



Discussion Papers In Economics And Business

The Persistence of Differences in Productivity,
Wages, Skill Mixes and Profits Between Firms

Katsuya Takii

Discussion Paper 07-10

Graduate School of Economics and
Osaka School of International Public Policy (OSIPP)
Osaka University, Toyonaka, Osaka 560-0043, JAPAN

The Persistence of Differences in Productivity, Wages, Skill Mixes and Profits Between Firms

Katsuya Takii

Discussion Paper 07-10

April 2007

この研究は「大学院経済学研究科・経済学部記念事業」
基金より援助を受けた、記して感謝する。

Graduate School of Economics and
Osaka School of International Public Policy (OSIPP)
Osaka University, Toyonaka, Osaka 560-0043, JAPAN

The Persistence of Differences in Productivity, Wages, Skill Mixes and Profits Between Firms*

Katsuya Takii[†]

April

Abstract

In this paper, we construct a dynamic assignment model that can provide a unified explanation of several observed features of persistent differences in productivity, wages, skill mixes and profits between firms. Large organization capital (high firm-specific knowledge) attracts skilled workers, who can create further organization capital in the future. This positive feedback brings about persistent differences in these variables. We also analyze how the real and perceived values of a firm's organization capital interactively influence persistence. We estimate parameters and simulate the model. Our results show that a positive assignment mechanism accounts for a large part of the observed persistence.

JEL Code: J24 L25

Key Words: Organization Capital, Assignment, Persistence

*I would like to thank R. Anton Braun, Julen Esteban-Pretel, Yuichi Fukuta, Shigeru Iwata, Charles Yuji Horioka, Hidehiko Ichimura, Boyan Jovanovic, Akihisa Shibata, Futoshi Yamauchi and participants at numerous seminars. Financial support from the Murata Science Foundation and the Japanese Ministry of Education, Culture, Sports, Science and Technology (a Grant-in-Aid for Scientific Research) are gratefully acknowledged.

[†]Address: Osaka School of International Public Policy, Osaka University, 1-31, Machikameyama, Toyonaka, Osaka, 560-0043, Japan. E-mail address: takii@osipp.osaka-u.ac.jp

1 Introduction

Why are some firms persistently more productive than others? Evidence repeatedly reveals that there are substantial and persistent differences in productivity between plants and between firms [e.g., Baily et al. (1992)]. Apparently, productivity is not the only variable that exhibits persistent differences. Evidence also shows that skill compositions and wage payments exhibit persistent differences between firms [e.g., Haltiwanger et al. (2000)]. Moreover, persistent differences in profits are pervasive [e.g., McGahan (1999)].

The coexistence of persistent differences in these variables is not coincidental. Productive firms employ skilled workers and pay high wages [e.g., Haltiwanger et al. (1999)]. In addition, skills and the market value of a firm are positively correlated [Abowd et al. (2004)]. Evidence implies that the persistence of differences in productivity, skills, wages and profits may have the same source.

Although economists pay it little attention, there is another well-known aspect of productivity differences. In Table 1, future relative productivity is regressed on present and past relative productivity; relative productivity is estimated by the logarithm of labor productivity relative to an industry and year average. The table shows that after controlling for current relative productivity, the sequence of past relative productivity levels still influences future relative productivity. This means that even a previously lower ranked firm that suddenly climbs up the ladder would find it difficult to maintain its success in the future. That is, one-off large temporal shocks are less likely to account for continuous success. To predict the movement of future relative productivity, we need to know the history of relative productivity. However, why does current relative productivity fail to summarize an influence from its past? A useful theory of productivity differences must explain not only why productivity differences persist between firms, but also why a firm's history matters for future persistence.

In this paper, we aim to provide a unified explanation for these findings. We

The Dependent variable is $D \ln y_{t+1}$.

constant	$D \ln y_t$	$D \ln y_{t-1}$	$D \ln y_{t-2}$	$D \ln y_{t-3}$
-0.001	0.671**	0.134**	0.060**	0.048**
(0.002)	(0.007)	(0.009)	(0.009)	(0.007)
# of observations	21154		Adjusted-R ²	0.763

Table 1: AR 4

The variable, $D \ln y_t$, is the logarithm of labor productivity relative to an industry average in year t . The variables are constructed from an industry annual dataset in COMPUSTAT for the period from 1975 to 2004. Data construction is explained in Appendix 1. Standard errors are in parentheses. ** means significant at the 0.5 percent level.

develop a dynamic assignment model for the relationship between the skills of workers and firm-specific knowledge, which we term a firm's organization capital.¹ We also examine how the assignment model explains the observed evidence.

Two assumptions are emphasized: organization capital and skills are complementary to each other; and skill is an input for the accumulation of organization capital. Because skill is complementary to organization capital, skilled workers are assigned to firms with large organization capital. On the other hand, because skill is an input for the accumulation of organization capital, the employment of skillful workers promotes the accumulation of organization capital. Hence, the dynamics exhibit positive feedback. We investigate the extent to which this positive feedback mechanism raises the persistence of a firm's organization capital. Given that the productivity, wages, skill composition and profits of a firm are shown to be strictly increasing functions of organization capital, the model can explain not only the persistence of these variables, but can also explain positive correlations between productivity, skills, wages

¹More specifically, we define organization capital as all types of intangible assets embodied in an organization. It might consist of organizational structure, daily practices, routines, information held by an organization, corporate culture, reputation and so on.

and profits.

The theory predicts that a rise in the heterogeneity of skills increases the persistence of organization capital. When the variance of skills is high, the top organization has the most advantages because it can attract the best workers who can provide the firm with the best knowledge and promote the accumulation of organization capital. Hence, the larger is the variance of skill, the longer the top organization can enjoy its relative advantage. It is shown that when there are no idiosyncratic shocks, every firm's rank stays the same forever and firms' relative advantages (and disadvantages) persist indefinitely.

Although this simple logic can explain why productivity differences between firms persist and why there are positive correlations between productivity, wages, skills and profits, it cannot explain why current relative productivity fails to summarize the influence of the past. In order to explain this evidence, we need a third assumption: organization capital is unobservable. When organization capital is unobservable, there is assignment between the quality of workers and perceived organization capital. If a firm's organization capital is believed to be high, this belief attracts skillful workers, who provide the firm with better knowledge. Because the current belief is constructed from past observations, historical observations can influence future organization capital by changing perceptions about its current level. This mechanism may explain why a firm's history is important.

The model of unobserved organization capital has further theoretical and empirical advantages. Theoretically, this extension allows us to analyze the dynamic interaction between a firm's real capacity and its perceived capacity. On the one hand, when organization capital is believed to be high, this belief attracts skilled workers who can develop superior knowledge in the firm. On the other hand, if actual organization capital is high, the firm's current performance is likely to be good, which generates the perception that the firm has a high level of organization capital. Hence, we can analyze how this dynamic positive feedback between actual capacity

and perceived capacity influences the persistence of productivity, skill mixes, wage payments and profits.

In particular, we examine how the noisiness of information influences persistence. Suppose that a firm’s current output reveals information about its organization capital. If the revealed information is noisy, managers learn little from the new observations, and thus do not change their beliefs drastically. If there is assignment between these beliefs and the quality of workers, the quality of assigned workers changes little and, therefore, so does accumulated actual organization capital. In particular, when output has no predictive power for organization capital, the belief never changes. In that case, we show that the firm’s rank remains the same on average forever and that the dynamics of actual organization capital exhibit temporal deviations from the constant belief.

Extending the model to incorporate unobserved organization capital has another advantage. It allows us to structurally estimate the parameters of our model from productivity dynamics. Exploiting this additional advantage, we differentiate two sources of persistence – positive assortative assignment and noisy information – from others by using an industry annual dataset from COMPUSTAT covering 1970 to 2004.

The estimated parameters are all significant and their signs are consistent with theoretical predictions. In particular, our empirical results show that two-year lagged relative wages have a positive impact on current relative productivity after controlling for one- and two-year lags of relative productivity; note that, in this paper, “relative” refers to the logarithm of each value relative to industry and year averages. This evidence is consistent with the hypothesis that skilled workers help firms to accumulate assets and raise future productivity. Our empirical results also show that perceived relative productivity, which is constructed by using sequences of past relative productivity, has a positive impact on future relative productivity. This is consistent with the hypothesis that people learn about a firm’s capacity from its past perfor-

mance and form beliefs that influence its employment of skilled workers and future performance.

By using the estimated parameters, we simulate not only the autocorrelations of relative productivity, relative wages and expected relative profits per worker, but also the correlation between relative productivity and relative wages. All simulated autocorrelations replicate the observed autocorrelations quite well. The model is also able to explain the observed high correlations between relative productivity and relative wages. That is, our model can quantitatively account for not only the persistence of these variables, but also the high positive correlation between relative productivity and relative wages.

We use our model to conduct two counterfactual experiments. They show that if there were no skill difference between workers and, therefore, if there were no assignment problem, firms' relative advantages (disadvantages) would almost disappear in about five years. In addition, the correlation between relative productivity and relative wages would diminish substantially, while even if output perfectly predicted the level of organization capital, there is only a minor influence on variables' persistence and the correlation between relative productivity and relative wages. These exercises consistently suggest that a positive assignment mechanism accounts for a large part of the observed persistence of variables. The difficulty of estimating organization capital plays only an auxiliary role.

It has long been recognized that an individual firm possesses particular resources [e.g., Kaldor (1934), Robinson (1934) and Lucas (1978)]. As a source of its specific resources, many economists emphasize the importance of firm-specific knowledge accumulated through experience [e.g., Penrose (1959) and Rosen (1972)]. Prescott and Visscher (1980) refer to this accumulated specific knowledge as a firm's organization capital. Recently, interest in organization capital has reemerged. Jovanovic and Rousseau (2001), Atkeson and Kehoe (2005) and Samaniego (2006) quantify the macroeconomic effects of organization capital. Faria (2003) explains merger waves

by using a model of assignment between organization capital and skills. However, no paper has addressed the question of why some firms succeed in accumulating organization capital, whereas others do not. This is the main aim of this paper.

Unlike previous researchers, we model organization capital as a form of the vintage human capital analyzed by Chari and Hopenhayn (1991). For any organization, ancestors determine a particular routine, culture, organizational structure, set of rules and how to arrange machines and structures that successors inherit and modify. Hence, the workers employed in the past influence the organization's future. This modeling strategy allows us to investigate how the assignment of workers to organizations has long-run effects on organization.

Positive assortative assignment models also have a long history. Becker (1973) originally derived a condition for positive assortative matching in a marital market. Sattinger (1979) analyzed a positive assortative assignment equilibrium between physical capital and skills. More recently, economists have rediscovered the importance of assignment models [e.g., Kremer (1993), Costrell and Loury (2004) and Shimer (2005)]. However, in most papers, the distribution of assigned variables is treated as given.

Notable exceptions are Acemoglu (1997) and Jovanovic (1998). Acemoglu (1997) endogenizes the distribution of skills and physical capital and Jovanovic (1998) endogenizes the distribution of skills and technology. Both authors examine persistent income inequality. Unlike them, we endogenize the distribution of organization capital and examine persistent differences in productivity, skills, wages and profits.

A key assumption behind positive assortative matching is that organization capital and skills are complementary. There is supporting evidence for this assumption. Chandler (1977) reports historical evidence that the development of administrative hierarchies is essential for monitoring and coordinating resources in modern business firms. Chandler (1977) demonstrates that this organizational structure demands skilled workers for processing information. Recent evidence obtained by Bresnahan,

Brynjolfsson and Hitt (2002) shows that reorganization associated with IT investment demands more skilled workers. For example, the use of flexible machinery often requires workers to have greater discretion, which in turn requires them to have data analysis skills and problem solving abilities.

Learning is another important feature of the model. As Jovanovic (1982) explains, a firm gradually learns its own productive capacity. However, unlike Jovanovic (1982), we assume that a firm's productive capacity is a product of its active investment and is affected by uncertainty arising from that investment. Ericson and Pakes (1995) incorporate similar dynamics for productive capacity. Hence, our model can be seen as a hybrid of the passive learning model of Jovanovic (1982) and the active investment model of Ericson and Pakes (1995).

Several models generate an equilibrium distribution of wages [e.g., Burdett and Mortensen (1998)] and an equilibrium distribution of productivity [e.g., Eeckhout and Jovanovic (2002)]. These researchers show how ex ante homogeneous agents can generate ex post heterogeneity. Unlike them, we assume an ex ante heterogeneity of skills and then generate distributions of organization capital, productivity, wages and profits. An advantage of our approach is that our unique stationary distribution is globally stable. Hence, our distribution is robust to disturbances.

The paper is organized as follows. In the next section, we set up a dynamic positive assortative assignment model under the assumption that organization capital is perfectly observable. In this section, we clarify the mechanism through which skill differences enhance persistence in the model. In Section 3, we extend the model to incorporate the imperfect observation of organization capital. We analyze how perceived capacity and actual capacity have interactive effects on persistence. In Section 4, we discuss the identification of the parameters from the data and report our empirical results. In Section 5, we simulate our model by using the estimated structural parameters. In Section 6, we discuss extensions and conclude the paper.

2 A Dynamic Assignment Model in which Organization Capital is Observable

In this section, it is assumed that organization capital is observable. We establish a positive assortative assignment equilibrium between organization capital and skills. The assumption that organization capital is observable is used to clarify the intuition about how an assignment model can explain persistent differences in productivity, wages, skills and profits. Our standard assignment model is based on that of Sattinger (1979). We extend the model to incorporate dynamics by endogenizing the distribution of organization capital.

The economy is represented by a continuum of workers and firms. The population of both firms and workers is normalized to unity. Each firm has organization capital of k_t^o , and a set of jobs, the total mass of which is also normalized to unity. We assume that the i th job in a firm that has organization capital of k_t^o employs one worker who has quality of q_{it} and produces output of y_{it} according the following production function:

$$y_{it} = A (k_t^o)^\alpha q_{it}^\psi, \quad \alpha > 0, \psi > 0,$$

where A , α and ψ are constant parameters. Because the total mass of jobs is 1, we interpret $\int_0^1 y_{it} di$ as both a firm's total output and its labor productivity.

Assume that the i th job pays competitive wages of $w(\ln q_t)$. As discussed later, the employment decision is made for each job and job supervisors are assumed to maximize the profits made from the job. The profit maximization problem by the managers in the i th job is written as

$$\chi_i(\ln k_t^o) = \arg \max_{\ln q_{it}} \{ \exp [\ln A + \alpha \ln k_t^o + \psi \ln q_{it}] - w(\ln q_{it}) \}, \quad \forall i, \ln k_t^o. \quad (1)$$

We express wages and policy as functions of $\ln k_t^o$ and $\ln q_t$ to simplify the algebra.

Suppose that $\ln k_t^o$ is normally distributed with a mean of μ_{kt} and a standard deviation of σ_{kt} at the date t . Assume also that $\ln q_t$ is normally distributed with a

mean of μ_q and a standard deviation of σ_q at any date. For simplicity, we assume that jobs and workers have reservation values of 0. Because the number of jobs is the same as the number of workers, nobody chooses the outside option and every agent can find a partner. Hence, these assumptions make it possible to focus on the assignment problem.

We focus on the positive assortative equilibrium. This means that the top x percent of $\ln k_t^o$ is assigned to the top x percent of $\ln q_t$ for any x . Let $\Phi(\cdot)$ denote the standard normal distribution. Given that $\frac{\ln k_t^o - \mu_{kt}}{\sigma_{kt}}$ and $\frac{\ln q_t - \mu_q}{\sigma_q}$ are distributed as standard normal variables, a positive assortative equilibrium implies that

$$1 - \Phi\left(\frac{\ln k_t^o - \mu_{kt}}{\sigma_{kt}}\right) = 1 - \Phi\left(\frac{\chi_i(\ln k_t^o) - \mu_q}{\sigma_q}\right), \quad \forall i, \ln k_t^o. \quad (2)$$

Equations (1) and (2) characterize a static market equilibrium.

Definition 1 *A market equilibrium with observed organization capital consists of $\chi_i(\cdot)$ and $w(\cdot)$ that satisfy equations (1) and (2).*

We aim to find a policy function and a wage function that are consistent with this definition of equilibrium. Equation (2) states that the policy function must satisfy

$$\chi(\ln k_t^o) \equiv \chi_i(\ln k_t^o) = \frac{\sigma_q}{\sigma_{kt}} [\ln k_t^o - \mu_{kt}] + \mu_q, \quad \forall \ln k_t^o.$$

Hence, all jobs in a firm are filled by workers of the same quality. This policy function means that, in equilibrium, highly qualified workers must be assigned to a firm that has a high level of organization capital. For this policy function to be consistent with the definition of equilibrium, the policy function must solve equation (1). Consider a firm that has organization capital of $\ln k_t^o = \frac{\sigma_{kt}}{\sigma_q} [\ln q_t - \mu_q] + \mu_{kt} \equiv \chi^{-1}(\ln q_t)$. For all jobs in this firm, $\ln q_t$ must be the optimal choice. Hence, marginal cost at $\ln q_t$ must be equal to the marginal product of $\ln q_t$, as follows:

$$w'(\ln q_t) = \psi y(\chi^{-1}(\ln q_t), \ln q_t), \quad \forall \ln q_t,$$

where $y(\ln k_t^o, \ln q_t) = \exp(\ln A + \alpha \ln k_t^o + \psi \ln q_t)$. Moreover, because the reservation value of workers is 0, $w(-\infty) = 0$. The following wage function is derived from the marginal condition and the boundary condition:

$$w(\ln q_t) = \frac{\frac{\psi \sigma_q}{\alpha \sigma_{\mu\infty}} y(\chi^{-1}(\ln q_t), \ln q_t)}{1 + \frac{\psi \sigma_q}{\alpha \sigma_{\mu\infty}}}.$$

It is easy to check that the second-order condition is satisfied by this wage function. Hence, the policy function and the wage function are consistent with the definition of equilibrium. By construction, the equilibrium is unique. Note that wage payments increase in $\ln q_t$, which is also an increasing function of $\ln k_t^o$. Hence, a firm that has high organization capital pays high wages.

The firm's profits are strictly increasing in $\ln k_t^o$.

$$\pi(\ln k_t^o) = \int_0^1 [y(\ln k_t^o, \chi_i(\ln k_t^o)) - w(\chi_i(\ln k_t^o))] di = \frac{y(\ln k_t^o, \chi(\ln k_t^o))}{1 + \frac{\psi \sigma_q}{\alpha \sigma_{\mu\infty}}}.$$

Hence, labor productivity, the quality of workers, wage payments and profits per worker are positively correlated, and if $\ln k_t^o$ is persistent these variables are persistent. To determine what influences the persistence of these variables, we analyze the dynamics of $\ln k_t^o$ below.

Dynamics: Following Atkeson and Kehoe (2005) and Samaniego (2006), we assume that organization capital is acquired by learning by doing. In the spirit of Arrow (1962), learning by doing is modeled as an unintended result of production.

More specifically, we assume that an individual worker cannot change a particular routine or culture in a firm, but a group of workers can. Because top managers cannot evaluate the qualities of individual workers, they must rely on evaluation by supervisors in each job. Although supervisors can evaluate the quality of each worker with respect to production in a particular job, they are unaware of how interaction between individual workers can change the firm's routines or culture. We implicitly

assume that communication cannot perfectly resolve this issue. Because skilled workers are likely to learn the mechanism of production well and have better ideas, the employment of skilled workers has indirect external effects that are not initially acknowledged.

We model this process by assuming that the average quality of employed workers improves organization capital in the next period:²

$$k_{t+1}^o = B (k_t^o)^\phi (q_t^e)^\gamma e^{\varepsilon_t}, 0 \leq \phi < 1, \gamma > 0, \quad (3)$$

where $q_t^e = \int_0^1 q_{it} di$, B , ϕ and γ are constant parameters and ε_t is a random variable, which is normally distributed with a mean of $-\frac{\sigma_\varepsilon^2}{2}$ and a standard deviation of σ_ε . The parameter ϕ measures the technological persistence of organization capital. Because some organization capital depreciates, we assume that a fraction, ϕ , of organization capital can be carried over to the next period.

The assumption about learning by doing might be unreasonable if top managers' talents are the most influential inputs for creating organization capital. When a firm employs top managers, it expects them to change the firm's structure and norms. Hence, the firm's maximization problem must also take into account equation (3). In Appendix 2, we assume that a firm solves a dynamic optimization problem by taking into account the dynamics of organization capital. It constructs a recursive positive assortative equilibrium and examines its properties. This shows that the dynamics of organization capital are the same as those obtained when one assumes

²Given that every job is filled by workers of the same quality, equation (3) generates dynamics that are the same as those from the transition equation,

$$k_{t+1}^o = B_1 (k_t^o)^{\phi_1} (y_t)^{\gamma_1} e^{\varepsilon_t},$$

where $y_t = A (k_t^o)^\alpha q_t^\psi$, and B_1 , ψ_1 and γ_1 are parameters. This equation implies that $\ln k_{t+1}^o$ is expressed as a weighted sum of $\{\ln y_{t-s}\}_{s=0}^t$. As discussed by Bahk and Gort (1993), in empirical studies, cumulative gross output is used as a proxy of experience accumulated through learning by doing. Hence, our assumption is consistent with the standard learning-by-doing assumption.

that there is learning by doing. Differences arise in the wage and profit functions.

We maintain the learning-by-doing assumption for three reasons. First, as convincingly argued by Simon (1997) and Nelson and Winter(1982), it is reasonable to assume that an individual in a firm would find it hard to change a firm's routines or culture. Second, the learning-by-doing assumption simplifies the model, but conveys the main logic of the paper. Hence, most of our analysis avoids the technical difficulties associated with dynamic optimization problems. Third, the wage and profit functions derived on the basis of the learning-by-doing assumption are useful for our empirical work. We discuss the unique outcomes generated by the the learning-by-doing assumption later.

Because all jobs in a firm are filled by workers of the same quality, in which case, $\ln q_t = \frac{\sigma_q}{\sigma_{kt}} (\ln k_t^o - \mu_{kt}) + \mu_q$, in equilibrium, the dynamics of organization capital can be written as

$$\ln k_{t+1}^o = \ln B + \phi \ln k_t^o + \gamma \left[\frac{\sigma_q}{\sigma_{kt}} (\ln k_t^o - \mu_{kt}) + \mu_q \right] + \varepsilon_t. \quad (4)$$

Because $\ln k_t^o$ and ε_t are normally distributed, $\ln k_{t+1}^o$ is also normally distributed. The dynamics of μ_{kt} and σ_{kt}^2 can be derived from equation (4) as follows:

$$\mu_{kt+1} = \ln B + \phi \mu_{kt} + \gamma \mu_q - \frac{\sigma_\varepsilon^2}{2}, \quad \sigma_{kt+1} = \sqrt{\left(\phi + \frac{\gamma \sigma_q}{\sigma_{kt}} \right)^2 \sigma_{kt}^2 + \sigma_\varepsilon^2}.$$

These two equations characterize the dynamics of the aggregate state variables. By using the dynamics of μ_{kt} , equation (4) can be rewritten as

$$\ln k_{t+1}^o - \mu_{kt+1} = \left(\phi + \frac{\gamma \sigma_q}{\sigma_{kt}} \right) (\ln k_t^o - \mu_{kt}) + \varepsilon_t^*,$$

where $\varepsilon_t^* = \varepsilon_t + \frac{\sigma_\varepsilon^2}{2}$ is normally distributed with a mean of 0 and a standard deviation of σ_ε . This equation implies that when $\ln k_t^o$ is larger than its industry mean μ_{kt} , the fraction, $\phi + \frac{\gamma \sigma_q}{\sigma_{kt}}$ of this relative advantage is carried over to the next period. The parameter ϕ captures the assumed persistence, and the second term, $\frac{\gamma \sigma_q}{\sigma_{kt}}$, represents the result of positive assortative assignment. When the ratio of the standard deviation of

skills to that of organization capital is large, organization capital is more persistent. When the ratio is large, the firm with the most organization capital derives the most benefits because this leading firm attracts the most talented workers, who provide the firm with the best knowledge. Therefore, relative advantages persist longer.

This intuitive result is analyzed more rigorously below. First, we show that the distribution of $\ln k_t^o$ converges to a stationary distribution.

Proposition 2 *The mean, μ_{kt} , and standard deviation, σ_{kt} , of $\ln k_t^o$ converge to the stationary points, $\mu_{k\infty}$ and $\sigma_{k\infty}$, respectively, where $\mu_{k\infty}$ and $\sigma_{k\infty}$ are*

$$\mu_{k\infty} = \frac{\ln B + \gamma\mu_q - \frac{\sigma_\varepsilon^2}{2}}{1 - \phi}, \quad (5)$$

$$\sigma_{k\infty} = \frac{\gamma\phi\sigma_q + \sqrt{(\gamma\sigma_q)^2 + (1 - \phi^2)\sigma_\varepsilon^2}}{1 - \phi^2}. \quad (6)$$

Moreover, when the distribution converges to the stationary distribution, the dynamics of organization capital follow a first-order autoregressive (AR(1)) process, as follows:

$$D \ln k_{t+1}^o = \left(\phi + \frac{\gamma\sigma_q}{\sigma_{k\infty}} \right) D \ln k_t^o + \varepsilon_t^*, \quad (7)$$

where $D \ln k_t^o = \ln k_t^o - \mu_{k\infty}$.

The proposition states that there is a unique, globally stable stationary distribution. Hence, an economy converges to the stationary distribution in the long run. In the remainder of this section, we discuss the dynamics of equation (7). Because persistence is only influenced by $\phi + \frac{\gamma\sigma_q}{\sigma_{k\infty}}$, we refer to this as the persistence parameter in what follows.

Persistence when $\sigma_\varepsilon = 0$: In order to have a clear intuition, we first analyze a deterministic model.

Proposition 3 *Suppose that $\sigma_\varepsilon = 0$. Then, there is still a stationary distribution, $\mu_{k\infty} = \frac{\ln B + \gamma\mu_q}{1-\phi}$, $\sigma_{k\infty} = \frac{\gamma\sigma_q}{1-\phi}$, and the level of organization capital remains constant: $\ln k_{t+1}^o = \ln k_t^o$.*

The proposition states that, if there is no shock, the ranking of organization capital is constant. When there are no idiosyncratic shocks, the top organization always attracts the best workers, who, in turn, equip the firm with the best knowledge. Hence, it remains at the top and maintains exactly the same level of organization capital in the long run.

Note that $\sigma_{k\infty}$ is not 0. Hence, organization capital varies between firms and so too, therefore, do productivity, wages, skill mixes and profits. Because σ_q is always positive, when σ_{kt} is small, the assignment effect, $\frac{\gamma\sigma_q}{\sigma_{kt}}$, is large. Hence, a firm with a high level of organization capital benefits substantially. This mechanism increases σ_{kt} . Ultimately, $\sigma_{k\infty}$ does not converge to 0.

Persistence when $\sigma_\varepsilon > 0$. When we introduce idiosyncratic shocks into the accumulation of organization capital, reversion to the mean occurs. Idiosyncratic shocks make changes in rankings possible. A firm that receives a positive shock climbs the rankings, which enables it to attract higher quality workers. This means that top organizations cannot remain the best. Top organizations might slip down the rankings.

Substituting equation (6) into the persistence parameter reveals that this parameter is a strictly increasing function of $\frac{\gamma\sigma_q}{\sigma_\varepsilon}$.

$$\phi + \frac{\gamma\sigma_q}{\sigma_{k\infty}} = \phi + \frac{1 - \phi^2}{\phi + \sqrt{1 + (1 - \phi^2) \left(\frac{\gamma\sigma_q}{\sigma_\varepsilon}\right)^{-2}}} \in (\phi, 1), \quad (8)$$

$$\frac{d\left(\phi + \frac{\gamma\sigma_q}{\sigma_{k\infty}}\right)}{d\frac{\gamma\sigma_q}{\sigma_\varepsilon}} > 0, \quad \lim_{\frac{\gamma\sigma_q}{\sigma_\varepsilon} \rightarrow \infty} \left(\phi + \frac{\gamma\sigma_q}{\sigma_{k\infty}}\right) = 1, \quad \lim_{\frac{\gamma\sigma_q}{\sigma_\varepsilon} \rightarrow 0} \left(\phi + \frac{\gamma\sigma_q}{\sigma_{k\infty}}\right) = \phi.$$

When $\sigma_\varepsilon > 0$, $\frac{\gamma\sigma_q}{\sigma_\varepsilon}$ is finite. Therefore, the persistence parameter is always less than 1. This means that equation (7) is covariance stationary. Hence, the dynamics exhibit reversion to the mean and eventually initial advantages disappear. In this case, it is easy to show that

$$\rho_{\ln k_j} \equiv \frac{E [D \ln k_t^o D \ln k_{t-j}^o]}{\sigma_{k_\infty}^2} = \left(\phi + \frac{\gamma\sigma_q}{\sigma_{k_\infty}} \right)^j.$$

That is, the larger the persistence parameter, the larger is the autocorrelation. Hence, an increase in $\frac{\gamma\sigma_q}{\sigma_\varepsilon}$ raises the autocorrelation.

Two remarks are worth making. First, the parameter $\frac{\gamma\sigma_q}{\sigma_\varepsilon}$ represents the importance of skill variation relative to the reshuffling effect. Hence, the importance of skill variation to persistence is not absolute. Small skill variations can induce substantial persistence in variables when an industry is stable. Second, when $\frac{\gamma\sigma_q}{\sigma_\varepsilon}$ is infinite, the persistence parameter converges to unity. Hence, the results in Proposition 3 can be seen as the limit of this general case.

3 A Dynamic Assignment Model in which Organization Capital is not Observable

In the previous section, we provided a clear intuition about how assignment influences the persistence of variables and explained why productivity, skills, wages and profits are positively correlated. However, our analysis did not explain how historical relative productivity influences future relative productivity. Because the dynamics of organization capital in the previous section were represented by an AR(1) process, once we control for a current value, past values do not influence future ones.

In this section, we examine the same dynamic assignment model but incorporate organization capital that is not observable. We show that past organization capital can influence future organization capital even when current organization capital is controlled for. We also analyze how perceived organization capital and actual

organization capital have interactive effects on the persistence of capital.

Assume that k_t^o cannot be directly observed, but can be inferred from the realizations of output. To capture this idea, we add a firm-specific idiosyncratic shock to the production function for each job as follows:

$$y_{it} = e^{u_t} A (k_t^o)^\alpha q_{it}^\psi, \quad \alpha > 0, \psi > 0, \quad (9)$$

where u_t is normally distributed with a mean of $-\frac{\sigma_u^2}{2}$ and a variance of σ_u^2 .

When employment decisions are made about the i th job, output is not realized. Hence, a decision must be based on a conditional expectation given the prior belief about the level of organization capital. We assume that the prior distribution of $\ln k_t$ is normally distributed with a mean of μ_{kt} and a variance of σ_{kt}^2 . Then the expected output from the job is

$$E[y_{it} | \mu_{kt}, \ln q_{it}] = \exp\left(\ln A + \alpha \mu_{kt} + \frac{\alpha^2 \sigma_{kt}^2}{2} + \psi \ln q_{it}\right). \quad (10)$$

Similarly to the previous section, all firms are assumed to have the same σ_{kt} at date t . However, we deviate from the previous section by assuming that the belief, μ_{kt} , differs between firms. Given that all agents in an economy receive the same information, these agents hold the same beliefs about a firm's organization capital. That is, the belief, μ_{kt} , characterizes a firm's position in the economy. It is assumed that the belief, μ_{kt} , is normally distributed with a mean of μ_{kt}^e and a standard deviation of $\sigma_{\mu t}$. We examine a positive assortative assignment equilibrium between a belief, μ_{kt} , and a skill, $\ln q_t$.

Similarly to the previous problem, by assuming that the wage is a function of $\ln q_t$, $\hat{w}(\ln q_t)$, we can define a market equilibrium as follows.

Definition 4 *A market equilibrium with unobserved organization capital consists of a $\hat{\chi}_i(\cdot)$ and a $\hat{w}(\cdot)$ that satisfy the following two equations:*

$$\hat{\chi}_i(\mu_{kt}) = \arg \max_{\ln q_{it}} \{E[y_{it} | \mu_{kt}, \ln q_{it}] - \hat{w}(\ln q_{it})\}, \quad \forall i, \mu_{kt}. \quad (11)$$

$$1 - \Phi\left(\frac{\mu_{kt} - \mu_{kt}^e}{\sigma_{\mu t}}\right) = 1 - \Phi\left(\frac{\hat{\chi}_i(\mu_{kt}) - \mu_q}{\sigma_q}\right), \quad \forall i, \mu_{kt}. \quad (12)$$

Applying the same logic used in the previous section, we can derive a policy function and a wage function that are consistent with this definition of equilibrium. Because the derivation is as before, the proof of the following theorem is omitted.

Theorem 5 *There exists a unique equilibrium, which is characterized by the following functions:*

$$\hat{\chi}(\mu_{kt}) \equiv \hat{\chi}_i(\mu_{kt}) = \frac{\sigma_q}{\sigma_{\mu t}}(\mu_{kt} - \mu_{kt}^e) + \mu_q, \quad (13)$$

$$\hat{w}(\ln q_t) = \frac{\frac{\psi\sigma_q}{\alpha\sigma_{\mu\infty}} E[y_t | \hat{\chi}^{-1}(\ln q_t), \ln q_t]}{1 + \frac{\psi\sigma_q}{\alpha\sigma_{\mu\infty}}}, \quad (14)$$

where $E[y_t | \mu_{kt}, \ln q_t] = \exp\left(\log A + \alpha\mu_{kt} + \frac{\alpha^2\sigma_{kt}^2}{2} + \psi \ln q_t\right)$ and $\hat{\chi}^{-1}(\ln q_t) \equiv \frac{\sigma_{\mu t}}{\sigma_q}(\ln q_t - \mu_q) + \mu_{kt}^e$. Moreover, the expected profits of the firm are

$$\hat{\pi}^e(\mu_{kt}) = \frac{E[y_t | \mu_{kt}, \hat{\chi}(\mu_{kt})]}{1 + \frac{\psi\sigma_q}{\alpha\sigma_{\mu\infty}}}. \quad (15)$$

This theorem shows that the levels of skill and expected profits are strictly increasing functions of μ_{kt} and that the wage function is a strictly increasing function of $\ln q_t$. Hence, the dynamics for skills, wages and expected profits follow the dynamics of μ_{kt} . On the other hand, labor productivity, $\ln y_t$, is strictly increasing in $\ln k_t^o$ and $\ln q_t$. Hence, the dynamics of labor productivity are influenced by the dynamics of $\ln k_t^o$ and μ_{kt} . To understand the dynamics of productivity, wages, skills and profits, we analyze the dynamics of $\ln k_t^o$ and μ_{kt} below.

Dynamics: Similarly to the analysis of the previous section, we can derive the dynamics of $\ln k_t^o$ by substituting equation (13) into equation (3) as follows:

$$\ln k_{t+1}^o = \ln B + \phi \ln k_t^o + \gamma \left[\frac{\sigma_q}{\sigma_{\mu t}}(\mu_{kt} - \mu_{kt}^e) + \mu_q \right] + \varepsilon_t. \quad (16)$$

To derive the dynamics of μ_{kt} , we must describe the information structure of the model. After the firm employs a worker, output is produced. From the realized output, the firm knows $e^{u_t} (k_t^o)^\alpha$. Hence, a firm uses a signal, $s_t \equiv \ln k_t^o + u_t^*$, to infer $\ln k_t^o$, where $u_t^* = \frac{1}{\alpha} \left(u_t + \frac{\sigma_u^2}{2} \right)$ is normally distributed with a mean of 0 and a standard deviation of $\frac{\sigma_u}{\alpha}$. Because $\mu_{kt+1} = E [\ln k_{t+1}^o | s_t, \mu_{kt}, \sigma_{kt}]$ and $\sigma_{kt+1} = \sqrt{\text{Var} [\ln k_{t+1}^o | s_t, \mu_{kt}, \sigma_{kt}]}$, the dynamics of μ_{kt} and σ_{kt} can be written as follows:

$$\mu_{kt+1} = \ln B + \phi E [\ln k_t^o | s_t, \mu_{kt}, \sigma_{kt}] + \gamma \left[\frac{\sigma_q}{\sigma_{\mu t}} (\mu_{kt} - \mu_{kt}^e) + \mu_q \right] - \frac{\sigma_\varepsilon^2}{2}, \quad (17)$$

$$\sigma_{kt+1} = \sqrt{\phi^2 (1 - h_t) \sigma_{kt}^2 + \sigma_\varepsilon^2}, \quad (18)$$

where

$$E [\ln k_t^o | s_t, \mu_{kt}, \sigma_{kt}] = (1 - h_t) \mu_{kt} + h_t s_t = (1 - h_t) \mu_{kt} + h_t (\ln k_t^o + u_t^*), \quad (19)$$

$$h_t = \frac{\left(\frac{\alpha \sigma_{kt}}{\sigma_u} \right)^2}{1 + \left(\frac{\alpha \sigma_{kt}}{\sigma_u} \right)^2}. \quad (20)$$

Equation (19) shows that $E [\ln k_t^o | s_t, \mu_{kt}, \sigma_{kt}]$ is a weighted average of the prior belief, μ_{kt} , and new information s_t , where the variable h_t is the weight on new information. As shown in equation (20), h_t is negatively related to σ_u . If the variance of temporal shocks is large, it is difficult to infer $\ln k_t^o$ from s_t and thus place a small weight on s_t . In this way, the variable h_t measures the reliability of new information.³ Because σ_{kt} is the same in all firms, equation (18) shows that σ_{kt+1} is also the same in all firms. Similarly, because μ_{kt} and s_t are normally distributed, equation (17) shows that μ_{kt+1} is also normally distributed. Hence, the normality of the distribution

³In fact, h_t can be also rewritten as follows:

$$h_t = 1 - \frac{E [\text{Var} [\ln k_t^o | s_t, \mu_{kt}, \sigma_{kt}]]}{\sigma_{kt}^2}.$$

This equation shows that h_t would be larger if the average conditional variance were smaller relative to the prior variance. It measures the accuracy of information, as previously used by Takii (2003, 2007), as a tractable measure of prediction ability.

is preserved. The following mean and standard deviation of the belief in the next period can be derived:

$$\mu_{kt+1}^e = \ln B + \phi\mu_{kt}^e + \gamma\mu_q - \frac{\sigma_\varepsilon^2}{2}, \quad (21)$$

$$\sigma_{\mu t+1} = \sqrt{\left(\phi + \frac{\gamma\sigma_q}{\sigma_{\mu t}}\right)^2 \sigma_{\mu t}^2 + \phi^2 h_t \sigma_{kt}^2}. \quad (22)$$

The derivation of equation (22) is given in Appendix 1. Furthermore, by substituting equations (19) and (21) into equations (16) and (17), we can also rewrite the dynamics of $\ln k_t^o$ and μ_{kt} as follows:

$$\ln k_{t+1}^o - \mu_{kt+1}^e = \phi(\ln k_t^o - \mu_{kt}^e) + \frac{\gamma\sigma_q}{\sigma_{\mu t}}(\mu_{kt} - \mu_{kt}^e) + \varepsilon_t^*, \quad (23)$$

$$\mu_{kt+1} - \mu_{kt+1}^e = \phi h_t(\ln k_t^o - \mu_{kt}^e) + \left[\phi(1 - h_t) + \frac{\gamma\sigma_q}{\sigma_{\mu t}}\right](\mu_{kt} - \mu_{kt}^e) + \phi h_t u_t^* \quad (24)$$

where $\varepsilon_t^* = \varepsilon_t + \frac{\sigma_\varepsilon^2}{2}$ is normally distributed with a mean of 0 and a standard deviation of σ_ε .

Equation (23) shows the dynamics of $\ln k_t^o$. The first term of equation (23) is influenced by technological persistence, ϕ . That is, if organization capital is above average, the fraction ϕ of this relative advantage is carried over to the next period. On the other hand, the second term is influenced by positive assignment. If organization capital is believed to be above average, the firm attracts skilled workers that help the firm accumulate further organization capital.

Equation (24) shows the dynamics of μ_{kt} . The first term captures how new information influences the dynamics of the belief. Managers know that the fraction ϕ of current organization capital affects the next period's organization capital. However, current organization capital is not observable and must be inferred from current output. High output can be the result of either a large temporal shock or a high level of organization capital. Because managers put a weight h_t on new information, the fraction ϕh_t of current organization capital is believed to be translated into the next period's level. New information incorporates noise. Hence, the ϕh_t portion of u_t^*

also influences the posterior belief. This effect is captured by the third term, $\phi h_t u_t^*$, in equation (24).

The second term of equation (24) captures the effect of the prior belief on the posterior belief. There are two separate effects. Because there is assignment between the prior belief and worker quality, the higher the level of organization capital is believed to be, *a priori*, the higher is the quality of workers that the firm can employ. Given that skilled workers help the firm to accumulate organization capital, organization capital in the next period is believed to be high. This assignment effect is captured by $\frac{\gamma \sigma_q}{\sigma_{\mu t}}$ in the second term. On the other hand, because output provides only noisy information about organization capital, a weight of $1 - h_t$ is placed on the prior belief. Because the fraction ϕ of current organization capital is translated into organization capital for the next period, the fraction $\phi(1 - h_t)$ of the prior belief influences the posterior. Overall, the fraction $\phi(1 - h_t) + \frac{\gamma \sigma_q}{\sigma_{\mu t}}$ of the prior belief influences the posterior.

The equations (23) and (24) provide some intuition about the dynamics of $\ln k_t^o$ and μ_{kt} . First, the interpretation of equation (23) is that $\ln k_t^o$ exhibits reversion to the belief μ_{kt} and the speed of the reversion is influenced by the constant parameter ϕ . Hence, assignment does not influence the persistence of $\ln k_t^o$ unless it affects μ_{kt} . Second, given equation (24), the smaller is h_t , the less is μ_{kt} subjected to two types of shock, ε_t^* and u_t^* . Hence, the ranking of μ_{kt} is less likely to change. Because there is assignment between μ_{kt} and $\ln q_t$, fewer changes in the rankings imply that μ_{kt} becomes more persistent. That is, the noisier is the information, the more persistent is the belief.

More importantly, equation (24) implies that the current belief, μ_{kt} , can be influenced by sequences of real organization capital, $\{\ln k_s^o\}_{s=0}^{t-1}$. Because current beliefs influence the next period's actual levels of organization capital according to equation (23), past sequences of actual organization capital can influence future values. It is shown later that this mechanism can explain why higher order lags matter for

productivity dynamics.

To confirm these arguments, we first show that this economy converges to the stationary distribution. Then, we analyze the dynamics of organization capital in an aggregate economy that reaches the stationary distribution.

Proposition 6 *The aggregate economy converges to a unique stationary distribution.*

$$\begin{aligned}\mu_\infty^e &= \frac{\ln B + \gamma\mu_q - \frac{\sigma_\varepsilon^2}{2}}{1 - \phi} \\ \sigma_{\mu_\infty} &= \frac{\phi\gamma\sigma_q + \sqrt{(\gamma\sigma_q)^2 + (1 - \phi^2)\phi^2 h_\infty \sigma_{k_\infty}^2}}{(1 - \phi^2)} \\ \sigma_{k_\infty}^2 &= \frac{\alpha^2 \sigma_\varepsilon^2 - (1 - \phi^2) \sigma_u^2 + \sqrt{[\alpha^2 \sigma_\varepsilon^2 - (1 - \phi^2) \sigma_u^2]^2 + 4\alpha^2 \sigma_\varepsilon^2 \sigma_u^2}}{2\alpha^2} \\ h_\infty &= \frac{\left(\frac{\alpha\sigma_{k_\infty}}{\sigma_u^2}\right)^2}{1 + \left(\frac{\alpha\sigma_{k_\infty}}{\sigma_u^2}\right)^2}\end{aligned}$$

Moreover, the dynamics of an individual firm in the stationary distribution are described by the following vector autoregression (VAR):

$$\mathbf{k}_{t+1} = \mathbf{M}\mathbf{k}_t + \boldsymbol{\xi}_t, \quad (25)$$

where

$$\mathbf{M} = \begin{bmatrix} \phi, & \frac{\gamma\sigma_q}{\sigma_{\mu_\infty}} \\ \phi h_\infty, & \phi(1 - h_\infty) + \frac{\gamma\sigma_q}{\sigma_{\mu_\infty}} \end{bmatrix}, \quad \mathbf{k}_t = \begin{bmatrix} D \ln k_t^o \\ D\mu_{kt} \end{bmatrix}, \quad \boldsymbol{\xi}_t = \begin{bmatrix} \varepsilon_t^* \\ \phi h_\infty u_t^* \end{bmatrix}$$

and $D \ln k_t^o = \ln k_t^o - \mu_{k_\infty}^e$ and $D\mu_{kt} = \mu_{kt} - \mu_{k_\infty}^e$.

Because the stationary distribution is unique and globally stable, the economy converges to the stationary distribution in the long run. We investigate the properties of equation (25) and discuss what influences the persistence of organization capital.

Persistence when $\sigma_u = \infty$: It is instructive to start with an extreme case, in which $\sigma_u = \infty$. In this case, information is too noisy and the firm can learn nothing about the level of organization capital. Hence, $h_\infty = 0$ and the firm only puts weight on its prior belief. Therefore, its belief never changes. By substituting $\sigma_u = \infty$ and $h_\infty = 0$ into equations (18), (22), (23) and (24), the following proposition is easily derived. It shows that the movement of organization capital reverts to this constant belief.

Proposition 7 *Suppose that $\sigma_u^2 = \infty$. Then, the dynamics of the economy are given by*

$$\mu_\infty^e = \frac{\ln B + \gamma\mu_q - \frac{\sigma_\varepsilon^2}{2}}{1 - \phi}, \quad \sigma_{\mu\infty} = \frac{\gamma\sigma_q}{1 - \phi}, \quad \sigma_{k\infty}^2 = \frac{\sigma_\varepsilon^2}{1 - \phi^2}$$

and

$$\ln k_{t+1}^o = \phi \ln k_t^o + (1 - \phi)\mu_{kt} + \varepsilon_t^*, \quad \mu_{kt+1} = \mu_{kt}.$$

Because the firm cannot learn about its own organization capital, the firm never changes its own belief. Hence, the belief is constant. Because assignment is based on this belief, the firm that is believed to have a high level of organization capital attracts good workers and maintains its position. Moreover, as actual organization capital is subjected to shocks, the movement of organization capital temporally deviates from the firm's own belief. However, the level of organization capital remains the same on average because of the constant belief.

Persistence when $\sigma_u \in (0, \infty)$, $\sigma_\varepsilon \in (0, \infty)$ and $\phi \in (0, 1)$. Let us examine a more general case. First, we analyze the stability of equation (25). Then, we analyze what influences persistence. The following lemma provides the conditions for stability.

Lemma 8 *Let λ_1 and λ_2 denote the eigenvalues of the matrix \mathbf{M} . Then, equation (25) is covariance stationary if $\lambda_1 = \phi + \frac{\gamma\sigma_q}{\sigma_{\mu\infty}} < 1$ and $\lambda_2 = \phi(1 - h_\infty) < 1$.*

Note that λ_2 is less than unity. This means that stability is guaranteed if $\lambda_1 < 1$. The eigenvalue λ_1 corresponds to the persistence parameter of the previous section. It consists of assumed persistence, ϕ , and the assignment effect, $\frac{\gamma\sigma_q}{\sigma_{\mu\infty}}$. Hence, similarly to the analysis of the previous section, if the assignment effect is not too strong, the dynamics eventually converge to the mean.

Clearly, $\frac{\gamma\sigma_q}{\sigma_{\mu\infty}}$ and h_∞ are endogenous variables. Thus, there are more fundamental conditions for stability. The following lemma describes the relationship between the endogenous variables, h_∞ and $\frac{\gamma\sigma_q}{\sigma_{\mu\infty}}$, and the exogenous variables, $\frac{\sigma_u}{\alpha\sigma_\varepsilon}$ and $\frac{\gamma\sigma_q}{\sigma_\varepsilon}$.

Lemma 9 ⁴ Suppose that $\phi \in (0, 1)$, $\frac{\sigma_u}{\alpha\sigma_\varepsilon} \in (0, \infty)$ and $\frac{\gamma\sigma_q}{\sigma_\varepsilon} \in (0, \infty)$. There exists a function $\eta(\cdot)$ and $\Sigma(\cdot, \cdot)$ such that

$$h_\infty = \eta\left(\frac{\sigma_u}{\alpha\sigma_\varepsilon}\right) \in (0, 1),$$

where $\eta'\left(\frac{\sigma_u}{\alpha\sigma_\varepsilon}\right) < 0$, $\lim_{\frac{\sigma_u}{\alpha\sigma_\varepsilon} \rightarrow 0} \eta\left(\frac{\sigma_u}{\alpha\sigma_\varepsilon}\right) = 1$ and $\lim_{\frac{\sigma_u}{\alpha\sigma_\varepsilon} \rightarrow \infty} \eta\left(\frac{\sigma_u}{\alpha\sigma_\varepsilon}\right) = 0$, and

$$\frac{\gamma\sigma_q}{\sigma_{\mu\infty}} = \Sigma\left(\frac{\gamma\sigma_q}{\sigma_\varepsilon}, h_\infty\right) \in (0, 1 - \phi), \quad (29)$$

where $\Sigma_1\left(\frac{\gamma\sigma_q}{\sigma_\varepsilon}, h_\infty\right) > 0$, $\Sigma_2\left(\frac{\gamma\sigma_q}{\sigma_\varepsilon}, h_\infty\right) < 0$, $\lim_{\frac{\gamma\sigma_q}{\sigma_\varepsilon} \rightarrow 0} \Sigma\left(\frac{\gamma\sigma_q}{\sigma_\varepsilon}, h_\infty\right) = 0$, $\lim_{\frac{\gamma\sigma_q}{\sigma_\varepsilon} \rightarrow \infty} \Sigma\left(\frac{\gamma\sigma_q}{\sigma_\varepsilon}, h_\infty\right) = 1 - \phi$, $\lim_{h_\infty \rightarrow 1} \Sigma\left(\frac{\gamma\sigma_q}{\sigma_\varepsilon}, h_\infty\right) \in (0, 1 - \phi)$ and $\lim_{h_\infty \rightarrow 0} \Sigma\left(\frac{\gamma\sigma_q}{\sigma_\varepsilon}, h_\infty\right) = 1 - \phi$.

⁴Explicit solutions for $\eta(\cdot)$ and $\Sigma(\cdot, \cdot)$ in the Lemma 9 can be derived. Lemma 9 describes results based on the following explicit solutions. Technical parts of the derivation are in Appendix 1.

$$\eta\left(\frac{\sigma_u}{\alpha\sigma_\varepsilon}\right) = \frac{\left(\frac{\alpha\sigma_{k\infty}}{\sigma_u}\right)^2}{1 + \left(\frac{\alpha\sigma_{k\infty}}{\sigma_u}\right)^2}, \text{ where} \quad (26)$$

$$\left(\frac{\alpha\sigma_{k\infty}}{\sigma_u}\right)^2 = \frac{\left(\frac{\sigma_u}{\alpha\sigma_\varepsilon}\right)^{-2} - (1 - \phi^2) + \sqrt{\left[\left(\frac{\sigma_u}{\alpha\sigma_\varepsilon}\right)^{-2} - (1 - \phi^2)\right]^2 + 4\left(\frac{\sigma_u}{\alpha\sigma_\varepsilon}\right)^{-2}}}{2}, \quad (27)$$

and

$$\Sigma\left(\frac{\gamma\sigma_q}{\sigma_\varepsilon}, h_\infty\right) = \frac{(1 - \phi^2)}{\phi + \sqrt{1 + \frac{(1 - \phi^2)\phi^2 h_\infty}{1 - \phi^2 + \phi^2 h_\infty} \left(\frac{\gamma\sigma_q}{\sigma_\varepsilon}\right)^{-2}}}. \quad (28)$$

The first part of this lemma shows that h_∞ and $\frac{\sigma_u}{\alpha\sigma_\varepsilon}$ have a one-to-one relationship. Hence, in the steady state, without loss of generality, h_∞ can be treated as an exogenous parameter.

The parameter $\frac{\sigma_u}{\alpha\sigma_\varepsilon}$ represents the standard deviation of noise relative to that of shocks on the accumulation of organization capital. If the standard deviation of a noise term is relatively large, firms cannot learn much and h_∞ is small. If the noise term has a relatively small variance, the firm can learn a lot and h_∞ is large.

The second part of the lemma shows that for a given h_∞ and ϕ , $\frac{\gamma\sigma_q}{\sigma_{\mu\infty}}$ and $\frac{\gamma\sigma_q}{\sigma_\varepsilon}$ exhibit a one-to-one relationship. As explained in the previous section, if $\frac{\gamma\sigma_q}{\sigma_\varepsilon}$ is large, the top organization derives the greatest benefit from a positive assortative assignment. Hence, the dynamics exhibit more persistence.

More interestingly, $\frac{\gamma\sigma_q}{\sigma_{\mu\infty}}$ is decreasing in h_∞ . When information is noisier, rational agents rely more on their prior beliefs to make inferences about the current level of organization capital. Therefore, the posterior and prior beliefs are similar. Hence, μ_{kt} is more persistent. Because the dynamics of actual organization capital exhibit reversion to the firm's own belief, noisy information makes actual organization capital more persistent.

Note that the properties of the function $\Sigma(\cdot, \cdot)$ imply that when h_∞ converges to 0, the persistence parameter, λ_1 , converges to 1. This means that when $\frac{\sigma_u}{\alpha\sigma_\varepsilon}$ is large, the dynamics can be approximated by the previous results, with $\sigma_u = \infty$.

Note also that equation (29) implies that $\lambda_1 = \phi + \frac{\gamma\sigma_q}{\sigma_{\mu\infty}} < 1$. Hence, the following proposition can be stated.

Proposition 10 *Suppose that $\phi \in (0, 1)$, and that $\frac{\sigma_u}{\alpha\sigma_\varepsilon}$ and $\frac{\gamma\sigma_q}{\sigma_\varepsilon}$ are finite. Equation (25) is covariance stationary.*

The two eigenvalues, λ_1 and λ_2 , are important determinants of the persistence of the stochastic process. Let $\rho_{\ln kj}$ denote the autocorrelation between current organization capital and organization capital j periods before. In addition, let $\rho_{\mu j}$

denote the autocorrelation between a belief about current organization capital and a belief about organization capital j periods before: $\rho_{\ln k_j} \equiv \frac{E[D \ln k_t^\circ D \ln k_{t-j}^\circ]}{\text{Var}(D \ln k_t^\circ)}$, $\rho_{\mu_j} \equiv \frac{E[D \mu_{kt} D \mu_{kt-j}]}{\sigma_{\mu_\infty}^2}$. The following proposition derives the autocorrelation of $\ln k_t$ and μ_{kt} . The proof is established in Appendix 1.

Proposition 11 *The autocorrelations of $\ln k_t^\circ$ and μ_{kt} are functions of λ_1 and λ_2*

$$\rho_{\ln k_j} = (1 - \omega) \lambda_1^j + \omega \lambda_2^j, \quad \rho_{\mu_j} = \lambda_1^j,$$

$$\text{where } \omega = \frac{\frac{\gamma \sigma_q}{\sigma_{\mu_\infty}} (1 - \lambda_1^2)}{(\lambda_1 - \lambda_2) (\phi^2 h_\infty + 1 - \lambda_1^2)}.$$

The above proposition states that the autocorrelation of organization capital can be expressed as a weighted average of λ_1^j and λ_2^j ; the autocorrelation of the belief is λ_1^j . Note that $\lambda_1 > \lambda_2$. Hence, the following corollary results from the above proposition.

Corollary 12 *Suppose that $\frac{\gamma \sigma_q}{\sigma_\varepsilon} > 0$. The autocorrelation of the belief about a firm's organization capital exceeds that of its actual organization capital: $\rho_{\mu_j} > \rho_{\ln k_j}$, $\forall j$.*

The corollary states that the belief is more persistent than is the organization capital itself. Because idiosyncratic shocks directly influence the realization of random variables, the variance of the realized random variable is generally larger than the variance of the conditional expectation. The same logic applies in this case. Given that the belief is less volatile than is actual organization capital, the autocorrelation of the belief exceeds the actual value. This corollary confirms this intuition.

Next, we show how the parameters $\frac{\gamma \sigma_q}{\sigma_\varepsilon}$ and h_∞ (or $\frac{\sigma_u}{\alpha \sigma_\varepsilon}$) affect the autocorrelations. The proof is established in Appendix 1.

Proposition 13 *1) There exist j^* and j^{**} such that for all $j \geq j^*$, $\frac{d\rho_{\ln k_j}}{d\frac{\gamma \sigma_q}{\sigma_\varepsilon}} > 0$ and for all $j \geq j^{**}$, $\frac{d\rho_{\ln k_j}}{dh_\infty} < 0$. 2) For all j , $\frac{d\rho_{\mu_j}}{d\frac{\gamma \sigma_q}{\sigma_\varepsilon}} > 0$ and $\frac{d\rho_{\mu_j}}{dh_\infty} < 0$.*

This proposition implies that an increase in $\frac{\gamma \sigma_q}{\sigma_\varepsilon}$ and a decrease in h_∞ increase the autocorrelation about the belief. The same changes can increase the autocorrelation

of organization capital after enough time has passed. Because there is positive assignment between the belief and skills, there are direct effects on the autocorrelation about the belief. However, both influence the autocorrelation of actual organization capital because future actual organization capital is influenced by the firm's current belief. Actual organization capital can temporally deviate from the belief. However, as time passes, an increase in the persistence of the belief dominates the temporal disturbance and increases the persistence of organization capital itself.

4 Empirical Examination

In this section, we derive empirically testable equations and examine the validity of our model. We show that the predictions of our model are broadly supported by the data. The estimated parameters are used to identify the structure of our models: technological persistence, ϕ ; the effect of assignment on persistence, $\frac{\gamma\sigma_q}{\sigma_{\mu\infty}}$; and a measure of the accuracy of information, h_∞ . These structural parameters are inputs into the simulation exercises of the next section.

Because we cannot observe k_t^o , we must translate the results from the previous section into dynamics for observable variables. One such variable is output, y_t . Given that the number of workers is assumed to be unity, we estimate y_t by using labor productivity. The dynamics of a firm's labor productivity relative to the industry and year average and the expected relative productivity in the steady state are derived from equation (25), as follows:

$$D \ln y_{t+1} = b_1 D \ln y_t + b_2 E [D \ln y_t | \mu_{kt}] + v_t, \quad (30)$$

$$E [D \ln y_{t+1} | \mu_{kt+1}] = b_3 D \ln y_t + b_4 E [D \ln y_t | \mu_{kt}], \quad (31)$$

where $D \ln y_t = \ln y_t - E [\ln y]$, $b_1 = \phi + \frac{\psi\sigma_q}{\alpha\sigma_{\mu\infty}}\phi h_\infty$, $b_2 = \frac{\gamma\sigma_q}{\sigma_{\mu\infty}} - \frac{\psi\sigma_q}{\alpha\sigma_{\mu\infty}}\phi h_\infty$, $b_3 = \phi h_\infty + \frac{\psi\sigma_q}{\alpha\sigma_{\mu\infty}}\phi h_\infty$, $b_4 = b_1 + b_2 - b_3$ and $v_t = \alpha (\varepsilon_t^* - \phi u_t^* + u_{t+1}^*)$. The derivations of equations (30) and (31) are given in Appendix 1.

Before showing how to estimate these parameters, it is instructive to discuss how the estimated parameters are related to ϕ , $\frac{\gamma\sigma_q}{\sigma_{\mu\infty}}$ and h_∞ . Suppose that $\frac{\psi\sigma_q}{\alpha\sigma_{\mu\infty}}$ is known. The parameters ϕ , $\frac{\gamma\sigma_q}{\sigma_{\mu\infty}}$ and ϕh_∞ can be identified from the following three equations:

$$\phi = b_1 - \frac{\psi\sigma_q}{\alpha\sigma_{\mu\infty}}\phi h_\infty, \quad \frac{\gamma\sigma_q}{\sigma_{\mu\infty}} = b_2 + \frac{\psi\sigma_q}{\alpha\sigma_{\mu\infty}}\phi h_\infty, \quad \phi h_\infty = \frac{b_3}{1 + \frac{\psi\sigma_q}{\alpha\sigma_{\mu\infty}}}.$$

These equations imply that ϕ , $\frac{\gamma\sigma_q}{\sigma_{\mu\infty}}$ and h_∞ are primarily related to b_1 , b_2 and b_3 , respectively.

To clarify the intuition behind these relationships, note that equation (30) can be written as

$$D \ln y_{t+1} = \phi D \ln y_t + \frac{\gamma\sigma_q}{\sigma_{\mu\infty}} E [D \ln y_t | \mu_{kt}] + \frac{\psi\sigma_q}{\alpha\sigma_{\mu\infty}} \phi h_\infty [D \ln y_t - E [D \ln y_t | \mu_{kt}]] + v_t.$$

As in equation (23), the coefficient on $D \ln y_t$ represents technological persistence and the coefficient on $E [D \ln y_t | \mu_{kt}]$ represents the effect of assignment on persistence. That is, because the fraction ϕ of current organization capital is transformed into next period's organization capital, the fraction ϕ of current labor productivity is linked to labor productivity in the next period. On the other hand, because belief in a high level of organization capital attracts skillful workers, the coefficient on $E [D \ln y_{t-1} | \mu_{kt-1}]$ represents the effect of assignment on organization capital in the next period. The third term does not appear in equation (23). It arises because of a prediction error. When realized output exceeds expected output, people update their beliefs. The updated belief attracts better workers in the next period, which raises output. Hence, if the effects of prediction error are removed, the coefficients on $D \ln y_t$ and $E [D \ln y_t | \mu_{kt}]$ in equation (30) separately identify ϕ and $\frac{\gamma\sigma_q}{\sigma_{\mu\infty}}$.

Note that we can separately identify ϕ and $\frac{\gamma\sigma_q}{\sigma_{\mu\infty}}$ even if $h_\infty = 1$. When $h_\infty = 1$, a firm knows the exact level of current organization capital. However, future organization capital remains unknown because it is influenced by shocks. Hence, a firm's employment decision must be based on its beliefs, rather than on real values. That is, our model can separate $\frac{\gamma\sigma_q}{\sigma_{\mu\infty}}$ from ϕ not because of noisy information, but because it distinguishes the effects of firms' decisions from those of firms' capabilities.

The prediction error is clearly influenced by the accuracy of information. In our regression analysis, the parameter b_3 provides useful information about h_∞ . Note that

$$b_1 - b_3 = \phi(1 - h_\infty). \quad (32)$$

This expression shows that the difference between b_1 and b_3 provides information about h_∞ . In equation (30), v_t and $D \ln y_t$ are correlated because $E[u_t^* | D \ln y_t] \neq 0$. Because measured productivity is influenced not only by the level of organization capital, but also by current temporal shocks, observed productivity contains information about current shocks. When rational agents predict future productivity, they efficiently extract this information from current productivity. Hence, b_3 deviates from the fundamental parameter b_1 . When the variation of u_t^* is a large component of the variation in measured productivity, measured productivity is influenced more by u_t^* and less by $\ln k_t^o$. Hence, measured productivity is useful for predicting u_t^* , but not for predicting $\ln k_t^o$. Therefore, h_∞ is small and the difference between b_1 and b_3 is large.

Estimation Method 1: To implement the procedure discussed above, we must estimate $E[D \ln y_t | \mu_{kt}]$ from the data. We propose two methods for doing this. Because each strategy has its own strengths and weaknesses, it is hoped that the strategies complement each other.

The first method applies the following proposition, which is proven by equations (14) and (15). Although the dynamics of organization capital are not influenced by the assumption that organization capital is accumulated through learning by doing, equations (14) and (15) are affected by this assumption. Therefore, the following useful proposition represents a benefit of assuming that there is learning by doing.

Proposition 14 *Perceived relative productivity is equal to relative wages and expected relative profits per worker.*

$$E [D \ln y_t | \mu_{kt}] = D \ln \hat{w} (\hat{\chi} (\mu_{kt})) = D \ln \hat{\pi}^e (\mu_{kt}),$$

where $D \ln \hat{w} (\hat{\chi} (\mu_{kt})) = \ln \hat{w} (\hat{\chi} (\mu_{kt})) - E [\ln \hat{w} (\hat{\chi} (\mu_{kt}))]$ and $D \ln \hat{\pi}^e (\mu_{kt}) = \ln \hat{\pi}^e (\mu_{kt}) - E [\ln \hat{\pi}^e (\mu_{kt})]$.

Proposition 14 states that perceived relative productivity can be estimated by using relative wages. Hence, the following testable equation is derived from equations (30) and (31):

$$D \ln y_t = \theta_1 D \ln y_{t-1} + \theta_2 D \ln y_{t-2} + \theta_3 D \ln w_{t-2} + v_{t-1}, \quad (33)$$

where $\theta_1 = b_1$, $\theta_2 = b_2 b_3$ and $\theta_3 = b_2 (b_1 + b_2 - b_3) > 0$. By using the estimated value of θ_1 , θ_2 and θ_3 , we can identify b_1 , b_2 and b_3 , which allows us to estimate ϕ , h_∞ and $\frac{\gamma \sigma_q}{\sigma_{\mu_\infty}}$.

Equation (33) shows that after controlling for the first and second lags of relative productivity, the second lag of relative wages must have a positive impact on current relative productivity. Given that skilled workers equip firms with better firm-specific knowledge, the theory predicts that there is a positive association between past wages and current productivity.

One econometric issue exists. Because v_{t-1} contains u_{t-1}^* , it is correlated with $D \ln y_{t-1}$. Hence, we need an instrument for this variable. Proposition 14 provides a suitable instrument. Because the firm makes employment decisions without observing realized output, relative wages, $D \ln w_{t-1} \equiv \ln w_{t-1} - E [\ln w]$, are not influenced by the realization of the noise term, u_{t-1}^* , but are correlated with $D \ln y_{t-1}$ because of positive assignment. Hence, $D \ln w_{t-1}$ can be used as the instrument.

Estimation Method 2: The next estimation method is more complex. However, it allows us to examine a different prediction of our theory. This is that the belief, which

is constructed from sequences of past relative productivity, influences future relative productivity. Furthermore, to apply this alternative method, we need not assume that there is learning by doing. Hence, even if top managers' skills are important elements of organization capital, this estimation method can be used to identify parameters. To construct $E[D \ln y_t | \mu_{kt}]$ from the data, we derive the following regression equation from equation (31):

$$D \ln y_t = b_3 \sum_{i=0}^{t-1} (b_4)^i D y_{t-1-i} + b_4^t E[D \ln y_0 | \mu_{k0}] + \varpi_t, \quad (34)$$

where $\varpi_t = D \ln y_t - E[D \ln y_t | \mu_{kt}]$. Note that ϖ_t is not correlated with $D \ln y_{t-1-i}$ for all $i \geq 0$ and $E[D \ln y_0 | \mu_{k0}]$.

As already discussed, the parameters b_1 and b_3 differ because $E[v_t | D \ln y_t] \neq 0$. Hence, to separate b_1 from b_3 , we apply ordinary least squares (OLS) to equation (34) and use an instrumental variables (IV) approach to estimate equation (30). The IV estimate provides a consistent estimator of the parameter b_1 and the OLS estimate provides a biased estimator of b_1 , which is b_3 . Hence, the difference between the IV estimates and the OLS estimates indicates the extent to which labor productivity provides information about the error term and identifies h_∞ . We use this technique in applying the second estimation procedure discussed below.

Assume that there is a proxy for $E[D \ln y_0 | \mu_{k0}]$. First, we choose an arbitrary value of b_4 , and construct $\sum_{i=0}^{t-1} (b_4)^i D y_{t-1-i}$ and $(b_4)^t E[D \ln y_0 | \mu_{k0}]$ from the data. Second, equation (34) is estimated under the constraint that the coefficient of $(b_4)^t E[D \ln y_0 | \mu_{k0}]$ is 1. This yields \hat{b}_3 , where \hat{b}_3 is the estimated value of b_3 . Third, using \hat{b}_3 and b_4 , we estimate $E[D \ln y_t | \mu_{kt}]$ by $\hat{b}_3 \sum_{i=0}^{t-1} (b_4)^i D y_{t-1-i} + b_4^t E[D \ln y_0 | \mu_{k0}]$. Fourth, using the estimated value of $E[D \ln y_t | \mu_{kt}]$, we estimate equation (30) by using the IV regression. We use $D \ln y_{t-1}$ and $D \ln w_t$ as instruments for $D \ln y_t$ and $E[D \ln y_t | \mu_{kt}]$. We need an additional instrument for $E[D \ln y_t | \mu_{kt}]$ because \hat{b}_3 contains a measurement error. This IV estimation procedure yields \hat{b}_1 and \hat{b}_2 , where \hat{b}_1 and \hat{b}_2 are the estimated values of b_1 and b_2 . Fifth, because there is a regulatory

relationship, according to which $b_4 = b_1 + b_2 - b_3$, we replace b_4 by $\hat{b}_1 + \hat{b}_2 - \hat{b}_3$ and repeat the same procedure until the estimated b_4 converges to the assumed b_4 .

Data: We use COMPUSTAT industry annual data from 1970 to 2004 for estimation. COMPUSTAT provides data on an unbalanced panel of publicly traded firms in the U.S. It contains information from balance sheets, and information on incomes, cash flows and financial variables. The variables value added per worker and the average wage rate are constructed for each firm and each year. Details of our data construction procedure and summary statistics on the constructed variables are given in Appendix 1.

We estimate $D \ln y_{ft}$ and $D \ln w_{ft}$ by $\ln y_{ft} - \frac{\sum_j^{m_t} \ln y_{ft}}{m_t}$ and $\ln w_{ft} - \frac{\sum_j^{m_t} \ln w_{ft}}{m_t}$, where y_{ft} is value added divided by the number of workers and labor expenses per worker in the f th firm in year t , respectively, and m_t is the number of firms in the corresponding four-digit industry in year t . We estimate each firm’s initial prior belief, $E[D \ln y_0 | \mu_{k0}]$, from the average value of $D \ln y_{ft}$ over the five consecutive years following the firm’s initial appearance in COMPUSTAT after 1970. Therefore, the following regression is estimated by using data for 1975–2004.

Results: First, in Table 2, we report the regression results from the first estimation method.

Because only few companies report labor and related expenses in COMPUSTAT, we estimate labor costs for companies that do not report this information. (The estimation method is described in Appendix 1.) To investigate the potential bias arising from the use of this estimation method, we also report regression results based on the sample of companies that report labor and related expenses. The “Small Sample” in Table 2 includes only companies that report labor and related expenses. The “Large Sample” includes companies whose labor costs we have estimated.

All coefficients in Table 2 are significant and positive, which is consistent with our

The dependent variable is $D \ln y_t$.

	Small Sample	Small Sample	Large Sample	Large Sample
$D \ln y_{t-1}$	0.708**	0.696**	0.708**	0.671**
	(0.027)	(0.027)	(0.008)	(0.008)
$D \ln y_{t-2}$	0.120**	0.132**	0.158**	0.084**
	(0.026)	(0.026)	(0.009)	(0.008)
$D \ln w_{t-2}$	0.116**	0.101**	0.040**	0.087**
	(0.018)	(0.018)	(0.006)	(0.006)
$D \ln k_t$		0.038**		0.106**
		(0.007)		(0.002)
# of observations	3113	3113	30135	20119

Table 2: Estimation Method 1

The variables $D \ln y_t$, $D \ln w_t$ and $D \ln k_t$ are relative labor productivity, relative wage payments and the relative capital–labor ratio, respectively. The “Small Sample” includes only companies that report labor and related expenses. The “Large Sample” also includes companies whose labor costs we have estimated. The variable $D \ln w_{t-1}$ is used as the instrument for this regression. Standard errors are reported in parentheses. ** denotes significance at the 0.5 percent level.

theoretical predictions. Moreover, the results do not depend on the sample size.

More interestingly, two-year lagged relative wage payments have a positive impact on current relative productivity even after conditioning the first and second lags of relative productivity. The elasticity of two-year lagged relative wage payments is 0.12 in the small sample and 0.04 in the large sample. The coefficient is smaller in the large sample. However, the results from both samples are significant and demonstrate the quantitative effect. The results support the hypothesis that skilled workers improve a firm’s assets.

In this regression, we implicitly assume that there are no adjustment costs of investment in physical capital. Given this assumption, physical capital can be derived

as a function of organization capital. Organization capital not only directly increases labor productivity, it also increases the physical capital stock, which in turn raises labor productivity. Because we are interested in the total effect of organization capital on labor productivity, we ignore the physical capital stock.

However, if adjustment costs of investment in physical capital are important, a high current level of labor productivity can partially be explained by the initial physical capital stock per worker. The omission of physical capital might have biased our estimates. To investigate this possibility, we add relative physical capital per worker, $D \ln k_t$. We estimate $D \ln k_t$ by using $\ln k_{ft} - \frac{\sum_f^{m_t} \ln k_{ft}}{m_t}$, where k_{ft} is the initial capital stock per worker in the f th firm in year t .

The inclusion of $D \ln k_t$ hardly changes the coefficients in the small sample, but raises the elasticity of $D \ln w_{t-2}$ and lowers that of $D \ln y_{t-2}$ in the large sample. This indicates that adjustment costs of investment might have biased our results in the large sample. This is a potential problem. However, this minimal bias is unlikely to affect our simulation results. We discuss this point later.

We report the regression results obtained by using the second estimation method in Tables 3 and 4. The initial value of b_4 is chosen to be 0.5. The result is not sensitive to this choice. The results in these tables are based on the estimated b_4 matching the assumed b_4 . Table 3 reports the results from the regression equation (34). Table 4 reports the results from the regression equation (30).

Table 3 shows that b_3 (the coefficient on $\sum_{i=0}^{t-1} (b_4)^i D \ln y_{t-1-i}$) is 0.64 in the small sample and 0.72 in the large sample. The large sample produces a slightly larger value of b_3 . To check whether the constrained regression produces a bias, we also ran an unconstrained regression. This regression yields a similar value of b_3 . This suggests that our estimates are not sensitive to the constraint.

The unconstrained regression also reveals an interesting feature of the data: the weighted initial prior has a persistent effect on labor productivity. This means that the effect of initial values declines over time, but does not fade out altogether. The

theory predicts a coefficient on the weighted initial prior of 1, but this is not supported by data. However, the coefficients are not far from 1. In particular, the coefficient in the small sample is close to 1, 0.93. These results indicate that the model is a useful first-order approximation of the data.

As already discussed, if adjustment costs of investment in physical capital are important, our results might be biased. Hence, we also include $D \ln k_t$ in our regressions. This does not materially change the coefficients of the regressions. Hence, our results are robust in this respect.

Table 4 shows that, after controlling for current relative productivity, the constructed belief about relative productivity continues to influence relative productivity in the next year. Note that $E [D \ln y_t | \mu_{kt}]$ is constructed from past observations. Our regression results are consistent with the hypothesis that people learn about a firm's capacity from its past performance and form a belief that influences the employment of skilled workers and future performance.

Table 4 shows that b_1 (the coefficient on $D \ln y_t$) is 0.84 in the small sample and 0.72 in the large sample. Given that b_3 is 0.64 in the small sample and 0.72 in the large sample, b_1 exceeds b_3 in the small sample, but both are similar in the large sample. Hence, equation (32) implies that $h_\infty < 1$ in the small sample, while $h_\infty = 1$ in the large sample. That is, according to the results from the large sample, labor productivity is useful for predicting organization capital.

Adding relative physical capital stock per worker hardly changes the coefficients in the small sample, but causes the coefficient of $E [D \ln y_t | \mu_{kt}]$ to decrease in the large sample. This indicates that the large-sample regression results might overstate the effects of assignment if adjustment costs of investment in physical capital are important. However, as is discussed later, this potential problem is unlikely to affect our simulation results.

The dependent variable is $D \ln y_t$.

	Small	Small	Small	Large	Large	Large
	Const	Unconst	Const	Const	Unconst	Const
$\sum_{i=0}^{t-1} (b_4)^i D \ln y_{t-1-i}$	0.637**	0.639**	0.627**	0.715**	0.719**	0.667**
	(0.009)	(0.009)	(0.009)	(0.003)	(0.003)	(0.003)
$b_4^t E [D \ln y_0 \mu_{k0}]$	1	0.931**	1	1	0.847**	1
		(0.026)			(0.006)	
$D \ln k_t$			0.074**			0.082**
			(0.007)			(0.002)
# of observations	3645	3645	3638	32211	32211	32114

Table 3: Estimation Method 2 – the First Stage

The dependent variable is $D \ln y_{t+1}$.

	Small	Small	Large	Large
$D \ln y_t$	0.838**	0.798**	0.716**	0.757**
	(0.041)	(0.040)	(0.013)	(0.013)
$E [D \ln y_t \mu_{kt}]$	0.090*	0.118**	0.211**	0.103**
	(0.043)	(0.042)	(0.014)	(0.013)
$D \ln k_{t+1}$		0.049**		0.085**
		(0.008)		(0.003)
# of observations	2772	2771	23019	23012

Table 4: Estimation Method 2 – the Second Stage

We report regression results in which the estimated b_4 matches the assumed b_4 in two tables. “Small” refers to the small sample, which includes only companies that report labor and related expenses. “Large” refers to the large sample that includes companies whose labor costs we have estimated. “Const” denotes the constrained regression and “Unconst” denotes the unconstrained regression. Table 3 reports

OLS results, and Table 4 reports IV results. Standard errors are reported in parentheses. * denotes significance at the 5 percent level. ** denotes significance at

the 0.5 percent level.

	Small Sample	Large Sample
$\frac{E[w_t]}{E[y_t]}$	0.429	0.608

Table 5: Labor Share

$\frac{E[w_t]}{E[y_t]} = \frac{\sum_i^I \sum_t^T \left[\frac{\sum_f^{m_{it}} w_{fit}}{\sum_f^{m_{it}} y_{fit}} \right]}{IT}$, where w_{fit} and y_{fit} are the wage payments and labor productivity of the f th firm in the i th industry in year t , m_{it} is the number of firms operating in the i th industry in year t , I is the number of industries and T is the number of years. “Small Sample” includes only the companies that report labor and related expenses and “Large Sample” includes also companies for which we have estimated labor costs.

5 Simulation Results

In this section, we report our estimates of ϕ , h and $\frac{\gamma\sigma_q}{\sigma_{\mu\infty}}$. Using the estimated parameters, we simulate our model and examine the extent to which assignment and the noisiness of information affect the persistence of relative productivity, relative wages and relative profits per worker.

Calibration of $\frac{\psi\sigma_q}{\alpha\sigma_{\mu\infty}}$: To identify the parameters, we need to know the value of $\frac{\psi\sigma_q}{\alpha\sigma_{\mu\infty}}$. For this purpose, we use a calibration technique. The wage function (14) implies that $\frac{\psi\sigma_q}{\alpha\sigma_{\mu\infty}}$ can be estimated from $\frac{E[w_t]}{E[y_t]}$, and we estimate $\frac{E[w_t]}{E[y_t]}$ from $\frac{\sum_i^I \sum_t^T \left[\frac{\sum_f^{m_{it}} w_{fit}}{\sum_f^{m_{it}} y_{fit}} \right]}{IT}$, where w_{fit} and y_{fit} are the wage payments and labor productivity of the f th firm in the i th industry in year t , m_{it} is the number of firms operating in the i th industry in year t , I is the number of industries and T is the number of years.

Table 5 reports our estimate of $\frac{E[w_t]}{E[y_t]}$. This is 0.43 in the small sample and 0.61 in the large sample. Because $\frac{E[w_t]}{E[y_t]}$ approximates the labor share, the estimate of 0.43 is fairly small. This may be because COMPUSTAT only includes publicly traded firms, which are relatively capital intensive and above average in size. In particular, because

only relatively large firms in COMPUSTAT report labor and related expenses, “Small Sample” contains only fairly large firms. In 2000, the average capital stock of firms not reporting labor and related expenses was 421 million dollars, whereas that of firms reporting labor and related expenses was 787 million dollars.

Note that our estimated values are similar to those obtained by previous studies based on COMPUSTAT data. Dhawan and Gerdes (1997) report an estimated labor share from COMPUSTAT of 0.3. Summary statistics in Bresnahan, Brynjolfsson and Hitt (2002) indicate a labor share for their selected sample from COMPUSTAT of 0.53.

Although our estimated $\frac{E[w_t]}{E[y_t]}$ varies between samples, fortunately our simulation results are not particularly sensitive to these variations. Below, we use our estimates to estimate ϕ , h_∞ and $\frac{\gamma\sigma_q}{\sigma_{\mu_\infty}}$.

Estimated Structural Parameters: The results from our regression analysis yield the following parameters of interest: $\frac{\psi\sigma_q}{\alpha\sigma_{\mu_\infty}}$, ϕ , h_∞ and $\frac{\gamma\sigma_q}{\sigma_{\mu_\infty}}$. Table 6 reports the results.

Our estimates generally differ between the small and large samples. The large difference in $\frac{\psi\sigma_q}{\alpha\sigma_{\mu_\infty}}$ arises because of different estimated values of $\frac{E[w_t]}{E[y_t]}$. The value of $\frac{\psi\sigma_q}{\alpha\sigma_{\mu_\infty}}$ measures the relative importance of worker quality to production. Hence, if $\frac{\psi\sigma_q}{\alpha\sigma_{\mu_\infty}}$ is large, a firm’s productivity is affected more by assignment. This partially explains why $\frac{\gamma\sigma_q}{\sigma_{\mu_\infty}}$ is relatively large and ϕ is relatively small in the large sample. Similarly, our estimates of h_∞ are also sensitive to the sample size.⁵ Output is useful for predicting organization capital in the large sample, but not in the small sample.

While different sample sizes yield large differences in the estimated parameters, different estimation methods produce similar results. In particular, the results in

⁵Because h_∞ cannot exceed unity, if the estimated value of h_∞ is greater than 1 we set $h_\infty = 1$ for the purpose of simulation.

	Small Sample	Small Sample	Large Sample	Large Sample
	Estimation 1	Estimation 2	Estimation 1	Estimation 2
$\frac{\psi\sigma_q}{\alpha\sigma_{\mu\infty}}$	0.751	0.751	1.549	1.549
ϕ	0.499	0.565	0.279	0.282
h_∞	0.556	0.644	1	0.993
$\frac{\gamma\sigma_q}{\sigma_{\mu\infty}}$	0.456	0.363	0.646	0.645
$\lambda_1 = \phi + \frac{\gamma\sigma_q}{\sigma_{\mu\infty}}$	0.955	0.928	0.925	0.927

Table 6: The Estimated Structural Parameters

$\frac{\psi\sigma_q}{\alpha\sigma_{\mu\infty}}$ measures the relative contribution of skills to current labor productivity. ϕ measures technological persistence. h_∞ measures the accuracy of the information contained in realized labor productivity for predicting the level of organization capital. $\frac{\gamma\sigma_q}{\sigma_{\mu\infty}}$ measures the importance of assignment for persistence. “Small Sample” includes only companies that report labor and related expenses. “Large Sample” also includes companies for which labor costs are estimated.

the large sample are almost identical. The robustness of the estimates to different estimation methods suggests that the results are reliable.

More importantly, although regressions from different samples produce different values, the estimated persistence parameters, $\lambda_1 = \phi + \frac{\gamma\sigma_q}{\sigma_{\mu\infty}}$, are remarkably stable. They range from 0.93 to 0.96. Note that Proposition 11 suggests that λ_1 is the most important parameter for the persistence of organization capital. It is shown later that λ_1 is also the most important parameter for the persistence of relative productivity, relative wages and expected relative profits. The remarkable stability of the persistence parameter explains why our simulation results are not particularly sensitive to variations in sample size and estimation method.

Persistence of Productivity, Wages and Profits per Workers: To understand

the effects of assignment on persistence, we calculate autocorrelations for productivity and expected productivity. Let us define the autocorrelations as $\rho_{\ln y_j} \equiv \frac{E[D \ln y_t D \ln y_{t-j}]}{\text{Var}(\ln y_t)}$ and $\rho_{E[\ln y|\mu]_j} \equiv \frac{E[E[D \ln y_t|\mu_t]E[D \ln y_{t-j}|\mu_{t-j}]]}{\text{Var}(E[\ln y_t|\mu_t])}$. The following proposition, which is proved in Appendix 1, generates the theoretical prediction.

Proposition 15 *The autocorrelations of relative productivity and expected relative productivity are*

$$\rho_{\ln y_j} = \frac{\phi h_\infty \left(1 + \frac{\psi \sigma_q}{\alpha \sigma_{\mu_\infty}}\right) \lambda_1^{j-1} \left[\phi h_\infty \left(1 + \frac{\psi \sigma_q}{\alpha \sigma_{\mu_\infty}}\right) \lambda_1 + 1 - \lambda_1^2\right]}{\left(1 + \frac{\psi \sigma_q}{\alpha \sigma_{\mu_\infty}}\right)^2 (\phi h_\infty)^2 + 1 - \lambda_1^2}, \quad \rho_{E[\ln y|\mu]_j} = \lambda_1^j,$$

where $\lambda_1 = \phi + \frac{\gamma \sigma_q}{\sigma_{\mu_\infty}}$ and $j \geq 1$.

Note that the predicted autocorrelations can be calculated by using the estimated structural parameters, $\frac{\psi \sigma_q}{\alpha \sigma_{\mu_\infty}}$, ϕ , h_∞ and $\frac{\gamma \sigma_q}{\sigma_{\mu_\infty}}$. Proposition 15 shows that λ_1 is the most important determinant of these autocorrelations. Given that there is a stable estimate of λ_1 , we do not expect the predicted autocorrelations to depend greatly on the sample size and estimation method. This expectation is confirmed below.

We compare the simulated correlations with the correlations observed in the data. Proposition 14 states that perceived relative productivity is equal to the relative wage, which is also equal to expected relative profits per worker. Hence, for wages and expected profits per worker, we can use the autocorrelations of expected relative productivity to compare the simulated correlations with the observed ones. Note that, in this model, autocorrelations are equivalent to correlations between current relative values and relative values from j periods previously. Hence, we estimate the observed correlations by using

$$\rho_{\ln x_j} = \frac{\sum_t^T \sum_i^I \sum_f^{m_{it}} \left[\ln x_{fit} - \frac{\sum_f^{m_{it}} \ln x_{fit}}{m_{it}} \right] \left[\ln x_{fi(t-j)} - \frac{\sum_f^{m_{it}} \ln x_{fi(t-j)}}{m_{it}} \right]}{\sqrt{\sum_t^T \sum_i^I \sum_f^{m_{it}} \left[\ln x_{fit} - \frac{\sum_f^{m_{it}} \ln x_{fit}}{m_{it}} \right]^2} \sqrt{\sum_t^T \sum_i^I \sum_f^{m_{it}} \left[\ln x_{fi(t-j)} - \frac{\sum_f^{m_{it}} \ln x_{fi(t-j)}}{m_{it}} \right]^2}}, \quad (35)$$

where x_{fit} represents either the labor productivity, labor expenses per worker or the operating income per worker of the f th firm in the i th industry in year t , m_{it} is the number of firms operating in the i th industry in year t , I is the number of industries and T is the number of years.

Figures 1, 2 and 3 summarize the results of our simulations. Figure 1 compares the simulated correlation for relative productivity with the one estimated from the data. As already discussed, the results are similar despite differences in sample size and estimation method. All predicted correlations fit the data quite well. In particular, the simulation results in the large sample are remarkably good. All results indicate that the model can quantitatively account for the observed persistence of productivity differences.

Figure 2 conducts the same exercises – for the relative wage. All results suggest that the model’s predictions are consistent with the data. Hence, the results in Figure 1 are unlikely to be the result of coincidence. Our model can also explain the persistent differences in wage payments.

Figure 3 compares the predicted correlation for expected relative profits per worker with the correlation of relative profits per worker. Although the predicted correlation is much larger than the observed correlation, this result is expected. Our theory explains the correlation for expected relative profits per worker, but not the one for real relative profit per worker. Real profits per worker are affected by unpredictable idiosyncratic shocks. Hence, this correlation is expected to be smaller. This reasoning is consistent with the pattern observed in Figure 3.

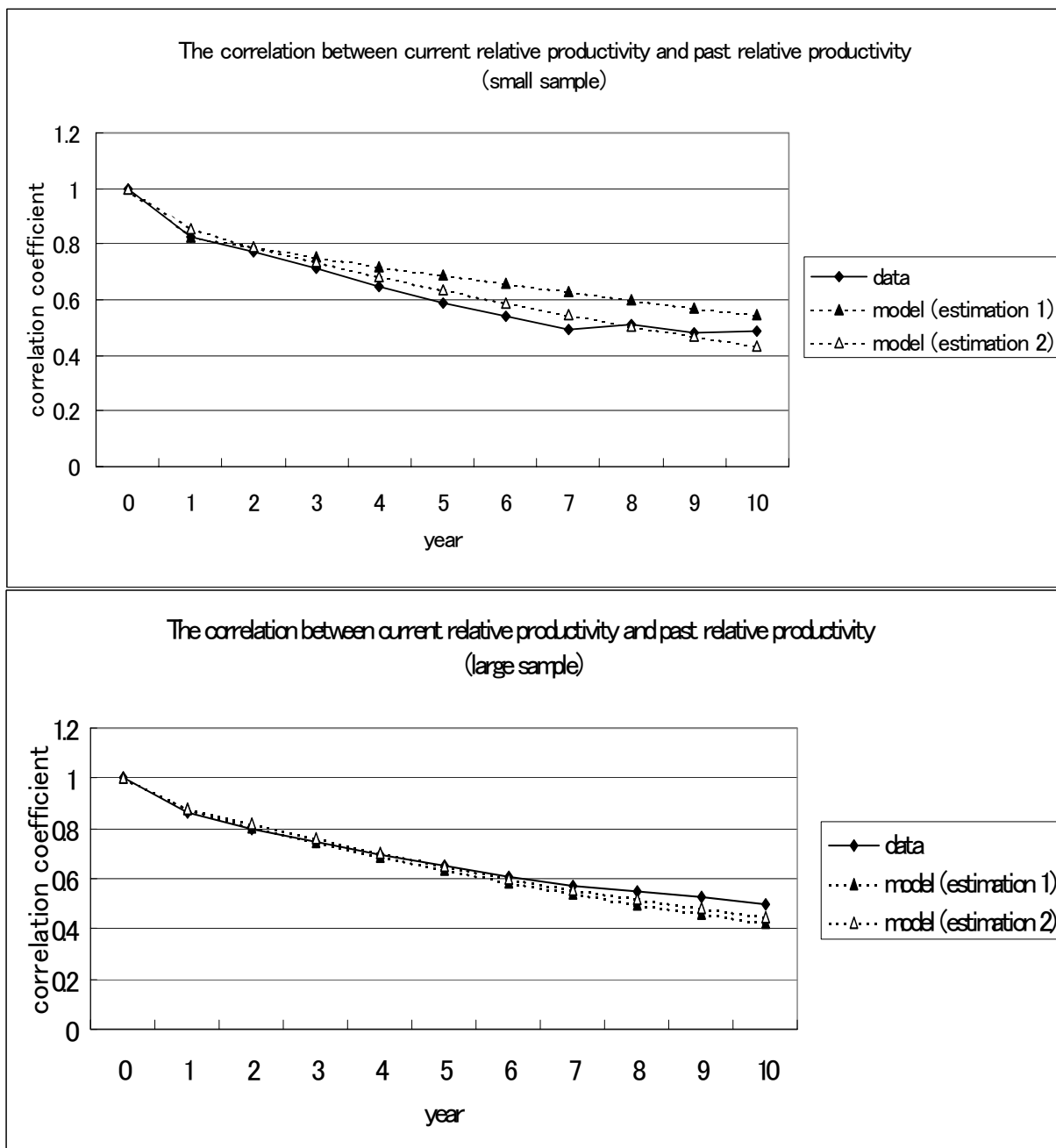


Figure 1:

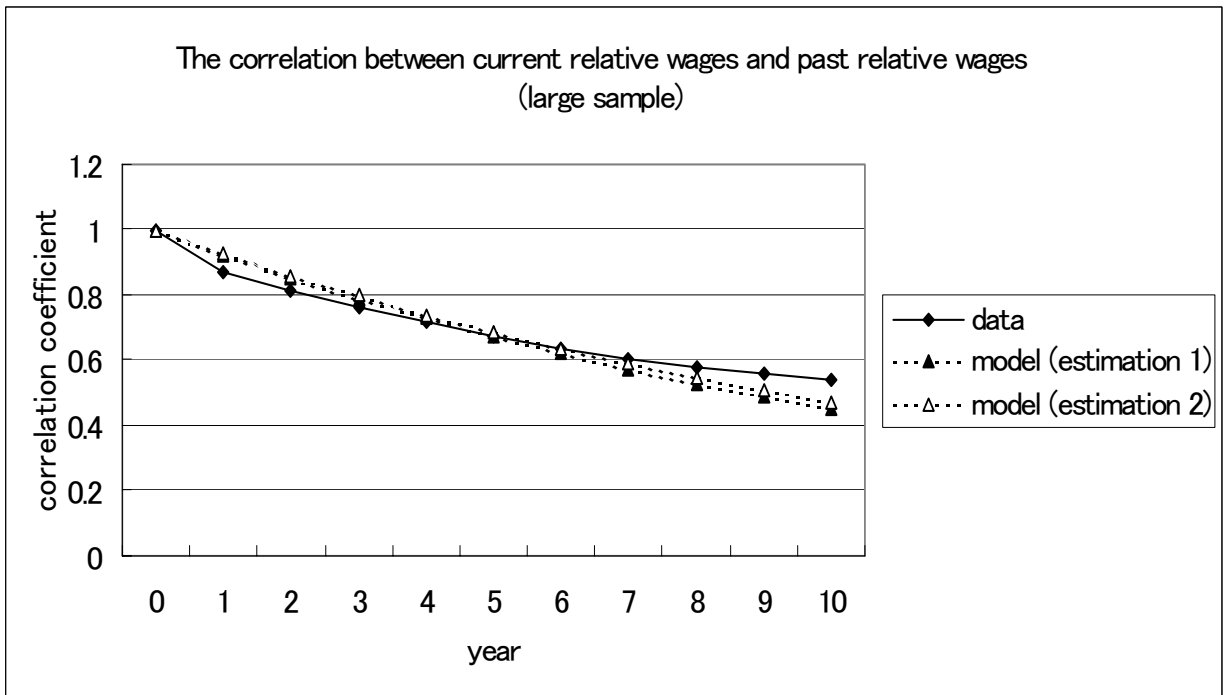
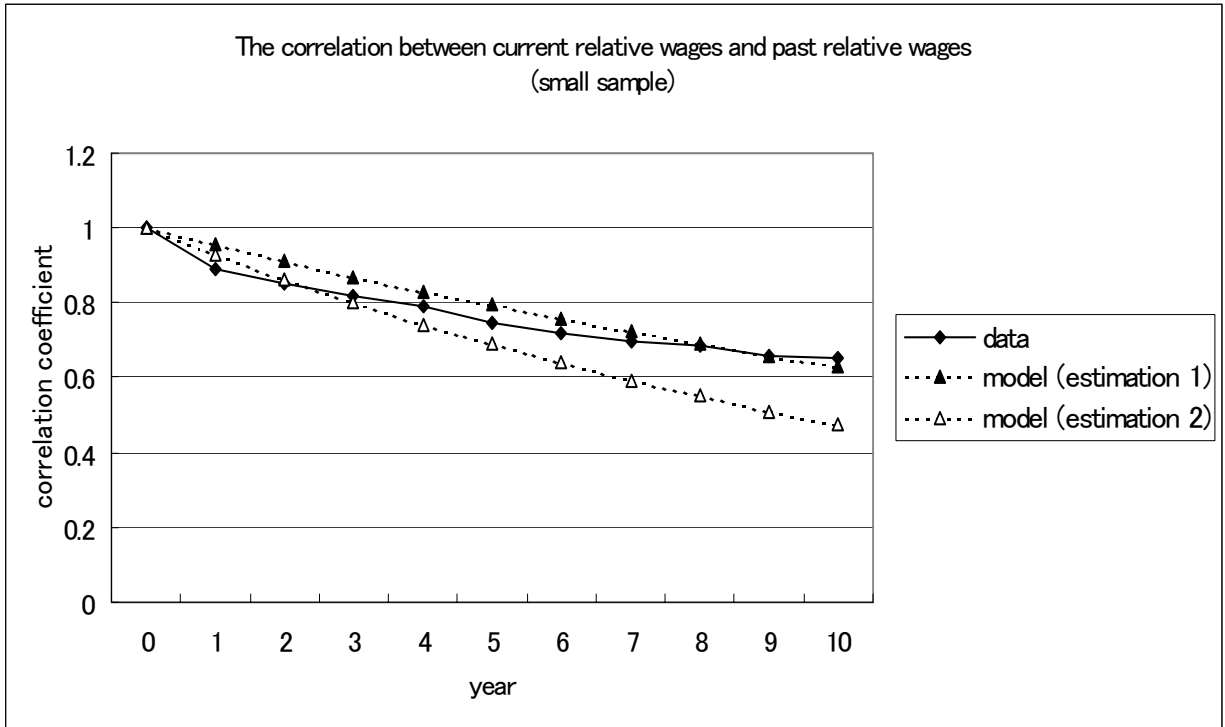


Figure 2:

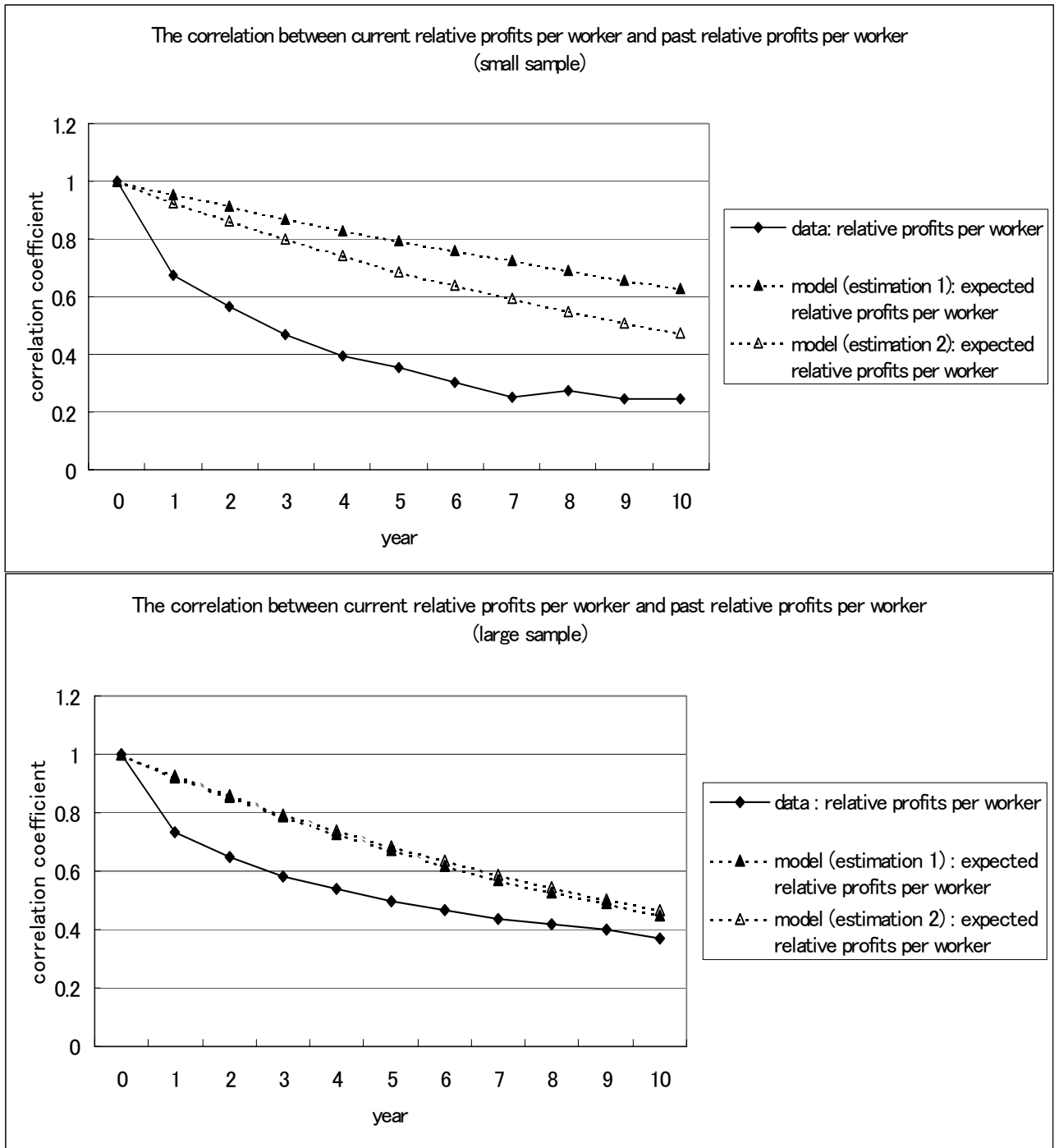


Figure 3:

Let us conduct a counterfactual experiment. We first ask “What would happen if people were homogenous and, therefore, there were no assignments in the economy?” This experiment can be done by assuming that $\sigma_q = 0$ and the other parameters are constant. The assumption of $\sigma_q = 0$ implies a zero labor share, $\frac{\psi\sigma_q}{\alpha\sigma_{\mu\infty}} = 0$, and implies that there is no assignment effect, $\frac{\gamma\sigma_q}{\sigma_{\mu\infty}} = 0$. Figures 4 and 5 report the results of this experiment.

Figure 4 shows that, if $\sigma_q = 0$, the autocorrelations for relative productivity diminish to about 0 after five years. This result does not depend on either the sample sizes or the estimation method. All results show that relative temporal advantages disappear quickly if there are no benefits from positive assignment.

This point is confirmed by Figure 5. It shows that if $\sigma_q = 0$, the autocorrelations of perceived relative productivity (which is equivalent to relative wages⁶ and expected relative profits per worker) become 0 after between four and six years. This result is not affected by either sample size or estimation methods. This means that positive assignment accounts for much of the observed persistence in wage payments and profits.

We also ask “What would happen if $h_\infty = 1$?”. Recently, many economists have recognized the importance of intangible assets to firms and have tried to estimate them [e.g., Hall (2001)]. This exercise may provide valuable information on how better estimation of organization capital might affect the persistence of variables.

Table 6 shows that labor productivity is quite informative in the large sample. This means that the noisiness of information is not the main source of persistence in the large sample. Hence, we conduct this exercise only for the small sample and investigate whether the small sample confirms the findings from the large sample.

⁶There is a caution for the interpretation of Figure 5. Equation (14) says that if $\sigma_q = 0$, wages must be 0 for all firms. Hence, the relative wage is always 0. In order to maintain a link between the relative wage and expected relative productivity, σ_q has to be slightly larger than 0. Hence, when we discuss the persistence of the relative wages, the results in Figure 5 have to be interpreted as an approximation.

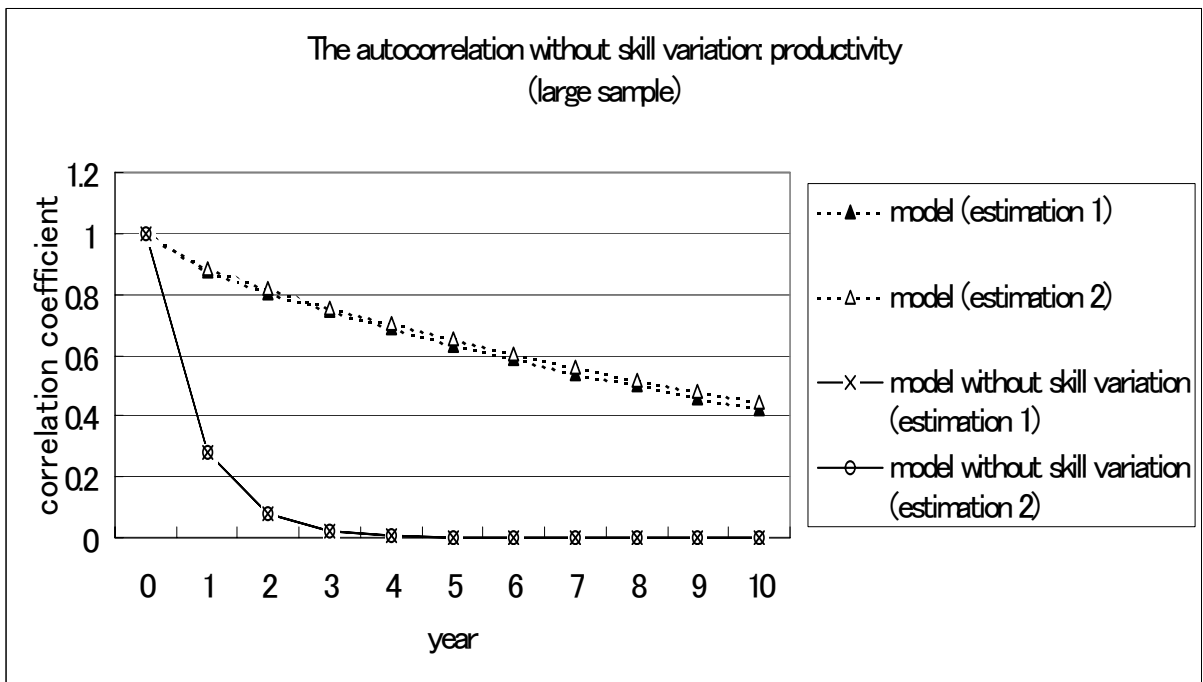
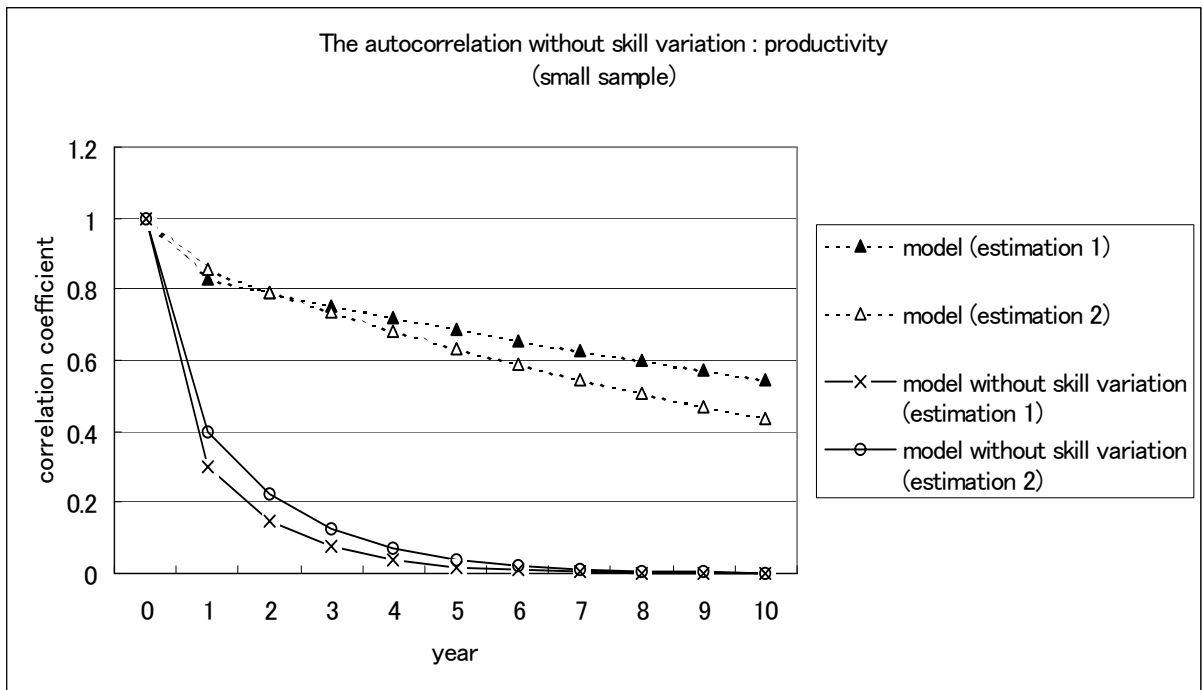


Figure 4:

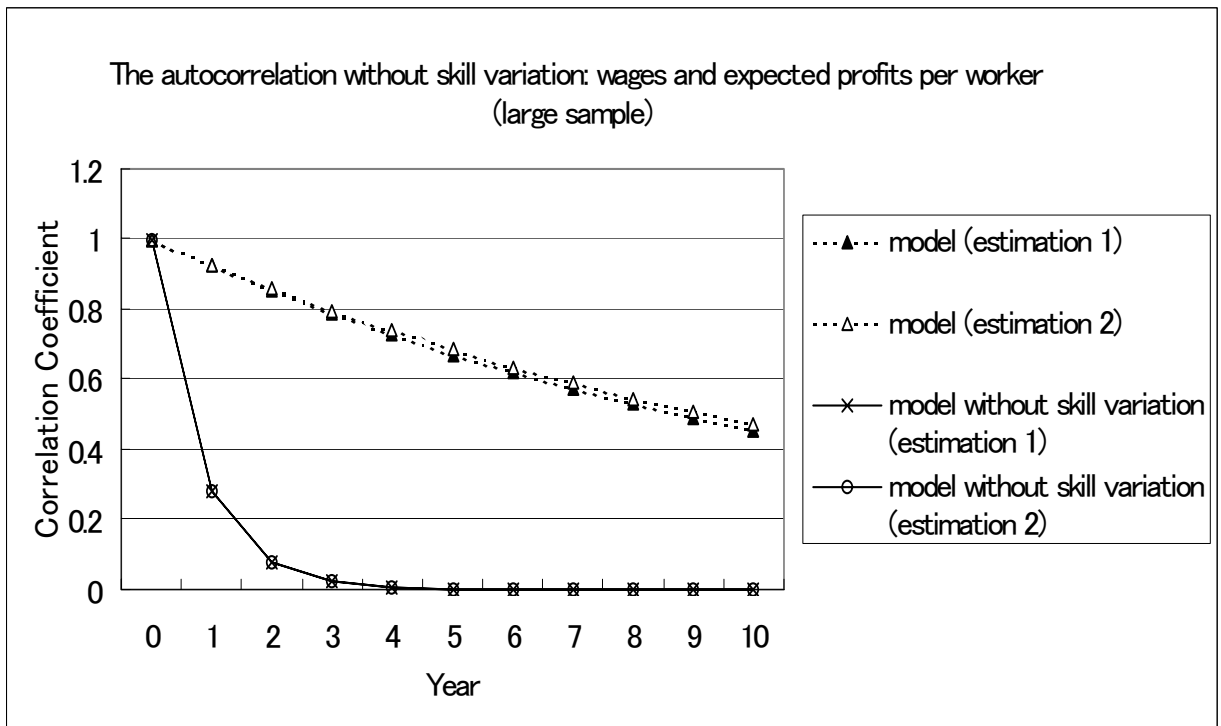
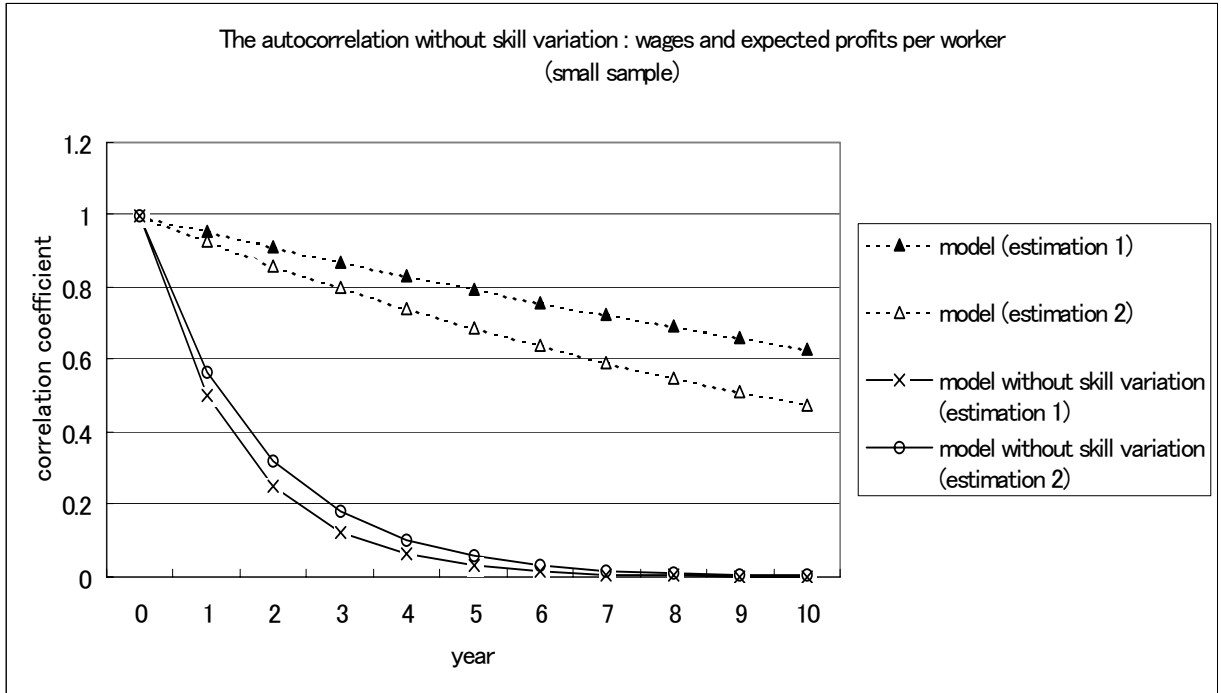


Figure 5:

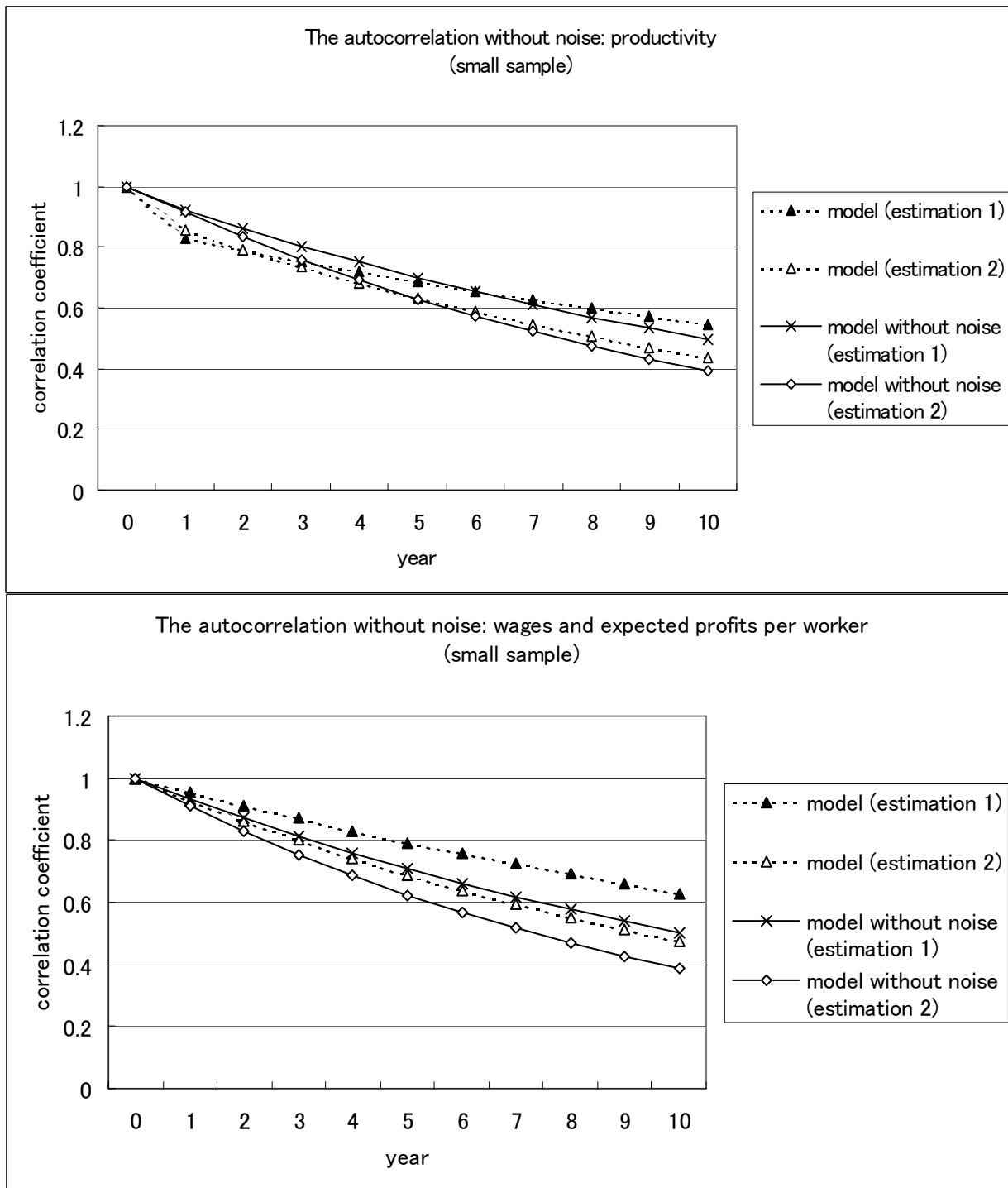


Figure 6:

When h_∞ changes, Lemma 9 shows that $\frac{\gamma\sigma_q}{\sigma_{\mu_\infty}}$ changes through equation (29), which in turn influences $\frac{\psi\sigma_q}{\alpha\sigma_{\mu_\infty}}$. This is because $\frac{\psi\sigma_q}{\alpha\sigma_{\mu_\infty}} = \frac{\psi}{\alpha\gamma} \frac{\gamma\sigma_q}{\sigma_{\mu_\infty}}$. These combined effects are reported in Figure 6. This shows that an improvement in information causes only slight changes in the persistence of productivity, even in a small sample.

Figure 6 also illustrates an interesting finding. When $h_\infty = 1$, the autocorrelations of relative wages and expected relative profits are smaller than those from the benchmark model; that of relative productivity is larger in the short run and smaller in the long run. When $h_\infty = 1$, because beliefs are frequently updated the autocorrelations of relative wages and expected relative profits per worker are small. Although the same effect influences the autocorrelation of relative productivity, there is an additional effect. When $h_\infty = 1$, $\sigma_u = 0$. Hence, labor productivity is not affected by temporal shocks and the autocorrelation may be larger. Although temporal shocks negatively affect the persistence of relative productivity in the short run, their effect is dominated by that of beliefs in the long run.

In summary, these exercises consistently suggest that positive assortative assignment accounts for much of the observed persistence of a firm's relative advantages (disadvantages), whereas the noisiness of information plays a relatively minor role.

Positive Correlation Between Relative Productivity and Relative Wages:

Our model can predict the correlation between relative productivity and relative wages, $\rho_{\ln y \ln w}$, where $\rho_{\ln y \ln w} \equiv \frac{E[D \ln y_t D \ln w_t]}{\sqrt{\text{Var}(D \ln y_t) \text{Var}(D \ln w_t)}}$. The following proposition is proved in Appendix 1.

Proposition 16 *The correlation between $D \ln y_t$ and $D \ln w_t$ can be simulated by*

$$\rho_{\ln y \ln w} = \frac{\left(1 + \frac{\psi\sigma_q}{\alpha\sigma_{\mu_\infty}}\right) \phi h_\infty}{\sqrt{\left(1 + \frac{\psi\sigma_q}{\alpha\sigma_{\mu_\infty}}\right)^2 \phi^2 h_\infty^2 + 1 - \lambda_1^2}}.$$

This proposition states that the correlation can be predicted by using our estimated parameters. We compare the simulated correlations with the observed ones.

	Small	Small	Large	Large
	Estimation 1	Estimation 2	Estimation 1	Estimation 2
Data	0.774	0.774	0.842	0.842
Model	0.853	0.863	0.879	0.886
Model without Skill Variation	0.305	0.404	0.278	0.280
Model without Noise	0.923	0.919	0.879	0.886

Table 7: The Correlation Between Relative Productivity and Relative Wages
“Small” includes only companies that report labor and related expenses. “Large” also includes companies for which we have estimated labor costs.

The observed correlation between relative productivity and relative wages is estimated by using a method similar to that used to estimate equation (35).

Table 7 reports the results. The model predicts a slightly higher correlation than the observed one: the model predicts a correlation of between 0.85 and 0.89, whereas the one recorded by the data is between 0.77 and 0.84. However, 0.77 and 0.84 are still high. Hence, the high predicted correlation reasonably captures the feature of the actual correlation.

There is an explanation of why the correlations predicted by the model exceed the observed correlations. To prove Proposition 16, we assume that organization capital is accumulated through learning by doing. However, some workers may be intentionally employed to develop organization capital. As Appendix 2 suggests, if a firm’s maximization problem takes into account the accumulation of organization capital, the wage is related not only to current expected productivity, but also to the market value of a firm. Hence, the predicted correlations would be lower than those reported in Table 7.

Similarly to the previous argument, if we assume that $\sigma_q = 0$ the correlation is between 0.28 and 0.4.⁷ This means that the observed positive correlation between

⁷When $\sigma_q = 0$, the wage is 0. Hence, the result obtained from the model that does not

relative productivity and the relative wage largely arises because of positive assortative assignment between organization capital and the quality of workers. On the other hand, if $h_\infty = 1$ the correlation is slightly larger. This indicates that the noisiness of information contributes little to reducing the correlation.

In summary, according to the theory, a large assignment effect and the generation of fairly accurate information for inferring organization capital from output combine to explain the high observed correlation between labor productivity and wages.

6 Conclusion

In this paper, we developed a theory that accounts for persistent differences in productivity, wages, skill mixes and profits between firms. Firms with a high level of organization capital attract skilled workers, who increase organization capital further. This positive feedback brings about persistent differences in these variables. It is also shown that history can be important for predicting future persistence because organization capital is unobservable. Our model is consistent with empirical evidence. Our simulation results show that positive assortative assignment explains a large proportion of observed persistence.

Some points are worth discussing. We defined organization capital as all the intangible assets embodied in an organization. As explained by Prescott and Visscher (1980), a source of organization capital is firm-specific human capital. Firm-specific human capital may raise an issue not explicitly considered in this paper. Following Atkeson and Kehoe (2005) and Samaniego (2006), in this paper, we assumed that firms receive benefits from organization capital. However, because firm-specific human capital is valuable only to a particular firm, it is difficult to determine who benefits from it.

Fortunately, the dynamics of organization capital are not affected by the explicit

incorporate skill variations, in Table 7, can be interpreted as the correlation when $\sigma_q \approx 0$.

incorporation of firm-specific human capital. In the context of this paper, organization capital is essentially vintage human capital, which is influenced by the history of skilled workers in a firm. Hence, one could consider a model in which a sequence of senior workers develops the organization capital of the firm and in which there is assignment between organization capital and worker skills. Although this would alter the wage function derived in the paper, it would not materially affect persistence. This is because profits and wages would continue to depend on organization capital, and the dynamics of organization capital would continue to be influenced by assignment. Because incorporating an internal labor market raises several separate issues and complicates the model, it would represent an interesting extension of our model.

We were unable to address questions about the dynamics of firm size because our assignment model requires that the number of workers is fixed. However, it would be possible to extend the model in order to analyze firm size. Assuming that assignment between top managers and organization capital determines the total factor productivity (TFP) of a firm, other factors such as physical assets and the number of workers can be derived as functions of TFP. That is, the larger is TFP, the higher are the levels of capital and labor. This approach can be used to generate theoretical predictions about the dynamics of firm size. In this case, assuming that there is learning by doing is questionable and, thus, one would rely on the results in Appendix 2. This interesting extension is left for future research.

Finally, it would be interesting to extend the model to incorporate entry and exit by firms. We ignored entry and exit by firms to focus on effects on the persistence of variables. Incorporating entry and exit would inevitably introduce nonlinearity and make it difficult to find an analytical solution. Hence, one would use computational exercises for this analysis. Because equation (12) implies that a firm's position, relative to the top, is important in an assignment model, one would expect that the cut-off points at the bottom of distribution would not greatly affect the theoretical

prediction of our model. Nonetheless, it would be interesting to examine how assignment affects entry and exit by firms. We plan to investigate this issue in future research.

7 Appendix 1

The Derivation of Equation (22): Because $E[\mu_{kt} E[(\ln k_t^o - \mu_{kt}) | \mu_{kt}]] = 0$, the covariance of μ_{kt} and $\ln k_t^o - \mu_{kt}$ is 0. Hence,

$$\begin{aligned}
\sigma_{\mu t+1}^2 &= \left(\phi + \frac{\gamma \sigma_q}{\sigma_{\mu t}} \right)^2 \sigma_{\mu t}^2 + (\phi h_t)^2 \text{Var}(\ln k_t^o - \mu_{kt}) + \left(\frac{\phi h_t}{\alpha} \right)^2 \sigma_u^2 \\
&= \left(\phi + \frac{\gamma \sigma_q}{\sigma_{\mu t}} \right)^2 \sigma_{\mu t}^2 + (\phi h_t)^2 E[E[(\ln k_t^o - \mu_{kt})^2 | \mu_{kt}]] + \left(\frac{\phi h_t}{\alpha} \right)^2 \sigma_u^2 \\
&= \left(\phi + \frac{\gamma \sigma_q}{\sigma_{\mu t}} \right)^2 \sigma_{\mu t}^2 + (\phi h_t)^2 \sigma_{kt}^2 + \left(\frac{\phi h_t}{\alpha} \right)^2 \sigma_u^2 \\
&= \left(\phi + \frac{\gamma \sigma_q}{\sigma_{\mu t}} \right)^2 \sigma_{\mu t}^2 + \left(\frac{\phi h_t}{\alpha} \right)^2 \frac{\alpha^2 \sigma_{kt}^2}{h_t} \\
&= \left(\phi + \frac{\gamma \sigma_q}{\sigma_{\mu t}} \right)^2 \sigma_{\mu t}^2 + \phi^2 h_t \sigma_{kt}^2.
\end{aligned}$$

The Derivation of Equations (28), (26) and (27): The derivations of equations (26) and (27) follow from Proposition 6. Hence, we only discuss the derivation of equation (28). From Proposition 6, it follows that

$$\frac{\gamma \sigma_q}{\sigma_{\mu \infty}} = \frac{(1 - \phi^2)}{\phi + \sqrt{1 + (1 - \phi^2) \phi^2 h_\infty \left(\frac{\sigma_{k\infty}}{\gamma \sigma_q} \right)^2}}. \quad (36)$$

Note that equation (18) implies $\sigma_{k\infty}^2 = \phi^2 (1 - h_t) \sigma_{k\infty}^2 + \sigma_\varepsilon^2$ in the steady state. Hence, it follows that

$$\sigma_{k\infty}^2 = \frac{\sigma_\varepsilon^2}{1 - \phi^2 (1 - h_\infty)}. \quad (37)$$

Substituting equation (37) into equation (36) yields equation (28).

The Proof of Proposition 11: Because equation (25) implies that

$$\mathbf{k}_t = \sum_{i=1}^j \mathbf{M}^{i-1} \boldsymbol{\xi}_{t-i} + \mathbf{M}^j \mathbf{k}_{t-j},$$

hence, we can show that

$$E[\mathbf{k}_t \mathbf{k}'_{t-j}] = E\left[\left[\sum_{i=1}^j \mathbf{M}^{i-1} \boldsymbol{\xi}_{t-i} + \mathbf{M}^j \mathbf{k}_{t-j}\right] \mathbf{k}'_{t-j}\right] = \mathbf{M}^j E[\mathbf{k}_{t-j} \mathbf{k}'_{t-j}]. \quad (38)$$

Note that Lemma 8 shows that the two eigenvalues of the matrix \mathbf{M} are $\lambda_1 = \phi + \frac{\gamma\sigma_q}{\sigma_{\mu\infty}}$ and $\lambda_2 = \phi(1 - h_\infty)$. Hence, we can easily calculate the following corresponding eigenvectors:

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix} z_1, \forall z_1 \in \mathbf{R} \text{ for } \lambda_1, \begin{bmatrix} -\frac{\gamma\sigma_q}{\sigma_{\mu\infty}} \\ \phi h_\infty \end{bmatrix} z_2, \forall z_2 \in \mathbf{R} \text{ for } \lambda_2.$$

Define two matrices, $\mathbf{\Lambda}$ and \mathbf{Z} :

$$\mathbf{\Lambda} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}, \mathbf{Z} = \begin{bmatrix} z_1 & -\frac{\gamma\sigma_q}{\sigma_{\mu\infty}} z_2 \\ z_1 & \phi h_\infty z_2 \end{bmatrix}, \forall z_1, z_2 \in \mathbf{R}. \quad (39)$$

Then, the matrix \mathbf{M} can be decomposed into $\mathbf{\Lambda}$ and \mathbf{Z} : $\mathbf{M} = \mathbf{Z}\mathbf{\Lambda}\mathbf{Z}^{-1}$. Hence, it follows from equation (38) that

$$E[\mathbf{k}_t \mathbf{k}'_{t-j}] = \mathbf{Z}\mathbf{\Lambda}^j \mathbf{Z}^{-1} E[\mathbf{k}_{t-j} \mathbf{k}'_{t-j}],$$

$$\text{where } \mathbf{Z}\mathbf{\Lambda}^j \mathbf{Z}^{-1} = \frac{1}{\phi h_\infty + \frac{\gamma\sigma_q}{\sigma_{\mu\infty}}} \begin{bmatrix} \phi h_\infty \lambda_1^j + \frac{\gamma\sigma_q}{\sigma_{\mu\infty}} \lambda_2^j & \frac{\gamma\sigma_q}{\sigma_{\mu\infty}} (\lambda_1^j - \lambda_2^j) \\ \phi h_\infty (\lambda_1^j - \lambda_2^j) & \frac{\gamma\sigma_q}{\sigma_{\mu\infty}} \lambda_1^j + \phi h_\infty \lambda_2^j \end{bmatrix}. \quad (40)$$

Because we can easily calculate

$$E[\mathbf{k}_{t-j} \mathbf{k}'_{t-j}] = \begin{bmatrix} \sigma_{\mu\infty}^2 + \sigma_{k\infty}^2 & \sigma_{\mu\infty}^2 \\ \sigma_{\mu\infty}^2 & \sigma_{\mu\infty}^2 \end{bmatrix},$$

an autocovariance matrix can be calculated as follows:

$$E [\mathbf{k}_t \mathbf{k}'_{t-j}] = \begin{bmatrix} \lambda_1^j \sigma_{\mu\infty}^2 + \left[\frac{\phi h_\infty \lambda_1^j + \frac{\gamma \sigma_q}{\sigma_{\mu\infty}} \lambda_2^j}{\phi h_\infty + \frac{\gamma \sigma_q}{\sigma_{\mu\infty}}} \right] \sigma_{k\infty}^2, & \lambda_1^j \sigma_{\mu\infty}^2 \\ \frac{\phi h_\infty}{\phi h_\infty + \frac{\gamma \sigma_q}{\sigma_{\mu\infty}}} (\lambda_1^j - \lambda_2^j) \sigma_{k\infty}^2 + \lambda_1^j \sigma_{\mu\infty}^2, & \lambda_1^j \sigma_{\mu\infty}^2 \end{bmatrix}.$$

By using the first row and first column element $E [\mathbf{k}_t \mathbf{k}'_{t-j}]$ and the second row and second column element of $E [\mathbf{k}_t \mathbf{k}'_{t-j}]$, the following autocorrelations of $\ln k_t^o$ and μ_{kt} are derived:

$$\begin{aligned} \rho_{kj} &= (1 - \omega) \lambda_1^j + \omega \lambda_2^j, \quad \rho_{\mu j} = \lambda_1^j, \\ \omega &\equiv \frac{\frac{\gamma \sigma_q}{\sigma_{\mu\infty}}}{(\lambda_1 - \lambda_2) \left(1 + \left(\frac{\sigma_{\mu\infty}}{\sigma_{k\infty}} \right)^2 \right)}. \end{aligned}$$

Note that equation (22) implies $\sigma_{\mu\infty}^2 = (\phi \sigma_{\mu\infty} + \gamma \sigma_q)^2 + \phi^2 h_\infty \sigma_{k\infty}^2$ in the steady state.

This means that

$$\left(\frac{\sigma_{\mu\infty}}{\gamma \sigma_q} \right)^2 = \lambda_1^2 \left(\frac{\sigma_{\mu\infty}}{\gamma \sigma_q} \right)^2 + \phi^2 h_\infty \left(\frac{\sigma_{k\infty}}{\sigma_{\mu\infty}} \right)^2 \left(\frac{\sigma_{\mu\infty}}{\gamma \sigma_q} \right)^2.$$

Hence, it follows from this equation that

$$\left(\frac{\sigma_{\mu\infty}}{\sigma_{k\infty}} \right)^2 = \frac{\phi^2 h_\infty}{1 - \lambda_1^2}. \quad (41)$$

Substituting this equation into the definition of ω , we have

$$\omega = \frac{\frac{\gamma \sigma_q}{\sigma_{\mu\infty}} (1 - \lambda_1^2)}{(\lambda_1 - \lambda_2) (1 - \lambda_1^2 + \phi^2 h_\infty)}.$$

The Proof of Proposition 13: The impacts on $\rho_{\mu j}$ are clear. Hence, we only derive the effects on $\rho_{\ln kj}$.

$$\begin{aligned} \frac{d\rho_{\ln kj}}{d\left(\frac{\gamma \sigma_q}{\sigma_\varepsilon}\right)} &= \lambda_1^{j-1} \left[j(1 - \omega) - \left(1 - \left(\frac{\lambda_2}{\lambda_1} \right)^j \right) \lambda_1 \frac{d\omega}{d\left(\frac{\gamma \sigma_q}{\sigma_{\mu\infty}}\right)} \right] \Sigma_1 \left(\frac{\gamma \sigma_q}{\sigma_\varepsilon}, h_\infty \right) \\ &> \lambda_1^{j-1} \left[j(1 - \omega) - \left| \lambda_1 \frac{d\omega}{d\left(\frac{\gamma \sigma_q}{\sigma_{\mu\infty}}\right)} \right| \right] \Sigma_1 \left(\frac{\gamma \sigma_q}{\sigma_\varepsilon}, h_\infty \right) \end{aligned}$$

Hence, there exists a j^* such that $\frac{d\rho_{\ln k_j}}{d\left(\frac{\gamma\sigma_q}{\sigma_\varepsilon}\right)} > 0, \forall j \geq j^*$. Similarly

$$\begin{aligned} \frac{d\rho_{\ln k_j}}{dh_\infty} &= \lambda_1^{j-1} \left[j(1-\omega) \Sigma_2 \left(\frac{\gamma\sigma_q}{\sigma_\varepsilon}, h_\infty \right) - \lambda_1 \left(\frac{d\omega}{d\frac{\gamma\sigma_q}{\sigma_{\mu_\infty}}} \Sigma_2 \left(\frac{\gamma\sigma_q}{\sigma_\varepsilon}, h_\infty \right) + \frac{d\omega}{dh_\infty} \right) \right] \\ &\quad + \lambda_2^{j-1} \left\{ \left(\frac{d\omega}{d\frac{\gamma\sigma_q}{\sigma_{\mu_\infty}}} \Sigma_2 \left(\frac{\gamma\sigma_q}{\sigma_\varepsilon}, h_\infty \right) + \frac{d\omega}{dh_\infty} \right) \lambda_2 - j\omega\phi \right\}. \end{aligned}$$

Hence, there exists a j^{**} such that $\frac{d\rho_{k_j}}{dh_\infty} < 0, \forall j \geq j^{**}$.

The Derivation of Equations (30) and (31): Substituting equation (13) into equation (9) yields, in the steady state,

$$D \ln y_t = \alpha (\ln k_t^o - \mu_{kt}^e) + \frac{\psi\sigma_q}{\sigma_{\mu_\infty}} D\mu_{kt} + u_t + \frac{\sigma_u^2}{2}, \quad (42)$$

$$E [D \ln y_t | \mu_{kt}] = \left(\alpha + \frac{\psi\sigma_q}{\sigma_{\mu_\infty}} \right) D\mu_{kt}. \quad (43)$$

We define the vectors \mathbf{y}_t and $\boldsymbol{\zeta}_t$ and a matrix \mathbf{G} such that

$$\mathbf{y}_t = \begin{bmatrix} D \ln y_t \\ E [D \ln y_t | \mu_{kt}] \end{bmatrix}, \quad \mathbf{G} = \begin{bmatrix} \alpha, & \frac{\psi\sigma_q}{\sigma_{\mu_\infty}} \\ 0, & \alpha + \frac{\psi\sigma_q}{\sigma_{\mu_\infty}} \end{bmatrix}, \quad \boldsymbol{\zeta}_t = \begin{bmatrix} u_t + \frac{\sigma_u^2}{2} \\ 0 \end{bmatrix}. \quad (44)$$

Then,

$$\mathbf{y}_t = \mathbf{G}\mathbf{k}_t + \boldsymbol{\zeta}_t.$$

Combining this equation and equation (25) yields

$$\mathbf{y}_{t+1} = \mathbf{G}\mathbf{M}\mathbf{G}^{-1}\mathbf{y}_t + \mathbf{G}\boldsymbol{\xi}_t + \boldsymbol{\zeta}_{t+1} - \mathbf{G}\mathbf{M}\mathbf{G}^{-1}\boldsymbol{\zeta}_t. \quad (45)$$

The results follow immediately because it is easy to derive

$$\begin{aligned} \mathbf{G}\mathbf{M}\mathbf{G}^{-1} &= \mathbf{M} + \frac{\psi}{\alpha} \frac{\sigma_q}{\sigma_{\mu_\infty}} \phi h_\infty \mathbf{J}, \text{ where } \mathbf{J} = \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix}, \quad (46) \\ \mathbf{G}\boldsymbol{\xi}_t + \boldsymbol{\zeta}_{t+1} - \mathbf{G}\mathbf{M}\mathbf{G}^{-1}\boldsymbol{\zeta}_t &= \begin{bmatrix} \alpha \left(\varepsilon_t + \frac{\sigma_\varepsilon^2}{2} \right) - \phi \left(u_t + \frac{\sigma_u^2}{2} \right) + u_{t+1} + \frac{\sigma_u^2}{2} \\ 0 \end{bmatrix}. \end{aligned}$$

The Proof of Proposition 15: We define vectors \mathbf{y}_t and $\boldsymbol{\zeta}_t$ and a matrix \mathbf{G} in the same way as we defined equation (44). Using equation (45), we find

$$\mathbf{y}_t = \sum_{i=1}^{j-1} (\mathbf{G}\mathbf{M}\mathbf{G}^{-1})^{i-1} \mathbf{v}_{t-i} + (\mathbf{G}\mathbf{M}\mathbf{G}^{-1})^{j-1} \mathbf{y}_{t-(j-1)},$$

where $\mathbf{v}_{t-i} = \mathbf{G}\boldsymbol{\xi}_t + \boldsymbol{\zeta}_{t+1} - \mathbf{G}\mathbf{M}\mathbf{G}^{-1}\boldsymbol{\zeta}_t$. Hence the covariance matrix is

$$E(\mathbf{y}_t \mathbf{y}_{t-j}^T) = (\mathbf{G}\mathbf{M}\mathbf{G}^{-1})^{j-1} E(\mathbf{y}_{t-(j-1)} \mathbf{y}_{t-j}^T) = (\mathbf{G}\mathbf{M}\mathbf{G}^{-1})^{j-1} E(\mathbf{y}_{t+1} \mathbf{y}_t^T).$$

We define matrices $\boldsymbol{\Lambda}$ and \mathbf{Z} in the same way as we defined equation (39). Then, as stated in the proof of Proposition 11, $\mathbf{M} = \mathbf{Z}\boldsymbol{\Lambda}\mathbf{Z}^{-1}$. Hence,

$$(\mathbf{G}\mathbf{M}\mathbf{G}^{-1})^{j-1} = (\mathbf{G}\mathbf{Z}\boldsymbol{\Lambda}\mathbf{Z}^{-1}\mathbf{G}^{-1})^{j-1} = \mathbf{Z}_y \boldsymbol{\Lambda}^{j-1} \mathbf{Z}_y^{-1},$$

where

$$\mathbf{Z}_y = \mathbf{G}\mathbf{Z} = \begin{bmatrix} z_1^*, & \left(\frac{\gamma\sigma_q}{\sigma_{\mu\infty}} - \frac{\psi\sigma_q}{\alpha\sigma_{\mu\infty}} \phi h_\infty \right) z_2^* \\ z_1^*, & -\left(\phi h_\infty + \frac{\psi\sigma_q}{\alpha\sigma_{\mu\infty}} \phi h_\infty \right) z_2^* \end{bmatrix} \quad \forall z_1^*, z_2^* \in \mathbf{R}.$$

Hence, we can show that

$$(\mathbf{G}\mathbf{M}\mathbf{G}^{-1})^{j-1} = \mathbf{Z}\boldsymbol{\Lambda}^{j-1}\mathbf{Z}^{-1} + \frac{(\lambda_1^{j-1} - \lambda_2^{j-1})}{\phi h_\infty + \frac{\gamma\sigma_q}{\sigma_{\mu\infty}}} \frac{\psi\sigma_q}{\alpha\sigma_{\mu\infty}} \phi h_\infty \mathbf{J}, \quad (47)$$

where $\mathbf{Z}\boldsymbol{\Lambda}^j\mathbf{Z}^{-1}$ is defined in equation (40) and \mathbf{J} is defined in equation (46). Next, we calculate $E[\mathbf{y}_{t+1}\mathbf{y}_t^T]$. Because $\mathbf{y}_{t+1} = \mathbf{G}\mathbf{M}\mathbf{G}^{-1}\mathbf{y}_t + \mathbf{v}_t$, $E[\mathbf{y}_{t+1}\mathbf{y}_t^T]$ can be decomposed into $E(\mathbf{y}_t \mathbf{y}_t^T)$ and $E(\mathbf{v}_t \mathbf{y}_t^T)$.

$$E(\mathbf{y}_{t+1}\mathbf{y}_t^T) = \mathbf{G}\mathbf{M}\mathbf{G}^{-1}E(\mathbf{y}_t \mathbf{y}_t^T) + E(\mathbf{v}_t \mathbf{y}_t^T)$$

Because $E(\mathbf{y}_t \mathbf{y}_t^T)$ and $E(\mathbf{v}_t \mathbf{y}_t^T)$ are calculated as

$$E(\mathbf{y}_t \mathbf{y}_t^T) = \begin{bmatrix} \left(1 + \frac{\psi\sigma_q}{\alpha\sigma_{\mu\infty}}\right)^2 \alpha^2 \sigma_{\mu\infty}^2 + \alpha^2 \sigma_{k\infty}^2 + \sigma_u^2, & \left(1 + \frac{\psi\sigma_q}{\alpha\sigma_{\mu\infty}}\right)^2 \alpha^2 \sigma_{\mu\infty}^2 \\ \left(1 + \frac{\psi\sigma_q}{\alpha\sigma_{\mu\infty}}\right)^2 \alpha^2 \sigma_{\mu\infty}^2, & \left(1 + \frac{\psi\sigma_q}{\alpha\sigma_{\mu\infty}}\right)^2 \alpha^2 \sigma_{\mu\infty}^2 \end{bmatrix}, \quad (48)$$

$$E(\mathbf{v}_t \mathbf{y}_t^T) = \begin{bmatrix} -\phi\sigma_u^2, & 0 \\ 0, & 0 \end{bmatrix}, \quad (49)$$

we can derive

$$E(\mathbf{y}_{t+1}\mathbf{y}_t^T) = \alpha^2 \left(1 + \frac{\psi\sigma_q}{\alpha\sigma_{\mu\infty}}\right) \begin{bmatrix} \left(1 + \frac{\psi\sigma_q}{\alpha\sigma_{\mu\infty}}\right) \lambda_1 \sigma_{\mu\infty}^2 + \phi \sigma_{k\infty}^2, & \left(1 + \frac{\psi\sigma_q}{\alpha\sigma_{\mu\infty}}\right) \lambda_1 \sigma_{\mu\infty}^2 \\ \left(1 + \frac{\psi\sigma_q}{\alpha\sigma_{\mu\infty}}\right) \lambda_1 \sigma_{\mu\infty}^2 + \phi \sigma_{k\infty}^2, & \left(1 + \frac{\psi\sigma_q}{\alpha\sigma_{\mu\infty}}\right) \lambda_1 \sigma_{\mu\infty}^2 \end{bmatrix}. \quad (50)$$

Hence, equations (47) and (50) can be used to show that

$$E(\mathbf{y}_t \mathbf{y}_{t-j}^T) = \alpha^2 \sigma_{k\infty}^2 \left(1 + \frac{\psi\sigma_q}{\alpha\sigma_{\mu\infty}}\right) \lambda_1^{j-1} \begin{bmatrix} \left(1 + \frac{\psi\sigma_q}{\alpha\sigma_{\mu\infty}}\right) \lambda_1 \left(\frac{\sigma_{\mu\infty}}{\sigma_{k\infty}}\right)^2 + \phi, & \left(1 + \frac{\psi\sigma_q}{\alpha\sigma_{\mu\infty}}\right) \lambda_1 \left(\frac{\sigma_{\mu\infty}}{\sigma_{k\infty}}\right)^2 \\ \left(1 + \frac{\psi\sigma_q}{\alpha\sigma_{\mu\infty}}\right) \lambda_1 \left(\frac{\sigma_{\mu\infty}}{\sigma_{k\infty}}\right)^2 + \phi, & \left(1 + \frac{\psi\sigma_q}{\alpha\sigma_{\mu\infty}}\right) \lambda_1 \left(\frac{\sigma_{\mu\infty}}{\sigma_{k\infty}}\right)^2 \end{bmatrix}.$$

Using the first row and first column of this covariance matrix, we can derive the autocorrelation between $\ln y_t$ and $\ln y_{t-j}$,

$$\rho_{\ln yj} = \frac{\left(1 + \frac{\psi\sigma_q}{\alpha\sigma_{\mu\infty}}\right) \lambda_1^{j-1} \left[\left(1 + \frac{\psi\sigma_q}{\alpha\sigma_{\mu\infty}}\right) \lambda_1 \left(\frac{\sigma_{\mu\infty}}{\sigma_{k\infty}}\right)^2 + \phi\right]}{\left(1 + \frac{\psi\sigma_q}{\alpha\sigma_{\mu\infty}}\right)^2 \left(\frac{\sigma_{\mu\infty}}{\sigma_{k\infty}}\right)^2 + 1 + \left(\frac{\sigma_u}{\alpha\sigma_{k\infty}}\right)^2}.$$

Note that equation (41) implies $\left(\frac{\sigma_{\mu\infty}}{\sigma_{k\infty}}\right)^2 = \frac{\phi^2 h_\infty}{1 - \lambda_1^2}$ and that equation (26) implies $\left(\frac{\sigma_u}{\alpha\sigma_{k\infty}}\right)^2 = \frac{1 - h_\infty}{h_\infty}$. Hence, $\rho_{\ln yj}$ can be rewritten as

$$\rho_{\ln yj} = \frac{\phi h_\infty \left(1 + \frac{\psi\sigma_q}{\alpha\sigma_{\mu\infty}}\right) \lambda_1^{j-1} \left[\phi h_\infty \left(1 + \frac{\psi\sigma_q}{\alpha\sigma_{\mu\infty}}\right) \lambda_1 + 1 - \lambda_1^2\right]}{\left(1 + \frac{\psi\sigma_q}{\alpha\sigma_{\mu\infty}}\right)^2 (\phi h_\infty)^2 + 1 - \lambda_1^2}.$$

Similarly, using the second row and second column of the covariance matrix, we can derive the autocorrelation between $E[\ln y_t | \mu_{kt}]$ and $E[\ln y_{t-j} | \mu_{kt-j}]$,

$$\rho_{E[\ln y | \mu_k]j} = \lambda_1^j.$$

The Proof of Proposition 16: Proposition 14 implies $\ln \hat{w}(\hat{\chi}(\mu_{kt})) - E[\ln \hat{w}(\hat{\chi}(\mu_{kt}))] = E[\ln y_t | \mu_{kt}] - E[\ln y_t]$. Hence,

$$\rho_{\ln y \ln w} = \frac{Cov[[\ln y_t - E[\ln y_t]] E[\ln y_t | \mu_{kt}] - E[\ln y_t]]}{\sigma_y \sigma_{E[y|s]}}.$$

Using equation (48), it is easy to show that

$$\rho_{\ln y \ln w} = \frac{\left(1 + \frac{\psi \sigma_q}{\alpha \sigma_{\mu_\infty}}\right) \frac{\sigma_{\mu_\infty}}{\sigma_{k_\infty}}}{\sqrt{\left(1 + \frac{\psi \sigma_q}{\alpha \sigma_{\mu_\infty}}\right)^2 \left(\frac{\sigma_{\mu_\infty}}{\sigma_{k_\infty}}\right)^2 + 1 + \left(\frac{\sigma_u}{\alpha \sigma_{k_\infty}}\right)^2}}.$$

Note that equation (41) implies $\left(\frac{\sigma_{\mu_\infty}}{\sigma_{k_\infty}}\right)^2 = \frac{\phi^2 h_\infty}{1 - \lambda_1^2}$ and that equation (26) implies $\left(\frac{\sigma_u}{\alpha \sigma_{k_\infty}}\right)^2 = \frac{1 - h_\infty}{h_\infty}$. Hence, ρ_{yw} can be rewritten as

$$\rho_{\ln y \ln w} = \frac{\left(1 + \frac{\psi \sigma_q}{\alpha \sigma_{\mu_\infty}}\right) \phi h_\infty}{\sqrt{\left(1 + \frac{\psi \sigma_q}{\alpha \sigma_{\mu_\infty}}\right)^2 \phi^2 h_\infty^2 + 1 - \lambda_1^2}}.$$

Data Appendix: ($\#X$)_{ft} implies COMPUSTAT number X of f th firm in year t and ($\#X$)_{fit} implies COMPUSTAT number X of f th firm in i th industry in year t .)

- Selection of data: We used industry annual data from 1970 to 2004 from COMPUSTAT. However, because we constructed initial priors for each firm by using the initial five annual observations in COMPUSTAT, our regression is based on data for 1975–2004. We deleted observations for which either the estimated wage or value added was negative and deleted those for which the labor share exceeded unity. This was because such observations are not consistent with the model’s assumptions. Because we are interested in deviations from the industry average, we retain industries that have at least five firms throughout the years for which data are available. Industries are classified based on four-digit industry codes.
- Total expenses are defined as $(\#41)_{ft} + (\#189)_{ft}$, where $(\#41)_{ft}$ is the cost of goods sold and $(\#189)_{ft}$ measures administrative, selling and general expenses.
- Labor expenses: If a firm reports labor and related expenses, $(\#42)_{ft}$, that includes employee benefits, we use this as our measure of labor expenses. The

small sample comprises these firms. Otherwise, we estimate labor expenses as follows. First, if a firm reports labor and related expenses that exclude employee benefits, we replace labor expenses by

$$\left[\frac{\frac{\sum_{f \in Y_t} (\#42)_{ft} / (\#29)_{ft}}{n_{Y_t}}}{\frac{\sum_{f \in X_t} (\#42)_{ft} / (\#29)_{ft}}{n_{X_t}}} \right] (\#42)_{ft}, \forall t,$$

where $(\#29)_{ft}$ is the number of workers in the f th firm in year t and Y_t is the set of firms that includes employee benefits for year t , X_t is the set of firms that exclude employee benefits for year t , n_{Y_t} is the number of firms in set Y_t and n_{X_t} is the number of firms in set X_t . This is an estimate of labor and related expenses that includes employee benefits. Second, if a firm does not report labor and related expenses, we estimate these expenses by

$$\left[\frac{\sum_{f \in Z_{it}} (\#42)_{fit} / [(\#41)_{fit} + (\#189)_{fit}]}{n_{Z_{it}}} \right] [(\#41)_{fit} + (\#189)_{fit}], \forall t, i,$$

where Z_{it} is the set of firms that report labor and related expenses in the i th industry in year t and $n_{Z_{it}}$ is the number of firms in set Z_{it} . Note that $(\#41)_{fit} + (\#189)_{fit}$ is defined as total expenses. This is our estimate of labor expenses for firms in the large sample.

- y_{ft} : Value added divided by the number of employees $(\#29)_{ft}$. Value added is measured as sales $(\#12)_{ft}$ minus the value of materials, which is total expenses minus labor expenses.
- w_{ft} : Labor expenses divided by the number of employees $(\#29)_{ft}$.
- π_{ft} : Operating income $(\#13)_{ft}$ divided by the number of employees $(\#29)_{ft}$.
- k_{ft} : Total net value of property and plant and equipment at the end of the previous year $(\#8)_{ft-1}$ divided by the number of employees $(\#29)_{ft}$. Hence, we approximate the initial capital stock by using the value at the end of the previous year.

		y_{it}	w_{it}	π_{it}	k_{it}
1990	Mean	82.9	34.8	48.1	32.9
	Std. Dev.	120	14.0	110	29.9
	Median	67.6	31.4	36.9	27.9
	# of Obs	142	142	142	141
2000	Mean	145	52.2	92.8	47.1
	Std. Dev.	39.8	14.0	31.6	22.7
	Median	139	49.1	84.2	41.5
	# of Obs	86	86	86	85

Table 8: Summary Statistics for the Small Sample

		y_{it}	w_{it}	π_{it}	k_{it}
1990	Mean	77.2	33.8	43.4	103
	Std. Dev.	366	45.7	360	338
	Median	48.0	28.4	14.4	25.4
	# of Obs	1371	1371	1371	1353
2000	Mean	153	59.1	94.0	221
	Std. Dev.	533	169	488	1141
	Median	75.3	41.2	26.5	32.7
	# of Obs	1561	1561	1558	1524

Table 9: Summary Statistics for the Large Sample

The nominal value is reported. Dollar figures are in thousands.

8 Appendix 2

In this appendix, we show that even if a firm takes into account the dynamics of organization capital for its maximization problem, the dynamics of organization capital do not change. However, the wage function and the value function are affected.

We define a recursive positive assortative equilibrium with unobserved organization capital. To formally define the equilibrium, we distinguish prevailing uncertainty about $\ln k_t^o$, σ_{kt}^a from a firm's uncertainty about $\ln k_t^o$, σ_{kt} , which are, however, equal in equilibrium. Let $\mathbf{x}_t = (\mu_{kt}^e, \sigma_{\mu t}, \sigma_{kt}^a)^T$ denote the vector of these aggregate state variables. The firm's problem can be rewritten as

$$V^*(\mu_{kt}, \sigma_{kt} : \mathbf{x}_t) = \max_{\ln q_t} \left\{ \begin{array}{l} E[y_t | \mu_{kt}, \sigma_{kt}^2, \ln q_t] - w^*(\ln q_t : \mathbf{x}_t) \\ + \beta \int V^*(\mu_{kt+1}, \sigma_{kt+1} : \mathbf{x}_{t+1}) d\Gamma_s(s_t | \mu_{kt}, \sigma_{kt}) \end{array} \right\}, \quad (51)$$

$$\begin{aligned} s.t. E[y_t | \mu_{kt}, \sigma_{kt}^2, \ln q_t] &= \exp\left(\ln A + \alpha \mu_{kt} + \frac{\alpha^2 \sigma_{kt}^2}{2} + \psi \ln q_t\right), \\ \mu_{kt+1} &= \ln B + \phi[(1 - h_t) \mu_{kt} + h_t s_t] + \gamma \ln q_t - \frac{\sigma_\varepsilon^2}{2}, \\ \sigma_{kt+1} &= \sqrt{\phi^2(1 - h_t) \sigma_{kt}^2 + \sigma_\varepsilon^2}, \\ \mu_{kt+1}^e &= f(\mathbf{x}_t), \sigma_{\mu t+1} = g(\mathbf{x}_t), \sigma_{kt+1}^a = m(\mathbf{x}_t), \end{aligned}$$

where $\Gamma_s(s_t | \mu_{kt}, \sigma_{kt})$ is a conditional distribution function of a signal s_t given μ_{kt} and σ_{kt} , and functions $f(\cdot)$, $g(\cdot)$ and $m(\cdot)$ represent firms' expectations about the transition of the aggregate state variables. Our candidate equilibrium is a positive assortative equilibrium. Hence, similarly to the static problem, the policy function must satisfy

$$\chi^*(\mu_{kt}, \sigma_{kt} : \mathbf{x}_t) = \frac{\sigma_q}{\sigma_{\mu t}} [\mu_{kt} - \mu_{kt}^e] + \mu_q.$$

Given this policy function, the dynamics of μ_{kt+1}^e and $\sigma_{\mu t+1}$ are

$$\begin{aligned} \mu_{kt+1}^e &= \ln B + \phi \mu_{kt}^e + \gamma \mu_q - \frac{\sigma_\varepsilon^2}{2}, \\ \sigma_{\mu t+1} &= \sqrt{(\phi \sigma_{\mu t} + \gamma \sigma_q)^2 + \phi^2 h_t (\sigma_{kt}^a)^2}. \end{aligned}$$

Hence, a recursive positive assortative equilibrium with unobserved organization capital is defined as follows.

Definition 17 *A recursive positive assortative equilibrium with unobserved organization capital comprises values of $\chi^*(\mu_{kt}, \sigma_{kt} : \mathbf{x}_t)$, $V^*(\mu_{kt}, \sigma_{kt} : \mathbf{x}_t)$, $w^*(q_t : \mathbf{x}_t)$, $f(\mathbf{x}_t)$, $g(\mathbf{x}_t)$ and $m(\mathbf{x}_t)$ that satisfy the following conditions.*

1. *An individual firm solves its maximization problem (51).*
2. *The labor market is cleared:*

$$\chi^*(\mu_{kt}, \sigma_{kt} : \mathbf{x}_t) = \frac{\sigma_q}{\sigma_{\mu t}} [\mu_{kt} - \mu_{kt}^e] + \mu_q.$$

3. *Expectations are rational:*

$$f(\mathbf{x}_t) = \ln B + \phi \mu_{kt}^e + \gamma \mu_q - \frac{\sigma_\varepsilon^2}{2}, \quad (52)$$

$$g(\mathbf{x}_t) = \sqrt{(\phi \sigma_{\mu t} + \gamma \sigma_q)^2 + \phi^2 h_t (\sigma_{kt}^a)^2}, \quad (53)$$

$$m(\mathbf{x}_t) = \sqrt{\phi^2 (1 - h_t) (\sigma_{kt}^a)^2 + \sigma_\varepsilon^2}, \quad (54)$$

where

$$h_t = \frac{\left(\frac{\alpha \sigma_{kt}^a}{\sigma_u}\right)^2}{1 + \left(\frac{\alpha \sigma_{kt}^a}{\sigma_u}\right)^2}, \sigma_{kt}^a = \sigma_{kt}^2.$$

Because $\sigma_{kt}^a = \sigma_{kt}$ in equilibrium, σ_{kt} is used subsequently to denote the aggregate state variable. Note that if this equilibrium exists, the dynamics of k_t^o and μ_{kt} are the same as those presented in the main text. Hence, our main results are unaffected. Because the labor- market clearing condition and rational expectations assumption determine the functions $\chi^*(\mu_{kt}, \sigma_{kt} : \mathbf{x}_t)$, $f(\mathbf{x}_t)$, $g(\mathbf{x}_t)$ and $m(\mathbf{x}_t)$, we need to find a value function and a wage function that are consistent with the definition of equilibrium. The next theorem derives the value function and the wage function.

Theorem 18 Suppose that $\phi \in [0, 1)$, $\alpha \in (0, \infty)$, $\psi \in (0, \infty)$, $\gamma \in (0, \infty)$, $\sigma_\varepsilon \in [0, \infty)$, $\sigma_q \in [0, \infty)$ and $\sigma_u \in [0, \infty)$. There exists a unique recursive positive assortative equilibrium with unobserved organization capital. In equilibrium, the value function and the wage function are

$$V^*(\mu_{kt}, \sigma_{kt} : \mathbf{x}_t) = \sum_{i=0}^{\infty} \prod_{s=1}^i \frac{\beta\phi}{\phi + \frac{\gamma\sigma_q}{\sigma_{\mu t+s-1}}} \frac{E[y_{t+i} | \mu_{kt}, \sigma_{kt}]}{1 + \frac{\psi\sigma_q}{\alpha\sigma_{\mu t+i}}}, \quad (55)$$

$$w^*(\ln q_t : \mathbf{x}_t) = \frac{\frac{\psi\sigma_q}{\alpha\sigma_{\mu t}} E[y_t | \mu_{kt}, \sigma_{kt}]}{1 + \frac{\psi\sigma_q}{\alpha\sigma_{\mu t}}} \Big|_{\mu_{kt} = \frac{\sigma_{\mu t}}{\sigma_q} (\ln q_t - \mu_q) + \mu_{kt}^e} \quad (56)$$

$$+ \frac{\beta\frac{\gamma\sigma_q}{\sigma_{\mu t}}}{\phi + \frac{\gamma\sigma_q}{\sigma_{\mu t}}} \int V(\mu_{kt+1}, \sigma_{kt+1} : \mathbf{x}_{t+1}) d\Gamma_s(s_t | \mu_{kt}, \sigma_{kt}) \Big|_{\mu_{kt} = \frac{\sigma_{\mu t}}{\sigma_q} (\ln q_t - \mu_q) + \mu_{kt}^e},$$

where $\prod_{s=1}^0 \frac{\beta\phi}{\phi + \frac{\gamma\sigma_q}{\sigma_{\mu t+s-1}}} = 1$ and

$$E[y_{t+i} | \mu_{kt}, \sigma_{kt}] = \exp \left[\begin{array}{c} \ln A + \psi \left(\mu_q - \frac{\sigma_q}{\sigma_{\mu t+i}} \mu_{kt+i}^e \right) + \frac{\alpha^2 \sigma_{kt+i}^2}{2} + \\ \left(\alpha + \frac{\psi\sigma_q}{\sigma_{\mu t+i}} \right) E[\mu_{kt+i} | \mu_{kt}, \sigma_{kt}] + \frac{\left(\alpha + \frac{\psi\sigma_q}{\sigma_{\mu t+i}} \right)^2}{2} \text{Var}[\mu_{kt+i} | \mu_{kt}, \sigma_{kt}] \end{array} \right],$$

$$E[\mu_{kt+i} | \mu_{kt}, \sigma_{kt}^2] = \mu_{kt+i}^e + \prod_{\tau=1}^i \left(\phi + \frac{\gamma\sigma_q}{\sigma_{\mu t+i-\tau}} \right) (\mu_{kt} - \mu_{kt}^e),$$

$$\text{Var}[\mu_{kt+i} | \mu_{kt}, \sigma_{kt}^2] = \sum_{\tau=1}^i \prod_{s=1}^{\tau-1} \left(\phi + \frac{\gamma\sigma_q}{\sigma_{\mu t+i-s}} \right)^2 \phi^2 h_{t+i-\tau} \sigma_{kt+i-\tau}^2,$$

and

$$\mu_{kt+i}^e = \frac{1 - \phi^i}{1 - \phi} \left(\ln B + \gamma\mu_q - \frac{\sigma_\varepsilon^2}{2} \right) + \phi^{i-1} \mu_{kt}^e,$$

$$\sigma_{\mu t+i} = g^i(\mathbf{x}_t) \equiv g \left((\mu_{kt+i-1}^e, g^{i-1}(\mathbf{x}_t), m^{i-1}(\mathbf{x}_t))^T \right),$$

$$\sigma_{kt+i}^2 = m^i(\mathbf{x}_t) \equiv m \left((\mu_{kt+i-1}^e, g^{i-1}(\mathbf{x}_t), m^{i-1}(\mathbf{x}_t))^T \right),$$

and

$$g^1(\mathbf{x}_t) = g(\mathbf{x}_t) = \sqrt{(\phi\sigma_{\mu t} + \gamma\sigma_q)^2 + \phi^2 h_t (\sigma_{kt}^a)^2},$$

$$m^1(\mathbf{x}_t) = m(\mathbf{x}_t) = \sqrt{\phi^2 (1 - h_t^a) (\sigma_{kt}^a)^2 + \sigma_\varepsilon^2}.$$

Proof. Consider a mapping, T

$$TV^* = \max_{\ln q_t} \left\{ \begin{array}{l} E[y_t | \mu_{kt}, \sigma_{kt}, \ln q_t] - w^*(\ln q_t : \mathbf{x}_t) \\ + \beta \int V^*(\mu_{kt+1}, \sigma_{kt+1} : \mathbf{x}_{t+1}) d\Gamma_s(s_t | \mu_{kt}, \sigma_{kt}) \end{array} \right\},$$

$$E[y_t | \mu_{kt}, \sigma_{kt}, \ln q_t] = \exp \left(\log A + \psi \ln q_t + \alpha \mu_{kt} + \frac{\alpha^2 \sigma_{kt}^2}{2} \right),$$

$$\mu_{kt+1} = \ln B + \phi [(1 - h_t) \mu_{kt} + h_t s_t] + \gamma \ln q_t - \frac{\sigma_\varepsilon^2}{2}.$$

Suppose that the value function and the wage function are represented by equations (55) and (56). Define

$$MPQ(\mu_{kt}, \sigma_{kt}, \ln q_t) = \frac{dE[y_t | \mu_{kt}, \sigma_{kt}, \ln q_t]}{d \ln q_t} + \beta \int \frac{dV^*(\mu_{kt+1}, \sigma_{kt+1} : \mathbf{x}_{t+1})}{d \ln q_t} d\Gamma_s(s_t | \mu_{kt}, \sigma_{kt}).$$

Then, we can derive

$$w^{*'}(\ln q_t : \mathbf{x}_t) = MPQ \left(\frac{\sigma_{\mu t}}{\sigma_q} (\ln q_t - \mu_q) + \mu_{kt}^e, \sigma_{kt}, \ln q_t \right).$$

Hence, when a firm is endowed with $\mu_{kt} = \chi^{-1}(\ln q_t, \sigma_{kt} : \mathbf{x}_t) = \frac{\sigma_{\mu t}}{\sigma_q} (\ln q_t - \mu_q) + \mu_{kt}^e$, it can equate the marginal cost of $\ln q_t$, $w'(\ln q_t : \mathbf{x}_t)$ to the marginal benefit of $\ln q_t$, $MPQ(\mu_{kt}, \sigma_{kt}, \ln q_t)$ by choosing $\ln q_t$. It is also easy to check that

$$w^{*''}(\ln q_t : \mathbf{x}_t) > \frac{\partial MPQ(\mu_{kt}, \sigma_{kt}, \ln q_t)}{\partial \ln q_t} \Big|_{\mu_{kt} = \frac{\sigma_{\mu t}}{\sigma_q} (\ln q_t - \mu_q) + \mu_{kt}^e}.$$

Hence, the objective function of the firm for which $\mu_{kt} = \chi^{-1}(\ln q_t, \sigma_{kt} : \mathbf{x}_t)$ is strictly concave and $\ln q_t$ is a unique optimal decision. This decision is consistent with a

policy function of $\chi(\mu_{kt}, \sigma_{kt} : \mathbf{x}_t)$. Hence,

$$\begin{aligned}
TV^* &= E[y_t | \mu_{kt}, \sigma_{kt}] - w^* \left(\frac{\sigma_q}{\sigma_{\mu t}} [\mu_{kt} - \mu_{kt}^e] + \mu_q : \mathbf{x}_t \right) \\
&\quad + \beta \int V^*(\mu_{kt+1}, \sigma_{kt+1} : \mathbf{x}_{t+1}) d\Gamma_s(s_t | \mu_{kt}, \sigma_{kt}) \\
&= \frac{\alpha E[y_t | \mu_{kt}, \sigma_{kt}]}{\alpha + \frac{\psi\sigma_q}{\sigma_{\mu t}}} + \frac{\beta\phi}{\phi + \frac{\gamma\sigma_q}{\sigma_{\mu t}}} \int V^*(\mu_{kt+1}, \sigma_{kt+1} : \mathbf{x}_{t+1}) d\Gamma_s(s_t | \mu_{kt}, \sigma_{kt}) \\
&= \frac{\alpha E[y_t | \mu_{kt}, \sigma_{kt}]}{\alpha + \frac{\psi\sigma_q}{\sigma_{\mu t}}} + \frac{\beta\phi}{\phi + \frac{\gamma\sigma_q}{\sigma_{\mu t}}} \sum_{i=0}^{\infty} \Pi_{s=1}^i \frac{\beta\phi}{\phi + \frac{\gamma\sigma_q}{\sigma_{\mu t+s}}} \frac{\alpha E[y_{t+1+i} | \mu_{kt}, \sigma_{kt}]}{\alpha + \frac{\psi\sigma_q}{\sigma_{\mu t+1+i}}} \\
&= \frac{\alpha E[y_t | \mu_{kt}, \sigma_{kt}]}{\alpha + \frac{\psi\sigma_q}{\sigma_{\mu t}}} + \sum_{j=1}^{\infty} \Pi_{u=1}^j \frac{\beta\phi}{\phi + \frac{\gamma\sigma_q}{\sigma_{\mu t+u-1}}} \frac{\alpha E[y_{t+j} | \mu_{kt}, \sigma_{kt}]}{\alpha + \frac{\psi\sigma_q}{\sigma_{\mu t+j}}} \\
&= \sum_{j=0}^{\infty} \Pi_{u=1}^j \frac{\beta\phi}{\phi + \frac{\gamma\sigma_q}{\sigma_{\mu t+u-1}}} \frac{\alpha E[y_{t+j} | \mu_{kt}, \sigma_{kt}]}{\alpha + \frac{\psi\sigma_q}{\sigma_{\mu t+j}}}.
\end{aligned}$$

Hence, $TV^* = V^*$.

Finally, we show that $\sum_{i=0}^{\infty} \Pi_{s=1}^i \frac{\beta\phi}{\phi + \frac{\gamma\sigma_q}{\sigma_{\mu t+s-1}}} \frac{\alpha E[y_{t+i} | \mu_{kt}, \sigma_{kt}]}{\alpha + \frac{\psi\sigma_q}{\sigma_{\mu t+i}}}$ is bounded. Because $\frac{\beta\phi}{\phi + \frac{\gamma\sigma_q}{\sigma_{\mu t+u-1}}} < 1$ and Proposition 6 show that μ_{kt+i}^e , σ_{kt+i} and $\sigma_{\mu t+j}$ converge to finite values, it is enough to show that $E[\mu_{kt+i} | \mu_{kt}, \sigma_{kt}]$ and $\text{Var}[\mu_{kt+i} | \mu_{kt}, \sigma_{kt}]$ converge to finite values. We define

$$D_i = \Pi_{x=0}^{i-1} \left(\phi + \frac{\gamma\sigma_q}{\sigma_{\mu t+x}} \right).$$

Suppose that $\phi \in (0, 1)$, $\sigma_u < \infty$ and that $\sigma_\varepsilon \in (0, \infty)$. Then, Proposition 8 shows that $\lim_{x \rightarrow \infty} \left(\phi + \frac{\gamma\sigma_q}{\sigma_{\mu t+x}} \right) = \phi + \frac{\gamma\sigma_q}{\sigma_{\mu\infty}} < 1$. Hence, $\lim_{i \rightarrow \infty} D_i \rightarrow 0$. This means that $E[\mu_{kt+i} | \mu_{kt}, \sigma_{kt}]$ and $\text{Var}[\mu_{kt+i} | \mu_{kt}, \sigma_{kt}]$ converge to finite values. Now suppose that $\phi = 0$, $\sigma_u = \infty$ or $\sigma_\varepsilon = 0$. Then, $\text{Var}[\mu_{kt+i} | \mu_{kt}, \sigma_{kt}^2] = 0$. Note that if $\lim_{i \rightarrow \infty} D_i$ is bounded, $\lim_{i \rightarrow \infty} E[\mu_{kt+i} | \mu_{kt}, \sigma_{kt}]$ is bounded. We need to prove that $\lim_{i \rightarrow \infty} D_i$ is bounded. Suppose that $\phi = 0$; then, $\sigma_{\mu t+1} = \gamma\sigma_q$ for all t . This means that $\phi + \frac{\gamma\sigma_q}{\sigma_{\mu t+x}} = 1$ for all t . Hence, $D_i = 1$. Suppose that $\sigma_u^2 = \infty$ or $\sigma_\varepsilon^2 = 0$. Because $\sigma_{\mu t} = \phi\sigma_{\mu t} + \sigma_q\gamma$, $1 = \phi + \frac{\gamma\sigma_q}{\sigma_{\mu\infty}}$. Note that

$$\sigma_{\mu t+x} = \frac{1 - \phi^x}{1 - \phi} \sigma_q \gamma + \phi^x \sigma_{\mu t}.$$

Hence,

$$\begin{aligned}
D_i &= \prod_{x=0}^{i-1} \left\{ \phi + \frac{\gamma\sigma_q}{\frac{1-\phi^x}{1-\phi}\sigma_q\gamma + \phi^x\sigma_{\mu t}} \right\} \\
&= \prod_{x=0}^{i-1} \left\{ 1 + (\phi - 1) \left[1 - \frac{\gamma\sigma_q}{\sigma_q\gamma + \left[1 - \left(\phi + \frac{\sigma_q\gamma}{\sigma_{\mu t}} \right) \right] \sigma_{\mu t}\phi^x} \right] \right\} \\
&= \prod_{x=0}^{i-1} \left\{ 1 - \frac{(1-\phi) \left[1 - \left(\phi + \frac{\sigma_q\gamma}{\sigma_{\mu t}} \right) \right]}{1 - \phi + \frac{\sigma_q\gamma}{\sigma_{\mu t}} \left(\frac{1}{\phi^x} - 1 \right)} \right\}.
\end{aligned}$$

Note that for both $1 \geq \phi + \frac{\sigma_q\gamma}{\sigma_{\mu t}}$ and $1 < \phi + \frac{\sigma_q\gamma}{\sigma_{\mu t}}$, the following condition is satisfied:

$$\prod_{x=0}^{i-1} \left\{ 1 - \frac{(1-\phi) \left[1 - \left(\phi + \frac{\sigma_q\gamma}{\sigma_{\mu t}} \right) \right]}{1 - \phi + \frac{\sigma_q\gamma}{\sigma_{\mu t}} \left(\frac{1}{\phi^x} - 1 \right)} \right\} \leq \prod_{x=0}^{i-1} \left\{ 1 - \frac{(1-\phi) \left[1 - \left(\phi + \frac{\sigma_q\gamma}{\sigma_{\mu t}} \right) \right]}{1 - \phi + (1-\phi) \left(\frac{1}{\phi^x} - 1 \right)} \right\}.$$

Hence, it can be shown that

$$\begin{aligned}
D_i &\leq \prod_{x=0}^{i-1} \left\{ 1 - \phi^x \left[1 - \left(\phi + \frac{\sigma_q\gamma}{\sigma_{\mu t}} \right) \right] \right\} = \exp \sum_{x=0}^{i-1} \log \left\{ 1 + \phi^x \left[\left(\phi + \frac{\sigma_q\gamma}{\sigma_{\mu t}} \right) - 1 \right] \right\} \\
&\leq \exp \sum_{x=0}^{i-1} \phi^x \left[\left(\phi + \frac{\sigma_q\gamma}{\sigma_{\mu t}} \right) - 1 \right] = \exp \frac{1 - \phi^i}{1 - \phi} \left[\left(\phi + \frac{\sigma_q\gamma}{\sigma_{\mu t}} \right) - 1 \right].
\end{aligned}$$

The third inequality uses the fact that $\log(1+g)$ is concave in g . This means that

$$\lim_{i \rightarrow \infty} D_i \leq \exp \frac{\left(\phi + \frac{\sigma_q\gamma}{\sigma_{\mu t}} \right) - 1}{1 - \phi}.$$

Hence, D_i is bounded. ■

References

- [1] Abowd, John M., John Haltiwanger, Ron Jarmin, Julia Lane, Paul Lengermann, Kristin McCue, Kevin McKinney and Kristin Sandusky (2004), "The Relation among Human Capital, Productivity and Market Value: Building Up from Micro Evidence," mimeo.

- [2] Acemoglu, Daron (1997), “Matching, Heterogeneity and the Evolution of Income Distribution,” *Journal of Economic Growth*, Vol. 2, No. 1, 61–92.
- [3] Arrow, Kenneth J. (1962), “The Economic Implication of Learning by Doing,” *Review of Economic Studies*, Vol. 29, June, 155–173.
- [4] Atkeson, Andrew and Patrick J. Kehoe (2005), “Modeling and Measuring Organization Capital,” *Journal of Political Economy*, Vol. 113, No. 5, 1026–1053.
- [5] Bahk, Byong-Hyong and Michael Gort (1993), “Decomposing Learning by Doing in New Plants,” *Journal of Political Economy*, Vol. 101, No. 4, 561–583.
- [6] Baily, Martin, Neil C. Hulten and David Campbell (1992), “Productivity Dynamics in Manufacturing Plants,” *Brookings Papers on Economic Activity. Microeconomics*, 187–249.
- [7] Becker, Gary S. (1973), “A Theory of Marriage: Part 1,” *Journal of Political Economy*, 81, 813–846.
- [8] Bresnahan, Timothy F., Erik Brynjolfsson and Lorin M. Hitt (2002), “Information Technology, Workplace Organization, and the Demand for Skilled Labor: Firm-Level Evidence.” *Quarterly Journal of Economics*, Vol. 117, 339–376.
- [9] Burdett, Kenneth and Dale T. Mortensen (1998), “Wage Differentials, Employer Size, and Unemployment,” *International Economic Review*, Vol. 39, No. 2, May, 257–273.
- [10] Chandler, A. (1977), *The Visible Hand, the Managerial Revolution in American Business*, The Belknap Press of Harvard University Press, Cambridge, Massachusetts.
- [11] Chari, V. V. and Hugo Hopenhayn (1991), “Vintage Human Capital, Growth, and the Diffusion of New Technology,” *Journal of Political Economy*, Vol. 99, No. 6. December, 1142–1165.

- [12] Costrell, Robert M. and Glenn C. Loury (2004), “Distribution of Ability and Earnings in a Hierarchical Job Assignment Model,” *Journal of Political Economy*, Vol. 112, No. 6, December, 1322–1363.
- [13] Dhawan, Rajeev and Geoffrey Gerdes (1997), “Estimating Technological Change Using a Stochastic Frontier Production Function Framework: Evidence from U.S. Firm-Level Data,” *Journal of Productivity Analysis*, Vol. 8, 431–446.
- [14] Eeckhout, Jan and Boyan Jovanovic (2002), “Knowledge Spillovers and Inequality,” *American Economic Review*, Vol. 92, No. 5, 1290–1307.
- [15] Ericson, Richard and Ariel Pakes (1995), “Markov-Perfect Industry Dynamics: A Framework for Empirical Work,” *Review of Economic Studies*, Vol. 62, 53–82.
- [16] Faria André L. (2003), “Mergers and the Market for Organization Capital,” Mimeo.
- [17] Hall, Robert D. (2001), “The Stock Market and Capital Accumulation,” *American Economic Review*, Vol. 91, No. 5, December, 1185–1202.
- [18] Haltiwanger, John C., Julia I. Lane and James R. Spletzer (1999), “Productivity Differences across Employers: The Roles of Employer Size, Age, and Human Capital,” *American Economic Review*, Vol. 89, No. 2, May, 94–98.
- [19] Haltiwanger, John C., Julia I. Lane and James R. Spletzer (2000), “Wages, Productivity, and the Dynamic Interaction of Businesses and Workers,” NBER Working Paper 7994.
- [20] Jovanovic, Boyan (1982), “Selection and the Evolution of Industry,” *Econometrica*, Vol. 50, No. 3, May, 649–670.
- [21] Jovanovic, Boyan (1998), “Vintage Capital and Inequality,” *Review of Economic Dynamics*, Vol. 1, 497–530.

- [22] Jovanovic, Boyan and Peter L. Rousseau (2001), “Vintage Organization Capital,” NBER Working Paper 8166.
- [23] Kaldor, Nicholas (1934), “The Equilibrium of the Firm,” *Economic Journal*, Vol. 44, 60–76.
- [24] Kremer, Michael (1993), “The O-Ring Theory of Economic Development,” *Quarterly Journal of Economics*, Vol. 108, August, 551–575.
- [25] Lucas, Robert E. (1978), “On the Size Distribution of Business Firms,” *Bell Journal of Economics*, Vol. 9, 508–523.
- [26] McGahan, Anita M. (1999), “The Performance of US Corporations: 1981–1994,” *Journal of Industrial Economics*, Vol. 47, No. 4, December, 373–398.
- [27] Nelson, Richard R. and Sidney G. Winter (1982), *An Evolutionary Theory of Economic Change*, The Belknap Press of Harvard University Press, Cambridge, Massachusetts.
- [28] Penrose, E. (1959), *The Theory of the Growth of the Firm*, Oxford University Press, Oxford, U.K.
- [29] Prescott, E. and M. Visscher (1980), “Organization Capital,” *Journal of Political Economy*, Vol. 88, 446–461.
- [30] Robinson, Austin (1934), “The Problem of Management and the Size of Firms,” *Economic Journal*, Vol. 44, 242–257.
- [31] Rosen, Sherwin (1972), “Learning by Experience as Joint Production,” *Quarterly Journal of Economics*, Vol. 86, August, 366–382.
- [32] Samaniego, Roberto M. (2006), “Organizational Capital, Technology Adoption and the Productivity Slowdown,” *Journal of Monetary Economics*, Vol. 53, October, 1555–1569.

- [33] Sattinger, Michael (1979), “Differential Rents and the Distribution of Earnings,” Oxford Economic Papers, Vol. 31, March, 60–71.
- [34] Shimer, Robert (2005), “The Assignment of Workers to Jobs in an Economy with Coordination Frictions,” Journal of Political Economy, Vol. 113, No. 5, 996–1025.
- [35] Simon, Herbert A.(1997), *Administrative Behavior*, The Free Press, New York.
- [36] Takii, Katsuya (2003), “Prediction Ability,” Review of Economic Dynamics, Vol. 6, 80–98.
- [37] Takii, Katsuya (2007), “Fiscal Policy and Entrepreneurship,” forthcoming in Journal of Economic Behavior and Organization.