Career Paths in Hierarchies: Theory and Evidence from Chinese Officials *

Jinghong Li †, Zanhui Liu ‡ and Yang Yao §

December 7, 2023

Abstract

This paper studies the dynamic patterns of job durations in a hierarchical government. Using a large dataset of Chinese officials from 1994 to 2017, we find systematic evidence that the promotion rates of officials at different levels of government are negatively correlated, contrary to the common finding of *fast track* in the literature. To characterize the career paths in hierarchies, we develop a dynamic model of optimal job assignments with symmetric learning. The model highlights the role of seniority in coping with the Type-II error, that is, the risk of over-evaluating an official's ability. Only very capable officials are spared, and the fast track applies; relatively capable officials are subject to the over-evaluation risk, and their job durations are negatively correlated across levels of government. A negative correlation is also more likely to happen when the controller is more risk averse, which may explain mixed findings in the literature.

Keywords: career paths, fast track, seniority, meritocracy

JEL Codes: M51, D8, H70

* We are grateful for Songrui Liu, Tianyang Xi, and Jiaqi Zhao for their helpful comments. We thank seminar participants in the Central University of Finance and Economics, Zhejiang University, the 22nd China Economics Annual Conference, the 2023 Asian Meeting of the Econometric Society, and the 2023 Workshop of Political Economy and Governance in the National School of Development, Peking University for their helpful discussions. All remaining errors are our own.

† National School of Development, Peking University. Email: jhli2019@nsd.pku.edu.cn.

‡ School of Economics and Management, Tsinghua University. Email: zhliu2018@nsd.pku.edu.cn.

§ National School of Development, Peking University. Email: yyao@nsd.pku.edu.cn.

1 Introduction

Promotion involves errors when companies do not have full information about employees' abilities. In hierarchical organizations where promotion is sequential, the errors have an accumulative effect along the ladder of the hierarchy. Therefore, how hierarchical organizations allocate talents—who get promoted and when—is crucial to their resulting productivity. Most of the existent literature finds evidence consistent with the fast-track hypothesis, that is, people who lead others earlier tend to always stay ahead throughout their career (Baker et al., 1994; Gibbons and Waldman, 1999, 2006; Breaugh, 2011)[1]. In other words, promotion rates are positively correlated across career stages. One exception is the study by Belzil and Bognanno (2010), which discovers a negative correlation, but the authors are unable to rationalize the finding from the perspective of organizational strategy.

Although it has received little attention, negative correlation may be a result of rational organizational choices that deserve more careful studies. In addition to being viewed as an indicator of high potentials (for example, the capability to successfully perform at the next level), previous history of a person's rapid promotion (that is, having shorter job durations at lower levels of the organization) entails the risk of Type-II errors in the promotion process. As such, the person's ability may have been over-evaluated because of the contamination of external shocks. This issue can be a serious concern in organizations, such as the government, where demotion or dismissal is challenging. The cost of the risk increases at higher levels of the organization where officials have larger responsibility and influence. Consequently, a person with a higher level of perceived ability may be promoted faster than others at lower ranks, but promotion slows down at higher ranks, which results in a negative sequential correlation.

[1]Extensive support for this phenomenon can also be found in Brüderl et al. (1991), Bernhardt (1995), Podolny and Baron (1997), Ariga et al. (1999), Chiappori et al. (1999), Seltzer and Merrett (2000), and Treble et al. (2001).

In this paper, we first document the negative sequential correlation for Chinese officials from 1994 to 2017, and then we develop a formal model to rationalize the correlation. By doing so, we reconcile mixed findings in a unified framework.

To motivate our empirical study, we first formulate an illustrative model with a risk-averse controller and one official. The official works for the government, which has three levels. In each period, the controller decides at which level the official works. No downward mobility is allowed, which means the official cannot be assigned to levels lower than his current level, and promotion is sequential, which means it strictly follows the hierarchical order. The official's ability determines the output, and at higher levels, the marginal output of the ability is higher. The model borrows the spirit of the career concern model of Holmstrom (1999). Specifically, neither the controller nor the official knows the exact level of the official's ability, but the nature of the model gives them a prior belief about the distribution of the official's ability. Later, the official's true ability can be inferred from historical outputs. The controller becomes more certain about the official's ability over time. When the controller tries to maximize the utility (a function of output) of each period, the controller faces a merit-seniority trade-off: the controller would like to promote the official when previous outputs signal high ability but is also worried about the over-evaluation risk.

This simple model generates two intriguing predictions. First, meritocracy holds for the whole system—that is, officials will be promoted faster throughout their careers if true ability is higher. Second, two countervailing effects jointly determine the direction of the serial correlation of job durations across levels: the merit effect and the seniority effect. The merit effect dictates that an official with a higher level of perceived ability should be promoted consistently faster than others. The seniority effect arises from the over-evaluation risk, which dictates that a previously fast promoted official should be subsequently evaluated for a longer time at the next level. Positive (negative) correlation can be observed when the merit (seniority) effect dominates.

We use the CCER Officials Dataset (COD) to conduct our empirical study. Chinese officials make a worthwhile sample. China has one of the most politically centralized states in the world (Xu, 2011), and the Chinese Communist Party (CCP) is in charge of the personnel control for all officials in the Chinese government regarding their appointment, promotion, and rotation (Manion, 1985; Huang, 2002). Besides, the Chinese government consists of levels from township to the national government, similar to a Weberian hierarchical organization. Furthermore, many studies show that China's bureaucratic system works efficiently and has undoubtedly contributed to its economic success.[2] Thus, China provides a nearly perfect case for studying the underlying mechanisms of career paths in hierarchies.

The COD collected detailed biographies of Chinese officials at or above the prefecture level between 1994 and 2017, which allows us to trace the complete career paths (some are truncated) for all officials in the sample. We study prefectural and provincial officials in this paper. Following Yao and Zhang (2015), we first estimate the ability of each official by his or her fixed effect in a city or provincial growth equation. Then, we estimate a Cox proportional-hazards model to investigate officials' durations at the prefectural and provincial levels. Consistent with our illustrative model, we find that higher ability is positively associated with higher promotion hazards at both the prefectural level and the provincial level. Moreover, we find that officials who previously received speedy promotions stayed longer in a position before their subsequent promotion. This delay shows that the seniority effect dominates the merit effect.

Following our empirical study, we extend our illustrative model to a dynamic model with overlapping cohorts of officials who compete for slots in the hierarchy. We solve the model quantitatively and find clear evidence for the merit-seniority trade-off. In fact, since the model is extended to a dynamic setting, the controller takes the officials' future outputs into account when determining which officials to promote, necessitating a higher priority

[2] See Xu (2011) for a comprehensive review.

being placed on seniority than in the illustrative model. We also find a more nuanced relationship, an inverted-U shaped curve, between job durations across levels. The upward-sloping part of the curve is driven by very talented officials, and the downward-sloping part of the curve is driven by less talented officials. The curve shifts to the left as the controller becomes more risk averse, indicating that the negative correlation of job duration is more likely to happen.

Our paper enriches the study of career paths in organizations. We find systematic and robust evidence for the negative correlation of job durations across levels of a hierarchical government. To our best knowledge, this correlation is new in the literature.[3] In addition, we provide a unified framework to understand both the fast track and the negative correlation. Given the controller's risk preference, fast tracks appear when an individual's perceived ability is constantly very high, so the merit effect dominates the seniority effect in promotions. Negative correlations appear when an individual only shows mediocre levels of ability, so the seniority effect dominates the merit effect. By incorporating the over-evaluation risk and the controller's risk attitude into our analysis, we are able to find more nuanced patterns of career paths in an organization, among which is a negative correlation of job duration across levels of hierarchy that can be rationalized either when individuals have mediocre perceived ability or when the controller is very risk averse.

Our paper adds to the literature on fast tracks. Other than time-invariant unobserved heterogeneities, existing studies provide many explanations for the existence of fast tracks in organizations, such as human capital acquisition (Baker et al., 1994; Ariga et al., 1999; Gibbons and Waldman, 1999; Clemens, 2012), signaling of promotions (Bernhardt, 1995; Belzil and Bognanno, 2010) and bounded rationality (Meyer, 1991). Those studies hardly

[3]Belzil and Bognanno (2010) find a negative correlation of job duration across levels in a very restricted subsample and an average positive correlation in the whole sample. However, they do not offer any rigorous discussion on the negative correlation. Instead, they simply list some possible mechanical channels, such as the ceiling effect when hierarchies are upper-bounded and when individuals must retire. Our dynamic model experiments with different retirement ages and finds that the negative correlation is robust, negating this interpretation.

view the fast track from the dynamic process of talent identification in organizations. In our study, random production shocks taint perceived ability, but individuals with exceptionally high abilities are less susceptible to shocks and can stand out in a succession of promotions, leading to the fast track.

We also provide a fundamental view to understand why organizations favor seniority in promotion decisions. Mills (1985) and Cirone et al. (2021) argue that the seniority rule is used in promotion because it improves organizational stability and cohesion. We do not deny those explanations. Rather, we want to add another explanation purely based on the concern of over-evaluation risks. The fear of promoting the wrong persons can be a serious concern in organizations. It is even more so in the government and other public institutions (such as the court system) where the conduct of officials can make large impacts on society. The concern of over-evaluation risks gives seniority a role in promotion. However, here seniority exists not because it produces positive values, but because it avoids losses for the controller.

Our study also sheds light on the long-standing literature on political selection in China. Several empirical studies find that higher growth rates increase officials'promotion rates, which supports the meritocracy hypothesis (Li and Zhou, 2005; Landry et al., 2018). Our empirical work presents a richer picture of nonlinear career paths for officials. Yao and Zhang (2015) and Wang et al. (2022) improve the literature by estimating each official's ability from his or her fixed effect in a city or provincial growth equation. Our paper provides a theoretical foundation for this approach—the fixed effect of each official is equivalent to the perceived ability that is estimated from previous outputs.

The paper is organized as follows. We develop our illustrative model in Section 2. Section 3 describes the institutional background of China's political hierarchy and introduces the data, and Section 4 presents the empirical results for Chinese officials. Section 5 presents the dynamic model. Finally, we conclude in Section 6.

2 An illustrative model

2.1 The setup

Consider a controller and an official. The controller is in charge of appointing the official.[4] The controller is myopic and only cares about the official's output in one period, which eliminates any dynamic consideration in the controller's decision. The official has an innate ability denoted by $\theta \sim N(0, \sigma_0^2)$, which is time-invariant and unknown to the official and controller. Over time, the controller and the official update their belief about θ by observing the official's outputs in the current and previous periods.

The government consists of three levels denoted, from lower levels to higher levels, by 1, 2, and 3. If the official is assigned to level j $(j \in \{1, 2, 3\})$ in period s $(s = 0, 1, 2, \ldots)$, then the official supplies one unit of inelastic labor and produces the following:

$$y_{js} = c_j(\theta + \epsilon_s),$$

where $c_3 > c_2 > c_1 > 0$. $\epsilon_s \overset{i.i.d}{\sim} N(0, \sigma_\epsilon^2)$ and is independent of θ. In this setup, higher levels value abilities more than lower levels do, as reflected by a bigger c_j.

The distributions of θ and ϵ_s and the values of $c_j (j \in \{1, 2, 3\})$ are common knowledge. Output and job assignment in each period are observed by all. Therefore, the model shares the spirit of Holmstrom (1999) of common learning. Our model's new feature is to decide when to stop the learning (and to promote the official).

There is an implicit assumption in the setting that the model has no incentive consideration (as the output is irrelevant to factors other than the ability of the official). However, incorporating incentive consideration into our setting does not influence the learning process (that is, the inference of the innate ability θ), since the efforts in the equilibrium are known to the controller as Holmstrom (1999) shows. When the controller would like to

[4] We assume that there is no cost in appointing the official or assigning the official to different positions.

promote the most capable officials, whether to introduce incentive consideration would not influence our analysis.

2.2 Merit-seniority trade-off

We start with the simple case in which the controller is free to place the official at any level of the government, which allows us to highlight the merit-seniority trade-off. We assume that the controller is risk averse with a constant absolute risk aversion (CARA) preference. The problem is given by the following:

$$\max_j \mathbf{E}\left(\frac{1-exp(-ry_{js})}{r}|\Omega_{s-1}\right),$$

where Ω_{s-1} refers to the sequence of outputs from period 0 to period $s-1$. Thus, based on the information set Ω_{s-1}, the controller allocates the official to a certain level of the government to maximize utility.

The controller infers the innate ability of the official from the sequence of historical outputs, which is defined by $z_s \equiv \theta + \epsilon_s = \frac{y_{js}}{c_j}$. Through the observation of the sequence $\{z_s\}$, the controller learns about θ. Under the normality and independence assumptions, for the official at the beginning of period s (before production), the posterior distribution of θ stays normal with mean and variance given by the following:

$$\phi_s = \frac{\sigma_0^2 \sum_{t=0}^{s-1} z_t}{\sigma_\epsilon^2 + s\sigma_0^2};$$
$$\sigma_s^2 = \frac{\sigma_\epsilon^2 \sigma_0^2}{\sigma_\epsilon^2 + s\sigma_0^2}.$$

ϕ_s is the mean of perceived ability in period s, and σ_s^2 is the variance of perceived ability in the same period.

The problem of the controller in period s is equivalent to the following problem (due to certainty equivalence):

$$\max_{j} c_j(\phi_t - \frac{1}{2} r c_j(\sigma_s^2 + \sigma_\epsilon^2)).$$

Thus, we have the following proposition.

Proposition 1[5] The optimal job assignment rule is given by:

(i) If $\phi_s \leq \frac{1}{2} r(c_1 + c_2)(\sigma_s^2 + \sigma_\epsilon^2)$, the official should be assigned to level 1.

(ii) If $\frac{1}{2} r(c_1 + c_2)(\sigma_s^2 + \sigma_\epsilon^2) < \phi_s \leq \frac{1}{2} r(c_2 + c_3)(\sigma_s^2 + \sigma_\epsilon^2)$, the official should be assigned to level 2.

(iii) If $\phi_s > \frac{1}{2} r(c_2 + c_3)(\sigma_s^2 + \sigma_\epsilon^2)$, the official should be assigned to level 3.

Proposition 1 has the following implications. First, it suggests that the system is merit-based. The optimal job assignment favors an official with a larger perceived ability (ϕ_s). Moreover, as $\phi_s \xrightarrow{p} \theta$ ($s \to \infty$), an official with large innate ability (θ) is more likely to be promoted. Second, it suggests a trade-off between merit (ϕ_s) and seniority (σ_s): the controller would like to promote more capable officials but faces the risk of over-evaluating officials when their true ability lies below the controller's belief. Such risk (measured by σ_s) decreases with s increasing, so the proposition also implies that the threshold for promotion becomes lower for more senior officials. Third, seniority is more important at higher levels. As uncertainty is less tolerable at higher levels (reflected by the fact that σ_s^2 is multiplied by $c_2 + c_3$ at level 3 rather than $c_1 + c_2$ at level 2), seniority plays a greater role in the decision of promotion for higher levels. Last, the importance of seniority depends on the absolute risk-aversion coefficient r, which represents how sensitive the controller is to the over-evaluation risk. It also depends on the systematic risk (σ_0 and σ_ϵ) and the cost born by the controller c_j. Given the risk σ_0 and σ_ϵ, seniority becomes more important

[5] See Appendix for proof.

when c_j becomes larger, because the marginal productivity at different levels is associated with the potential loss of promoting an unable official (an official with innate ability $\theta < 0$). These discussions may explain different promotion patterns of different organizations in the existing literature, as the parameters may vary in different organizations.

2.3 Sequential correlation

We now move to the more realistic case in which the allocation of the official is subject to sequential rules. We suppose that the official is assigned to level 1 in period 0. The sequential rules are described by the following two assumptions:

Assumption 1 *Downward rigidity* (no demotion):

The official is not allowed to be assigned to level j at period s and to level j' at period $s + 1$ if $j' < j$.

Assumption 2 *Upward stability* (no jumps in promotion):

The official is not allowed to be assigned to level 1 at period s and to level 3 at period $s + 1$.

We define t_1 as the period when the official is promoted from level 1 to level 2, and t_2 as the period when the official is promoted from level 2 to level 3. t_1 (or t_2) equals ∞ if the official is never assigned to level 2 (or 3). Following this definition, the duration at level 1 is t_1, whereas the duration at level 2 equals $t_2 - t_1$. First-order stochastic dominance is denoted by $\succsim$. The following proposition shows the relationship between ability and duration:

Proposition 2[6] (i) $t_1|\theta \succsim t_1|\theta' \iff \theta \leq \theta'$. (ii) $t_2|(\phi_{t_1}, t_1) \succsim t_2|(\phi'_{t'_1}, t'_1)$ if $t_1 = t'_1$ and

[6]See Appendix for proof.

$\phi_{t_1} \leq \phi'_{t'_1}$.

Proposition 2 predicts that the official will be promoted faster when the official's ability rises, and the result also holds for perceived ability ϕ_s. The result is intuitive; compared with the official of low ability, the official of high ability has a greater probability of being identified with high ability despite the presence of noise. Proposition 2 characterizes the meritocratic feature of the system.

We turn to the sequential correlation between durations (t_1 and $t_2 - t_1$), for which the literature produces mixed findings. The "$\Rightarrow$" part of Proposition 2(i) predicts that the official promoted to level 2 faster (lower t_1) tends to have higher ability. Proposition 2(ii) suggests that the official of higher perceived ability (higher ϕ_{t_1}) is also promoted to level 3 faster (given his past duration). This is the merit effect. However, previous instances of faster promotion might be a result of over-evaluation, that is, the official's ability might be over-evaluated because of the existence of positive shocks (large positive realizations of ϵ). To correct this over-evaluation risk, faster promoted officials have to reach higher thresholds of ability for their next promotion so their duration at level 2 is expected to be longer. As can be seen, the ability threshold for promotion to level 3 in period $t_1 + k$ is $\frac{1}{2}r(c_2 + c_3)(\sigma^2_{t_1+k} + \sigma^2_\epsilon)$, which is larger for smaller t_1. This is the seniority effect. Whether the correlation between durations is positive or negative depends on whether the merit effect or the seniority effect dominates. For example, when an official's innate ability is large enough, then the perceived ability is most likely larger than $\frac{1}{2}r(c_2 + c_3)(\sigma^2_0 + \sigma^2_\epsilon)$. In this case, the official will be promoted as quickly as possible, and seniority will not be considered. Rather, when officials have mediocre abilities, seniority effect plays a role and negative correlation is predicted.

We summarize this discussion in the following conjecture:

Conjecture. The job durations of officials can be either positively (for high-ability officials) or negatively (for mediocre-ability officials) correlated across levels of government.

3 Institutional backgrounds and data

3.1 Institutional backgrounds

In China, the CCP controls personnel assignments. The literature has provided abundant material describing how the system works (see Xu (2011) and Landry et al. (2018)). In this paper, we only provide a brief description that highlights the role of ability in the CCP's evaluation of officials.

The Chinese bureaucracy is hierarchical. There are five levels of government: central, provincial, municipal, county (district), and township. At each level, there is a corresponding party committee whose secretary is the No. 1 person in the jurisdiction. The executive officer of the government is the No. 2 person whose task is to implement the decision of the party committee. The party committee, through its organizational department, manages the personnel of its level of government as well as the appointment of the two heads—the party secretary and the executive officer—one level below. In this paper, we only study the promotion of those two heads. When a vacancy (for example, a mayor position) appears, the organizational department one level above will begin a search. Qualified officials include the deputies (for example, deputy mayors and deputy party secretary not confined to the same city), governors of large districts, and mayors of smaller cities. Although there is no explicit competition, officials understand that their future promotion depends on their current performance.

In addition to party loyalty, merit is one of the most crucial qualities for promotion, which is made clear by the CCP's official documents.[7] Admittedly, merit has many facets; however, the ability to grow the economy is by far the most important because economic growth has been one of the sources for the CCP's political legitimacy. Other kinds of ability are likely to be correlated with economic growth. In addition, economic growth

[7]See Working Regulations for the Selection and Appointment of Party and Government Officials retrieved on May 29, 2023.

is potentially the easiest to quantify among all kinds of abilities. As a result, economic growth has become one of the strongest competitive fronts for officials. The literature finds that the organizational department identifies and credibly rewards officials who can nurture economic growth (Nathan, 2003; Li and Zhou, 2005; Landry et al., 2018). However, it remains a question whether the organizational department infers officials'ability to grow the economy from their observed performance or blindly uses their performance to decide on their promotion. Theoretically, the former makes more sense because observed performance may be contaminated by noises. Several studies have found that the organizational department does infer ability from officials'economic performance and makes promotion decisions based on that inferred ability (Yao and Zhang, 2015; Wang et al., 2022). Our work is premised on this empirical result as well as the theoretical consideration initiated by Holmstrom (1999).

3.2 Data and variables

The COD dataset compiles detailed biographies of all Chinese officials at or above the level of prefectures from 1994 to 2017.[8] It contains the following information for those officials: the starting and ending years of each term that they served, prefecture or province that they worked in, their positions or ranks, and their individual characteristics (gender, ethnicity, education, and so on). We can trace the complete career path for each official and build a working network for him or her.

In this paper, we study two samples of officials. The first sample comprises prefectural governors and party secretaries[9] recorded by the COD. They make up the level-1 officials in our theoretical model. There are 3,282 prefectural officials with a total of 20,103 observations of terms. The second sample comprises provincial governors and party secretaries who constitute the level-2 officials in our theoretical model. There are 314 provincial officials

[8]A detailed description of the dataset can be found in Wang et al. (2022).

[9]The former is the head of the prefectural government and the latter one is the head of the prefectural branch of the Communist Party.

with a total of 2,093 observations of terms.

Promotion For prefectural (provincial) officials, promotion is defined as being appointed to any provincial (national) position with higher administrative rank than prefectural (provincial) leaders.

Ability In our theoretical model, the controller infers an official's ability from the record of past performance. The controller makes decisions based on the perceived ability ϕ_s, which is essentially the precision-adjusted average of the official's past performances. We thus follow Yao and Zhang (2015) and Wang et al. (2022) to estimate prefectural officials'abilities from the following (relative) growth equation:

$$Rg_{ijt} = \alpha + X_{jt}\beta + \phi_i^1 + \eta_j + \mu_t + \epsilon_{ijt} \tag{1}$$

where Rg_{ijt} is the relative growth rate of prefecture j in year t when official i served in the prefecture. Rg_{ijt} is defined as the ratio of the growth rate of prefecture j in year t to the growth rate of the province k which that prefecture j belongs to in year t. X_{jt} is a set of variables that includes gross domestic product per capita in the first year of a term, population, and inflation rate. The three parameters ϕ_i^1, η_j, and μ_t are the fixed effect of official i, prefecture j, and year t, respectively. By the setup, ϕ_i^1 is the researcher's estimate for city official i's contribution to local growth net of city-specific confounding factors. Therefore, it corresponds to ϕ_s in our model when s indicates the ending year of the official's tenure as a city official. That is, it is the perceived ability of prefectural official i at the end of his or her tenure.[10] Yao and Zhang (2015) show that the relative size of ϕ_i^1 can be determined in a connected sample of cities in which there have been officials who were transferred among them. That is, officials in a connected sample are comparable in ability.[11]

[10]Therefore, our model provides a theoretical foundation for Yao and Zhang (2015)'s empirical strategy to identify officials' abilities.

[11]The largest connected sample contains 285 cities and 2,438 mayors and party secretaries (88 percent of the entire sample).

We adopt a similar specification to estimate the abilities of provincial officials.

$$Rg_{ikt} = \alpha + X_{kt}\beta + \phi_i^2 + \eta_k + \mu_t + \epsilon_{ikt} \tag{2}$$

where Rg_{ikt} is province k's growth rate relative to the national average in year t when official i worked in province k[12]. Like ϕ_i^1, ϕ_i^2 is the parameter of interest that measures provincial official i's perceived ability. Figure A1 in the appendix presents the distributions of the standardized estimates of abilities for prefectural and provincial officials. Both are close to the normal distribution.

4 Empirical evidence from Chinese officials

4.1 Meritocracy at the prefectural level

Our empirical work tests Proposition 2 and the Conjecture raised by our theory. In this subsection, we test Proposition 2(i) for meritocracy using the prefectural sample. The tests of Proposition 2(ii) and the Conjecture will be conducted in the next subsection.

Our first exercise is to estimate the nonparametric Kaplan-Meier failure functions for two groups of prefectural officials. Officials with above-the-median ability are grouped into *high ability* and the rest are grouped into *low ability*. Figure 1 presents the Kaplan-Meier failure functions for the two groups of officials. As seen in the figure, *high ability* stochastically dominates *low ability*, supporting Proposition 2(i) that more capable prefectural officials get promoted faster than their peers.

[Figure 1 about here]

Going beyond the simple comparison in Figure 1, we examine ability's impacts on officials'job durations and promotions by estimating the Cox proportional-hazards model

[12] All 31 provinces in (mainland) China are connected.

of the following form:

$$h(i,t) = h_0(t)\exp(\beta * Ability_i + \gamma Z_i + \delta_t + \mu_{j(i,t)}) \tag{3}$$

where i indexes officials, $h(i,t)$ is the hazard rate of promotion to the province in year t of official i, $h_0(t)$ is the baseline hazard function, and $Ability_i$ is official i's ability extracted from Equation (1). Following the semi-parametric approach devised by Cox (1972), we leave the baseline hazard function $h_0(t)$ completely unrestricted and estimate the rest of the coefficients by the partial maximum likelihood method. The estimated coefficient β captures the effect of perceived ability on promotion probability. A positive (negative) β indicates that an increase in the value of ability raises (lowers) the promotion hazard. The vector Z consists of other individual characteristics, including gender and ethnicity[13], that might confound our estimates. Furthermore, we include prefecture (province) fixed effects $\mu_{j(i,t)}$ and year fixed effects δ_t in our specification because the promotion rule also depends on regional heterogeneity as well as the political cycle.

Table 1 reports the results of the Cox proportional-hazards regressions with the sample of prefectural officials. Column (1) shows the results with a binary variable to indicate whether an official is of high ability (= 1 if above the median). High ability is positively related to the hazard rate of promotion and statistically significant at the 1 percent level. For officials of high ability, the hazard of promotion is exp(0.301) - 1 = 35%, which is higher relative to those of low ability, and it translates to 4 percent higher promotion probability on average at the prefecture level (4 percent is the marginal increment of promotion probability when the official switches from low ability to high ability and other variables are on the mean). The finding is unaffected when we control for gender and ethnicity in column (2). We directly use perceived ability as the key independent variable in columns (3) and (4). According to the coefficient of column (3), each 1-unit (also one standard deviation

[13] Ethnicity = 1 if official i is Han, = 0 otherwise.

as the ability is standardized) increment to ability from the average level increases the promotion hazard by approximately exp(0.313) -1 = 37%, which translates to 35 percent higher promotion probability on average at the prefecture level.

[Table 1 about here]

Our findings in Table 1 are robust against many tests. One concern is the functional form and the estimation method. The flexibility allowed for by the semiparametric estimation of the Cox model comes at a cost of efficiency. Therefore, we estimate a parametric Weibull hazard model instead.[14] We report the estimates from the Weibull hazard model with the sample of prefectural officials in column (1) of Table 2. Results still hold under the alternative specification.

Another concern is the existence of omitted variables. For example, an argument is that the controller prefers younger officials because younger officials could work longer in organizations and reduce replacement cost. To lesson this concern, we control for officials'age when they become prefectural officials in column (2) of Table 2. The coefficient of ability is almost unchanged compared to benchmark results and still significant at the 10 percent level.

In addition, one question is whether officials with strong political connections could be promoted faster than their peer cohorts. To lesson this concern, we control for political connection in column (3) of Table 2. Political connection is defined as a binary variable that is equal to 1 if an official had worked as a subordinate no more than two levels below the concurrent general party secretary by year t[15]. The coefficient of ability is near to benchmark results and still significant at the 10 percent level.

Nevertheless, remaining unobserved heterogeneity could still bias our estimates, no matter what the correlation between the included and excluded variables may be(Van den

[14] Our test rejects the exponential baseline hazard function. We choose Weibull model adhering to Akaike's information criterion (AIC).

[15] We focus on shared workplace connections because it is supported by a large amount literature (Jia et al., 2015; Jiang, 2018; Landry et al., 2018)

Berg, 2001). Therefore, we generalize the Weibull hazard model to allow for an unobserved gamma-distributed frailty. The results allowing for unobserved heterogeneity are shown in column (4) of Table 2. The estimates of the Weibull model controlling for unobserved heterogeneity are very similar to those of the Cox model. More important, allowing for unobserved heterogeneity increases the magnitude of the coefficient, which suggests an even higher promotion hazard than the benchmark.

[Table 2 about here]

4.2 Negative correlation of job durations

We now turn to test the Conjecture. We first take a glance at the relationships between the job durations at different levels for officials in the central government.[16] Figure A2 in the appendix shows that officials who stay for shorter periods in prefectural positions tend to stay longer in provincial positions. Also, the correlation coefficient suggests that with one fewer year at the prefecture level, an official would expect 0.45 years longer at the province level. In the previous subsection, we show that more capable prefectural officials are promoted faster. The two results we have obtained here indicate that although they are more prudently scrutinized at the provincial level, they eventually end up moving faster to the central government in their whole career.

As before, to evaluate the correlation of job durations between prefectural and provincial governments, we classify provincial officials into a *fast group* and a *slow group*, but this time based on whether officials' job durations at the prefecture level is shorter than the median. Figure 2 shows that officials in the *slow group* stochastically dominate officials in the *fast group*, which confirms the existence of negative correlation.

[Figure 2 about here]

[16]To prevent the truncation of career paths, we only include national officials in the sample.

Next, we provide a more rigorous examination on how prefectural job duration affects the promotion hazard at the provincial level. The hazard function we estimate is as follows:

$$h(i,t) = h_0(t)\exp(\alpha * PrefecturalDuration_i + \beta * Ability_i + \gamma Z_i + \delta_t + \mu_{j(i,t)}) \qquad (4)$$

where the hazard is defined on the promotion to the central level, $Prefectural\ Duration_i$ is provincial official i's job duration at the prefectural level, $Ability_i$ is official i's ability obtained from Equation (2) and the definitions of other variables remain as before. The coefficient of $Prefectural\ Duration_i$, α, provides a test for the Conjecture. A positive α indicates that the durations of the prefecture and the province are negatively correlated, so the seniority effect is likely to dominate. A negative α indicates that the durations of the prefecture and the province are positively correlated, so the merit effect is likely to dominate. Note that the seniority effect only kicks in for officials at the provincial level; at the prefectural level, merit is the only factor in consideration. Because we control $Prefectural\ Duration_i$ in Equation (4), the coefficient of $Ability_i$, β, provides a test of Proposition 2(ii)—that is, meritocracy at the provincial level.

Table 3 reports the results of the Cox proportional-hazards regressions with the sample of provincial officials. Column (1) shows the results only with $Prefectural\ Duration_i$ and $Ability_i$, as well as province and year fixed effects.[17] Both prefectural job duration and ability are positively related to the hazard rate of promotion and statistically significant at least at the 5 percent level. The positive result of job duration indicates that the seniority effect dominates the merit effect when provincial officials are considered for promotion. According to the coefficient of column (1), each one more year at the prefectural level increases the promotion hazard to the provincial level by approximately exp(0.092) -1 = 9.6%, which translates to 1.9 percent higher promotion probability to the national level.

[17]One advantage of putting both prefectural duration and ability as independent variables is that we can separate the effects of abilities on promotion hazards, and thus, the effect of prefectural duration on promotion hazards is purely from exogenous output shocks.

However, more capable officials are still able to forge ahead of their less capable peers, like Proposition 2(ii) states, as we also find a positive β. Moreover, the positive β in column (1) indicates that 1-unit increment (also one standard deviation as the ability is standardized) to ability from the average level will result in 15 percent higher promotion probability on average at the provincial level.

Our conclusion remains unaffected when we control for gender and ethnicity in column (2), and additionally when we control for political connection in column (3). Furthermore, in columns (4) and (5), we partition the sample into a *high ability* subsample for officials whose ability is higher than the median and a *low ability* subsample for officials whose ability is lower than the median. We find that seniority only significantly improves promotion hazards in the *low ability* subsample, which supports our discussion of the merit-seniority trade-off in section 2.3, whereas ability improves promotion hazards in both subsamples, supporting Proposition 2(ii).

[Table 3 about here]

4.3 Alternative explanations for negative correlation

The ceiling effect. A possible argument is that because of the existence of mandatory retirement, officials who stay longer at the prefectural level should automatically stay shorter at the provincial level if they are chosen for promotion to the central government. If such a ceiling effect exists, the negative correlation of job duration may be a mechanical result. First, we note that for the ceiling effect, individual job durations for the prefectural and provincial levels added together should be concentrated around a value. Figure A3 in the appendix demonstrates that effect. In both the sample of all provincial officials and the subsample of provincial officials never promoted to the national level, the sum of durations at the prefectural level and the provincial level is not concentrated at a particular value. Therefore, the ceiling effect is unlikely a driving factor behind the negative correlation. In the next section, our dynamic model tries different retirement ages and shows that the negative correlation still exists, which further alleviates the concern of the ceiling effect.

Human capital accumulation. In addition to the over-evaluation risk that we have proposed, the promotion rule may favor seniority for the consideration of accumulated human capital related to the length of service. However, for officials at the prefectural or higher levels, job tenure may not improve human capital because typical prefectural officials are already 48 years old when they take the job, and their human capital formation have been basically completed. To get a more concrete result, we plot provincial officials'perceived ability improvement on their prefectural perceived ability against their job durations at the prefectural level and provincial level. The results are shown in in Figure A4 in the appendix. No significant correlation is found. Therefore, human capital accumulation is unlikely a factor driving the negative correlation.

5 A dynamic model with a continuum of officials

5.1 Towards a dynamic stochastic model

The illustrative model in Section 2 shows the trade-off between merit and seniority and sheds light on the correlation of duration. It fails to capture two interesting features of hierarchies.

The first feature is the dynamic dimension of hierarchies. The controller has to take into account the possibility of eventually promoting an official to a higher level when the controller decides the official's promotion at a lower level. Both the Type-I and Type-II errors will lower the quality of the pool of the selected officials, among which higher-level officials are again selected. The hierarchical nature of the system dictates that the mistakes have an accumulative effect along the ladder of the hierarchy. Therefore, the controller should be more prudent when making promotion decisions for lower-level officials.

The second feature is competition among officials working at the same level of the hierarchy. Slots are limited at the next level of government. Often, the number of slots gets smaller moving up the hierarchical ladder, so the controller cannot promote all qualified officials and the competition among officials is ensured. After introducing competition, relative ability rather than absolute ability becomes the decisive factor for promotion.

In this section, we extend the model in Section 2 to a dynamic stochastic model for promotion by taking into account those two features. Although we take a minimalist approach to the modeling, we are unable to obtain an analytical solution to the model and have to solve it quantitatively.

5.2 The Model

5.2.1 Setup

Consider a permanent controller who is in charge of appointing all officials in a hierarchical government of three levels. As in Section 2, the levels of government are denoted by $j \in \{1, 2, 3\}$. The hierarchy is still as described by Assumptions 1 and 2. There is a continuum of slots in the hierarchy with measure 1. Each official works in the everlasting government for T periods. When any official reaches the retirement age and, thus, leaves the government, a new official enters level 1 to keep the measure of officials equal to the measure of slots.[18]

Similar to Section 2, officials are assumed to have innate abilities that follow a normal distribution with mean 0 and variance σ_0^2, and the production technology is linear with marginal product c_j ($j \in \{1, 2, 3\}$). The measure of slots at level j is α_j ($j \in \{1, 2, 3\}$) with $\alpha_1 + \alpha_2 + \alpha_3 = 1$.

The controller tries to maximize the sum of discounted expected utility through time. More specifically, in period s, for official i ($i \in [0, 1]$), the official's output is denoted by y_{is} (we do not specify which level official i is allocated to). The controller's instantaneous

[18] One official at level 1 needs to be promoted to level 2 if the retirement happens at level 2. If the retirement happens at level 3, then one official at level 2 has to be promoted to level 3.

utility in period s, u_s, generated from the outputs of all the officials in the hierarchy is given as follows:

$$u_s = \int_0^1 \frac{1 - exp(-ry_{is})}{r}\, di.$$

Therefore, the sum of discounted expected utility is $E(\sum_{s=0}^{\infty} \beta^s u_s)$, where β is the discount rate. The controller tries to maximize this sum of utility by allocating officials to different levels of governments in each period.

The timing of the model in period s ($s = 0, 1, 2, ...$) is as follows:

(i) Vacancies are created by retirement at all three levels of the government.

(ii) The controller decides which officials to promote to fill the vacancies of the next level based on beliefs about officials'abilities.

(iii) Nature selects new officials to fill the vacancies at level 1 of the government.

(iv) Officials produce at the new level according to the production technology.

(v) Outputs are realized and the controller updates beliefs about officials' abilities. The process of how the controller updates the beliefs is the same as in Section 2.

5.2.2 Value functions

According to the Fubini principle and the specific form of u_s, we can solve the controller's problem by constructing value functions for it. We denote $v(j, t, \phi, s)$ as the expected utility that an official at level j with work experience t and perceived ability ϕ can bring to the controller in period s. It might differ in different periods. Promoting this official to level $j + 1$ increases the controller's utility by $R(j, t, \phi, s) = v(j + 1, t, \phi, s) - v(j, t, \phi, s)$. To maximize total expected utility, the controller will promote officials who can bring higher $R(j, t, \phi, s)$. Because there is a limited number of slots to fill at any time, the controller will promote officials with the highest $R(j, t, \phi, s)$ with their number just equal to the number of vacancies. Therefore, there exist thresholds $v_{1,2}(s)$ and $v_{2,3}(s)$ such that only when $R(j,t,\phi,s) \geq v_{j,j+1}(s)$ will the official be promoted from level j to level $j+1$ ($j =$ 1,2). This is the optimal promotion rule. It should be noted that $v_{j,j+1}(s)$ is endogenous and can vary across period.

The following lemma characterizes the value functions:

Lemma 1[19]

(i) For $t \geq T$, $v(j,t,\phi,s) = 0$.

(ii) For $t < T$, $v(j,t,\phi,s)$ is given by the following:[20]

$$v(3,t,\phi,s) = \frac{1 - exp[-rc_3(\phi - \frac{1}{2}rc_3 f(t))]}{r} + \beta E[v(3,t+1,\phi',s+1)|\phi]$$

$$\begin{aligned} v(2,t,\phi,s) = & \frac{1 - exp[-rc_2(\phi - \frac{1}{2}rc_2 f(t))]}{r} \\ & + \beta E[\mathcal{I}_{\{v(3,t+1,\phi',s+1)-v(2,t+1,\phi',s+1) \geq v_{2,3}(s+1)\}} v(3,t+1,\phi',s+1)|\phi] \\ & + \beta E[\mathcal{I}_{\{v(3,t+1,\phi',s+1)-v(2,t+1,\phi',s+1) < v_{2,3}(s+1)\}} v(2,t+1,\phi',s+1)|\phi] \end{aligned}$$

$$\begin{aligned} v(1,t,\phi,s) = & \frac{1 - exp[-rc_1(\phi - \frac{1}{2}rc_1 f(t))]}{r} \\ & + \beta E[\mathcal{I}_{\{v(2,t+1,\phi',s+1)-v(1,t+1,\phi',s+1) \geq v_{1,2}(s+1)\}} v(2,t+1,\phi',s+1)|\phi] \\ & + \beta E[\mathcal{I}_{\{v(2,t+1,\phi',s+1)-v(1,t+1,\phi',s+1) < v_{1,2}(s+1)\}} v(1,t+1,\phi',s+1)|\phi], \end{aligned}$$

where $\phi' = \phi + g(t)\lambda$ ($\lambda \overset{i.i.d.}{\sim} N(0,1)$), $f(t) = \sigma_t^2 + \sigma_\epsilon^2$ and $g(t) = \dfrac{\sigma_\epsilon \sigma_0^2}{\sqrt{(\sigma_\epsilon^2 + (t+1)\sigma_0^2)(\sigma_\epsilon^2 + t\sigma_0^2)}}$.

5.2.3 Equilibrium and calibration

For tractability, we only consider the stationary case in which the measures of officials with different work experiences are the same and equal $\frac{1}{T}$. This stationary case is partly supported by Figure A5 in the appendix, which demonstrates that the number of officials who became mayors or prefectural party secretaries was nearly constant over political

[19] See Appendix for proof

[20] $\mathcal{I}$ is the indicator function.

cycles. We denote $\alpha_{jt}(s)$ as the measure of officials at level j with work experience t in period s. Thus, we can define the equilibrium as follows:

Definition 1 Stationary equilibrium:

A stationary equilibrium is characterized by $\{v_{1,2}(s), v_{2,3}(s), (\alpha_{jt}(s))_{j=1,2,3;t=0,1,...,T-1}\}_{s=0,1,2,...}$ and the value function $v(j,t,\phi,s)$, which satisfies the following:

(i) Given $v_{1,2}(s)$ and $v_{2,3}(s)$, $v(j,t,\phi,s)$ is given by Lemma 1.

(ii) Given the promotion rule based on $v_{1,2}(s)$, $v_{2,3}(s)$ and $v(j,t,\phi,s)$, combined with the initial condition $\alpha_{jt}(0)$, $\alpha_{jt}(s)(s>0)$ are derived.

(iii) $\sum\limits_{t} \alpha_{jt}(s) = \alpha_j \ (j=1,2,3)$.

(iv) Stationary condition: $v_{1,2}(s), v_{2,3}(s), (\alpha_{jt}(s))_{j=1,2,3;t=0,1,\cdots,T-1}$ and $v(j,t,\phi,s)$ are ergodic (constant with respect to s).

The model is characterized by 11 parameters: α_j $(j=1,2,3)$, c_j $(j=1,2,3)$, β, r, σ_0, σ_ϵ, and T. We calibrate the model to the data provided by the COD. Levels 1-3 in our model refer to prefectural, provincial, and national positions, respectively. To match the COD data, we set $\alpha_1 = \frac{20}{23}$, $\alpha_2 = \frac{2}{23}$, $\alpha_3 = \frac{1}{23}$ and $T = 23$. Following convention, we set $r = 1.5$ and $\beta = 0.96$. In addition, σ_0 is normalized to 1 and σ_ϵ is set to 0.6 to match the results of estimated officials' abilities (see Section 3). c_j reflect the size of over-evaluation risks and, therefore, determine job durations. We thus calibrate the three parameters c_j $(j=1,2,3)$ to match the following three moments: average duration in prefectural positions (conditional on promotion to provincial positions), average duration in provincial positions (conditional on promotion to national positions), and average duration in national positions. The calibrated values of c_j are: $c_1 = 0.82$, $c_2 = 3.81$ and $c_3 = 5.66$. Table A1 in the appendix compares the calculated moments of this model to the COD.

5.3 Results

5.3.1 The merit-seniority trade-off

As discussed in the illustrative model, the ability thresholds for promotion for more senior officials are lower, because seniority lowers the risk of over-evaluation. The dynamic model obtains an exact relationship. Figure 3 shows a negative relationship between work experience and ability thresholds and supports the merit-seniority trade-off.

[Figure 3 about here]

5.3.2 Meritocracy

The evidence of meritocracy is presented in Figure 4. It takes a shorter period for more capable officials to receive a promotion at each level and in their whole career. One feature is that promotion from level 1 to level 2 is more sensitive to ability (and, thus, less sensitive to seniority) compared to promotion from level 2 to level 3. Seniority is more important at higher levels, because the potential loss from over-evaluation is larger at higher levels. The meritocratic feature of the system is also captured by Figure A6 in the appendix in which more capable officials are shown to have larger probabilities of promotion. Also in the appendix, Figure A7 provides further evidence from another perspective: officials who have advanced to higher levels tend to have higher abilities. It is noteworthy that the distribution of ability becomes more concentrated at higher levels because of a prominent characteristic of meritocracy—only officials with high abilities are likely to be promoted.

[Figure 4 about here]

5.3.3 Career path

The joint work of merit and seniority determines officials'career paths. For officials with extremely high abilities, the merit effect is likely to dominate the seniority effect so they may move on the fast track and get ahead of their peers at all levels of the government. Officials of relatively high abilities are likely to be promoted faster from level 1 to level 2, but their ride may slow down afterward because seniority begins to play a greater role. Figure 5 presents the career paths of four groups of officials with different levels of ability. The most capable group—whose ability is 2.5, or 2.5 standard deviation above the mean—rises up very quickly to level 3. The least capable group—whose ability is 1.75, or 1.75 standard deviation above the mean—almost never makes it to level 3. The middle groups—whose abilities are 2.25 and 2—can reach level 3, but they show different patterns. Compared to officials of ability 2, though officials of ability 2.25 rises quicker from level 1 to level 2, they gradually lose their lead as work experience increases. That is a sign for a greater role of seniority.

[Figure 5 about here]

5.3.4 Inverted-U shape relationship between durations

The discussion in Section 2.3 hints at an inverted-U shaped relationship between job durations across levels, which is supported by Figure 6. Positive and negative correlation coexist because the relative strength of the two countervailing effects—merit effect and seniority effect—differs for officials with different levels of ability. Because of the complexity of the model, the interplay of the two effects is not readily clear. But we conjecture that for officials with extremely high abilities, the merit effect always dominates, and early faster promotion leads to faster promotion at a later level; for officials with modestly high abilities, the seniority effect dominates so a negative correlation of job duration ensures. We verify this finding in Figure A8 in the appendix. The figure depicts the ability distributions of two groups of officials, one on the upward side of the inverted-U curve and the other on the downward side of the inverted-U curve. The first distribution clearly dominates the second distribution.

[Figure 6 about here]

To show that the inverted-U shaped relationship is not a mechanical result stemming from the retirement age, we have experimented with both a shorter retirement age and a longer retirement age to make $T = 17$ and $T = 30$, respectively. The inverted-U shaped curve remains. We do not report the detailed graphs due to limited space.

5.3.5 Comparative statics with varying degrees of risk aversion

The results suggest an explanation for the mixed findings in the literature. In most cases, individuals having embarked on the track of promotion have already been selected several times, so their abilities are high and the fast track is found for them. However, in some cases, individuals with mediocre levels of ability are promoted, so a negative correlation of duration is found. But a more substantive factor driving the different results may be the controller's risk attitude. A more prudent controller is more averse to making mistakes, so the negative correlation is more likely to occur.

We experiment with two values of r against its benchmark value 1.5, one smaller (1.0) and one bigger (2.0). Figure 7 presents the inverted-U shaped curve for each of the three cases. It is clear that the curve shifts northwest as r rises—the positive portion is shortened and becomes steeper. That is, as the controller becomes more risk averse, the ability threshold is raised for officials to qualify for the fast track, and even for those officials on the fast track, they must stay at level 2 longer than before. Apparently, seniority becomes more crucial to prevent the over-evaluation risk. Moreover, because the curve shifts left as r rises, the negative portion of the curve becomes longer. Therefore, we are more likely to find a negative correlation of duration under a more risk-averse controller. The majority of existing empirical literature studies the business world. In that world, risk taking is part of entrepreneurs' daily routine, if not one of their appreciated virtues. Therefore, fast tracks prevail. Governments behave quite differently; prudence supersedes innovation because a failure would entail serious political consequences, which is probably why we find a negative correlation in this paper.

[Figure 7 about here]

6 Conclusion

Career dynamics has remained one of the central themes in organizational economics. Against the common finding of fast tracks, this paper finds a negative correlation of job durations between levels of government. We provide a new explanation for the sequential correlation of career paths in hierarchies. Central to our explanation is the over-evaluation risk, which produces the trade-off between the merit effect and the seniority effect. Our dynamic model pins down two cases in which fast tracks are more likely to be found. One is when most of the individuals under consideration are very capable, and the other is when the controller is not very risk averse. Conversely, there are also two cases in which a negative correlation of job duration is more likely to be found. One is when most of the individuals under consideration are not very capable, and the other is when the controller is very risk averse.

One of the drawbacks of this paper is that we assume that officials exert a fixed amount of effort. For that, we also exclude the possibility of strategic interactions among officials. Adding elastic effort but not allowing for strategic play does not change our results—the learning process and the key results that we have reached are the same as found by Holmstrom (1999). However, allowing for strategic play will bring substantive difficulties to our modeling. For example, knowing the negative correlation of job duration, a modestly capable person may not be willing to work very hard at level 1. But this shirking may engender a promotion to level 2 because the person's peers may take advantage of the shirking and work harder. The model can easily become intractable. Our model is a minimalist model that takes into account the necessary features of the system (dynamic, incomplete information, and many agents) and enables us to work out the key mechanisms behind the mixed empirical findings in the literature (including the empirical finding of this paper).

References

Ariga, K., Ohkusa, Y., and Brunello, G. (1999). Fast track: Is it in the genes? The promotion policy of a large Japanese firm. *Journal of Economic Behavior & Organization*, 38(4):385–402.

Baker, G., Gibbs, M., and Holmstrom, B. (1994). The internal economics of the firm: Evidence from personnel data. *The Quarterly Journal of Economics*, 109(4):881–919.

Belzil, C. and Bognanno, M. (2010). The promotion dynamics of American executives. In *Research in Labor Economics*, volume 30, pages 189–231. Emerald Group Publishing Limited, Leeds.

Bernhardt, D. (1995). Strategic promotion and compensation. *The Review of Economic Studies*, 62(2):315–39.

Breaugh, J. A. (2011). Modeling the managerial promotion process. *Journal of Managerial Psychology*, 26(4):264–77.

Brüderl, J., Diekmann, A., and Preisendörfer, P. (1991). Patterns of intraorganizational mobility: Tournament models, path dependency, and early promotion effects. *Social Science Research*, 20(3):197–216.

Chiappori, P.-A., Salanie, B., and Valentin, J. (1999). Early starters versus late beginners. *Journal of Political Economy*, 107(4):731–60.

Cirone, A., Cox, G. W., and Fiva, J. H. (2021). Seniority-based nominations and political careers. *American Political Science Review*, 115(1):234–51.

Clemens, A. (2012). Position-specific promotion rates and the "fast track" effect. In *Research in Labor Economics*, volume 36, pages 77–107. Emerald Group Publishing Limited, Leeds.

Cox, D. R. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society: Series B (Methodological)*, 34(2):187–202.

Gibbons, R. and Waldman, M. (1999). A theory of wage and promotion dynamics inside firms. *The Quarterly Journal of Economics*, 114(4):1321–58.

Gibbons, R. and Waldman, M. (2006). Enriching a theory of wage and promotion dynamics inside firms. *Journal of Labor Economics*, 24(1):59–107.

Holmstrom, B. (1999). Managerial incentive problems: A dynamic perspective. *The Review of Economic Studies*, 66(1):169–82.

Huang, Y. (2002). Managing Chinese bureaucrats: an institutional economics perspective. *Political Studies*, 50(1):61–79.

Jia, R., Kudamatsu, M., and Seim, D. (2015). Political selection in China: The complementary roles of connections and performance. *Journal of the European Economic Association*, 13(4):631–68.

Jiang, J. (2018). Making bureaucracy work: Patronage networks, performance incentives, and economic development in China. *American Journal of Political Science*, 62(4):982–99.

Landry, P. F., Lü, X., and Duan, H. (2018). Does performance matter? Evaluating political selection along the Chinese administrative ladder. *Comparative Political Studies*, 51(8):1074–105.

Li, H. and Zhou, L.-A. (2005). Political turnover and economic performance: The incentive role of personnel control in China. *Journal of Public Economics*, 89(9-10):1743–62.

Manion, M. (1985). The cadre management system, post-Mao: The appointment, promotion, transfer and removal of party and state leaders. *The China Quarterly*, 102:203–33.

Meyer, M. A. (1991). Learning from coarse information: Biased contests and career profiles. *The Review of Economic Studies*, 58(1):15–41.

Mills, D. Q. (1985). Seniority versus ability in promotion decisions. *ILR Review*, 38(3):421–25.

Nathan, A. J. (2003). Authoritarian resilience. *Journal of Democracy*, 14(1):6–6.

Podolny, J. M. and Baron, J. N. (1997). Resources and relationships: Social networks and mobility in the workplace. *American Sociological Review*, pages 673–93.

Seltzer, A. and Merrett, D. (2000). Personnel policies at the Union Bank of Australia: Evidence from the 1888–1900 entry cohorts. *Journal of Labor Economics*, 18(4):573–613.

Treble, J., Van Gameren, E., Bridges, S., and Barmby, T. (2001). The internal economics of the firm: further evidence from personnel data. *Labour Economics*, 8(5):531–52.

Van den Berg, G. J. (2001). Duration models: specification, identification and multiple durations. In *Handbook of Econometrics*, volume 5, pages 3381–460.

Wang, Z., Yao, Y., and Zhang, J. (2022). The competence-loyalty tradeoff in China's political selection. *Journal of Comparative Economics*, 50(2):369–93.

Xu, C. (2011). The fundamental institutions of China's reforms and development. *Journal of Economic Literature*, 49(4):1076–151.

Yao, Y. and Zhang, M. (2015). Subnational leaders and economic growth: Evidence from Chinese cities. *Journal of Economic Growth*, 20(4):405–36.

Figures and Tables

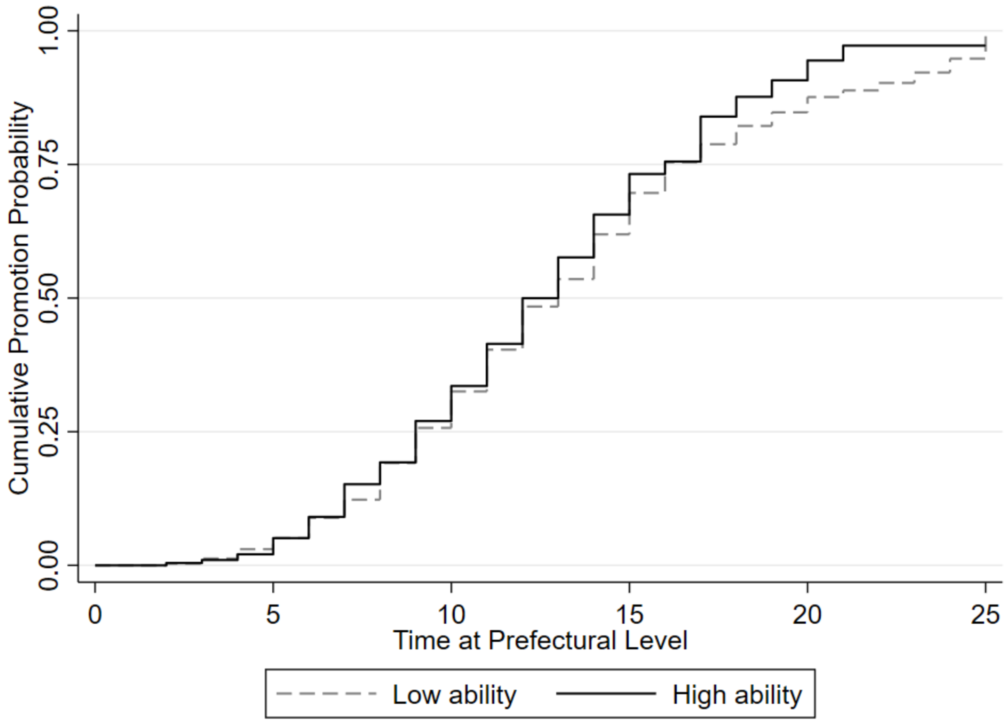


Figure 1: Kaplan-Meier Estimates of Failure Functions with the Sample of Prefectural Officials

Note: *High ability* is defined as the perceived ability of the official at prefecture level above median, while *low ability* is defined as below median.

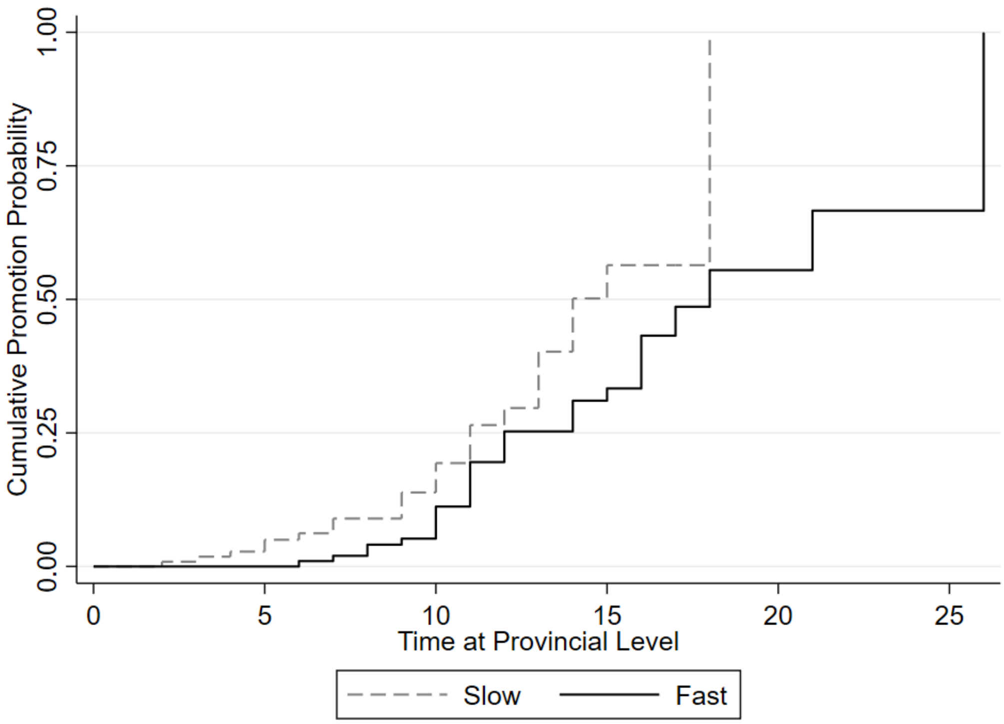


Figure 2: Failure Functions in Provincial Officials Sample by the Length of Duration

Note: This figure shows Kaplan-Meier failure functions with the sample of provincial officials. *Slow* is defined as the length of duration at prefecture level above median, while *fast* is defined as below median.

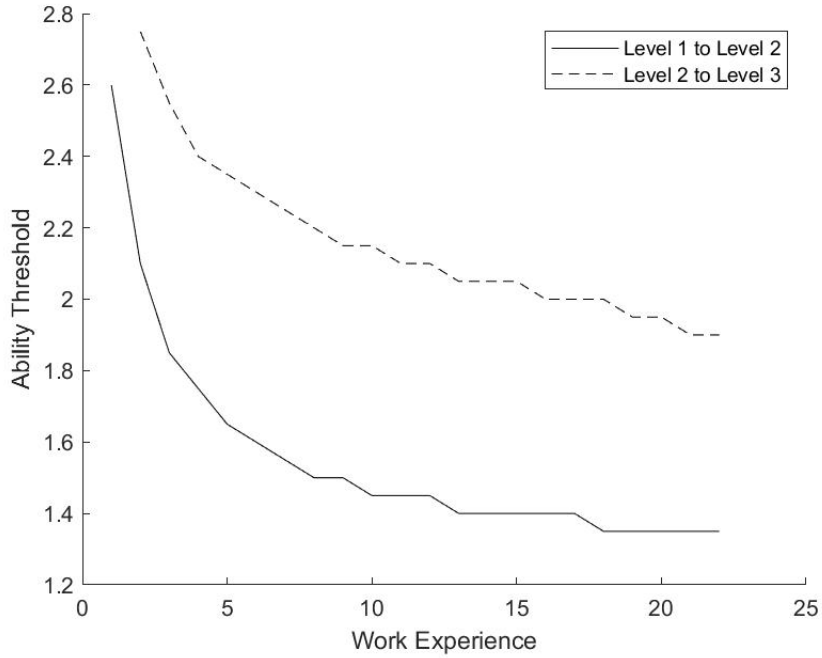


Figure 3: Ability Requirement for Officials with Different Work Experience

Note: This figure is generated from the dynamic model. The ability threshold concerns perceived abilities of officials.

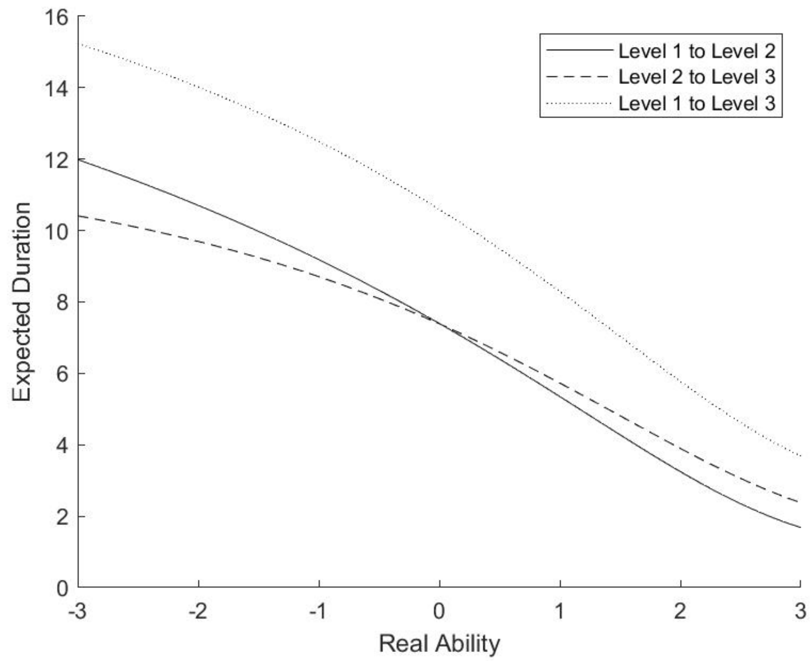


Figure 4: Duration before Promotion for Officials of Different Abilities

Note: This figure is generated from the dynamic model. Expected duration refers to the expected duration before promotion conditional on being promoted.

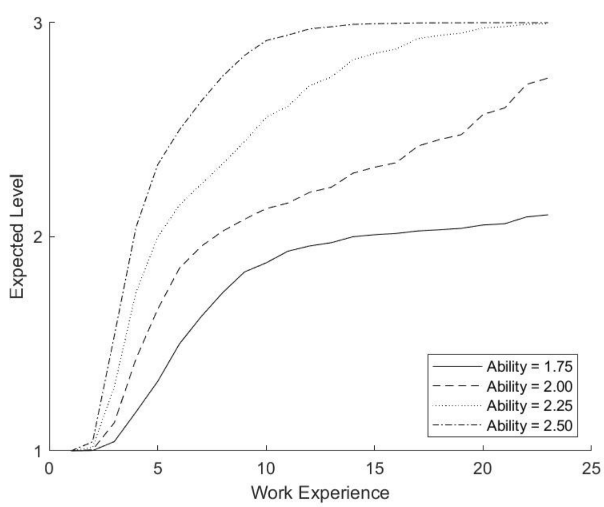


Figure 5: Career Paths of Officials of Different Abilities

Note: This figure is generated from the dynamic model. We only consider officials whose abilities are higher than 2 since those with lower abilities are not likely to be promoted to level 3 and are, therefore, not samples of interest.

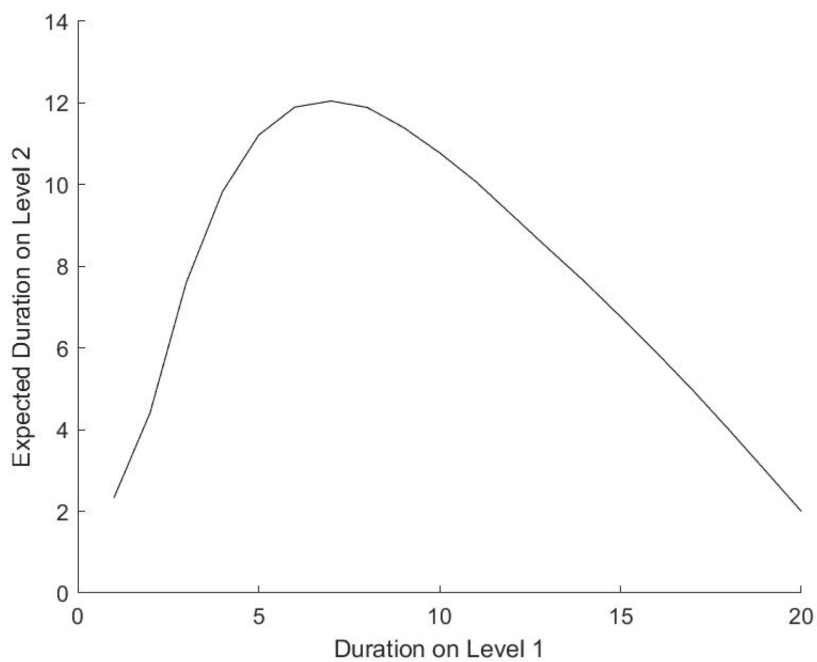


Figure 6: Correlation between Duration

Note: This figure is generated from the dynamic model. Expected duration at level 2 refers to the expected duration at level 2 before promotion conditional on being promoted to level 3.

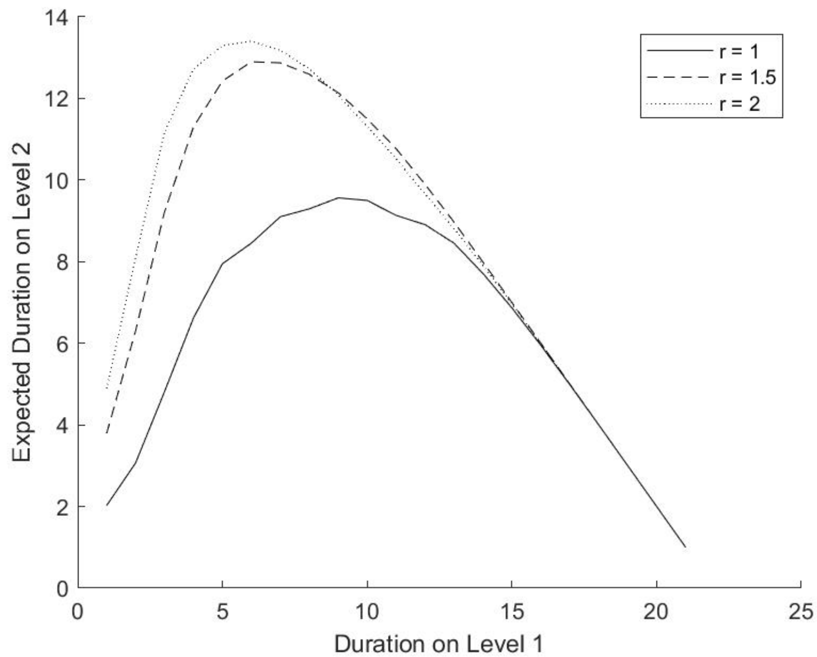


Figure 7: Comparative Statics on Absolute Risk Aversion Coefficient

Note: This figure is generated from the dynamic model. Expected duration at level 2 refers to the expected duration at level 2 before promotion conditional on being promoted to level 3.

Table 1: Ability on Promotion Hazard (Prefecture Sample)

	(1)	(2)	(3)	(4)
High Ability (dummy)	0.301***	0.307***		
	(0.093)	(0.094)		
Ability			0.313***	0.307***
			(0.067)	(0.067)
Female		0.403***		0.402***
		(0.154)		(0.154)
Ethnicity		0.327***		0.323**
		(0.124)		(0.124)
prefecture Fixed Effect	Yes	Yes	Yes	Yes
Year Fixed Effect	Yes	Yes	Yes	Yes
# observations	6,288	6,276	6,288	6,276
# subjects	2,572	2,564	2,572	2,564
# failures	849	849	849	849
Pseudo-R2	0.062	0.064	0.063	0.065
χ_2	694.684	710.512	704.835	719.707
Test for proportion	p = 0.76	p = 0.81	p = 0.99	p = 0.89
Δ failure function	0.04	0.04	0.35	0.34

Note: Table 1 reports the estimated coefficient from the Cox proportional-hazards model of Equation (3). We verify the proportional-hazards assumption based on Schoenfeld residuals and report the p-value of global chi-square test. The global chi-square test fails to reject the null hypothesis that the hazards were proportional across groups. Standard errors are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 2: Robustness Check

	(1)	(2)	(3)	(4)
Ability	0.339***	0.283***	0.301***	0.355***
	(0.068)	(0.067)	(0.067)	(0.074)
Female	0.452***	0.449***	0.404***	0.487***
	(0.156)	(0.154)	(0.154)	(0.171)
Ethnicity	0.389***	0.394***	0.326***	0.414***
	(0.125)	(0.127)	(0.124)	(0.136)
Age when Becoming Prefectural Official		-0.054***		
		(0.009)		
Political Connection			0.534	
			(0.326)	
City Fixed Effect	Yes	Yes	Yes	Yes
Year Fixed Effect	Yes	Yes	Yes	Yes
Model	Weibull	Cox	Cox	Weibull
Heterogeneity	No	No	No	Yes
# observations	6,276	6,138	6,276	6,276
# subjects	2,564	2,494	2,564	2,564
# failures	850	842	850	850
Pseudo-R2		0.071	0.065	
χ_2	711.720	775.659	728.948	700.943

Note: Weibull model is used in columns (1) and (4), and Cox proportional-hazards model is used in columns (2) and (3). Estimates within prefecture sample are from columns (1) to (4). Standard errors are in parentheses.* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3: Preceding Duration on Subsequent Promotion (Province Sample)

	(1)	(2)	(3)	(4)	(5)
Duration on Prefectural level	0.092**	0.094**	0.089**	0.090	0.785***
	(0.045)	(0.043)	(0.044)	(0.064)	(0.284)
Ability	0.552***	0.479**	0.452**	0.863**	2.898**
	(0.212)	(0.214)	(0.216)	(0.433)	(1.278)
Political Connection			0.912	0.518	-0.297
			(0.814)	(1.020)	(1.528)
Sample	Full	Full	Full	High ability	Low ability
Province Fixed Effect	Yes	Yes	Yes	Yes	Yes
Year Fixed Effect	Yes	Yes	Yes	Yes	Yes
Controls	No	Yes	Yes	Yes	Yes
# observations	699	699	699	354	345
# subjects	247	247	247	121	126
# failures	42	42	42	23	19
Pseudo-R2	0.226	0.238	0.241	0.289	0.533
chi2	79.589	83.678	84.915	48.307	68.82
phtest	p = 0.85	p = 0.99	p = 0.98	p = 0.79	p = 0.99

Note: Table 3 reports estimated coefficient from the Cox proportional-hazards model of Equation (4). We verify the proportional-hazards assumption based on Schoenfeld residuals and report the p-value of global chi-square test. The global chi-square test fails to reject the null hypothesis that the hazards were proportional across groups. Standard errors are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Appendix

Proof

Proof for Proposition 1

For i and j $(i > j)$, $c_i(\phi_s - \frac{1}{2}rc_i(\sigma_s^2 + \sigma_\epsilon^2)) \geq c_j(\phi_s - \frac{1}{2}rc_j(\sigma_s^2 + \sigma_\epsilon^2)) \iff \phi_s \geq \frac{1}{2}r(c_i + c_j)(\sigma_s^2 + \sigma_\epsilon^2)$. Proposition 1 follows directly. QED.

Proof for Proposition 2

We denote $a_t = \frac{1}{2}r(c_1 + c_2)(\sigma_t^2 + \sigma_\epsilon^2)$ and $b_t = \frac{1}{2}r(c_2 + c_3)(\sigma_t^2 + \sigma_\epsilon^2)$.

(i) For every positive integer k,

$$
\begin{aligned}
Pr(t_1 \leq k|\theta) =& 1 - Pr(t_1 > k|\theta) \\
=& 1 - Pr(\phi_1 \leq a_1, ..., \phi_k \leq a_k|\theta) \\
=& 1 - Pr(\frac{\sigma_0^2}{\sigma_\epsilon^2 + \sigma_0^2}\epsilon_0 \leq a_1 - \frac{\sigma_0^2}{\sigma_\epsilon^2 + \sigma_0^2}\theta, ..., \frac{\sigma_0^2}{\sigma_\epsilon^2 + k\sigma_0^2}\sum_{s=0}^{k-1}\epsilon_s \leq a_k - \frac{k\sigma_0^2}{\sigma_\epsilon^2 + k\sigma_0^2}\theta),
\end{aligned}
$$

which is increasing with θ increasing. Therefore, $t_1|\theta \gtrsim t_1|\theta' \iff \theta \leq \theta'$.

(ii) For $t > t_1$,

$$
\begin{aligned}
\phi_t \leq b_t \iff& \frac{\sigma_0^2}{\sigma_\epsilon^2 + t\sigma_0^2}\sum_{s=0}^{t-1} z_s \leq b_t \\
\iff& \frac{\sigma_0^2}{\sigma_\epsilon^2 + t\sigma_0^2}\sum_{s=t_1}^{t-1} z_s \leq b_t - \frac{\sigma_\epsilon^2 + t_1\sigma_0^2}{\sigma_\epsilon^2 + t\sigma_0^2}\sigma_0^2\phi_{t_1}.
\end{aligned}
$$

Therefore, $Pr(t_2 \leq k|t_1, \phi_{t_1})$ is increasing with ϕ_{t_1} increasing. Hence, $t_2|\phi_{t_1}, t_1 \gtrsim t_2|\phi'_{t'_1}, t'_1$ if $t_1 = t'_1$ and $\phi_{t_1} \leq \phi'_{t'_1}$. QED.

Proof for Lemma 1

For $t \geq T$, the official is retired and $v(i, t, \phi, s) = 0$.

Output y_j at level j follows a normal distribution with mean $c_j\phi$ and variance $c_j^2 f(t)$. $f(t)$ is the sum of estimation error σ_t^2 and the variance of random shock σ_ϵ^2. Therefore, $E(\frac{1-exp(-ry_j)}{r}) = \frac{1-exp[-rc_j(\phi - \frac{1}{2}rc_j f(t))]}{r}$. Combined with the Bellman equation, we get formulas in (ii).

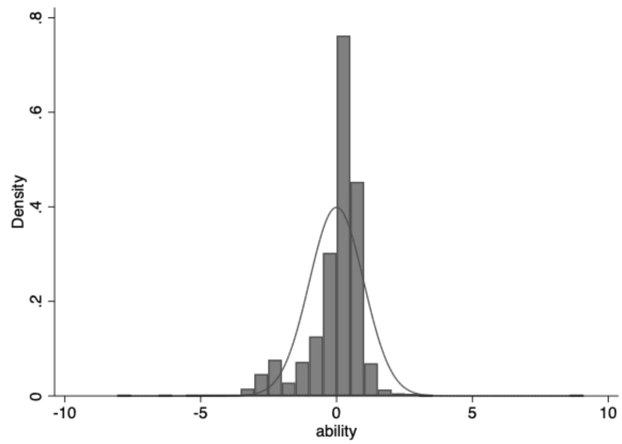
Now we characterize the path of ϕ.

$$\begin{aligned}
\phi_{t+1} &= \frac{\sigma_0^2 \sum_{s=0}^{t} z_s}{\sigma_\epsilon^2 + (t+1)\sigma_0^2} \\
&= \frac{\sigma_\epsilon^2 + t\sigma_0^2}{\sigma_\epsilon^2 + (t+1)\sigma_0^2}\phi_t + \frac{\sigma_0^2}{\sigma_\epsilon^2 + (t+1)\sigma_0^2} z_t.
\end{aligned}$$

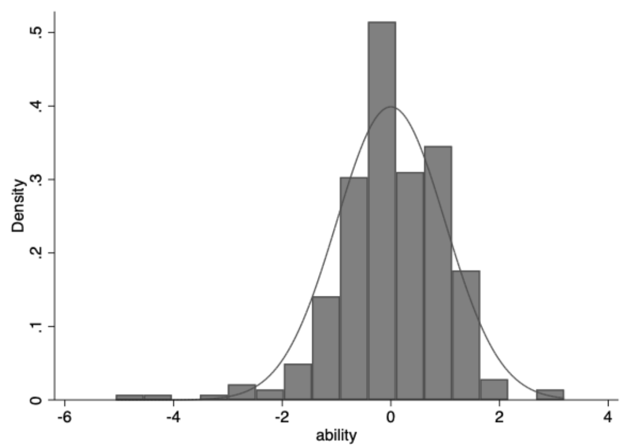
The conditional variance of z_t is given by the following:

$$variance(z_t|\phi_t) = variance(\theta|\phi_t) + \sigma_\epsilon^2 = \sigma_t^2 + \sigma_\epsilon^2.$$

Thus the posterior belief of z_t follows $N(\phi_t, \sigma_t^2 + \sigma_\epsilon^2)$. We reorganize the equation and we can get the state transition function: $\phi_{t+1} = \phi_t + g(t)\lambda(\lambda \overset{i.i.d.}{\sim} N(0,1))$. QED.



Panel (a)



Panel (b)

Figure A1: Distribution of Abilities

Note: Panel (a) shows the distribution of abilities of prefectural officials and panel (b) shows that of provincial officials.

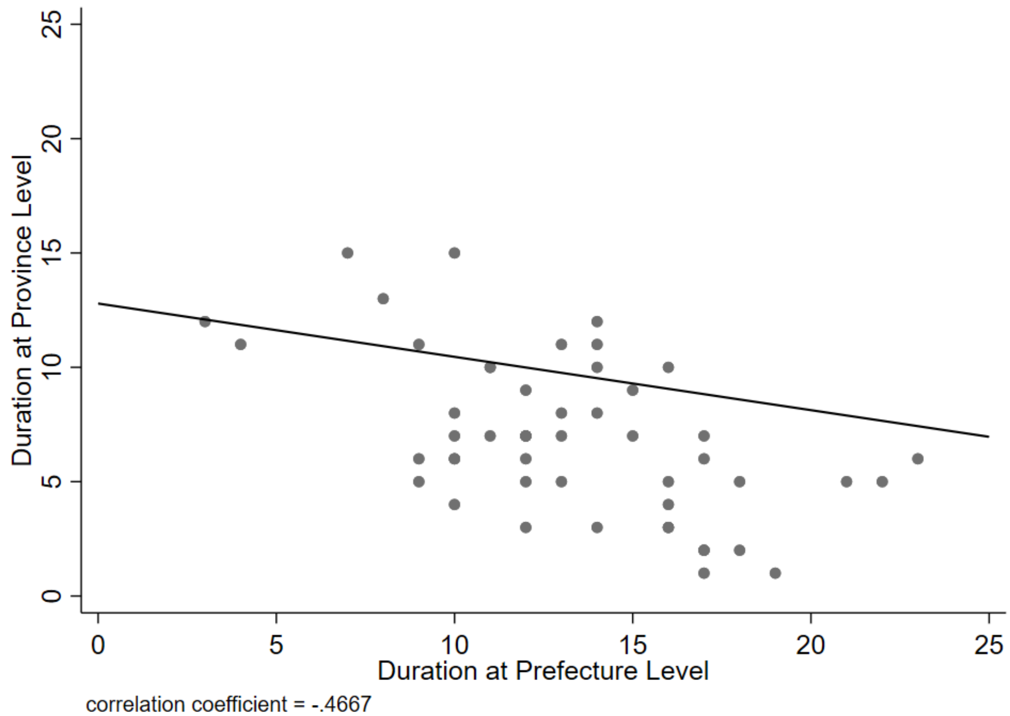
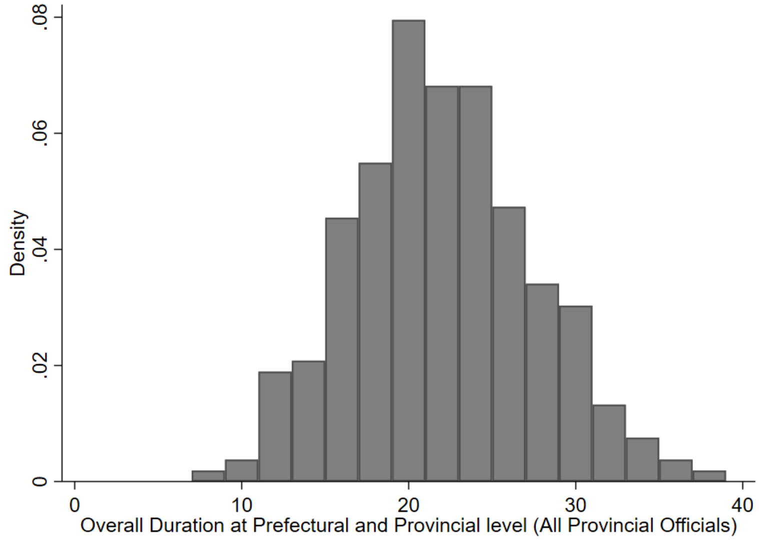
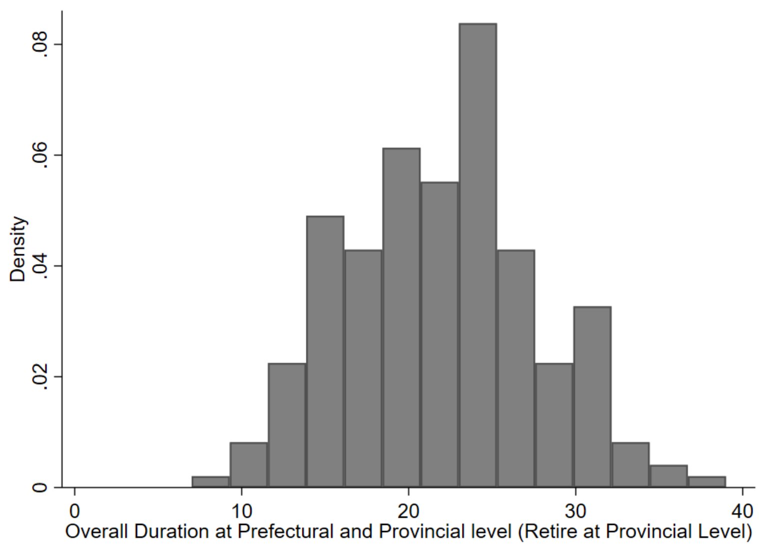


Figure A2: Correlation between Duration on Province Level and Prefecture Level

Note: The total of an official's duration at the prefectural and provincial levels may be close to fixed for those who are not promoted to the national level. For those officials, the relationship between duration at the prefecture and province level may be mechanical. To prevent this ceiling effect, we only include national officials in the sample.



Panel (a)



Panel (b)

Figure A3: The Distribution of Total Duration on Province Level and Prefecture Level

Note: Panel (a) shows the distribution of overall duration at prefectural and provincial levels for all provincial officials, and panel (b) shows the same distribution for provincial officials who are never promoted to the national level.

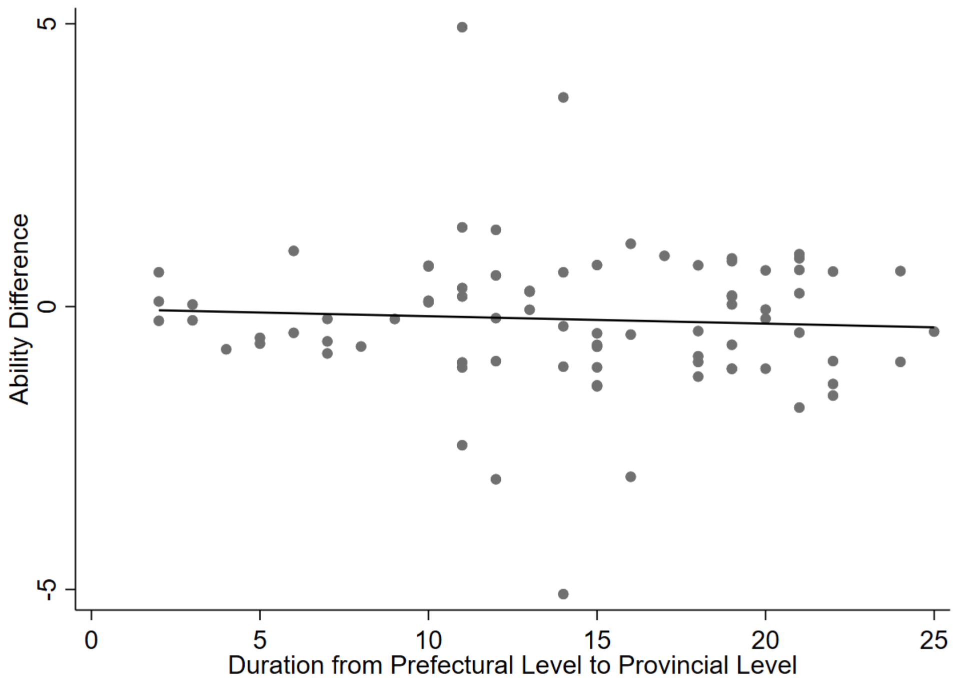


Figure A4: Whether Duration Improves Ability

Note: The duration from prefectural level to provincial level is defined as the time length from the year when the official first becomes prefectural governor or party secretary to the year when the official retires from the provincial level or is promoted to the national level. The ability difference is the difference between the perceived ability at the provincial level and the perceived ability at the prefectural level.

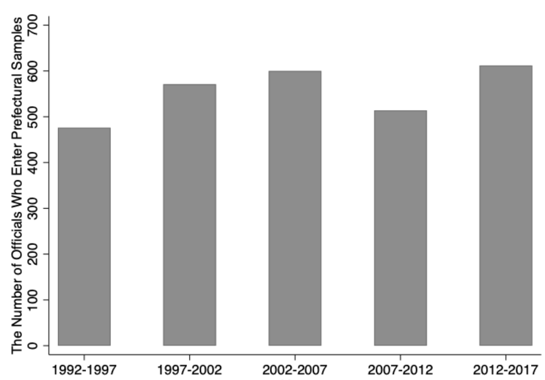


Figure A5: Number of Officials Who Enter Prefectural Samples

Note: The time interval 1992–1997 means [1992,1997), 1997–2002 means [1997,2002), etc. The personnel turnover is more frequent in the year when the CCP Congress is held. We plot the number of officials who become mayors or prefectural party secretaries in each political cycle to remove influence from the political cycle.

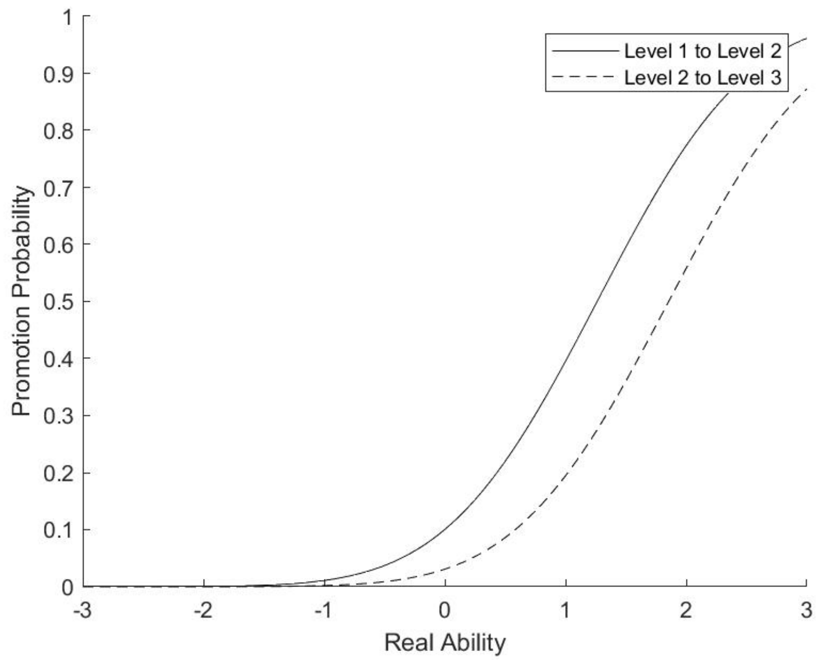


Figure A6: Promotion Probability for Officials of Different Abilities

Note: This figure is generated from the dynamic model.

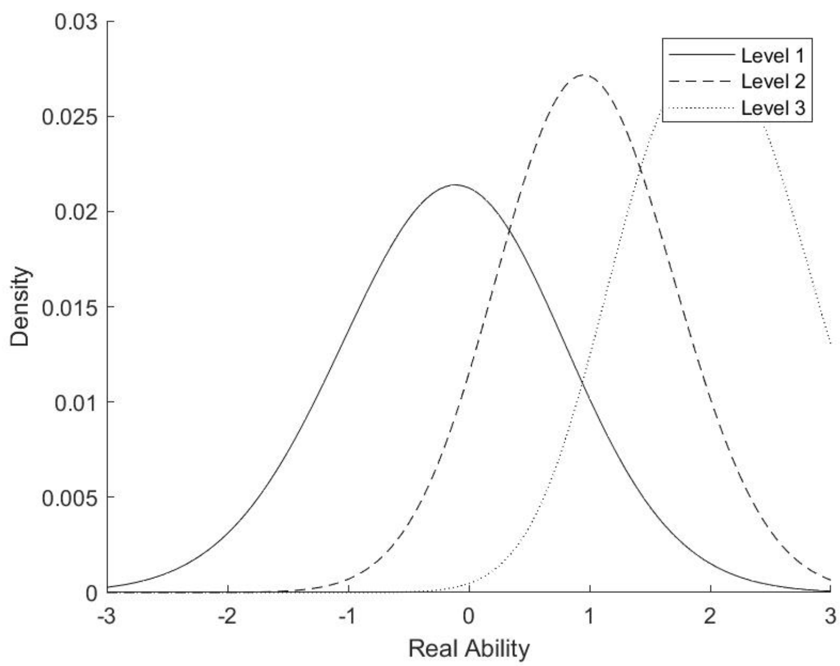


Figure A7: Distribution of Ability on Different Levels

Note: This figure is generated from the dynamic model.

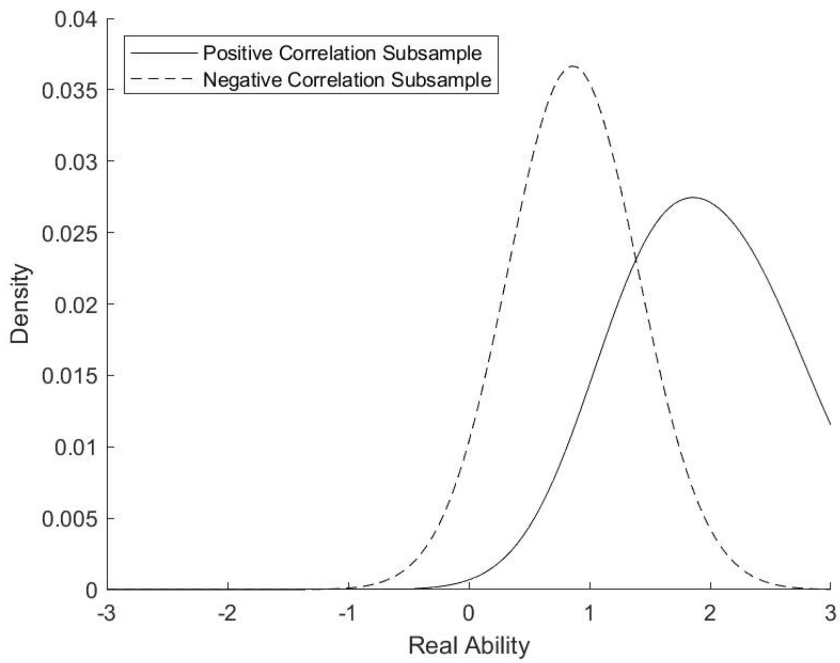


Figure A8: Distribution of Ability in Two Sides of inverted-U Shape

Note: This figure is generated from the dynamic model.

Table A1: Comparison between the Model and the Real Data

	Predicted Value	Real Value
Ratio of Slots in Prefectural Level	0.88	0.87
Ratio of Slots in Provincial Level	0.11	0.09
Ratio of Slots in National Level	0.01	0.04
Average Duration on Prefectural Level	7.28	6.29
Average Duration on Provincial Level	7.30	7.36
Average Duration on National Level	10.04	9.50