tilde{w})]E[Ry|y \geq \tilde{w}] > [1 - F(w^*)]E[Ry|y \geq w^*]$ (more performance-based rewards from the new project).

Notice that the optimal degree of adaptation is to support the new project with probability $F(E_g(y))$, while an uninformed politician always supports the new project. Rotation solves the problem of entrenched interests, although at a cost. A politician who is constantly rotated has no incentive to acquire local information, so she can abandon an old project that turns out to be especially valuable and support the new project precisely because she does not know the value of the old project. This generates excessive adaptation that is undesirable from the perspective of the principal and the society. But, of course, this is a “necessary evil”; otherwise, an informed and unchecked politician will just take everything and the entrepreneur has no incentive to invest in period $t = 1$.

4.3 Rotation, Performance Rewards, and Adaptation

In $t = 1$, the politician does not pay the cost of learning if:

$$(1 - \pi)\{(1 - p)U_1 + pU_2\} + \pi\tilde{U} \geq$$

$$(1 - \pi)\left\{ (1 - p)(1 + R)E(y) + p\{\Pi U_2 + (1 - \Pi)E[(1 + R)y|(1 + R)y \geq U_2]\} \right\} + \pi\tilde{U} - c. \quad (5)$$

where $U_1 \equiv \max_w [1 - F(w)]\{w + E_f[Ry|y \geq w]\}$, $U_2 \equiv \max_w [1 - G(w)]\{w + E_g[Ry|y \geq w]\}$, $\Pi = F(\frac{U_2}{1+R})$.

$(1 - \pi)\{(1 - p)U_1 + pU_2\} + \pi\tilde{U}$ is the payoff to a politician uninformed about the old project. With probability $(1 - \pi)(1 - p)$, the politician continues her tenure in $t = 2$ and no new projects arrive. The politician gets $U_1 = \max_w [1 - F(w)]\{w + E_f[Ry|y \geq w]\}$. With probability $(1 - \pi)p$, the continuing politician sees the arrival of a new project. Proposition 5 tells us that the uninformed politician always supports the new project, which gives her a payoff of $U_2 \equiv \max_w [1 - G(w)]\{w + E_g[Ry|y \geq w]\}$.

$(1 - \pi) \left\{ (1 - p) (1 + R) E(y) + p \{ \Pi U_2 + (1 - \Pi) E[(1 + R)y | (1 + R)y \geq U_2] \} \right\} + \pi \tilde{U} - c$ is the payoff to a politician informed about the old project. With probability $1 - \pi$, the politician continues her term. When the new project does not arrive (with probability $1 - p$), the politician gets $(1 + R)E[y]$ from the old project. When the new project arrives, the politician supports the new project if $U_2 = \max_w [1 - G(w)] \{ w + E_g[Ry | y \geq w] \} > (1 + R)y'$, which occurs with probability $\Pi = F\left(\frac{U_2}{1+R}\right)$. With probability $1 - \Pi$, the politician still supports the old project and gets $E[(1 + R)y | (1 + R)y \geq U_2]$. We want to see whether the results in Proposition 2 are robust with the possible arrival of the new project.

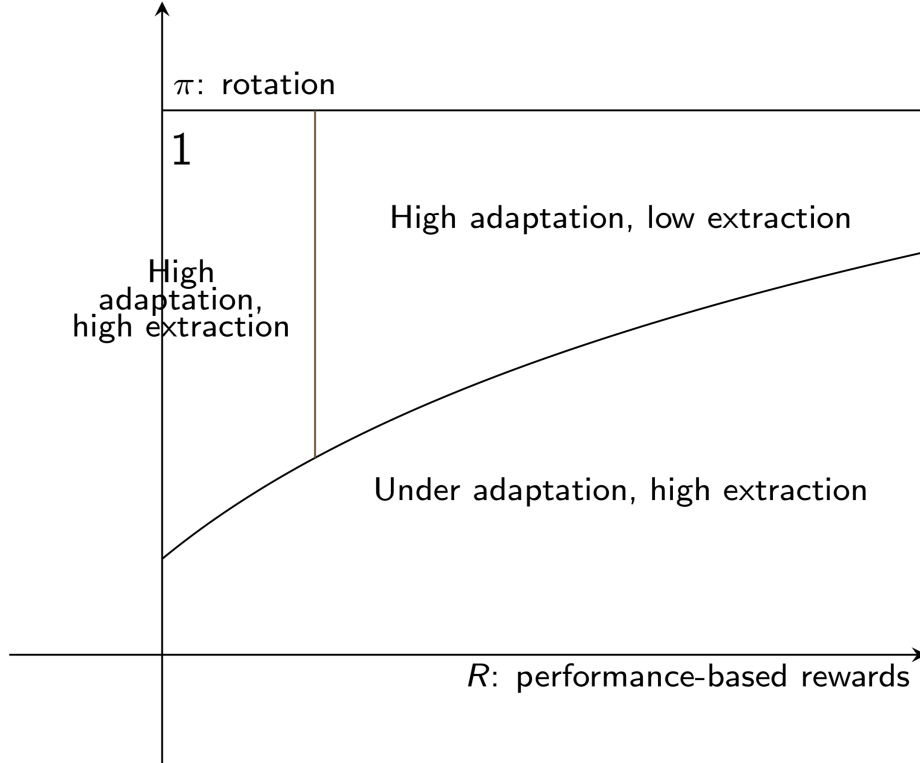


Figure 4: *ex ante* cost of investment $k > 0$ with prob. < 1

Proposition 6: 1. Sufficiently intense rotation and performance-based rewards induce an adaptive and benevolent local politician. The old and the new entrepreneurs invest accordingly.

In algebra, denote \hat{R} such that $\hat{R}(k) = \max\{R_1, R_2\}$, where R_1 satisfies $(1 - p)[1 - F(w^*(R_1))] \{ E_f(y | y \geq w^*(R_1)) - w^*(R_1) \} = k$ and R_2 satisfies $[1 - G(\tilde{w}(R_2))] [E_g(y | y \geq \tilde{w}(R_2)) - \tilde{w}(R_2)] = k$. Also, $\underline{\pi} = 1 - \frac{c}{(1-p)\{(1+R)E(y)-U_1\}+p\{(1-\Pi)[E_f[(1+R)y|(1+R)y \geq U_2]-U_2]\}}$.

If $R > \hat{R}$ and $\pi \geq \underline{\pi}$, the politician does not pay c , and both the old and the new entrepreneurs reap non-negative returns.

2. $\frac{\partial \pi(R,c,p)}{\partial R} > 0$: *the minimum rotation frequency increases when performance-based rewards are stronger. If the minimum rotation frequency does not change, more intense performance-based rewards incentivize the politician in $t = 1$ to pay c , and the politician in $t = 2$ will fully predate the old entrepreneur.*

The results are summarized in Figure 4, which shows that the complementarity between rotation and performance-based rewards persist under adaptation concerns. Specifically, with frequent rotation and strong performance-based rewards, the entrepreneur will invest. The local politician will facilitate “creative destruction” if a more productive project arrives. With stronger performance-based rewards, the equilibrium of high investment and high creative destruction can only be sustained with more frequent rotation. Thus, we formalize the notion that rotation and performance-based rewards provide property rights and facilitate “creative destruction” simultaneously, a remarkable feature rarely satisfied by *de facto* property rights based on reputation.

5. A Model with Opportunity of Bargaining before Rotation

A critique of above models is that rotation raises the bargaining power of the entrepreneur simply because the politician and the entrepreneur can only bargain in $t = 2$. If the politician has a chance to engage in multiple rounds of bargaining, she may gradually learn the productivity of the entrepreneur’s project by inferring from the entrepreneur’s strategy to reject or accept the politician’s proposed rent extraction. The politician may not need to invest in local knowledge. If so, anticipated rotation in the future is not helpful at all. In this section, I deal with this critique by showing that all intuitions in Section 3 are robust even if the players can bargain in $t = 1$.

The timeline is the same as that in Section 3.2 except that the politician can propose to extract w_1 in $t = 1$ before she decides whether to learn about y . If the entrepreneur accepts w_1 , the politician helps the entrepreneur to finish the project in $t = 2$ and gets her rent w_1 . The project’s value is still realized in $t = 2$ because it is a long-term investment. If the politician and the entrepreneur settle an agreement to extract w_1 , the politician has no incentive to pay c and learn about the entrepreneur’s project anymore. If w_1 is rejected, the politician can pay c to learn y , and the “subgame” is exactly the same as in Section 3.2 following a rejection of w_1 : the principal rotates the politician with probability π , the politician proposes to extract w_2 , and the entrepreneur accepts or rejects the proposal.

Assumption 4: Suppose that $w_1 = w_2$. If an entrepreneur with \hat{y} accepts w_1 , then any entrepreneur with $y \geq \hat{y}$ also accepts w_1 .

The assumption is crucial to reduce the complexity of the problem when we consider the politician's strategy to offer $w_1 = w_2$. Without Assumption 4, there are so many possible distributions of the entrepreneur that reject w_1 when $w_1 = w_2$. With Assumption 4, the politician in $t = 2$ faces a distribution censored from the right, which significantly simplifies the problem of the politician in $t = 2$.

I can prove the following key result:

Proposition 7: Denote:

$$\underline{\pi}_1(R) = 1 - \frac{c}{(1+R)E[y] - \max_w [1 - F(w)]\{w + E[Ry | y \geq w \]}}$$

$$\underline{\pi}_2(R) : (1 - \underline{\pi}_2(R)) \max_{w_2} [1 - F(w_2)]\{w_2 + E[Ry | y \geq w_2 \] =$$

$$\max_{w_1} \left\{ F(w_1) \left\{ (1-\pi) \left\{ (1+R)E[y | y \geq w_1] \right\} + \pi \tilde{U} - c \right\} + [1 - F(w_1)] \left\{ (1-\pi) \left\{ w_1 + E[Ry | y \geq w_1] \right\} + \pi \tilde{U} \right\} \right\}.$$

$$\hat{R} : k = [1 - F(w^*(\hat{R}))]\{E[y | y \geq w^*(\hat{R}) - w^*(R)]\}, \frac{w^*(\hat{R})f(w^*(\hat{R}))}{1 - F(w^*(\hat{R}))} = \frac{1}{1 + \hat{R}}.$$

1. If $\pi \geq \underline{\pi}_1(R)$, $\pi \geq \underline{\pi}_2(R)$, and $R \geq \hat{R}$, the politician does not pay c to learn about y , and the entrepreneur makes the investment k .

2.

$$\frac{\partial \underline{\pi}_1(R)}{\partial R} > 0, \frac{\partial \underline{\pi}_2(R)}{\partial R} > 0$$

So with stronger performance rewards, the minimum rotation frequency $\underline{\Pi}(R) = \max\{\underline{\pi}_1(R), \underline{\pi}_2(R)\}$ also rises to guarantee the entrepreneur's ex ante investment.

Although the algebra becomes much more complex, the core intuition of the paper remains unchanged. As shown in Figure 5, with stronger performance-based rewards, the minimum rotation frequency still rises. The first constraint $\pi \geq \underline{\pi}_1(R)$ is the same constraint as

in Section 3.2. It guarantees the sequential rationality of no information acquisition after w_1^* has been rejected. For the second constraint $\pi \geq \underline{\pi}_2(R)$, it guarantees that the payoff as an uninformed politician is higher than even the highest payoff that detailed information can deliver, which is:

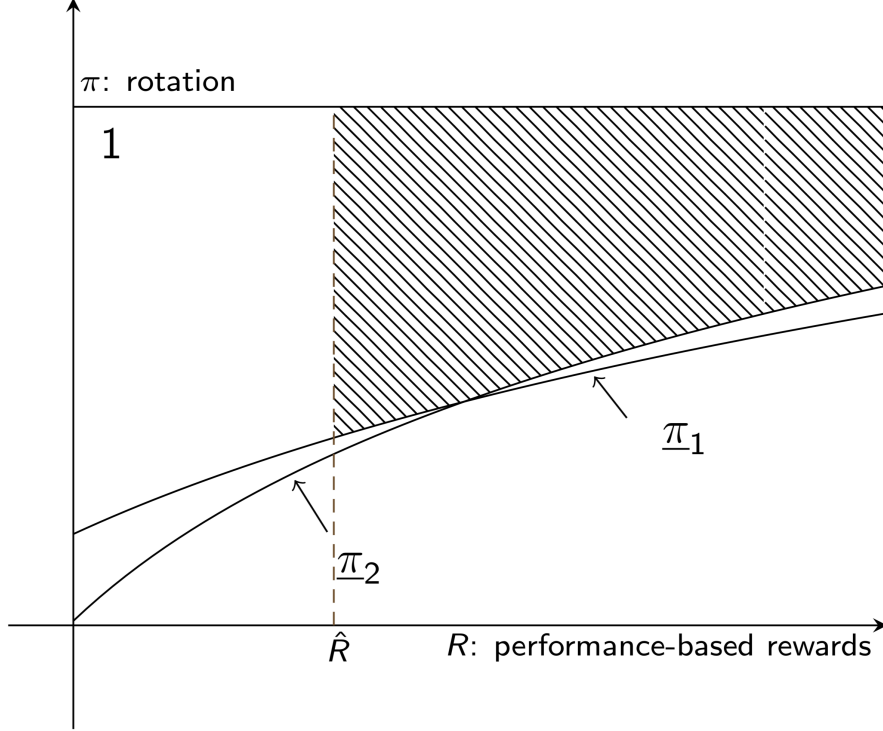


Figure 5

$$\max_{w_1} \left\{ F(w_1) \left\{ (1-\pi)(1+R)E[y|y \geq w_1] + \pi\tilde{U} - c \right\} + [1-F(w_1)] \left\{ (1-\pi) \left\{ w_1 + E[Ry | y \geq w_1] \right\} + \pi\tilde{U} \right\} \right\} \quad (6)$$

$$= \max_{w_1} \left\{ -F(w_1)c + \pi\tilde{U} + (1-\pi) \left\{ [1-F(w_1)]w_1 + F(w_1)E[y|y \leq w_1] + E[Ry] \right\} \right\}. \quad (7)$$

To understand Equation (6), note that if the politician proposes to extract w_1 in $t = 1$, the entrepreneur gets $y - w_1$ by accepting w_1 . If he rejects w_1 , the politician becomes informed in $t = 2$ and extracts all surplus. Thus, the entrepreneur rejects w_1 if $y - w_1 \leq 0$, which occurs with probability $F(w_1)$. In this case, the politician pays c in $t = 1$. If not rotated, the politician extracts all surplus, so the expected payoff is $U^1 \equiv (1-\pi)\{(1+R)E[y|y \geq$

$w_1\}} + \pi\tilde{U} - c$, where the unrotated politician gets $(1 + R)E[y|y \geq w_1]$ and the rotated politician gets \tilde{U} . If $y - w_1 \geq 0$, the entrepreneur accepts w_1 , which occurs with probability $1 - F(w_1)$. The payoff to the politician is $U^2 = (1 - \pi)\{w_1 + E[Ry|y \geq w_1]\} + \pi\tilde{U}$: the unrotated politician gets $w_1 + E[Ry|y \geq w_1]$. Thus, the payoff to a politician planning to pay c is $F(w_1)U^1 + [1 - F(w_1)]U^2 = F(w_1)\{(1 - \pi)(1 + R)E[y|y \geq w_1] + \pi\tilde{U} - c\} + [1 - F(w_1)]\{(1 - \pi)\{w_1 + E[Ry|y \geq w_1]\} + \pi\tilde{U}\}$.

Simple algebra gives us Equation (7), which shows that an unrotated and informed politician still reaps full benefit from performance-based rewards with bargaining opportunity in $t = 1$. Again, there will be no bargaining breakdown at all: either the players strike a bargain in $t = 1$, which delivers a non-negative surplus to the entrepreneur, or the players strike a bargain in $t = 2$, when the entrepreneur gets 0. But in either case, there will be no bargaining breakdown under any circumstance. By contrast, a substantial risk of bargaining breakdown persists even with multiple rounds of bargaining. If the uninformed politician is not rotated, the maximum utility is still $\max_w [1 - F(w)]\{w + E[Ry|y \geq w]\}$. It seems counter-intuitive that the additional bargaining opportunity cannot improve the payoff of an uninformed politician. She cannot learn anything from the bargaining opportunity in $t = 1$. This is because it is never optimal for a forward-looking entrepreneur to reveal information about his type early in the bargaining process. Given that he is patient, he can wait and exhaust all such opportunities for the politician. Thus, the politician cannot benefit from the additional bargaining opportunity and gets exactly the same utility as the case with only one opportunity to bargain. The key implication is that the unrotated and uninformed politician still only reaps partial benefit from performance-based rewards, which is $[1 - F(w^*)]E[Ry|y \geq w^*]$. In general, even if the players are not fully patient, the politician cannot learn everything from multiple rounds of bargaining (Sobel and Takahashi, 1986). This means that the politician who plans to not pay c remains less informed than the entrepreneur so that bargaining breakdown is always a possibility. Thus, in more general settings where the players are not fully patient, stronger performance-based rewards can still tempt the politician to invest in information acquisition, which has to be balanced by more frequent rotation.

6. Empirical Evidence

I examine how anticipated personnel events on local politicians affect firm's decisions in China. Beforehand, I need to introduce additional institutional background for the empirical analysis. Local governments have four tiers in China. There are 31 provincial-level units,

with 10 to 15 cities/prefectures in each province. A city usually has jurisdiction over several counties. The two key politicians for a city are the party secretary and the mayor. The party secretary is the chief politician in the city. The mayor is formally subordinate to the secretary because the mayor always holds the concurrent position of a deputy party secretary. However, there is a distinctive division of labor between the party secretary and the mayor. The party secretary wields political power unmatched by any other politician in the city. She guarantees such power through direct control of the organizational department and propaganda department, the two most powerful departments in city party committee. The organizational department in a city controls all political appointment at the county level. The propaganda department directs state-owned media and controls censorship of commercial media. However, everyday management of the city government is the sole responsibility of the mayor. Specifically, all economic departments, such as revenue, construction, and commerce, are under the direct leadership of the mayor. By contrast, the city party committee does not include any economic departments. This arrangement is in stark contrast to local politics in the Soviet Union, where the party secretary also wields direct economic power through economic departments in the party committee. The empirical strategy heavily relies on variation from the dual leadership.

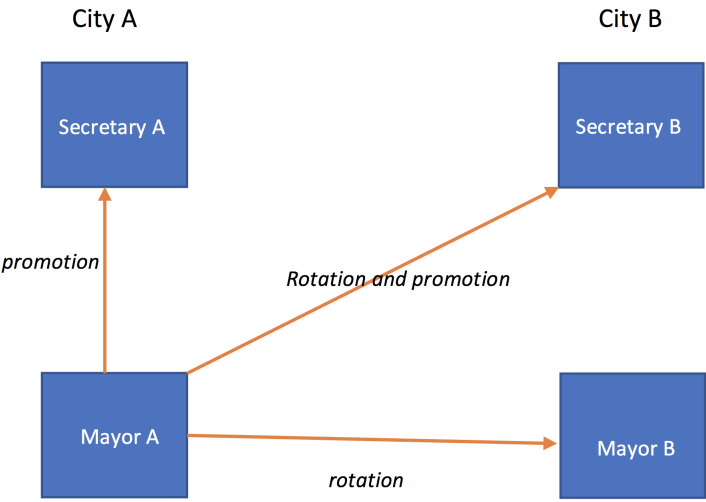


Figure 6

Figure 6 illustrates the definition of rotation and promotion that I will use in the empirical analysis. Rotation is defined as a lateral transfer for a mayor to another mayor position. A promotion event occurs when a mayor is appointed as the secretary of her own city. A large literature documents that promotion in China is strongly correlated with performance (Maskin et al. 2000; Li and Zhou, 2005; Jia et al. 2017), so a promotion opportunity is a

special form (and the most common form) of performance-based rewards in China. Rotation *and* promotion occur when a mayor is promoted as a secretary in another city.

The provincial organizational department devises and implements personnel management of all city-level politicians, including the appointment of both the party secretary and the mayor of a city. Hence, transfer of city politicians across provinces is very rare, which has important implications for the empirical analysis.

6.1 Testable Implications from the Theory

The main testable implication is the complementarity of rotation and performance-based rewards. The theory predicts that a firm invests more when the prospect of rotation and promotion are strong for the mayor of the city where the firm locates. The effect should be robust even after I account for the individual effects of rotation and promotion. The ideal experiment is to randomly assign mayors into four groups. The first group serves as the control group; the second group receives only promotion prospect; the third group receives only rotation prospect; the fourth group receives both. As such random assignment is rare, I want to utilize variation in anticipated rotation and promotion that are arguably exogenous.

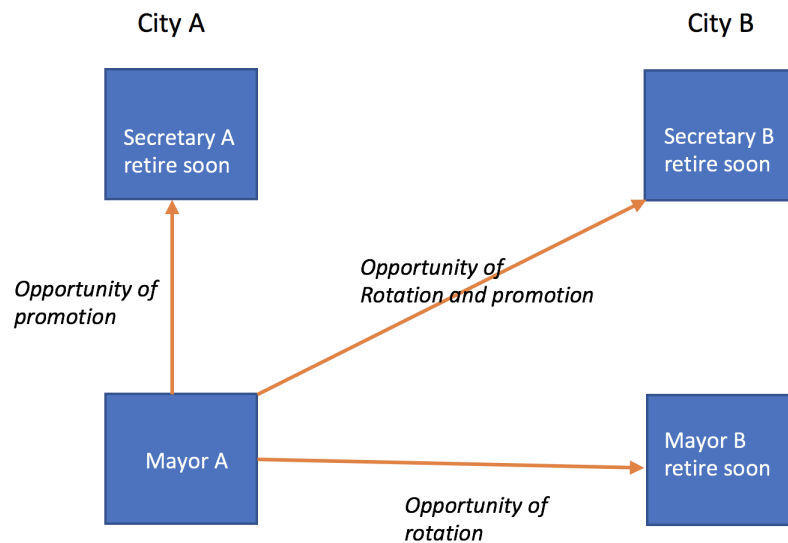


Figure 7

Specifically, I extensively explore the “jackknife” or “leave-one-out” variation (see Figure 7). A natural proxy for future rotation events is the anticipated retirement of mayors in

other cities ². City-level officials in China face a mandatory retirement age of 60. Before that, they are usually transferred to an honorary and powerless position in the legislature as a transition to full retirement at 60 (Wang, 2016; Xi et al. 2017). This event can happen any time after the city-level official turns 56. Thus, if the mayor of City B turns 56, all politicians in the province expect a job vacancy in the next four years. When the mayor of City B is actually assigned the honorary position or even retires from on her job, the mayor in City A is among the likely candidates to fill up the mayorship in City B. Hence, I proxy anticipated rotation by the fraction of mayors in *other* cities that are more than 56 years old.

For a promotion event, I look at whether the secretary in a specific city (e.g., City A) is more than 56 years old. In such a case, the mayor in city A anticipates a high likelihood of promotion if the city's economy grows fast, as the mayor of City A is the most likely candidate for the secretary position in City A.

We have a few proxies for the complementarity between rotation and promotion. The interactions between the above proxies for rotation and promotion can reasonably measure the complementarity effect. In addition, the anticipated retirement of secretaries in *other* cities improves the prospect of both rotation and promotion for a mayor, providing another variable that approximates the complementarity. Figure 7 illustrates the empirical variation I used to construct the proxies for anticipated personnel events.

To summarize: if my theory captures an important dimension of China's political economy, coefficients should be positive and precisely estimated for the interaction terms between proxies for rotation and promotion and for the fraction of secretaries in other cities who are more than 56 years old.

The effect should also be different for private firms vs. state-owned enterprises (SOEs). We should observe a much more pronounced response from private firms to better-protected property rights, but the effect should be smaller for SOEs that have already enjoyed substantial bargaining power against the mayor.

6.2 Relevance of Retirement in Other Cities on a Mayor's Own Tenure

I obtain personnel data on politicians from Chen (2016). For each city-year observation, the dataset identifies the governing secretary and mayor along with their age, gender, ethnic-

²I only look at mayors in other cities, but in the same province. This is because, as mentioned, personnel management at city-level is controlled by the provincial organizational department. So inter-provincial transfers of city-level politicians are extremely rare.

ity, education, and work experience. The dataset covers all cities in the 27 provinces and autonomous regions between 2000 and 2010³. For firm covariates, I obtain them from Annual Survey of Industrial Production (ASIP). The survey contains all firms that have annual sales above 5 million RMB (equivalent to 800,000 USD). For each firm-year observation, the survey records the location, sales, inventory, number of employees, total assets, fixed assets (physical capital), accumulated depreciation, liquid assets (such as cash and account receivables), intangible assets (such as intellectual property), liability, industry code, and ownership by paid-in capital. I use the balance sheet data to construct the measurement of physical capital and also many other firm characteristics as control variables. I merge the ASIP data with politician personnel data so that each firm-year observation is identified with politicians governing the city where the firm locates. Summary statistics for key variables are provided in Table 1.

Before testing the theoretical predictions, I document the relevance of retirement in other cities for mayor turnover in a specific city. The regression I run is:

$$n_{jt} = \mu_j + PC_t + \text{retirement of other secretaries}_{jt} + \text{retirement of other mayors}_{jt} + s_{jt} + m_{jt} + \varepsilon_{jt}.$$

n_{jt} is whether the mayor j 's term terminates in the year t . μ_j is mayor fixed effects. PC_t is the well-documented effect of "party congress": turnover rate is high in years before party congress (Xi et al. 2017). The retirement of other secretaries/mayors records the fraction of secretaries/mayors who are 59 years old in the province in the year $t - 1$. These politicians will all retire within the year t . In the process, their retirement generates many vacancies for the provincial organizational department to fill up.

s_{jt} denotes whether the secretary co-ruling with the mayor j is older than 56, and m_{jt} is whether mayor j is older than 56. Robustness checks with different age cutoffs are implemented, showing similar results. Alternative specifications using logit models produce similar estimates. They are omitted here. All standard errors are clustered at the mayor level.

Table 2 shows that retirement of other secretaries or mayors is strongly correlated with a mayor's turnover. When secretaries or mayors in other cities retire, it strongly predicts

³In other words, it only excludes the four cities that enjoy provincial status: Beijing, Shanghai, Tianjin, and Chongqing. It is reasonable to exclude these cities, as their mayors and party secretaries are (exceptionally important) politicians with provincial ranks. It is the Central Organizational Department (rather than the provincial branch) that manages such prestigious positions. In addition, the availability of these positions is irrelevant to the mayor and the party secretary of a typical city, who only enjoys city ranks. This is because even the strongest candidate with a city rank can only be promoted to a position with a deputy provincial rank.

termination of a mayor’s term. Other covariates also make economic sense. If the mayor or the secretary is old, turnover is more likely. An old mayor faces high “risk” of a transfer to an honorary position, while an old secretary increases the likelihood of promotion for the mayor. For the party congress effects, the omitted years are “party congress in 5 years”. As the Party Constitution requires a meeting of the party congress every 5 years, the omitted category represents the year immediately following a party congress. This year comes with the lowest turnover probability. We can see that the two years before the party congress, along with the year when party congress assembles, have pronounced higher likelihood of turnover. This finding is consistent with prior studies (Xi et al. 2017). In Columns (4) and (5), the retirement of other secretaries is defined as the fraction of other secretaries that were 58 (57) two (three) years ago. A similar definition applies to the retirement of other mayors. The results are very similar to Column (3).

6.3 Anticipated Retirement on Composition of Capital

The main empirical specification is:

$$y_{it} = \alpha_0 + \mu_i + \lambda_t + X_{it}\rho_0 + \rho_1s_{it} + \rho_2R_{it} + \beta s_{it} * R_{it} + \gamma C_{it} + \varepsilon_{it}, \quad (8)$$

where y_{it} is the ratio of fixed assets to total assets. In the Chinese Accounting Standards, fixed assets are defined as physical capital with long-term returns. μ_i is firm fixed effects, λ_t is year fixed effects, and s_i is a proxy for performance-based rewards. As said, studies on Chinese politics document that an age cutoff works very well to proxy for promotion opportunities (Wang, 2016; Xi et al. 2017). A secretary older than 56, who will retire in the next four years, is a precursor to a promotion opportunity for the mayor. Hence, I denote $s_{it} = 1$ if the secretary is older than 56 and 0 otherwise. R_{it} is the fraction of mayors in other cities older than 56. C_{it} records the fraction of secretaries in other cities older than 56. We are mainly interested in the coefficients on $s_{it} * R_{it}$ and C_{it} . X_{it} is a vector of control variables in all specifications of Tables 3 and 4. It includes rich covariates measuring characteristics about the firm, as well as the mayor and the party secretary of the city where the firm locates. For firm characteristics, I control (lag and logarithm of) output, the number of employees, value added, profit, management fee, inventory, firm age, and debt. For politician characteristics, I control (for both the mayor and the party secretary) age, gender, ethnicity, education, and work experience (includes whether the politician used to work in the Communist Youth League and whether the politician used to work as a personal

assistant or the director of the office for a senior politician)⁴.

Table 3 lists the main results of the paper. The first two rows are proxies for the complementarity of rotation and performance-based rewards. Column (1) is restricted to mayors who are in their first or second year in office. Notice that we have precisely estimated coefficients for the fraction of old secretaries in other cities, but the coefficient on $s_{it} * R_{it}$ is not precisely estimated. Columns (2) and (3) split the sample into young mayors and old mayors, with the cutoff at 53 years old⁵. We can see that the effects of anticipated rotation and promotion are much stronger for young mayors than for old mayors. In Column (2), the two proxies are both precisely estimated and economically significant for young mayors. If the fraction of old secretaries in other cities increase by 1%, a firm raises the ratio of fixed assets to total assets by 0.114%. The two proxies become either imprecisely estimated or have the wrong sign for old mayors in Column (3). The results are reasonable, as anticipated job vacancy is much less relevant for old mayors than for the young mayors. As a consequence, the anticipated rotation and promotion induced by expected job vacancy have a much stronger impact on young mayors than on old mayors. In other words, Column (3) serves as a placebo test, while Column (2) is the preferred specification because anticipated job vacancy provides the strongest incentives for young mayors new in office.

Column (4) is another placebo test, as it examines young mayors serving in office for more than three years. As the average term of a mayor is 3.8 years, these mayors face a higher hazard of immediate turnover. Presumably, they have less incentive to promote long-term investment, which generates payoffs after these mayors finish their tenures. The results in Column (4) show that the coefficient on $s_{it} * R_{it}$ is similar in scale but imprecisely estimated, while the coefficient on C_{it} has a wrong sign.

Column (5) and Column (6) split the sample of young mayors new in office to test heterogeneous effects on private firms vs. SOEs. We can see that the coefficient on C_{it} is very precisely estimated for both private firms and SOEs, but the point estimator for private firms is about 40% higher than that for SOEs. The second proxy for the complementarity effect $s_{it} * R_{it}$ is precisely estimated for private firms, but it is imprecisely estimated with the wrong sign for SOEs. Taken together, Columns (5) and (6) lend additional support to the theory. We do find that the effects of anticipated rotation and promotion are smaller for SOEs, presumably because SOEs already enjoy large bargaining power against local

⁴Whenever possible, I also include whether the mayor is older than 56, and its interaction with R_{it} . They are always insignificant with very small point estimates. They are also dropped from most specifications as the sample is usually restricted to either young or old mayors.

⁵This cutoff allows me to cluster standard errors at city level for Column (3). If I use cutoffs larger than 54, there are two few cities to cluster for Column (3). The qualitative results are quite similar, however, if I employ different cutoffs.

politicians.

6.4 Timing of Anticipated Retirement on Composition of Capital

Table 4 summarizes the results to test whether there are any dynamically heterogeneous effects. Specifically, I define three separate C_{it} : C_{it}^1 is the fraction of secretaries in other cities that are more than 58 years old; C_{it}^2 is the fraction of secretaries in other cities between 56 and 57 years old; and C_{it}^3 is the fraction of secretaries in other cities between 53 and 55 years old. R_{it}^1 , R_{it}^2 , R_{it}^3 are defined in a similar manner. Thus, I run:

$$y_{it} = \alpha_0 + \mu_i + \lambda_t + X_{it}\rho_0 + \rho_1 s_{it} + \rho_2 R_{it} + \beta_1 s_{it} * R_{it}^1 + \beta_2 s_{it} * R_{it}^2 + \beta_3 s_{it} * R_{it}^3 + \gamma_1 C_{it}^1 + \gamma_2 C_{it}^2 + \gamma_3 C_{it}^3 + \varepsilon_{it}.$$

β_1 to β_3 and γ_1 to γ_3 capture the dynamic heterogeneity of anticipated personnel events, if any. Column (1) of Table 4 again includes all mayors who are new in office. We can see that the coefficient on C_{it}^2 is precisely estimated, while coefficients on C_{it}^1 and C_{it}^3 are either very small or have a negative sign. None of the coefficients on $s_{it} * R_{it}^2$, $s_{it} * R_{it}^3$, or $s_{it} * R_{it}^1$ are precisely estimated. In the preferred specification, Column (2), both the coefficients on C_{it}^2 and $s_{it} * R_{it}^2$ are precisely estimated, while all other coefficients are either imprecisely estimated or have negative signs. Columns (3) and (4) are placebo tests using old mayors and mayors in their 3rd or 4th year in office. Columns (5) and (6) again split the sample into private firms and SOEs, and we can see that the precisely estimated coefficients on C_{it}^2 and $s_{it} * R_{it}^2$ are mostly driven by private firms.

Hence, the most relevant personnel events are those during 3 to 4 years from now. Those that will happen within 2 years have few effects, while those that will happen in 5 to 7 years have, at most, negative effects. Anticipated personnel events within 2 years cannot be influenced much by long-term investment today, which usually generates payoffs in the more distant future. In other words, it is largely determined who will fill up the job vacancy available in 2 years. Thus, mayors have less incentive to promote long-term investment to bid for job vacancies available almost immediately. For those personnel events that occur more than 5 years from today, it is also less relevant for the mayors who have on average 3.8 years of tenure. The presence of many mayors and secretaries between the age 53 and 55 can even intensify competition for the most lucrative positions for the next few years, as these politicians are fully qualified for personnel shifts. As these politicians have one last big chance for a meaningful promotion before retirement and as they are also the most experienced politicians that can be promoted, they constitute the most formidable competitors for higher positions. This explains why we find a negative sign for β_3 and γ_3 .

By contrast, job vacancy available 3 to 4 years from now will be released at exactly the time when the new mayor, with an average tenure of 3.8 years, anticipates that her tenure will be finished. Thus, the availability of jobs elsewhere precisely 3 to 4 years from now is especially meaningful for the mayors who just took office. This explains why we see especially pronounced effects for job vacancies anticipated in the “intermediate” future.

7. Conclusion

The insights of this paper are applicable to generic organizations, as we can interpret the politician as a division manager and the entrepreneur as her subordinate. Thus, the models integrate key elements of personnel management, such as rotation, performance-based rewards, and authority relationships within a division. Starting from the simple intuition that a politician/manager can accumulate power over her tenure, the analysis reaches the surprising yet intuitive conclusion that strong performance-based rewards can be destructive. This is precisely because performance-based rewards encourage the politician/manager to accumulate power. The theory gives a simple rationale behind the management practice to implement intense rotation and performance-based rewards at the same time, which restores their desired disciplinary roles. Rotation-based solutions also deliver both flexibility and commitment, enabling organizations to achieve static and dynamic efficiencies.

There are several directions for future research. The paper presupposes high bureaucratic capacity with impersonal rules over politicians and shows that institutionalized rotation and performance-based rewards can energize the private economy. Where does such bureaucratic capacity come from? Specifically, is there any feedback from the private economy to bureaucratic capacity? The traditional view is that bureaucratic capacity is hurt by a strong private economy, as the best talents are recruited by the private sector (Caselli and Morelli, 2004). However, a strong market economy and impersonal bureaucracy arose together in many episodes (Fukuyama, 2011; Tackett, 2014). This begs the question of whether there is any positive feedback from a strong private sector to bureaucratic capacity. We also have many other episodes where strong bureaucracy lags significantly behind economic achievements (Fukuyama, 2011). It will be fruitful to construct a unified model and use the model for two exercises: 1) deriving conditions that determine the sign of the feedback’s direction 2) mapping key elements of the theory to real-world episodes.

Highly related to the above research agenda, it is intriguing to understand the historical origins and evolutions of bureaucratic capacity. I am now pursuing this line of research through a textual analysis of historical records of Imperial China. Specifically, a large number

of historical bureaucratic positions are identified from a dictionary of them, and correlations of these bureaucratic positions in a large database of historical records are being constructed. This provides a rich empirical platform to investigate institutional complementarity and substitution that can inspire future theoretical research on the origins and evolutions of bureaucratic capacity.

References

- [1] Acemoglu, Daron. “Why not a Political Coase Theorem? Social Conflict, Commitment, and Politics.” *Journal of Comparative Economics* 31, no. 4 (2003): 620-652.
- [2] Acemoglu, Daron, Philippe Aghion, and Fabrizio Zilibotti. “Distance to Frontier, Selection, and Economic Growth.” *Journal of the European Economic Association* 4, no. 1 (2006): 37-74.
- [3] Aghion, Philippe, and Peter Howitt. “A Model of Growth through Creative Destruction”. *Econometrica* 60, (1992): 323–351.
- [4] Aghion, Philippe, Ufuk Akcigit, and Peter Howitt. “What do We Learn from Schumpeterian Growth Theory?”. No. w18824. National Bureau of Economic Research, 2013.
- [5] Aghion, Philippe, and Peter Howitt. *Endogenous Growth Theory*. MIT Press, 1998.
- [6] Bai, Chong-En, Chang-Tai Hsieh, and Zheng Michael Song. “Crony Capitalism with Chinese Characteristics.” The University of Chicago working paper (2014).
- [7] Besley, Timothy, and Torsten Persson. *Pillars of Prosperity: The Political Economics of Development Clusters*. Princeton University Press, 2011.
- [8] Brandt, Loren, Johannes Van Biesebroeck, and Yifan Zhang. “Creative Accounting or Creative Destruction? Firm-level Productivity Growth in Chinese Manufacturing.” *Journal of Development Economics* 97, no. 2 (2012): 339-351.
- [9] Caselli, Francesco, and Massimo Morelli. “Bad Politicians.” *Journal of Public Economics* 88, no. 3 (2004): 759-782.
- [10] Chassang, Sylvain. “Building Routines: Learning, Cooperation, and the Dynamics of Incomplete Relational Contracts.” *The American Economic Review* 100, no. 1 (2010): 448-465.

- [11] Chen, Shuo. "From Governance to Institutionalization: Political Selection from the Perspective of Central-local Relations in China—Past and Present (1368-2010)" Department of Economics, Fudan University Working Paper (2016).
- [12] Dal Bo, Ernesto, and Martin A. Rossi. "Term Length and the Effort of Politicians." *The Review of Economic Studies* 78, no. 4 (2011): 1237-1263.
- [13] Evans, Peter, and James E. Rauch. "Bureaucracy and Growth: A Cross-National Analysis of the Effects of 'Weberian' State Structures on Economic Growth." *American Sociological Review* (1999): 748-765.
- [14] Eriksson, Tor, and Jaime Ortega. "The Adoption of Job Rotation: Testing the Theories." *Industrial and Labor Relations Review* 59, no. 4 (2006): 653-666.
- [15] Ferguson, Thomas, and Hans-Joachim Voth. "Betting on Hitler—The Value of Political Connections in Nazi Germany." *The Quarterly Journal of Economics* 123, no. 1 (2008): 101-137.
- [16] Fukuyama, Francis. *The Origins of Political Order: From Pre-Human Times to the French Revolution*. Farrar, Straus and Giroux, 2011.
- [17] Gibbons, Robert. "Piece-rate Incentive Schemes." *Journal of Labor Economics* 5, no. 4, Part 1 (1987): 413-429.
- [18] Gordon, Roger, and Wei Li. "Tax Structures in Developing Countries: Many Puzzles and a Possible Explanation." *Journal of Public Economics* 93, no. 7 (2009): 855-866.
- [19] Grossman, Sanford J., and Oliver D. Hart. "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration." *Journal of Political Economy* 94, no. 4 (1986): 691-719.
- [20] Hart, Oliver. *Firms, Contracts, and Financial Structure*. Clarendon Press, 1995.
- [21] Holmstrom, Bengt, and Paul Milgrom. "Multitask Principal-agent Analyses: Incentive Contracts, Asset Ownership, and Job Design." *Journal of Law, Economics, and Organization* 7 (1991): 24-52.
- [22] Holmstrom, Bengt, and Jean Tirole. "The Theory of the Firm." *Handbook of Industrial Organization* 1 (1989): 61-133.

- [23] Jin, Xin, and Michael Waldman. "Lateral Moves, Promotions, and Task-specific Human Capital: Theory and Evidence." (2017).
- [24] Khwaja, Asim Ijaz, and Atif Mian. "Tracing the Impact of Bank Liquidity Shocks: Evidence from an Emerging Market." *The American Economic Review* 98, no. 4 (2008): 1413-1442.
- [25] Kung, James Kai-sing, and Shuo Chen. "The tragedy of the Nomenklatura: Career Incentives and Political Radicalism during China's Great Leap Famine." *American Political Science Review* 105, no. 1 (2011): 27-45.
- [26] Laffont, Jean-Jacques. *Incentives and Political Economy*. Oxford University Press, 2000.
- [27] Lazear, Edward P., and Sherwin Rosen. "Rank-Order Tournaments as Optimum Labor Contracts." *Journal of Political Economy* 89, no. 5 (1981): 841-864.
- [28] Lazear, Edward P., and Paul Oyer. "Personnel Economics". In Robert Gibbons, and John Roberts. eds *Handbook of Organizational Economics*. (2012): 479-519.
- [29] Levin, Jonathan. "Relational Incentive Contracts." *The American Economic Review* 93, no. 3 (2003): 835-857.
- [30] Lieberthal, Kenneth. *Governing China: From Revolution to Reform (Second Edition)*. W. W. Norton & Company (2003).
- [31] Mann, Michael. *The Sources of Social Power, Vol. I: A History of Power from the Beginning to 1760 AD*. Cambridge University Press (1986).
- [32] Maskin, Eric, Yingyi Qian, and Chenggang Xu. "Incentives, Information, and Organizational Form." *The Review of Economic Studies* 67, no. 2 (2000): 359-378.
- [33] North, Douglass C., and Barry R. Weingast. "Constitutions and Commitment: The Evolution of Institutions Governing Public Choice in Seventeenth-Century England." *The Journal of Economic History* 49, no. 4 (1989): 803-832.
- [34] Olson, Mancur. *The Rise and Decline of Nations: Economic Growth, Stagflation and Social Rigidities*. Yale University Press, 1982.
- [35] Olson, Mancur. "Dictatorship, Democracy, and Development." *American Political Science Review* 87.03 (1993): 567-576.

- [36] Ortega, Jaime. "Job Rotation as a Learning Mechanism." *Management Science* 47, no. 10 (2001): 1361-1370.
- [37] Persson, Torsten, and Guido Tabellini. *Political Economics: Explaining Public Policy* (2000). The MIT Press.
- [38] Qian, Yingyi. "Incentives and Loss of Control in an Optimal Hierarchy." *The Review of Economic Studies* 61, no. 3 (1994): 527-544.
- [39] Raith, Michael, and Guido Friebel. "Managers, Training, and Internal Labor Markets." (2014).
- [40] Roland, Gerard. *Transition and Economics: Politics, Markets, and Firms*. MIT press, 2000.
- [41] Rose-Ackerman, Susan, and Bonnie J. Palifka. *Corruption and Government: Causes, Consequences, and Reform*. Cambridge University Press, 2016.
- [42] Sanchez de la Sierra, Raul. "On the Origins of the State: Stationary Bandits and Taxation in Eastern Congo". (2017).
- [43] Svulik, Milan W. *The Politics of Authoritarian Rule*. Cambridge University Press, 2012.
- [44] Tackett, Nicolas. *The Destruction of the Medieval Chinese Aristocracy*. Harvard University Press, 2014.
- [45] Wang, Shaoda. "Fiscal Competition and Coordination: Evidence from China." (2016).
- [46] Wedeman, Andrew. "Crossing the River by Feeling the Stones or Carried Across the Current? The Transformation of the Chinese Automotive Sector", in *Beyond the Middle Kingdom. Comparative Perspectives on China's Capitalist Transformation*, edited by Scott Kennedy. Stanford: Stanford University Press (2011): 66-88.
- [47] Williamson, Oliver E. *The Economic Institutions of Capitalism*. Simon and Schuster, 1985.
- [48] Xi, Tianyang, Yang Yao, and Muyang Zhang. "Bureaucratic Capability and Political Opportunism: An Empirical Investigation of City Officials in China." Working Paper. (2017).
- [49] Young, Alwyn. "The Razor's Edge: Distortions and Incremental Reform in the People's Republic of China." *The Quarterly Journal of Economics* 115, no. 4 (2000): 1091-1135.

Table 1: Summary Statistics Part I

	count	mean	sd	min	max
Firm characteristics					
ratio of fixed assets	2211700	.3529688	.228051	0	.9999979
fixed assets	2211700	8.574127	1.751689	0	19.15294
"flexible" assets	1810983	8.817254	1.753437	0	18.03701
output	2051817	10.24835	1.414884	0	19.70661
debt	2200074	9.178852	1.746334	0	18.85835
inventory	2097618	7.728052	1.879163	0	17.45033
management	2191117	6.900411	1.634475	0	17.41659
#employees	2058589	4.800024	1.148959	0	13.25281
firm age	2211601	12.15931	11.67262	1	363
ownership:					
state	2571858	.0718267	.2582007	0	1
private	2571858	.7132773	.452231	0	1
foreign	2571858	.2148731	.4107343	0	1
Mayor characteristics					
gender	2167305	1.950869	.2161407	1	2
ethnicity	2152417	.9724881	.1635695	0	1
age	2165908	50.48117	4.196176	36	60
age \leq 56	2179957	.9093711	.2870807	0	1
Work experience					
Youth league	2150124	.3670756	.4820075	0	1
Director of office	2149246	.5895435	.4919167	0	1

Table 1: Summary Statistics Part II

	count	mean	sd	min	max
Secretary characteristics					
gender	1782536	1.980095	.1396751	1	2
ethnicity	1734758	.9361456	.2444935	0	1
age	1770591	52.07351	3.766545	39	60
Secretary > 56	2179957	.2959815	.4564828	0	1
Work experience:					
Youth League	1781359	.345736	.4756078	0	1
Director of office	1790342	.6562059	.4749735	0	1
Anticipated retirement					
Fraction of other secretaries older than 56	1888847	.1735954	.1465631	0	1
Fraction of other mayors older than 56	1888847	.117085	.1096847	0	1

The following variables are in log scale: fixed assets, "flexible" assets, output, debt, inventory, management, number of employees. "Flexible" assets sums up liquid assets and intangible assets. Ownership variables are indicators: whether the firm is an SOE, private firm, or foreign firm. The ownership characterization is based on paid-in capital. For gender, 1 indicates female, 2 indicates male. For ethnicity, 0 indicates ethnic minority, 1 indicates *Han* Chinese. Work experience includes two dummies: whether the politician used to serve in Communist Youth League; whether the politician used to work as director of the office.

Table 2: Relevance of Others' Retirement on Own Turnover

	(1)	(2)	(3)	(4)	(5)
	Mayor Turnover				
Retirement of other secretaries	0.629** (0.301)	1.222*** (0.389)	0.830*** (0.317)	0.615*** (0.212)	0.784*** (0.185)
Retirement of other mayors	1.838*** (0.462)	2.702*** (0.634)	2.133*** (0.627)	1.808*** (0.509)	1.550*** (0.377)
Secretary age > 56	0.0737*** (0.0181)	0.298*** (0.0337)	0.255*** (0.0370)	0.254*** (0.0370)	0.257*** (0.0370)
Mayor age > 56	0.176*** (0.0395)	0.400*** (0.0680)	0.398*** (0.0647)	0.398*** (0.0649)	0.396*** (0.0672)
Party congress in: this year			0.152*** (0.0322)	0.163*** (0.0316)	0.162*** (0.0315)
2 years			0.157*** (0.0271)	0.170*** (0.0267)	0.164*** (0.0263)
3 years			0.0232 (0.0213)	0.0248 (0.0216)	0.00781 (0.0217)
Mayor FE	No	Yes	Yes	Yes	Yes
<i>N</i>	2884	2884	2884	2884	2884
adj. R^2	0.022	0.088	0.112	0.112	0.116

Standard errors are in parentheses and are clustered at mayor level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.
The dependent variable is whether the mayor's term is terminated in year t .

Table 3: Anticipated Job Vacancy on Capital Composition

	(1)	(2)	(3)	(4)	(5)	(6)
	fixed assets	fixed assets	fixed assets	fixed assets	fixed assets	fixed assets
fraction of other	0.120***	0.114***	0.0591	-0.0322	0.118***	0.0737***
secretaries old (C_{it})	(0.0232)	(0.0231)	(0.0496)	(0.0570)	(0.0242)	(0.0178)
own secretary old*	0.0611	0.157**	-0.581*	0.202*	0.184**	-0.0560
other old mayors ($s_{it} * R_{it}$)	(0.0778)	(0.0773)	(0.328)	(0.104)	(0.0782)	(0.0807)
fraction of other	-0.00108	-0.00411	0.107*	0.0974*	-0.00393	-0.0548
mayors old (R_{it})	(0.0468)	(0.0373)	(0.0579)	(0.0548)	(0.0393)	(0.0416)
own secretary old (s_{it})	0.000457	-0.0104	-0.0183	0.00745	-0.0128	0.00994
	(0.0106)	(0.0103)	(0.0196)	(0.00890)	(0.0103)	(0.0137)
Sample:						
Mayor age	all	young	old	young	young	young
Mayor term	early	early	early	late	early	early
Firm ownership	all	all	all	all	private+foreign	SOEs
N	376654	279576	78940	110072	258531	16117
adj. R^2	0.659	0.667	0.703	0.702	0.659	0.731

Standard errors are in parentheses and are clustered at city level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. The dependent variable is the ratio of fixed assets to total assets. All specifications control firm fixed effects, year fixed effects, the firm's characteristics, and the politicians' characteristics. For firm characteristics, I control (lag and logarithm of) output, the number of employees, value added, profit, management fee, inventory, firm age, and debt. For politician characteristics, I control (for both the mayor and the party secretary) age, gender, ethnicity, education, work experience (includes whether the politician used to work in the Communist Youth League and whether the politician used to work as a personal assistant or the director of the office for a senior politician).

Table 4: Timing of Anticipated Job Vacancy on Capital Composition

	(1)	(2)	(3)	(4)	(5)	(6)
	fixed assets	fixed assets	fixed assets	fixed assets	fixed assets	fixed assets
fraction of other secretaries	0.00644	0.0663	0.0790	0.00853	0.0737	0.0655*
retire in 2 years (C_{it}^1)	(0.0456)	(0.0488)	(0.0802)	(0.0689)	(0.0526)	(0.0376)
own secretary old * fraction of other	-0.0170	-0.150	1.560*	-0.0592	-0.134	-0.261***
mayors retire in 2 years ($s_{it} * R_{it}^1$)	(0.0950)	(0.128)	(0.853)	(0.219)	(0.136)	(0.0912)
fraction of other secretaries	0.117***	0.111***	0.0247	-0.0426	0.114***	0.0820***
retire 3 to 4 years from now (C_{it}^2)	(0.0296)	(0.0287)	(0.0475)	(0.0681)	(0.0301)	(0.0263)
own secretary old * fraction of other	0.0680	0.280**	-1.032***	0.101	0.307**	0.157
mayors retire 3 to 4 years ($s_{it} * R_{it}^2$)	(0.127)	(0.134)	(0.244)	(0.125)	(0.134)	(0.181)
fraction of other secretaries	-0.0331**	-0.0202	-0.0361	-0.00935	-0.0201	-0.0141
retire 5 to 7 years (C_{it}^3)	(0.0163)	(0.0202)	(0.0414)	(0.0294)	(0.0211)	(0.0201)
own secretary old * other	-0.0382	-0.107**	0.386	-0.158*	-0.104**	-0.133**
mayors retire 5 to 7 years ($s_{it} * R_{it}^3$)	(0.0383)	(0.0473)	(0.302)	(0.0802)	(0.0478)	(0.0659)
Sample:						
Mayor age	all	young	old	young	young	young
Mayor term	early	early	early	late	early	early
Firm ownership	all	all	all	all	private+foreign	SOEs
N	376654	279576	78940	110072	258531	16117
adj. R^2	0.660	0.668	0.705	0.703	0.661	0.732

Standard errors in parentheses and are clustered at city level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. The dependent variable is the ratio of fixed assets to total assets. All specifications control firm fixed effects, year fixed effects, the firm's characteristics, and the politicians' characteristics. For firm characteristics, I control (lag and logarithm of) output, the number of employees, value added, profit, management fee, inventory, firm age, and debt. For politician characteristics, I control (for both the mayor and the party secretary) age, gender, ethnicity, education, work experience (includes whether the politician used to work in the Communist Youth League and whether the politician used to work as a personal assistant or the director of the office for a senior politician).

Appendix

A1. Proofs for Propositions

Proposition 1 *The entrepreneur will invest under sufficiently strong performance-based rewards.*

Algebraically, denote \hat{R} such that $1 - F(w^*(\hat{R}))\{E(y|y \geq w^*(\hat{R})) - w^*(\hat{R})\} = k$. The entrepreneur expects a non-negative profit if $R \geq \hat{R}$.

Proof: By Assumption 1, $k > [1 - F(w^*(R = 0))]\{E(y|y \geq w^*(R = 0)) - w^*(R = 0)\}$. Recall that $\frac{w^*f(w^*)}{1-F(w^*)} = \frac{1}{1+R}$. As $R \rightarrow \infty$, $\frac{w^*f(w^*)}{1-F(w^*)} \rightarrow 0$. So $\lim_{R \rightarrow \infty} w^*(R) = 0$. $\lim_{R \rightarrow \infty} [1 - F(w^*(R))]\{E(y|y \geq w^*(R)) - w^*(R)\} = [1 - F(0)]\{E[y|y \geq 0] - 0\} = E[y] > k$ by Assumption 1.

Notice that:

$$\begin{aligned} & \frac{\partial}{\partial R} [1 - F(w^*)]\{E(y|y \geq w^*(R)) - w^*\} = \\ & \frac{\partial}{\partial R} \left\{ \int_{w^*}^{\bar{y}} z f(z) dz - [1 - F(w^*(R))]w^* \right\} \\ & = -w^*(R)f(w^*)\frac{dw^*}{dR} - \{[1 - F(w^*(R)) - w^*(R)f(w^*)]\frac{dw^*(R)}{dR}\} \\ & = -wf(w^*)\frac{dw^*}{dR} > 0 \text{ as } \frac{dw^*}{dR} < 0 \end{aligned}$$

So $\exists! \hat{R}$ such that $\forall R \geq \hat{R}$, $k \leq [1 - F(w^*)]\{E(y|y \geq w^*) - w^*\}$. This covers the sunk cost of investment for the entrepreneur.

Q.E.D.

Proposition 2 *1. Sufficiently strong rotation and performance-based rewards incentivizes the entrepreneur to invest.*

In algebra, denote $\hat{R}(k)$ and $\underline{\pi}(R, c)$ such that $[1 - F(w^*(\hat{R}))]\{E(y|y \geq w^*(\hat{R})) - w^*(\hat{R})\} = k$ and $\underline{\pi}(R, c) = 1 - \frac{c}{(1+R)E[y] - [1 - F(w^*(R))]\{w^* + E[Ry | w^*(R) \leq y]\}}$. If $R > \hat{R}(k)$ and $\pi \geq \underline{\pi}(R, c)$, the entrepreneur will reap non-negative return.

2. $\frac{\partial \underline{\pi}(R, c)}{\partial R} > 0$: the minimum rotation frequency increases with stronger performance-based rewards. If minimum rotation frequency does not change, stronger performance-based rewards incentivize politician in $t = 1$ to pay c , and politician in $t = 2$ will fully predate entrepreneur.

Proof: The first claim: If $\pi < \underline{\pi}$, the politician pays c , and the entrepreneur anticipates a net profit of $-k$ if he invests.

If $\pi \geq \underline{\pi}$, the politician will not pay c . The entrepreneur anticipates a net profit of $[1 - F(w^*(R))]\{E(y|y \geq w^*(R)) - w^*(R)\} - k$.

Proposition 1 already shows that under an uninformed politician, $\exists! \hat{R}$ such that $\forall R \geq \hat{R}$, $k \leq [1 - F(w^*)]\{E(y|y \geq w^*) - w^*\}$. This covers the sunk cost of investment for the entrepreneur.

The second claim: $\underline{\pi} = 1 - \frac{c}{(1+R)E[y] - [1-F(w)]\{w + E[Ry | y \geq w \}]}$. So it is sufficient to prove that $\Delta \equiv (1 + R)E[y] - \max_w [1 - F(w)]\{w + E[Ry | y \geq w \}]$ is an increasing function of R . Apply Envelope Theorem,

$$\begin{aligned} \frac{\partial \Delta}{\partial R} &= E[y] - [1 - F(w)]E[y|y \geq w^*] \\ &= \int_{\underline{y}}^{\bar{y}} z f(z) dz - [1 - F(w)] \frac{\int_{w^*}^{\bar{y}} z f(z) dz}{[1 - F(w)]} = \int_{\underline{y}}^{w^*} z f(z) dz > 0 \end{aligned}$$

Q.E.D.

Proposition 3 Denote $\underline{\pi}(R) \equiv 1 - \frac{c}{(1+R)E[y] - [1-F(w^*(R))]\{w^*(R) + E[Ry | y \geq w^*(R) \]}}$, $\bar{\pi}(R) \equiv 1 - \frac{\eta k}{[1-F(w^*(R))]\{w^*(R) + E[Ry | y \geq w^*(R) \]}}$. For $R \geq 0$, $\exists \pi \in [\underline{\pi}(R), \bar{\pi}(R)]$, such that the local politician neither acquire information nor steal capital.

Moreover, if $R' < R''$, then $\underline{\pi}(R') < \underline{\pi}(R'')$, and $\bar{\pi}(R') < \bar{\pi}(R'')$, In other words, the complementarity between rotation and performance-based rewards persist.

Proof:

To have $\underline{\pi} \leq \bar{\pi}$, we need:

$$1 - \frac{c}{(1 + R)E[y] - [1 - F(w^*)]\{w^* + E[Ry | y \geq w^* \]}} \leq 1 - \frac{\eta k}{[1 - F(w^*)]\{w^* + E[Ry | y \geq w^* \]}}$$

$$D \equiv (c + \eta k)([1 - F(w^*)]\{w^* + E[Ry | y \geq w^* \]}) - \eta k(1 + R)E[y] \geq 0$$

for $R = 0$, we have $D(R = 0) = (c + \eta k)(1 - F(w^*(0)))w^*(0) - \eta k E[y] > 0$, or $\frac{c}{\eta k} > \frac{E[y]}{[1 - F(w^*(0))]w^*(0)} - 1$.

Now we need $\Delta(R)$ to be an increasing function:

$$\Delta'(R) = (c + \eta k)([1 - F(w^*)]E[y|y \geq w^*]) - \eta k E[y] > 0$$

$$\frac{c}{\eta k} > \frac{\int_y^{w^*} z f(z) dz}{\int_{w^*}^y z f(z) dz} = \frac{F(w^*)E[y|y \leq w^*]}{[(1 - F(w^*))E[y|y \geq w^*]}$$

It is easy to see that $\max_R \frac{\int_y^{w^*} z f(z) dz}{\int_{w^*}^y z f(z) dz} = \frac{\int_y^{w^*(R=0)} z f(z) dz}{\int_{w^*(R=0)}^y z f(z) dz}$. So it is sufficient that:

$$\frac{c}{\eta k} > \frac{F(w^*(R=0))E[y|y \leq w^*(R=0)]}{[(1 - F(w^*(R=0)))E[y|y \geq w^*(R=0)]}$$

The second claim can be proved by noting that both $\underline{\pi}(R)$ and $\bar{\pi}(R)$ are increasing function of R . $\underline{\pi}'(R) > 0$ has been shown in Theorem 1. For $\bar{\pi}'(R) > 0$, it is shown in the text.

Q.E.D.

Proposition 4: *Suppose the politician paid c and is thus informed about the old project that has been invested.*

1. *She will endorse the new project with probability $F(\frac{1}{1+R}[1 - G(\tilde{w})]\{\tilde{w} + E_g[Ry|y \geq \tilde{w}]\})$, which increases with R : performance-based rewards encourage adaptation.*

2. $\forall R < \infty$, $F(\frac{1}{1+R}[1 - G(\tilde{w})]\{\tilde{w} + E_g[Ry|y \geq \tilde{w}]\}) < F(E_g(y))$, *the first-best probability of adaptation. An informed politician holds entrenched interests in the old project, not matter how strong performance-based rewards are.*

Proof: The politician will get $(1 + R)y$ from the old project, while get an expected value of $[1 - G(\tilde{w})]\{\tilde{w} + E_g[Ry|y \geq \tilde{w}]\}$ from the new project. He will endorse the new project if $(1 + R)y \leq [1 - G(\tilde{w})]\{\tilde{w} + E_g[Ry|y \geq \tilde{w}]\}$, which happens with probability $F(\frac{1}{1+R}[1 - G(\tilde{w})]\{\tilde{w} + E_g[Ry|y \geq \tilde{w}]\})$.

For the first claim:

$$\partial[F(\frac{1}{1+R}[1 - G(\tilde{w})]\{\tilde{w} + E_g[Ry|y \geq \tilde{w}]\})]/\partial R =$$

$$\begin{aligned}
&= f\left\{-\frac{1}{(1+R)^2}[1-G(\tilde{w})]\{\tilde{w}+E_g[Ry|y \geq \tilde{w}]\} + \frac{1}{1+R} \int_{\tilde{w}}^{\bar{y}} z g(z) dz\right\} \\
&= f\left\{\frac{(1+R) \int_{\tilde{w}}^{\bar{y}} z g(z) dz - [1-G(\tilde{w})]\tilde{w} - R \int_{\tilde{w}}^{\bar{y}} z g(z) dz}{(1+R)^2}\right\} = f\frac{\int_{\tilde{w}}^{\bar{y}} (z-\tilde{w})g(z) dz}{(1+R)^2} > 0
\end{aligned}$$

So the upper bound of $F(\frac{1}{1+R}[1-G(\tilde{w})]\{\tilde{w}+E_g[Ry|y \geq \tilde{w}]\})$ is $\lim_{R \rightarrow \infty} F(\frac{1}{1+R}[1-G(\tilde{w})]\{\tilde{w}+E_g[Ry|y \geq \tilde{w}]\}) = F(\lim_{R \rightarrow \infty} \frac{[1-G(\tilde{w})]\{\tilde{w}+E_g[Ry|y \geq \tilde{w}]\}}{1+R})$.

Using L'Hospital's Rule, $\lim_{R \rightarrow \infty} \frac{[1-G(\tilde{w})]\{\tilde{w}+E_g[Ry|y \geq \tilde{w}]\}}{1+R} = \lim_{R \rightarrow \infty} \int_{\tilde{w}(R)}^{\bar{y}} z g(z) dz = \int_{\underline{y}}^{\bar{y}} z g(z) dz = E_g(y)$. Notice that $E_g(y) < \bar{y}$, so $F(E_g(y)) < 1$.

Q.E.D.

Proposition 5 *Suppose the politician didn't pay the cost of learning and a new project arrives with $G(y) \leq F(y), \forall y$. Then the politician will always endorse the new project.*

Proof: By definition,

$$U_2 \equiv \max_w [1-G(w)]\{w+E_g[Ry|y \geq w]\} = [1-G(\tilde{w})]\{\tilde{w}+E_g[Ry|y \geq \tilde{w}]\} \geq$$

$$[1-G(w^*)]\{w^*+E_g[Ry|y \geq w^*]\} = [1-G(w^*)]w^* + R \int_{w^*}^{\bar{y}} z g(z) dz$$

Notice that $\int_{w^*}^{\bar{y}} z g(z) dz \geq \int_{w^*}^{\bar{y}} z f(z) dz$ because of F.O.S.D.⁶ Also, because $G(y) \leq F(y) \forall y$, $[1-G(w^*)]w^* \geq [1-F(w^*)]w^*$. So we have:

$$U_2 \geq [1-G(w^*)]w^* + R \int_{w^*}^{\bar{y}} z g(z) dz \geq [1-F(w^*)]w^* + R \int_{w^*}^{\bar{y}} z f(z) dz \equiv U_1$$

So it is optimal for the politician always to endorse the new project.

Q.E.D.

⁶This comes from an equivalent definition of first order stochastic dominance. Specifically, " G f.o.s.d. F " \Leftrightarrow " $G(y) \leq F(y) \forall y$ " \Leftrightarrow "For every weakly increasing utility function u , $\int u(x)dG \geq \int u(x)dF$."

Proposition 6 1. Sufficiently frequent rotation plus sufficiently strong performance-based rewards will induce a benevolent local government. The old entrepreneur will invest accordingly.

2. If $p < \frac{\int_{\underline{y}}^{w^*(0)} z f(z) dz}{\int_{\underline{y}}^{w^*(0)} z f(z) dz + \int_{\underline{y}}^{\bar{y}} z g(z) dz}$, then complementarity between rotation and performance-based rewards retains. Specifically, if x increases, then $\underline{\pi}$ increases, where:

$$\underline{\pi} = 1 - \frac{c}{(1-p)\{(1+R)E(y) - U_1\} + p\{(1-\Pi)[E_f[(1+R)y|(1+R)y \geq U_2] - U_2]\}} \quad (9)$$

Proof: Assume the politician doesn't learn the profitability of the first-period project. To induce investment from the first-period entrepreneur with $k > 0$, it must be that $k \leq (1-p)[1 - F(w^*)]\{E(y|y \geq w^*) - w^*\}$. Proof is similar to in the first part of Proposition 2, though the conditions on \hat{R} is stronger.

From (5), we immediately get $\pi \geq 1 - \frac{c}{(1-p)\{(1+R)E(y) - U_1\} + p\{(1-\Pi)[E_f[(1+R)y|(1+R)y \geq U_2] - U_2]\}} \equiv \underline{\pi}$.

To prove second part, it is sufficient to show that $Q \equiv (1-p)\{(1+R)E(y) - U_1\} + p\{(1-\Pi)[E_f[(1+R)y|(1+R)y \geq U_2] - U_2]\}$ is an increasing function of R .

$$\begin{aligned} \frac{\partial Q}{\partial R} &= (1-p)\{E(y) - \int_{w^*}^{\bar{y}} z f(z) dz\} + p \frac{\partial}{\partial R} \left\{ (1+R) \int_{\frac{U_2(R)}{1+R}}^{\bar{y}} z f(z) dz - (1 - F(\frac{U_2}{1+R})) U_2 \right\} \\ &= (1-p) \int_{\underline{y}}^{w^*} z f(z) dz + p \frac{\partial}{\partial R} \left\{ (1+R) \int_{\frac{U_2(R)}{1+R}}^{\bar{y}} z f(z) dz - (1 - F(\frac{U_2}{1+R})) U_2 \right\} \end{aligned}$$

Denote $M \equiv \frac{\partial}{\partial R} \left\{ (1+R) \int_{\frac{U_2(R)}{1+R}}^{\bar{y}} z f(z) dz - (1 - F(\frac{U_2}{1+R})) U_2 \right\}$:

$$M = \int_{\frac{U_2(R)}{1+R}}^{\bar{y}} z f(z) dz - (1+R) \frac{U_2(R)}{1+R} f\left[\frac{U_2(R)}{1+R}\right] \frac{\partial}{\partial R} \left[\frac{U_2(R)}{1+R}\right] - \frac{\partial}{\partial R} F\left[\frac{U_2(R)}{1+R}\right] U_2(R) - (1 - F\left[\frac{U_2(R)}{1+R}\right]) \int_{\tilde{w}}^{\bar{y}} z g(z) dz$$

$$\int_{\frac{U_2(R)}{1+R}}^{\bar{y}} z f(z) dz - U_2(R) f\left[\frac{U_2(R)}{1+R}\right] \frac{\partial}{\partial R} \left[\frac{U_2(R)}{1+R}\right] - f\left[\frac{U_2(R)}{1+R}\right] U_2(R) \frac{\partial}{\partial R} \left[\frac{U_2(R)}{1+R}\right] - (1 - F\left[\frac{U_2(R)}{1+R}\right]) \int_{\tilde{w}}^{\bar{y}} z g(z) dz$$

$$= \int_{\frac{U_2(R)}{1+R}}^{\bar{y}} z f(z) dz - (1 - F[\frac{U_2(R)}{1+R}]) \int_{\tilde{w}}^{\bar{y}} z g(z) dz \geq - \int_{\underline{y}}^{\bar{y}} z g(z) dz$$

So:

$$\frac{\partial Q}{\partial R} \geq (1-p) \int_{\underline{y}}^{w^*} z f(z) dz - p \int_{\underline{y}}^{\bar{y}} z g(z) dz \geq (1-p) \int_{\underline{y}}^{w^*(0)} z f(z) dz - p \int_{\underline{y}}^{\bar{y}} z g(z) dz$$

We can immediately see that when $p < \frac{\int_{\underline{y}}^{w^*(0)} z f(z) dz}{\int_{\underline{y}}^{w^*(0)} z f(z) dz + \int_{\underline{y}}^{\bar{y}} z g(z) dz}$, $\frac{\partial Q}{\partial R} \geq 0$.

Q.E.D.

Proposition 7 Denote:

$$\underline{\pi}_1(R) = 1 - \frac{c}{(1+R)E[y] - \max_w [1 - F(w)]\{w + E[Ry | y \geq w \}]}$$

$$\underline{\pi}_2(R) : (1 - \underline{\pi}_2(R)) \max_{w_2} [1 - F(w_2)]\{w_2 + E[Ry | y \geq w_2 \}] =$$

$$\max_{w_1} \{-F(w_1)c + (1 - \underline{\pi}_2(R))\{[1 - F(w_1)]\{w_1 + E[Ry | y \geq w_1 \}] + F(w_1)(1+R)E[y | y \geq w_1]\}\}$$

$$\hat{R} : k = [1 - F(w^*(\hat{R}))]\{E[y | y \geq w^*(\hat{R})] - w^*(\hat{R})\}, \frac{w^*(\hat{R})f(w^*(\hat{R}))}{1 - F(w^*(\hat{R}))} = \frac{1}{1+R}$$

1. If $\pi \geq \underline{\pi}_1(R)$, $\pi \geq \underline{\pi}_2(R)$, and $R \geq \hat{R}$, the politician does not pay c to learn about y , and the entrepreneur makes the investment k .

2.

$$\frac{\partial \underline{\pi}_1(R)}{\partial R} > 0, \frac{\partial \underline{\pi}_2(R)}{\partial R} > 0$$

So with stronger performance-based rewards, minimum rotation frequency $\underline{\Pi}(R) = \max\{\underline{\pi}_1(R), \underline{\pi}_2(R)\}$ also rises to guarantee the entrepreneur's ex ante investment.

Proof:

Denote:

$$w^* = \operatorname{argmax}_w [1 - F(w)] \{w + E[Ry | y \geq w]\}$$

w^* is the optimal extraction at $t = 2$ for an uninformed politician with no entrepreneur accepting w_1 .

Notice that at $t = 2$ newly appointed politician also honors w'_1 , the contract that the veteran politician made in $t = 1$. Also, if the entrepreneur rejects w'_1 , the information set for a newly appointed politician is exactly the same as the veteran politician. They also have the same utility. So they must offer the same w'_2 .

In $t = 2$, the dominant strategy is for entrepreneur to accept w'_2 if $y \geq w'_2$ and entrepreneur has not accepted in $t = 1$.

Back to $t = 1$. For all $w'_1 \in [\underline{y}, \bar{y}]$, For all $w'_1 \in [\underline{y}, \bar{y}]$, I now derive the set of strategies that can satisfy sequential rationality. This enables the politician to predict her payoff if she chooses a specific w'_1 .

1. Fix a $w'_1 \geq w^*$: There are two cases: the politician finds it sequentially rational to pay c or pay 0. The strategy at learning stage is a function of w'_1 . Denote $\operatorname{learn}(w'_1) = 1$ if the politician learns after w'_1 and $\operatorname{learn}(w'_1) = 0$ otherwise.

1.1 Suppose $\operatorname{learn}(w'_1) = 1$: Given that $\operatorname{learn}(w'_1) = 1$ is sequentially rational, the entrepreneur accepts any $y - w'_1 \geq 0$ as in $t = 2$ he gets 0 surplus.

Given that $\operatorname{learn}(w'_1) = 1$ is sequentially rational, the politician asks to extract $w'_2 = y$ for any entrepreneur who did not accept w'_1 .

The utility of w'_1 if $\operatorname{learn}(w'_1) = 1$ is sequentially rational is:

$$U(\{w'_1, \operatorname{learn}(w'_1) = 1\}) = -F(w'_1)c + (1-\pi)\{[1-F(w'_1)]\{w'_1 + E[Ry | y \geq w'_1]\} + (1+R)E[y | y \leq w'_1]\} + \pi \hat{U}$$

1.1 Suppose $\operatorname{learn}(w'_1) = 0$: If $y - w'_1 > y - w'_2$ or $w'_1 < w'_2$, entrepreneur accepts w'_1 if $y \geq w'_1$. This is impossible as:

$$w'_2 = \operatorname{argmax}_w \operatorname{Prob}(w \leq y | y \leq w'_1) \{w + E[Ry | w \leq y, \text{ given } y \leq w'_1]\} < w'_1$$

If $y - w'_1 \leq y - w'_2$ or $w'_1 \geq w'_2$, entrepreneur rejects w'_1 . In this case:

$$w'_2 = \operatorname{argmax}_w [1 - F(w)] \{w + E[Ry | y \geq w]\} = w^*$$

Indeed $w'_1 \geq w'_2 = w^*$, consistent with the assumption that $w'_1 \geq w^*$.

So given that $learn(w'_1) = 0$ is sequentially rational:

$$U(\{w'_1, learn(w'_1) = 0\}) = (1 - \pi)[1 - F(w^*)]\{w^* + E[Ry | y \geq w^*]\} + \pi \hat{U}$$

So by choosing $w'_1 \geq w^*$, the maximum utility the politician can get is:

$$\begin{aligned} \max \{ & -F(w'_1)c + (1 - \pi)\{(1 + R)E[y | y \leq w'_1] + [1 - F(w'_1)]\{w'_1 + E[Ry | y \geq w'_1]\} + \pi \hat{U}\} \\ & , (1 - \pi)[1 - F(w^*)]\{w^* + E[Ry | y \geq w^*]\} + \pi \hat{U} \} \end{aligned}$$

2. Fix a $w'_1 < w^*$:

2.1 Suppose $learn(w'_1) = 1$: The algebra is the same as 1.1. We have:

$$U(\{w'_1, learn(w'_1) = 1\}) = -F(w'_1)c + (1 - \pi)\{[1 - F(w'_1)]\{w'_1 + E[Ry | y \geq w'_1]\} + (1 + R)E[y | y \leq w'_1]\} + \pi \hat{U}$$

1.1 Suppose $learn(w'_1) = 0$: If $y - w'_1 > y - w'_2$ or $w'_1 < w'_2$, entrepreneur accepts w'_1 if $y \geq w'_1$. This is impossible as:

$$w'_2 = \operatorname{argmax}_w \operatorname{Prob}(w \geq w | w \leq w'_1) \{w + E[Ry | w \leq y, \text{ given } y \leq w'_1]\} < w'_1$$

If $y - w'_1 < y - w'_2$ or $w'_1 > w'_2$, entrepreneur rejects w'_1 . In this case:

$$w'_2 = \operatorname{argmax}_w [1 - F(w)] \{w + E[Ry | y \geq w]\} = w^*$$

Another contradiction as we assume $w'_1 > w'_2$ and $w^* > w'_1$.

The only possible case is $w'_1 = w'_2$. Given Assumption 4, suppose that the cutoff is \hat{w} such that an entrepreneur with $y \geq \hat{w}$ accepts w'_1 . We have $\hat{w} > w'_1$. Entrepreneur with $w'_1 \leq y < \hat{w}$ accepts w'_2 . To make $w'_2 = w'_1$ sequentially rational, it must be that \hat{w} satisfies:

$$w'_1 = w'_2 = \operatorname{argmax}_w \operatorname{prob}(w \leq y, \text{ given } y \leq \hat{w}) \{w + E[Ry | w \leq y, \text{ given } y \leq \hat{w}]\}$$

The probability density function of $f(y | y \leq \hat{w}) = \frac{f(y)}{F(\hat{w})} \equiv g(y, \hat{w})$ with support on $[y, \hat{w}]$. We have:

$$\text{prob}(w \leq y, \text{ given } y \leq \hat{w}) = \int_w^{\hat{w}} f(z|z \leq \hat{w})dz = \frac{\int_w^{\hat{w}} f(z) dz}{1 - F(\hat{w})} = \frac{F(\hat{w}) - F(w)}{F(\hat{w})}$$

$$E[y|w \leq y, \text{ given } y \leq \hat{w}] = \int_w^{\hat{w}} zg(z, \hat{w}|z \geq w)dz$$

$$g(z, \hat{w}|z \geq w) = \frac{g(z, \hat{w})}{1 - G(w; \hat{w})} = \frac{f(z)/F(\hat{w})}{1 - \int_y^w \frac{f(v)}{F(\hat{w})}dv} = \frac{f(z)/F(\hat{w})}{1 - \frac{F(w)}{F(\hat{w})}} = \frac{f(z)}{F(\hat{w}) - F(w)}$$

Thus we have:

$$E[y|w \leq y, \text{ given } y \leq \hat{w}] = \frac{\int_w^{\hat{w}} zf(z)dz}{F(\hat{w}) - F(w)}$$

So:

$$\text{prob}(w \leq y, \text{ given } y \leq \hat{w})\{w + RE[y|w \leq y, \text{ given } y \leq \hat{w}]\}$$

$$= \frac{F(\hat{w}) - F(w)}{F(\hat{w})}w + R \frac{F(\hat{w}) - F(w)}{F(\hat{w})} \frac{\int_w^{\hat{w}} zf(z)dz}{F(\hat{w}) - F(w)}$$

$$= \frac{[F(\hat{w}) - F(w)]w + R \int_w^{\hat{w}} zf(z)dz}{F(\hat{w})}$$

As $w'_1 = w'_2 = \text{argmax}_w = \frac{[F(\hat{w}) - F(w)]w + \int_w^{\hat{w}} zf(z)dz}{F(\hat{w})}$, F.O.C. tells us that:

$$\frac{w'_1 f(w'_1)}{F(\hat{w}) - F(w'_1)} = \frac{1}{1 + R}$$

As $\frac{1}{1+R} = \frac{w^* f(w^*)}{1 - F(w^*)}$, it is indeed the case that $w'_1 < w^*$: $\frac{1}{1+R} = \frac{w^* f(w^*)}{1 - F(w^*)} = \frac{w'_1 f(w'_1)}{F(\hat{w}) - F(w'_1)} > \frac{w'_1 f(w'_1)}{1 - F(w'_1)}$.

⁷Note that when $w'_1 \geq w^*$, I did not discuss this equilibrium where $w'_1 = w'_2$, $y \geq \hat{w}$ accepts w'_1 , and $w'_2 \leq y < w'_1$ accepts w'_2 . Such equilibrium does not exist, as it requires:

$$w'_1 = w'_2 = \text{argmax}_w \frac{[F(\hat{w}) - F(w)]w + R \int_w^{\hat{w}} zf(z)dz}{F(\hat{w})}$$

So if $learn(w'_1) = 0$ is sequentially rational following a $w'_1 < w^*$, then:

$$U(\{w'_1, w'_1(learn) = 0\}) = (1 - \pi)\{[1 - F(\hat{w})]\{w'_1 + E[Ry | y \geq \hat{w}]\} \\ + F(\hat{w})\frac{[F(\hat{w}) - F(w'_1)]w'_1 + R \int_{w'_1}^{\hat{w}} z f(z) dz}{F(\hat{w})}\} + \pi \tilde{U}$$

$$U(\{w'_1, learn(w'_1) = 0\}) = (1 - \pi)\{[1 - F(\hat{w})]\{w'_1 + E[Ry | y \geq \hat{w}]\} + [F(\hat{w}) - F(w)]w + R \int_w^{\hat{w}} z f(z) dz\} + \pi \tilde{U}$$

The maximum utility of the politician from a $w'_1 < w^*$ is:

$$max\left\{-F(w'_1)c + (1 - \pi)\{[1 - F(w'_1)]\{w'_1 + E[Ry | y \geq w'_1]\} + (1 + R)E[y | y \leq w'_1] + \pi \hat{U}\},\right.$$

$$\left.(1 - \pi)\{[1 - F(\hat{w})]\{w'_1 + E[Ry | y \geq \hat{w}]\} + [F(\hat{w}) - F(w)]w + R \int_w^{\hat{w}} z f(z) dz\} + \pi \hat{U}\right\}$$

We have derived the function of politician's payoff at $t = 1$ for any w'_1 and learning decision. We can use this function to find the sequential equilibrium strategy.

I want to ensure that $learn(w'_1) = 0$ is the strategy on equilibrium path. That gives a utility of:

$$U(w'_1, learn(w'_1) = 0) =$$

$$\begin{cases} (1 - \pi)[1 - F(w^*)]\{w^* + E[Ry | y \geq w^*]\} + \pi \hat{U} & w'_1 \geq w^* \\ (1 - \pi)\{[1 - F(\hat{w})]\{w_1 + E[Ry | y \geq \hat{w}]\} + \{[F(\hat{w}) - F(w)]w + R \int_w^{\hat{w}} z f(z) dz\} + \pi \hat{U} & w'_1 < w^* \end{cases}$$

Notice that:

$$[1 - F(\hat{w})]\{w_1 + E[Ry | y \geq \hat{w}]\} + \{[F(\hat{w}) - F(w)]w + R \int_w^{\hat{w}} z f(z) dz\}$$

We still need $\frac{w^* f(w^*)}{1 - F(w^*)} = \frac{1}{1+x} = \frac{w'_1 f(w'_1)}{F(\hat{w}) - F(w'_1)} > \frac{w'_1 f(w'_1)}{1 - F(w'_1)}$, so $w'_1 < w^*$. A contradiction to the assumption that $w'_1 \geq w^*$.

$$\begin{aligned}
&= w'_1 \int_{\hat{w}}^{\bar{y}} f(z)dz + R \int_{\hat{w}}^{\bar{y}} zf(z)dz + w'_1 \int_{w'_1}^{\hat{w}} f(z)dz + R \int_{w'_1}^{\hat{w}} zf(z)dz \\
&= w'_1 \int_{w'_1}^{\bar{y}} f(z)dz + R \int_{w'_1}^{\bar{y}} zf(z)dz = [1 - F(w'_1)]\{w'_1 + E[Ry | y \geq w'_1]\} \\
&\leq [1 - F(w^*)]\{w^* + E[Ry | y \geq w^*]\}
\end{aligned}$$

So it is always optimal to choose $w'_1 \geq w^*$.

In that case, I want to make sure that at the stage of learning, it is not optimal to learn:

$$-c + (1 - \pi)(1 + R)E[y] \leq (1 - \pi)[1 - F(w^*)]\{w^* + E[Ry | y \geq w^*]\}$$

And at the stage to decide w'_1 , the utility from not learning is also larger than the highest utility from learning:

$$(1 - \pi) \max_{w_2} [1 - F(w_2)]\{w_2 + E[Ry | y \geq w_2]\} \geq$$

$$\max_{w_1} \{-F(w_1)c + (1 - \pi)\{[1 - F(w_1)]\{w_1 + E[Ry | y \geq w_1]\} + F(w_1)(1 + R)E[y | y \geq w_1]\}\}$$

$\underline{\pi}(R)$ is defined as:

$$(1 - \underline{\pi}(R)) \max_{w_2} [1 - F(w_2)]\{w_2 + E[Ry | y \geq w_2]\} =$$

$$\max_{w_1} \{-F(w_1)c + (1 - \underline{\pi}(R))\{[1 - F(w_1)]\{w_1 + E[Ry | y \geq w_1]\} + F(w_1)(1 + R)E[y | y \geq w_1]\}\}$$

By Envelope Theorem and Implicit Function Theorem:

$$(1 - \underline{\pi}(R))[1 - F(w^*)][E[y | y \geq w^*] - \underline{\pi}'(R)U^*] = (1 - \underline{\pi}(R))\{[1 - F(w_1^*)]E[y | y \geq w_1] + F(w_1^*)E[y | y \leq w_1^*]\} - \underline{\pi}'(R)\tilde{U}^*$$

where $U^* = [1 - F(w^*)]\{w^* + E[Ry | y \geq w^*]\}$, $\tilde{U}^* = [1 - F(w_1^*)]\{w_1^* + E[Ry | y \geq w_1^*]\} + F(w_1^*)(1 + R)E[y | y \geq w_1^*]$, we need $\tilde{U}^* - U^* > 0$ for the analysis to be meaningful. So:

$$\underline{\pi}'(R) = \frac{1 - \underline{\pi}(R)}{\tilde{U}^* - U^*} \{[1 - F(w_1^*)]E[y | y \geq w_1] + F(w_1^*)E[y | y \leq w_1^*] - [1 - F(w^*)][E[y | y \geq w^*]]\}$$

$$= \frac{1 - \underline{\pi}(R)}{\tilde{U}^* - U^*} \left[\int_{\underline{y}}^{\bar{y}} z f(z) dz - \int_{\underline{y}}^{w^*} z f(z) dz \right] = \frac{1 - \underline{\pi}(R)}{\tilde{U}^* - U^*} \int_{\underline{y}}^{w^*} z f(z) dz > 0$$

$\underline{\Pi}(R) = \max\{\underline{\pi}_1(R), \underline{\pi}_2(R)\}$ is an increasing function of R because $\underline{\pi}_1'(R) > 0$ and $\underline{\pi}_2'(R) > 0$.

Q.E.D.

A2. The Model when the New Politician is never Informed

In this appendix, I validate my claim that it is not necessary for the newly appointed politician to know everything about the project and we can still uncover the complementarity, though under more restrictive conditions. Suppose that the cost of learning is a random variable $c \sim H(c)$. To illustrate the extreme case, assume that the newly appointed politician knows nothing about the project except $F(y)$, even if the old politician invests c . The old politician will invest c if:

$$c < (1 - \pi)\{(1 + R)E[y] - [1 - F(w^*(R))]\{w^* + E[Ry | w^*(R) \leq y]\}\} = (1 - \pi)\Delta(R).$$

Hence, the expected surplus of the entrepreneur should satisfy:

$$\{\pi + (1 - \pi)\{1 - H[(1 - \pi)\Delta(R)]\}\} \overbrace{[1 - F(w^*(R))]\{E(y|y \geq w^*(R)) - w^*(R)\}}^{\equiv S(R)} > k.$$

With probability π , the politician is replaced with a new one who is uninformed, and the entrepreneur reaps a surplus of $S(R)$. With probability $1 - \pi$, the same politician governs in $t = 1$ and $t = 2$, and she does not pay c with probability $1 - H[(1 - \pi)\Delta(R)]$. Thus, the expected surplus for the entrepreneur is $\{\pi + (1 - \pi)\{1 - H[(1 - \pi)\Delta(x)]\}\}S(R) = \{1 - (1 - \pi)\{1 - H[(1 - \pi)\Delta(R)]\}\}S(R)$.

The left-hand side is monotonically increasing in π . Define $\underline{\pi}$ such that:

$$\underbrace{\{1 - (1 - \underline{\pi})H[(1 - \underline{\pi})\Delta(R)]\}}_{M(\underline{\pi}, R)} [1 - F(w^*(R))]\{E(y|y \geq w^*(R)) - w^*(R)\} = k.$$

We have:

$$HS(R)d\underline{\pi} + (1 - \underline{\pi})h\Delta(R)S(R)d\underline{\pi} - (1 - \underline{\pi})^2h\Delta'(R)S(R)dx + M(\underline{\pi}, R)S'(R)dR = 0$$

$$\frac{d\pi}{dR} = \frac{(1 - \pi)^2 h \Delta'(R) S(R) - M(\pi, R) S'(R)}{HS(R) + (1 - \pi) h \Delta(R)}.$$

With $\Delta'(R) > 0$, $S'(R) > 0$, we have two competing effects. $S'(R) > 0$ formalizes the notion that performance-based rewards induce an uninformed politician to extract less. However, stronger performance-based rewards increase the temptation to become informed in the first place, as $\Delta'(R) > 0$ shows. The sign of $\frac{d\pi}{dR}$ depends on the relative strength of the two mechanisms. Nonetheless, the main intuition is robust: the politician will be more likely to invest in information acquisition with stronger performance rewards, and the complementarity between rotation and performance rewards is preserved if the information acquisition effect is sufficiently strong.

For Section 4.2, we have similar results. Again, assume that the old politician's cost of information acquisition $c \sim H(c)$. She will pay c with probability $H((1 - \pi)\{(1 - p)(1 + R)E(y) + p\{\Pi U_2 + (1 - \Pi)E[(1 + R)y | (1 + R)y \geq U_2]\} - \{(1 - p)U_1 + pU_2\}\}) \equiv H((1 - \pi)\tilde{\Delta}(R))$. The second claim of Proposition 6 shows that $\tilde{\Delta}'(R) > 0$ (as $\pi = 1 - \frac{c}{\tilde{\Delta}(R)}$ and Proposition 6 shows that $d\pi/dR > 0$). The minimum rotation frequency to induce investment from the entrepreneur is:

$$k = (1 - p) \underbrace{\{[1 - (1 - \pi)H[(1 - \pi)\tilde{\Delta}(R)]]\}}_{\equiv \tilde{M}(\pi, R)} \overbrace{\{[1 - F(w^*(R))\{E_f[y | y \geq w^*(R)] - w^*(R)\}]}}^{=S(R)}.$$

The entrepreneur with project $y \sim F(y)$ gets support from the uninformed politician only if the new project does not arrive. If a new project arrives, an uninformed politician always supports the new project, while an informed politician awards zero rent to the entrepreneur with project $y \sim F(y)$ regardless of which project she chooses to endorse. However, when the new project does not arrive, we have the same situation as in Section 3.2. Because $\tilde{\Delta}'(R) > 0$, $S'(R) > 0$, the two competing effects determine whether rotation complements performance or not.

A3. General Functional Forms for Performance-based Rewards

We can construct a more elaborated model for the reduced-form formulation of the key parameter x to facilitate its interpretation. Suppose now that the game is infinitely repeated, with a discount factor $\beta \in (0, 1)$. In each period, the politician has a chance of promotion with probability $V(y, \alpha)$, where $\frac{\partial V}{\partial y} > 0$, $\frac{\partial V}{\partial \alpha} > 0$, and $V \in (0, 1)$: higher output means

higher likelihood of promotion, and α is an exogenous shifter that increase the probability of promotion for any output level. Also, $V(0, \alpha) = 0$: there is no promotion opportunity for the politician without any achievements. Suppose in equilibrium the utility of the politician is U , and a promotion event adds a utility of R to that. Suppose that the rotation frequency $\pi = 0$. By definition:

$$U = \max_w [1 - F(w)]w + \beta \left\{ \int_{w^*}^{\bar{y}} \{V(z, \alpha)[U + R] + (1 - V(z, \alpha))U\} f(z) dz + \int_{\underline{y}}^{w^*} U f(z) dz \right\}$$

$$= [1 - F(w^*)]w^* + \beta U + \beta \int_{w^*}^{\bar{y}} V(z, \alpha) f(z) dz$$

$$U = \frac{1}{1 - \beta} [1 - F(w^*)]w^* + \frac{\beta}{1 - \beta} R \int_{w^*}^{\bar{y}} V(z, \alpha) f(z) dz$$

The politician derives a utility from a stream of future rents, which is $\frac{1}{1 - \beta} [1 - F(w^*)]w^*$; she also values the promotion opportunity, which gives her an additional utility of $\frac{\beta}{1 - \beta} R \int_{w^*}^{\bar{y}} V(z, \alpha) f(z) dz$.

The F.O.C. yields:

$$1 - F(w^*) - f(w^*)w^* - \beta R V(w^*, \alpha) f(w^*) = 0$$

$$w^* + \beta R V(w^*, \alpha) = \frac{1 - F(w^*)}{f(w^*)} \equiv s(w^*)$$

$s' < 0$ by Monotone Hazard Rate Property. We can do standard comparative statics of w^* with respect to β :

$$dw^* + R V d\beta + \beta R \frac{\partial}{\partial w^*} V(w^*, \alpha) = s' dw^*$$

$$\frac{\partial w^*}{\partial \beta} = - \frac{R V(w^*, \alpha)}{1 - s'(w^*) + \beta R \frac{\partial}{\partial w^*} V(w^*, \alpha)} < 0$$

Similarly, we can derive:

$$\frac{\partial w^*}{\partial R} = - \frac{\beta V(w^*, \alpha)}{1 - s'(w^*) + \beta R \frac{\partial}{\partial w^*} V(w^*, \alpha)} < 0$$

$$\frac{\partial w^*}{\partial \alpha} = -\frac{\beta \frac{\partial}{\partial \alpha} V(w^*, \alpha)}{1 - s'(w^*) + \beta R \frac{\partial}{\partial \beta} V(w^*, \alpha)} < 0$$

So an uninformed politician extracts fewer rents if: promotion opportunity is valuable; the politician is patient; the probability of promotion increases (because, for example, more higher positions are available). These are all very intuitive results.

A3.1 The Complementarity between Rotation and Performance-based Rewards

This section proves similar results as Proposition 2. If the politician choose to pay c , her payoff will be:

$$\begin{aligned} U' &= E[y] + \beta \int_{\underline{y}}^{\bar{y}} \{V(z, \alpha)[U + R] + (1 - V(z, \alpha))U\} f(z) dz \\ &= E[y] + \beta U + \beta \int_{\underline{y}}^{\bar{y}} V(z, \alpha) R f(z) dz \end{aligned}$$

$$\Delta = U' - U = E[y] + \beta R \int_{\underline{y}}^{\bar{y}} V(z, \alpha) f(z) dz - \{[1 - F(w^*)]w^* + \beta U + \beta R \int_{w^*}^{\bar{y}} V(z, \alpha) f(z) dz\}$$

We find that:

$$\frac{\partial U}{\partial R} = \beta \int_{\underline{y}}^{w^*} V(z, \alpha) f(z) dz > 0, \quad \frac{\partial U}{\partial \beta} = R \int_{\underline{y}}^{w^*} V(z, \alpha) f(z) dz > 0, \quad \frac{\partial U}{\partial \alpha} = \beta R \int_{\underline{y}}^{w^*} \frac{\partial V(z, \alpha)}{\partial \alpha} f(z) dz > 0$$

So the politician wants to accumulate local knowledge if: promotion opportunity is valuable; the politician is patient; the probability of promotion increases. The traditional parameters that should improve welfare all increase the benefit of local knowledge and exacerbate the temptation to learn.

With probability π of rotation, the politician does not pay c if:

$$(1 - \pi) \{[1 - F(w^*)]w^* + \beta U + \beta \int_{w^*}^{\bar{y}} V(z, \alpha) f(z) dz\} + \pi \tilde{U} \geq (1 - \pi) \{E[y] + \beta U + \beta \int_{\underline{y}}^{\bar{y}} V(z, \alpha) R f(z) dz\} + \pi \tilde{U} - c$$

$$\pi \geq 1 - \frac{c}{E[y] + \beta R \int_{\underline{y}}^{\bar{y}} V(z, \alpha) f(z) dz - \{[1 - F(w^*)]w^* + \beta U + \beta R \int_{w^*}^{\bar{y}} V(z, \alpha) f(z) dz\}} = 1 - \frac{c}{\Delta} \equiv \underline{\pi}$$

$$\frac{\partial \pi}{\partial R} > 0, \frac{\partial \pi}{\partial \beta} > 0, \frac{\partial \pi}{\partial \alpha} > 0$$

By taking a reduced-form approach in Section 3, we can neatly summarize the key finding that $\frac{\partial \pi}{\partial R} > 0$, $\frac{\partial \pi}{\partial \beta} > 0$, and $\frac{\partial \pi}{\partial \alpha} > 0$ in one equation. We can do so because the mechanisms that drives $\frac{\partial \pi}{\partial R} > 0$, $\frac{\partial \pi}{\partial \beta} > 0$, and $\frac{\partial \pi}{\partial \alpha} > 0$ are similar. Higher values of R , β , or α all raise the benefit of local knowledge that necessitates intense rotation.

The same with Section 3.2, $\pi \geq \underline{\pi}$ is a necessary condition for the entrepreneur to invest. We also need that the entrepreneur expects more surplus than the cost of investment: $S \equiv [1 - F(w^*)]\{E[y|y \geq w^*] - w^*\} \geq k$. We can show that $\frac{\partial S}{\partial R} > 0$, $\frac{\partial S}{\partial \beta} > 0$, $\frac{\partial S}{\partial \alpha} > 0$:

$$S = [1 - F(w^*)]\{E[y|y \geq w^*] - w^*\} = \int_{w^*}^{\bar{y}} z f(z) dz - w^* \int_{w^*}^{\bar{y}} f(z) dz$$

$$\frac{\partial S}{\partial R} = - \int_{w^*}^{\bar{y}} f(z) dz \frac{\partial w^*}{\partial R} > 0$$

Proofs for $\frac{\partial S}{\partial \beta} > 0$, $\frac{\partial S}{\partial \alpha} > 0$ are similar. We can summarize this section by (without loss of generality, set $\beta = 1$):

Proposition A2: 1. *Sufficiently strong rotation and performance-based rewards incentivize the entrepreneur to invest.*

In algebra, denote $\hat{R}(k, \alpha)$ and $\underline{\pi}(R, \alpha, c)$ such that $[1 - F(w^*(\hat{R}))]\{E(y|y \geq w^*(\hat{R})) - w^*(\hat{R})\} = k$ and $\underline{\pi}(R, \alpha, c) = 1 - \frac{c}{E[y+V(y,\alpha)R] - [1-F(w^*)]\{w^*+E[V(y,\alpha)R|w^* \leq y]\}}$. If $R > \hat{R}(k, \alpha)$ and $\pi \geq \underline{\pi}(R, \alpha, c)$, the politician does not pay c , and the entrepreneur will reap a non-negative return.

2. $\frac{\partial \underline{\pi}(R, \alpha, c)}{\partial R} > 0$ and $\frac{\partial \underline{\pi}(R, \alpha, c)}{\partial \alpha} > 0$: *the minimum rotation frequency increases when performance-based rewards are stronger. If the minimum rotation frequency does not change, more intense performance-based rewards incentivize the politician in $t = 1$ to pay c , and the politician in $t = 2$ will fully predate the entrepreneur.*

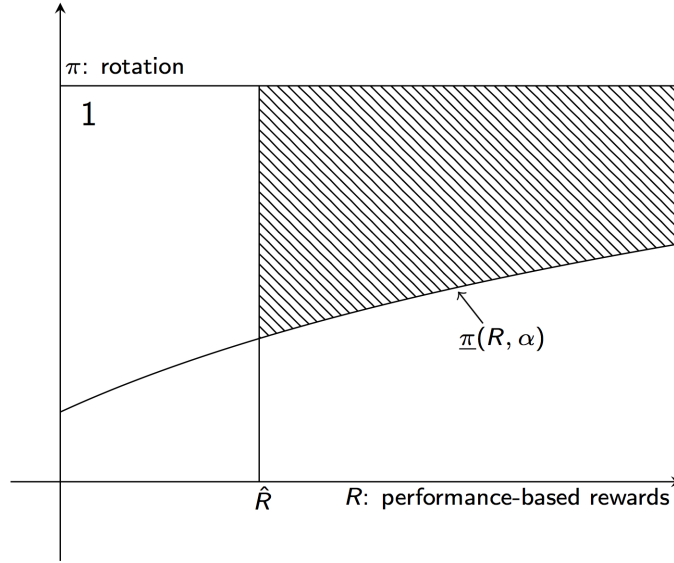


Figure A1: fixed α .

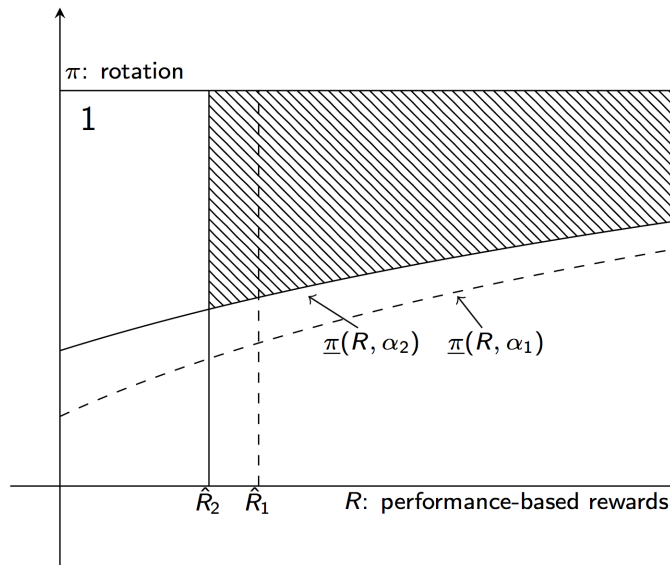


Figure A2: $\alpha_1 > \alpha_2$.

Notice that these key results impose minimal assumption on $V(y, \alpha)$: we only require that $V(y, \alpha) \in [0, 1]$ is an increasing function on y and α . This is true for all the following sections (except Section A3.4). Specifically, we don't need to make any assumptions on the curvatures of V or the cross derivatives. This demonstrates that the complementarity between rotation and performance-based rewards is highly generic.

A3.2 Roving bandits and the Complementarity between Rotation and Performance-based Rewards

Without loss of generality, let us re-focus on the two-period model and assumes $\beta = 1$. At the end of her term, the politician either gets a promotion (with payoff R) or enters retirement (with payoff 0). The politician will not steal capital if:

$$\eta k + \pi * \tilde{U} + (1 - \pi) * 0 \leq \pi * \tilde{U} + (1 - \pi) * \{[1 - F(w^*)]w^* + R \int_{w^*}^{\bar{y}} V(z, \alpha) f(z) dz\}$$

$$\pi \leq 1 - \frac{\eta k}{[1 - F(w^*)]w^* + R \int_{w^*}^{\bar{y}} V(z, \alpha) f(z) dz} \equiv \bar{\pi}(R, \alpha)$$

Notice that:

$$\frac{\partial \bar{\pi}}{\partial R} = \frac{\eta k}{\{[1 - F(w^*)]w^* + R \int_{w^*}^{\bar{y}} V(z, \alpha) f(z) dz\}^2} \int_{w^*}^{\bar{y}} V(z, \alpha) f(z) dz > 0$$

Also:

$$\frac{\partial \bar{\pi}}{\partial \alpha} = \frac{\eta k}{\{[1 - F(w^*)]w^* + R \int_{w^*}^{\bar{y}} V(z, \alpha) f(z) dz\}^2} R \int_{w^*}^{\bar{y}} \frac{\partial V(z, \alpha)}{\partial \alpha} f(z) dz > 0$$

So a rise in R , or α raises the stake of promotion opportunity, making the politician less tempted to steal private capital. With similar regularity assumption (as Assumption 3), we have:

Proposition A3: Denote $\underline{\pi} \equiv 1 - \frac{c}{E[y+V(y,\alpha)R]-[1-F(w^*)]\{w^*+E[V(y,\alpha)R|y \geq w^*]\}}$, $\bar{\pi} \equiv 1 - \frac{\eta k}{[1-F(w^*)]\{w^*+E[V(y,\alpha)R|y \geq w^*]\}}$. $\forall R \geq 0, \exists \pi \in [\underline{\pi}, \bar{\pi}]$, such that the local politician neither acquires information nor steals capital.

Moreover, if $R' < R''$, then $\underline{\pi}(R', \alpha) < \underline{\pi}(R'', \alpha)$, and $\bar{\pi}(R', \alpha) < \bar{\pi}(R'', \alpha)$, Stronger performance-based rewards complement and are complemented by more frequent rotation.

A3.3 Adaptation: the Informed Politician

This section proves results analogous to Proposition 4. This is the only section where the second derivative matters. Specifically, I need to assume that V is convex in y . A new project arrives with probability p . The productivity of the project follows $G(\cdot)$ that *f.o.s.d.* $F(\cdot)$. The results are for R . They also (trivially) applies to α if α enters $V(y, \alpha)$ linearly.

Proposition A4: Suppose $V(y, \alpha)$ is convex in y . If the politician paid c and is thus informed about the old project:

She will endorse the new project with probability $F(\hat{y})$, where \hat{y} satisfies $\hat{y} + V(\hat{y}, \alpha)R = [1 - G(\tilde{w})]\{\tilde{w} + E_g[V(y, \alpha)R|y \geq \tilde{w}]\} \equiv U_2$. $F(\hat{y})$ increases with R : performance-based rewards encourages adaptation.

Proof: Suppose that the old entrepreneur in $t = 1$ has invested. The informed politician supports the new project if:

$$y' + RV(y', \alpha) \leq [1 - G(\tilde{w})]\tilde{w} + R \int_{\tilde{w}}^{\bar{y}} V(z, \alpha)g(z)dz.$$

\tilde{w} maximizes the right hand side of the inequality. y' is a draw from $F(\cdot)$. As the left hand side is monotone in y' , $\exists!$ \hat{y} such that $\hat{y} + RV(\hat{y}, \alpha) = [1 - G(\tilde{w})]\tilde{w} + R \int_{\tilde{w}}^{\bar{y}} V(z, \alpha)g(z)dz$.

The probability that the informed politician supports the new project is thus $F(\hat{y}(R, \alpha))$. We are mainly concerned whether the increase in the value of promotion opportunity R encourages adaptation. As $\frac{dF(\hat{y}(R, \alpha))}{dR} = f \frac{d\hat{y}(R, \alpha)}{dR}$, the sign of $\frac{dF(\hat{y}(R, \alpha))}{dR}$ depends on $\frac{d\hat{y}(R, \alpha)}{dR}$. We have:

$$d\hat{y} + R \frac{\partial V}{\partial \hat{y}} d\hat{y} + V dR = \int_{\tilde{w}}^{\bar{y}} V(z, \alpha)g(z)dz dR$$

$$\frac{d\hat{y}}{dR} = \frac{\int_{\tilde{w}}^{\bar{y}} V(z, \alpha)g(z)dz - V(\hat{y}, \alpha)}{1 + R \frac{\partial V}{\partial \hat{y}}}$$

The sign depends on $\int_{\tilde{w}}^{\bar{y}} V(z, \alpha)g(z)dz - V(\hat{y}, \alpha)$.

By contradiction, suppose that $\int_{\tilde{w}}^{\bar{y}} V(z, \alpha)g(z)dz - V(\hat{y}, \alpha) < 0$. As by definition $\hat{y} + RV(\hat{y}, \alpha) = [1 - G(\tilde{w})]\tilde{w} + R \int_{\tilde{w}}^{\bar{y}} V(z, \alpha)g(z)dz$, we have $\hat{y} - [1 - G(\tilde{w})]\tilde{w} = \int_{\tilde{w}}^{\bar{y}} V(z, \alpha)g(z)dz - V(\hat{y}, \alpha) < 0$, or $\hat{y} < [1 - G(\tilde{w})]\tilde{w}$.

Notice that $\int_{\tilde{w}}^{\bar{y}} V(z, \alpha)g(z)dz = [1 - G(\tilde{w})]E_g[V(y, \alpha)|y \geq \tilde{w}]$. By conditional Jensen's Inequality, as $V(y, \alpha)$ is convex in y , we have:

$$[1 - G(\tilde{w})]V(E[y|y \geq \tilde{w}], \alpha) \leq [1 - G(\tilde{w})]E_g[V(y, \alpha)|y \geq \tilde{w}]$$

Together with $[1 - G(\tilde{w})]E_g[V(y, \alpha)|y \geq \tilde{w}] < V(\hat{y}, \alpha)$ and $\hat{y} < [1 - G(\tilde{w})]\tilde{w}$:

$$[1 - G(\tilde{w})]V(E[y|y \geq \tilde{w}], \alpha) < [1 - G(\tilde{w})]E_g[V(y, \alpha)|y \geq \tilde{w}] < V(\hat{y}, \alpha) < V([1 - G(\tilde{w})]\tilde{w}, \alpha).$$

Thus, $[1 - G(\tilde{w})]V(E[y|y \geq \tilde{w}], \alpha) < V([1 - G(\tilde{w})]\tilde{w}, \alpha)$.

As $V(y, \alpha)$ is convex in y , by definition $V([1 - G(\tilde{w})]\tilde{w} + G(\tilde{w}) * 0, \alpha) < [1 - G(\tilde{w})]V(\tilde{w}, \alpha) + G(\tilde{w})V(0, \alpha) = [1 - G(\tilde{w})]V(\tilde{w}, \alpha)$.

So we have $[1 - G(\tilde{w})]V(E[y|y \geq \tilde{w}], \alpha) < [1 - G(\tilde{w})]V(\tilde{w}, \alpha)$, or $E[y|y \geq \tilde{w}] < \tilde{w}$, this is a contradiction. So it must be that $\int_{\tilde{w}}^{\bar{y}} V(z, \alpha)g(z)dz - V(\hat{y}, \alpha) > 0$.

Q.E.D.

As before, the first best from the perspective of the principal is to support the old project if and only if $y' \geq E_g(y)$.

The asymptotic result as $R \rightarrow \infty$ becomes: the politician supports the old project if and only if $V(y', \alpha) \geq \int_{\underline{y}}^{\bar{y}} V(z, \alpha)g(z)dz = E_g[V(y, \alpha)]$. So in general, there is still distortion even as $R \rightarrow \infty$, as the politician takes the curvature of promotion function $V(y, \alpha)$ into consideration.

Specifically, as $V(y, \alpha)$ is convex in y , $E_g(V(y, \alpha)) > V(E_g(y), \alpha)$. So if the politician supports the old project (i.e., $V(y', \alpha) \geq E_g[V(y, \alpha)]$), we have $V(y', \alpha) > V(E_g(y), \alpha)$ or $y' > E_g(y)$: the principal also prefers the old project. Thus, there are cases where the principal prefers the old project, but the politician supports the new project: the informed politician excessively adapts when $R \rightarrow \infty$.

A3.4 Adaptation: the Uninformed Politician

This section is analogous to Proposition 5. Again, she always supports the new project:

$$U_2 \equiv \max_w [1 - G(w)]\{w + \beta RE_g[V(y, \alpha)|y \geq w]\} = [1 - G(\tilde{w})]\{\tilde{w} + \beta RE_g[y|V(y, \alpha) \geq \tilde{w}]\} \geq$$

$$[1 - G(w^*)]\{w^* + \beta RE_g[V(y, \alpha)|y \geq w^*]\} = [1 - G(w^*)]w^* + \beta R \int_{w^*}^{\bar{y}} V(z, \alpha)g(z)dz$$

Notice that $\int_{w^*}^{\bar{y}} V(z, \alpha)g(z)dz \geq \int_{w^*}^{\bar{y}} V(z, \alpha)f(z)dz$ because of F.O.S.D.⁸ Also, because $G(y) \leq F(y) \forall y$, $[1 - G(w^*)]w^* \geq [1 - F(w^*)]w^*$. So we have:

$$U_2 \geq [1 - G(w^*)]w^* + \beta R \int_{w^*}^{\bar{y}} V(z, \alpha)g(z)dz \geq [1 - F(w^*)]w^* + \beta R \int_{w^*}^{\bar{y}} V(z, \alpha)f(z)dz \equiv U_1$$

⁸Again, “ G f.o.s.d. F ” \Leftrightarrow “ $G(y) \leq F(y) \forall y$ ” \Leftrightarrow “For every weakly increasing utility function u , $\int u(x)dG \geq \int u(x)dF$.”

So it is optimal for the politician always to endorse the new project. Notice that an uninformed politician is still more much adaptive than an informed politician even as $R \rightarrow \infty$. Proposition A5 is exactly the same as Proposition 5:

Proposition A5: *Suppose the politician didn't pay the cost of learning and a new project arrives with $G(y) \leq F(y)$, $\forall y$. The politician will always endorse the new project.*

A3.5 Adaptation: the Complementarity between Rotation and Performance Rewards

This section is analogous to Proposition 6. The politician does not pay c if:

$$(1 - \pi)\{(1 - p)U_1 + pU_2\} + \pi\tilde{U} \geq$$

$$(1 - \pi)\{(1 - p)\{E_f[y + RV(y, \alpha)]\} + p\Pi U_2 + (1 - \Pi)\{E_f[y + RV(y, \alpha) | y + RV(y, \alpha) \geq U_2]\}\} + \pi\tilde{U} - c$$

where $\Pi = F(\hat{y})$, $\hat{y} + RV(\hat{y}, \alpha) = U_2$.

Proposition A6: *1. Sufficiently intense rotation and performance-based rewards induce an adaptive and benevolent local politician. The old and the new entrepreneurs invest accordingly.*

In algebra, denote $\hat{R}(k, \alpha)$ such that $\hat{R}(k, \alpha) = \max\{R_1, R_2\}$, where R_1 satisfies $(1 - p)[1 - F(w^*(R_1, \alpha))]\{E_f(y | y \geq w^*(R_1, \alpha)) - w^*(R_1, \alpha)\} = k$ and R_2 satisfies $[1 - G(\tilde{w}(R_2, \alpha))][E_g(y | y \geq \tilde{w}(R_2, \alpha)) - \tilde{w}(R_2, \alpha)] = k$. Also, $\underline{\pi} = 1 - \frac{c}{(1 - p)\{E(y + V(y, \alpha)R) - U_1\} + p\{(1 - \Pi)[E_f[y + V(y, \alpha)R | y + V(y, \alpha)R \geq U_2] - U_2]\}}$.

If $R > \hat{R}$ and $\pi \geq \underline{\pi}$, the politician does not pay c , and both the old and the new entrepreneurs reap non-negative returns.

2. Denote $w^*(0) = \operatorname{argmax}_w [1 - F(w)]w$. If $p < \frac{\int_y^{w^*(0)} V(z, \alpha) f(z) dz}{\int_y^{w^*(0)} V(z, \alpha) f(z) dz + \int_y^{\tilde{y}} V(z, \alpha) g(z) dz}$

and $p < \frac{\int_y^{w^*(0)} \frac{\partial V(z, \alpha)}{\partial \alpha} f(z) dz}{\int_y^{w^*(0)} \frac{\partial V(z, \alpha)}{\partial \alpha} f(z) dz + \int_y^{\tilde{y}} \frac{\partial V(z, \alpha)}{\partial \alpha} g(z) dz}$, then $\frac{\partial \underline{\pi}}{\partial R} > 0$ and $\frac{\partial \underline{\pi}}{\partial \alpha} > 0$.

Proof: We need:

$$\pi \geq 1 - \frac{c}{\{(1 - p)E_f[y + RV(y, \alpha)]\} - U_1 + p\{(1 - \Pi)\{E_f[y + RV(y, \alpha) | y + RV(y, \alpha) \geq U_2] - U_2\}\}}$$

$$\equiv 1 - \frac{c}{\Delta}.$$

Note that $\Delta = (1 - p)\Delta_1 + \pi\Delta_2$. $\Delta_1 = E_f[y + \beta RV(y, \alpha)] - U_1$, so:

$$\frac{\partial \Delta_1}{\partial R} = E_f[V(y, \alpha)] - \int_{w^*}^{\bar{y}} V(z, \alpha) f(z) dz = \int_{\underline{y}}^{w^*} V(z, \alpha) f(z) dz.$$

As before, denote \hat{y} such that $\hat{y} + RV(\hat{y}, \alpha) = [1 - G(\tilde{w})]\tilde{w} + R \int_{\tilde{w}}^{\bar{y}} V(z, \alpha) g(z) dz \equiv U_2$. $\Delta_2 = (1 - \Pi)\{E_f[y + RV(y, \alpha) | y + RV(y, \alpha) \geq U_2] - U_2\} = \int_{\hat{y}}^{\bar{y}} (z + RV(z, \alpha)) f(z) dz - [1 - F(\hat{y})]U_2$. So:

$$\frac{\partial \Delta_2}{\partial R} = \int_{\hat{y}}^{\bar{y}} V(z, \alpha) f(z) dz - [\hat{y} + RV(\hat{y}, \alpha)] f(\hat{y}) \frac{\partial \hat{y}}{\partial R} + f(\hat{y}) \frac{\partial \hat{y}}{\partial R} U_2 - [1 - F(\hat{y})] \int_{\tilde{w}}^{\bar{y}} V(z, \alpha) g(z) dz$$

As $\hat{y} + RV(\hat{y}, \alpha) = U_2$, $[\hat{y} + RV(\hat{y}, \alpha)] f(\hat{y}) \frac{\partial \hat{y}}{\partial R} = f(\hat{y}) \frac{\partial \hat{y}}{\partial R} U_2$, so:

$$\frac{\partial \Delta_2}{\partial R} = \int_{\hat{y}}^{\bar{y}} V(z, \alpha) f(z) dz - [1 - F(\hat{y})] \int_{\tilde{w}}^{\bar{y}} V(z, \alpha) g(z) dz \geq - \int_{\underline{y}}^{\bar{y}} V(z, \alpha) g(z) dz$$

To have $\frac{\partial \Delta}{\partial R} > 0$, a sufficient condition can be derived:

$$\begin{aligned} \frac{\partial \Delta}{\partial R} &= (1 - p) \frac{\partial \Delta_1}{\partial R} + p \frac{\partial \Delta_2}{\partial R} \geq (1 - p) \int_{\underline{y}}^{w^*} V(z, \alpha) f(z) dz - p \int_{\underline{y}}^{\bar{y}} V(z, \alpha) g(z) dz \\ &\geq (1 - p) \int_{\underline{y}}^{w^*(0)} V(z, \alpha) f(z) dz - p \int_{\underline{y}}^{\bar{y}} V(z, \alpha) g(z) dz \geq 0, \end{aligned}$$

where $w^*(0) = \operatorname{argmax}_w [1 - F(w)]w$. So if $p < \frac{\int_{\underline{y}}^{w^*(0)} V(z, \alpha) f(z) dz}{\int_{\underline{y}}^{w^*(0)} V(z, \alpha) f(z) dz + \int_{\underline{y}}^{\bar{y}} V(z, \alpha) g(z) dz}$, we have

$$\frac{\partial \Delta}{\partial R} \geq 0.$$

Proof for $\frac{\partial \Delta}{\partial \alpha} > 0$ is similar:

$$\frac{\partial \Delta_1}{\partial \alpha} = R \int_{\underline{y}}^{w^*} \frac{\partial V(z, \alpha)}{\partial \alpha} f(z) dz.$$

$$\begin{aligned}
\frac{\partial \Delta_2}{\partial \alpha} &= R \int_{\hat{y}}^{\bar{y}} \frac{\partial V(z, \alpha)}{\partial \alpha} f(z) dz - [\hat{y} + RV(\hat{y}, \alpha)] f(\hat{y}) \frac{\partial \hat{y}}{\partial \alpha} + f(\hat{y}) \frac{\partial \hat{y}}{\partial \alpha} U_2 - [1 - F(\hat{y})] R \int_{\bar{w}}^{\bar{y}} \frac{\partial V(z, \alpha)}{\partial \alpha} g(z) dz \\
&= R \int_{\hat{y}}^{\bar{y}} \frac{\partial V(z, \alpha)}{\partial \alpha} f(z) dz - R[1 - F(\hat{y})] \int_{\bar{w}}^{\bar{y}} \frac{\partial V(z, \alpha)}{\partial \alpha} g(z) dz > -R \int_{\underline{y}}^{\bar{y}} \frac{\partial V(z, \alpha)}{\partial \alpha} g(z) dz
\end{aligned}$$

A sufficient condition for $\frac{\partial \Delta}{\partial \alpha} > 0$:

$$\begin{aligned}
\frac{\partial \Delta}{\partial \alpha} &= (1 - p) \frac{\partial \Delta_1}{\partial \alpha} + p \frac{\partial \Delta_2}{\partial \alpha} \\
&> (1 - p) R \int_{\underline{y}}^{w^*} \frac{\partial V(z, \alpha)}{\partial \alpha} f(z) dz - p R \int_{\underline{y}}^{\bar{y}} \frac{\partial V(z, \alpha)}{\partial \alpha} g(z) dz \\
&> (1 - p) R \int_{\underline{y}}^{w^{*(0)}} \frac{\partial V(z, \alpha)}{\partial \alpha} f(z) dz - p R \int_{\underline{y}}^{\bar{y}} \frac{\partial V(z, \alpha)}{\partial \alpha} g(z) dz > 0 \\
p &< \frac{\int_{\underline{y}}^{w^{*(0)}} \frac{\partial V(z, \alpha)}{\partial \alpha} f(z) dz}{\int_{\underline{y}}^{w^{*(0)}} \frac{\partial V(z, \alpha)}{\partial \alpha} f(z) dz + \int_{\underline{y}}^{\bar{y}} \frac{\partial V(z, \alpha)}{\partial \alpha} g(z) dz}.
\end{aligned}$$

Q.E.D.

A3.6 Two Rounds of Bargaining

This section is in analogous to Proposition 7:

Proposition A7: Denote:

$$\underline{\pi}_1 = 1 - \frac{c}{E[y + V(y, \alpha)R] - \max_w [1 - F(w)]\{w + E[V(y, \alpha)R|y \geq w]\}}.$$

$$\underline{\pi}_2 : (1 - \underline{\pi}_2)\max_{w_2} [1 - F(w_2)]\{w_2 + E[V(y, \alpha)R|y \geq w_2]\} =$$

$$\max_{w_1} \{-F(w_1)c + (1 - \underline{\pi}_2)\{[1 - F(w_1)]\{w_1 + E[V(y, \alpha)R|y \geq w_1]\} + F(w_1)E[y + V(y, \alpha)R|y \geq w_1]\}\}.$$

$$\hat{R} : k = [1 - F(w^*(R, \alpha))\{E[y|y \geq w^*(R, \alpha)] - w^*(R, \alpha)\}, w^* + V(w^*, \alpha)R] = \frac{1 - F(w^*)}{w^* f(w^*)}.$$

1. If $\pi \geq \underline{\pi}_1$, $\pi \geq \underline{\pi}_2$, and $R > \hat{R}$, the politician does not pay c to learn about y , and the entrepreneur makes the investment k .

2.

$$\frac{\partial \underline{\pi}_1}{\partial R} > 0, \frac{\partial \underline{\pi}_2}{\partial R} > 0, \frac{\partial \underline{\pi}_1}{\partial \alpha} > 0, \frac{\partial \underline{\pi}_2}{\partial \alpha} > 0$$

So with stronger performance rewards, the minimum rotation frequency $\underline{\Pi}(x) = \max\{\underline{\pi}_1, \underline{\pi}_2\}$ also rises to guarantee the entrepreneur's ex ante investment.

Proof: Denote that

$$w^* = \operatorname{argmax}_w [1 - F(w)]\{w + E[V(y, \alpha)R|y \geq w]\}.$$

w^* is the optimal extraction at $t = 2$ for an uninformed politician with no entrepreneur accepting w_1 .

Notice that at $t = 2$ newly appointed politician also honors w'_1 , the contract that the veteran politician made in $t = 1$. Also, if the entrepreneur rejects w'_1 , the information set for a newly appointed politician is exactly the same as the veteran politician. They also have the same utility. So they must offer the same w'_2 .

In $t = 2$, the dominant strategy is for entrepreneur to accept w'_2 if $y \geq w'_2$ and entrepreneur has not accepted in $t = 1$.

Back to $t = 1$. For all $w'_1 \in [y, \bar{y}]$, I now derive the set of strategies that can satisfy sequential rationality. This enables the politician to predict her payoff if she chooses a specific w'_1 .

1. Fix a $w'_1 \geq w^*$: There are two cases: the politician finds it sequentially rational to pay c or pay 0. The strategy at learning stage is a function of w'_1 . Denote $learn(w'_1) = 1$ if the politician learns after w'_1 and $learn(w'_1) = 0$ otherwise.

1.1 Suppose $learn(w'_1) = 1$: Given that $learn(w'_1) = 1$ is sequentially rational, the entrepreneur accepts any $y - w'_1 \geq 0$ as in $t = 2$ he gets 0 surplus.

Given that $learn(w'_1) = 1$ is sequentially rational, the politician asks to extract $w'_2 = y$ for any entrepreneur who did not accept w'_1 .

The utility of w'_1 if $learn(w'_1) = 1$ is sequentially rational is:

$$U(\{w'_1, learn(w'_1) = 1\}) = -F(w'_1)c + (1 - \pi)\{[1 - F(w'_1)]\{w'_1 +$$

$$E[V(y, \alpha)R|y \geq w'_1]\} + E[y + V(y, \alpha)R|y \leq w'_1]\} + \pi\hat{U}$$

1.2 Suppose $learn(w'_1) = 0$: If $y - w'_1 > y - w'_2$ or $w'_1 < w'_2$, entrepreneur accepts w'_1 if $y \geq w'_1$. This is impossible as:

$$w'_2 = \operatorname{argmax}_w \operatorname{Prob}(w \leq y | y \leq w'_1) \{w + E[V(y, \alpha)R|w \leq y, \text{given } y \leq w'_1]\} < w'_1$$

If $y - w'_1 \leq y - w'_2$ or $w'_1 \geq w'_2$, entrepreneur rejects w'_1 . In this case:

$$w'_2 = \operatorname{argmax}_w [1 - F(w)] \{w + E[V(y, \alpha)R|y \geq w]\} = w^*$$

Indeed $w'_1 \geq w'_2 = w^*$, consistent with the assumption that $w'_1 \geq w^*$

So given that $learn(w'_1) = 0$ is sequentially rational:

$$U(\{w'_1, learn(w'_1) = 0\}) = (1 - \pi)[1 - F(w^*)]\{w^* + E[V(y, \alpha)R|y \geq w^*]\} + \pi\hat{U}$$

So by choosing $w'_1 \geq w^*$, the maximum utility the politician can get is:

$$\begin{aligned} \max \{ & -F(w'_1)c + (1-\pi)\{E[y+V(y, \alpha)R|y \leq w'_1] + [1-F(w'_1)]\{w'_1 + E[V(y, \alpha)R|y \geq w'_1]\} + \pi\hat{U}\} \\ & , (1-\pi)[1-F(w^*)]\{w^* + E[V(y, \alpha)R|y \geq w^*]\} + \pi\hat{U} \} \end{aligned}$$

2. Fix a $w'_1 < w^*$:

2.1 Suppose $learn(w'_1) = 1$: The algebra is the same as 1.1. We have:

$$\begin{aligned} U(\{w'_1, learn(w'_1) = 1\}) = & -F(w'_1)c + (1-\pi)\{E[y + V(y, \alpha)R|y \leq w'_1] + \\ & [1 - F(w'_1)]\{w'_1 + E[V(y, \alpha)R|y \geq w'_1]\} + \pi\hat{U} \end{aligned}$$

2.2 Suppose $learn(w'_1) = 0$: If $y - w'_1 > y - w'_2$ or $w'_1 < w'_2$, entrepreneur accepts w'_1 if $y \geq w'_1$. This is impossible as:

$$w'_2 = \operatorname{argmax}_w \operatorname{Prob}(w \geq w | w \leq w'_1) \{w + E[V(y, \alpha)R | w \leq y, \text{given } y \leq w'_1]\} < w'_1$$

If $y - w'_1 < y - w'_2$ or $w'_1 > w'_2$, entrepreneur rejects w'_1 . In this case:

$$w'_2 = \operatorname{argmax}_w [1 - F(w)] \{w + E[V(y, \alpha)R | y \geq w]\} = w^*$$

Another contradiction as we assume $w'_1 > w'_2$ and $w^* > w'_1$.

The only possible case is $w'_1 = w'_2$. Given Assumption 4, suppose that the cutoff is \hat{w} such that an entrepreneur with $y \geq \hat{w}$ accepts w'_1 . We have $\hat{w} > w'_1$. Entrepreneur with $w'_1 \leq y < \hat{w}$ accepts w'_2 . To make $w'_2 = w'_1$ sequentially rational, it must be that \hat{w} satisfies:

$$w'_1 = w'_2 = \operatorname{argmax}_w \operatorname{prob}(w \leq y, \text{given } y \leq \hat{w}) \{w + \beta E[V(y, \alpha)R | w \leq y, \text{given } y \leq \hat{w}]\}$$

The probability density function of $f(y|y \leq \hat{w}) = \frac{f(y)}{F(\hat{w})} \equiv g(y, \hat{w})$ with support on $[y, \hat{w}]$. We have:

$$\operatorname{prob}(w \leq y, \text{given } y \leq \hat{w}) = \int_w^{\hat{w}} f(z|z \leq \hat{w}) dz = \frac{\int_w^{\hat{w}} f(z)}{1 - F(\hat{w})} dz = \frac{F(\hat{w}) - F(w)}{F(\hat{w})}$$

$$E[V(y, \alpha)|w \leq y, \text{ given } y \leq \hat{w}] = \int_w^{\hat{w}} V(z, \alpha)g(z, \hat{w}|z \geq w)dz$$

$$g(z, \hat{w}|z \geq w) = \frac{g(z, \hat{w})}{1 - G(w; \hat{w})} = \frac{f(z)/F(\hat{w})}{1 - \int_y^w \frac{f(v)}{F(\hat{w})}dv} = \frac{f(z)/F(\hat{w})}{1 - \frac{F(w)}{F(\hat{w})}} = \frac{f(z)}{F(\hat{w}) - F(w)}$$

Thus we have:

$$E[V(y, \alpha)|w \leq y, \text{ given } y \leq \hat{w}] = \frac{\int_w^{\hat{w}} V(z, \alpha)f(z)dz}{F(\hat{w}) - F(w)}$$

So:

$$\text{prob}(w \leq y, \text{ given } y \leq \hat{w})\{w + E[V(y, \alpha)|w \leq y, \text{ given } y \leq \hat{w}]\}$$

$$= \frac{F(\hat{w}) - F(w)}{F(\hat{w})}w + \frac{F(\hat{w}) - F(w)}{F(\hat{w})} \frac{\int_w^{\hat{w}} V(z, \alpha)Rf(z)dz}{F(\hat{w}) - F(w)}$$

$$= \frac{[F(\hat{w}) - F(w)]w + \int_w^{\hat{w}} V(z, \alpha)Rf(z)dz}{F(\hat{w})}$$

As $w'_1 = w'_2 = \text{argmax}_w \frac{[F(\hat{w}) - F(w)]w + \int_w^{\hat{w}} V(z, \alpha)Rf(z)dz}{F(\hat{w})}$, F.O.C. tells us that:

$$w'_1 + RV(w'_1, \alpha) = \frac{F(\hat{w}) - F(w'_1)}{F(w'_1)}$$

Now let us verify that $w'_1 < w^*$. By contradiction, suppose that $w'_1 \geq w^*$. $w^* = \text{argmax}_w \{[1 - F(w)]w + R \int_y^{w^*} V(z, \alpha)f(z)dz\}$. So w^* satisfies:

$$\frac{1 - F(w^*)}{f(w^*)} = w^* + V(w^*, \alpha)R$$

As we assume that $w'_1 \geq w^*$, and that $\frac{1 - F(w)}{f(w)}$ is a decreasing function, we have:

$$w'_1 + RV(w'_1, \alpha) = \frac{F(\hat{w}) - F(w'_1)}{F(w'_1)} < \frac{1 - F(w'_1)}{F(w'_1)} \leq \frac{1 - F(w^*)}{F(w^*)} = w^* + RV(w^*, \alpha)$$

As $\{w + RV(w, \alpha)\}$ is an increasing function in w , $w'_1 < w^*$, we obtain the contradiction. So we must have $w'_1 < w^*$.

Thus, if $learn(w'_1) = 0$ is sequentially rational following a $w'_1 < w^*$:

$$U(\{w'_1, w'_1(learn) = 0\}) = (1 - \pi)\{[1 - F(\hat{w})]\{w'_1 + E[V(y, \alpha)R|y \geq \hat{w}]\} +$$

$$F(\hat{w})\frac{[F(\hat{w}) - F(w'_1)]w'_1 + \int_{w'_1}^{\hat{w}} V(z, \alpha)Rf(z)dz}{F(\hat{w})}\} + \pi\tilde{U}$$

$$U(\{w'_1, learn(w'_1) = 0\}) = (1 - \pi)\{[1 - F(\hat{w})]\{w'_1 +$$

$$E[V(y, \alpha)R|y \geq \hat{w}]\} + [F(\hat{w}) - F(w)]w + \int_w^{\hat{w}} V(z, \alpha)Rf(z)dz\} + \pi\tilde{U}$$

The maximum utility of the politician from a $w'_1 < w^*$ is:

$$\max\{-F(w'_1)c + (1 - \pi)\{[1 - F(w'_1)]\{w'_1 + E[V(y, \alpha)|y \geq w'_1]\} + F(w'_1)E[y + V(y, \alpha)R|y \leq w'_1] + \pi\hat{U}\},$$

$$(1 - \pi)\{[1 - F(\hat{w})]\{w'_1 + E[V(y, \alpha)R|y \geq \hat{w}]\} + [F(\hat{w}) - F(w)]w + \int_w^{\hat{w}} V(z, \alpha)Rf(z)dz\} + \pi\tilde{U}\}$$

We have derived the function of politician's payoff at $t = 1$ for any w'_1 and learning decision. We can use this function to find the sequential equilibrium strategy.

I want to ensure that $learn(w'_1) = 0$ is the strategy on equilibrium path. That gives a utility of:

$$U(w'_1, learn(w'_1) = 0) =$$

$$\begin{cases} (1 - \pi)[1 - F(w^*)]\{w^* + E[V(y, \alpha)R|y \geq w^*]\} + \pi\hat{U} & w'_1 \geq w^* \\ (1 - \pi)\{[1 - F(\hat{w})]\{w_1 + E[V(y, \alpha)R|y \geq \hat{w}]\} + \{[F(\hat{w}) - F(w)]w + \int_w^{\hat{w}} V(z, \alpha)Rf(z)dz\} + \pi\hat{U} & w'_1 < w^* \end{cases}$$

Notice that:

$$\begin{aligned}
& [1 - F(\hat{w})]\{w_1 + E[V(y, \alpha)R|y \geq \hat{w}]\} + \{[F(\hat{w}) - F(w)]w + \int_w^{\hat{w}} V(z, \alpha)Rf(z)dz\} \\
&= w'_1 \int_{\hat{w}}^{\bar{y}} f(z)dz + \int_{\hat{w}}^{\bar{y}} V(z, \alpha)Rf(z)dz + w'_1 \int_{w'_1}^{\hat{w}} f(z)dz + \int_{w'_1}^{\hat{w}} V(z, \alpha)Rf(z)dz \\
&= w'_1 \int_{w'_1}^{\bar{y}} f(z)dz + \int_{w'_1}^{\bar{y}} V(z, \alpha)Rf(z)dz = [1 - F(w'_1)]\{w'_1 + E[V(y, \alpha)R|y \geq w'_1]\} \\
&\leq [1 - F(w^*)]\{w^* + E[V(y, \alpha)R|y \geq w^*]\}
\end{aligned}$$

So it is always optimal to choose $w'_1 \geq w^*$.

In that case, I want to make sure that at the stage of learning, it is not optimal to learn:

$$-c + (1 - \pi)E[y + \beta V(y, \alpha)R] \leq (1 - \pi)[1 - F(w^*)]\{w^* + E[V(y, \alpha)R|y \geq w^*]\}$$

And at the stage to decide w'_1 , the utility from not learning \geq the highest utility from learning:

$$\begin{aligned}
& (1 - \pi)max_{w_2}[1 - F(w_2)]\{w_2 + E[V(y, \alpha)R|y \geq w_2]\} \geq \\
& max_{w_1}\{-F(w_1)c + (1 - \pi)\{[1 - F(w_1)]\{w_1 + E[V(y, \alpha)R|y \geq w_1]\} + F(w_1)E[y + V(y, \alpha)R|y \geq w_1]\}\} \\
& (1 - \underline{\pi})max_{w_2}[1 - F(w_2)]\{w_2 + E[V(y, \alpha)R|y \geq w_2]\} = \\
& max_{w_1}\{-F(w_1)c + (1 - \underline{\pi})\{[1 - F(w_1)]\{w_1 + E[V(y, \alpha)R|y \geq w_1]\} + F(w_1)E[y + V(y, \alpha)R|y \geq w_1]\}\}
\end{aligned}$$

defines $\underline{\pi}$. By Envelope Theorem and Implicit Function Theorem:

$$\begin{aligned}
& (1 - \underline{\pi})[1 - F(w^*)]E[V(y, \alpha)|y \geq w^*]dR - U^*d\underline{\pi} = \\
& (1 - \underline{\pi})\{[1 - F(w_1^*)]E[V(y, \alpha)|y \geq w_1] + F(w_1^*)E[V(y, \alpha)|y \leq w_1^*]\}dR - \tilde{U}^*d\underline{\pi} \\
& \text{where } U^* = [1 - F(w^*)]\{w^* + E[V(y, \alpha)R|y \geq w^*]\}, \tilde{U}^* = [1 - F(w_1^*)]\{w_1^* + E[V(y, \alpha)R|y \geq
\end{aligned}$$

$w_1^*] + F(w_1^*)E[y + V(y, \alpha)R|y \geq w_1^*]$, we need $\tilde{U}^* - U^* > 0$ for the analysis to be meaningful.
So:

$$\begin{aligned} \frac{\partial \pi}{\partial R} &= \frac{1 - \pi}{\tilde{U}^* - U^*} \{ [1 - F(w_1^*)]E[V(y, \alpha)|y \geq w_1^*] + F(w_1^*)E[V(y, \alpha)|y \leq w_1^*] - [1 - F(w^*)]E[V(y, \alpha)|y \geq w^*] \} \\ &= \frac{1 - \pi}{\tilde{U}^* - U^*} \left[\int_{\underline{y}}^{\bar{y}} V(z, \alpha) f(z) dz - \int_{w^*}^{\bar{y}} V(z, \alpha) f(z) dz \right] = \frac{1 - \pi}{\tilde{U}^* - U^*} \int_{\underline{y}}^{w^*} V(z, \alpha) f(z) dz > 0 \end{aligned}$$

$\underline{\Pi} = \max\{\pi_1, \pi_2\}$ is an increasing function of R because $\frac{\pi_1}{\partial R} > 0$ and $\frac{\pi_2}{\partial R} > 0$.

Similarly, we can show that $\underline{\Pi}$ is an increasing function of α , as π_1 increases with α , and:

$$\frac{\partial \pi_2}{\partial \alpha} = \frac{1 - \pi}{\tilde{U}^* - U^*} R \int_{\underline{y}}^{w^*} \frac{\partial V(z, \alpha)}{\partial \alpha} f(z) dz > 0$$

Q.E.D.