The Persistence of Differences in Productivity, Wages, Skill Mixes and Profits Between Firms in a Rapidly Changing Environment

February 20, 2008

Katsuya Takii

Associate Professor, Osaka School of International Public Policy (OSIPP)

【キーワード】J24, L25, Organization Capital, Assignment, Persistence

【要約】In this paper, we construct a dynamic assignment model that can provide a unified explanation of persistent differences in productivity, wages, skill mixes and profits between firms in a changing and uncertain environment. Large expected organization capital (firm-specific knowledge) attracts skilled workers, who help to accumulate organization capital. Accumulated large organization capital, in turn, confirms high expectations. This positive feedback brings about persistent differences in these variables in an uncertain environment. We estimate parameters and simulate the model. Our results show that a positive assignment mechanism accounts for a large part of the observed persistence; the difficulty of estimating organization capital plays only an auxiliary role.

This paper is the revised version of a previous working paper, Takii (2007b). I would like to thank R. Anton Braun, Julen Esteban-Pretel, Yuichi Fukuta, Ryo Horii, Shigeru Iwata, Charles Yuji Horioka, Hidehiko Ichimura, Boyan Jovanovic, Dan Sasaki, Akihisa Shibata, Futoshi Yamauchi and seminar participants at numerous seminars and conferences for helpful comments. 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. *Osaka School of International Public Policy, Osaka University, 1-31, Machikaneyama, Toyonaka, Osaka, 560-0043, Japan. E-mail: 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. (2007)]. 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.

As suggested by Haltiwanger et al. (2007), the assignment model provides a potential explanation for these observed persistent heterogeneities. If a quasi-fixed firm-specific resource and workers' skills are complementary to each other, a firm endowed with the large resource is willing to pay high wages to attract skilled workers. Such a firm achieves high productivity and earns large profits.

However, this seemingly plausible explanation does not provide a complete answer. First, why do some firms succeed in investing and maintaining their specific resources while others do not? Evidence shows that the pace of job creation and job destruction is quite rapid and that idiosyncratic factors are the main source of the observed gross job flows in the US economy [e.g., Davis and Haltiwanger (1999)]. This indicates that firms always confront idiosyncratic changes that may destroy some firm-specific resources. What is the mechanism that enables productive firms to maintain their core resources and prevents unproductive firms from investing these resources in a changing environment?

Second, can the assignment model provide a reasonable explanation even if we cannot observe firm-specific resources? Evidence shows that unobserved heterogeneity explains a large part of the variations in productivity [e.g., Bartelsman and Doms (2000)]. This indicates that intangible assets are likely to be the main component of firm-specific resources. Because intangible assets are, by definition, difficult to estimate, assignment based on intangible assets must rely on perceived values. How do speculative beliefs influence the persistence of variables? More importantly, to what extent is the observed persistence influenced by the discrepancy between beliefs and fundamental values? Because researchers disagree about the productive importance of intangible assets [e.g., Bond and Cummins (2000) and Hall (2001).], this question is important for understanding persistent inequalities in the era of the knowledge economy.

In this paper, we aim to answer these questions and provide a unified explanation of observed persistence in changing and uncertain environments. We propose a dynamic assignment model for the relationship between the skills of workers and unobserved firm-specific knowledge, which we term a firm's organization capital.[1] There are three key assumptions in our model. 1) Skill and organization capital are complementary inputs. This means that skilled workers can better utilize the available knowledge in a firm. 2) Skill is an input for the accumulation of organization capital. In particular, we model the firm's organization capital as a variant of its vintage human capital. 3) Although we cannot directly observe the amount of firm-specific knowledge, we can infer it from the firm's output. Hence, assignment is based on the beliefs about organization capital that are formed from the firm's past performance. This allows us to analyze how not only the assignment mechanism but also the discrepancy between beliefs and fundamental values influences the persistence of observed variables.

The main logic can be explained as follows. If a firm's organization capital is be-

[1] More 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.

lieved to be high, this belief attracts skilled workers. On the other hand, because skill is an input for the accumulation of organization capital, the employment of skilled workers promotes the accumulation of organization capital. A firm that accumulates more organization capital can be expected to improve its performance, which generates the perception that the firm has a higher level of organization capital. Hence, this persistence is induced by two positive feedback mechanisms: feedback between the accumulation of organization capital and the employment of skilled workers, and feedback between the fundamental capability of a firm and the beliefs about the capability.

From these two feedback mechanisms, we identify two sources of persistence: the heterogeneity of skills and the difficulty of measuring organization capital. 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. We show that if there is no stochastic disturbance, every firm's rank remains the same and firms' relative advantages (and disadvantages) persist indefinitely.

The theory also predicts that a rise in the noisiness of information increases the persistence of organization capital. If the revealed information is noisy, managers learn little from the new observations, and thus do not change their beliefs drastically. Because 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.

To examine the quantitative importance of the two sources, we estimate parameters and simulate the model. Because it is perceived organization capital that attracts skilled workers who help to accumulate organization capital, the effect of the perceived value on persistence provides information about the role of assignment in persistence in our framework. Exploiting this information, 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. 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; note that, in this paper, "relative" refers to the logarithm of each value relative to industry and year averages. 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 the stylized facts of interest.

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 [e.g., Becker (1973) and Sattinger (1979)]. More recently, Kremer (1993) demonstrates that the model of positive assortative matching among workers can explain a variety of evidence. Similarly to his model, our model incorporates positive assortative matching among workers. In addition, although Kremer's (1993) model is static,[2] our model conveys the sprit of Kremer's (1993) idea in a dynamic framework; that is, current skilled

[2]Most assignment models are static and 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.

workers attract skilled successors.

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 itself changes because of learning by doing and because of idiosyncratic shocks that change the usefulness of the accumulated knowledge. Hence, even mature firms must continue to learn about their capability. We suggest that this modeling strategy mimics the nature of firms' behavior in a changing and uncertain environment.

Although estimating organization capital is difficult, the key assumptions made in this paper are broadly consistent with the evidence. Evidence shows that productive organizational arrangement demands skill [e.g., Chandler (1977), Caroli and Van Reenen (2001) and Bresnahan et al. (2002)]. This is consistent with our assumption of complementarity between organization capital and skill. In addition, evidence from Caroli and Van Reenen (2001) suggests that firms need the intangible assets accumulated by skilled workers to make organizational changes productive. Evidence also shows that the intangible assets accumulated by skilled workers are an important determinant of technology adoption [e.g., Doms et al. (1997)]. Hence, the evidence consistently indicates that organization capital, as modeled in this paper, plays an important role in improving productivity by stimulating technological and organizational changes.

The paper is organized as follows. In the next section, we set up a dynamic positive assortative assignment model. We clarify the mechanism through which skill differences and noisy information enhance persistence in the model. In Section 3, we use regression analysis to estimate the parameters of our model. In Section 4, we simulate our model by using the estimated structural parameters. In Section 5, we discuss extensions and conclude the paper. Proofs of propositions and technical aspects of the derivation of equations are presented in Takii (2007b).

2 A Dynamic Assignment Model

In this section we establish a positive assortative assignment equilibrium between unobserved organization capital and skills. Our static assignment model is based on that of Sattinger (1979). We extend the model to incorporate dynamics, uncertainty and learning to describe the nature of assignment in a dynamically changing environment.

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 ith 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[3]:

$$y_{it} = e^{u_t} A \left(k_t^o\right)^{\alpha} q_{it}^{\psi}, \ \alpha > 0, \psi > 0, \tag{1}$$

where A, α and ψ are constant parameters and u_t is a firm-specific productivity shock and is normally distributed with a mean of $-\frac{\sigma_u^2}{2}$ and a variance of σ_u^2. We call this shock, u_t, noise because its only role is to make organization capital difficult to observe. 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 k_t^o cannot be directly observed, but can be inferred from the realizations of output. When employment decisions are made about the ith 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

[3] Alternatively, we can assume the following production function without changing our results:

$$y_{it} = e^{u_t} A \left(k_t^o\right)^{\alpha} \left[\int_0^1 q_{ijt}^{\beta} dj\right]^{\frac{\psi}{\beta}}, \ \alpha > 0, \psi > 0, \beta < 1,$$

where q_{ijt} is the quality of the jth worker in the ith job at date t. This assumption captures Kremer's (1993) idea of team production.

distribution of $\ln k_t^o$ is normally distributed with a mean of μ_{kt} and a variance of σ_{kt}^2. Then the expected output from the job is

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

All firms are assumed to have the same σ_{kt} at date t. However, 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.

Assume that the ith 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 ith job is written as

$$\chi_i(\mu_{kt}) \equiv \arg\max_{\ln q_{it}} \left\{E\left[y_{it}|\mu_{kt}, \ln q_{it}\right] - w(\ln q_{it})\right\}, \ \forall i, \mu_{kt}. \tag{3}$$

Assume that $\ln q_t$ is normally distributed with a mean of μ_q and a standard deviation of σ_q at any date. 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, $\ln q_t$.

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

$$1 - \Phi\left(\frac{\mu_{kt} - \mu_{kt}^e}{\sigma_{\mu t}}\right) = 1 - \Phi\left(\frac{\chi_i(\mu_{kt}) - \mu_q}{\sigma_q}\right), \ \forall i, \mu_{kt}. \tag{4}$$

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, equations (3) and (4) characterize a static market equilibrium.

Definition: A static market equilibrium consists of $\chi_i(\cdot)$ and $w(\cdot)$ that satisfy equations (3) and (4).

We aim to find a policy function and a wage function that are consistent with this definition of equilibrium. Equation (4) states that the policy function must satisfy

$$\chi(\mu_{kt}) \equiv \chi_i(\mu_{kt}) = \frac{\sigma_q}{\sigma_{\mu t}}(\mu_{kt} - \mu_{kt}^e) + \mu_q.$$

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 perceived organization capital. For this policy function to be consistent with the definition of equilibrium, the policy function must solve equation (3). Consider a firm that has $\mu_{kt} = \frac{\sigma_{\mu t}}{\sigma_q}\left[\ln q_t - \mu_q\right] + \mu_{kt}^e \equiv \chi^{-1}(\ln q_t)$. 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:

$$w'(\ln q_t) = \psi E\left[y_t | \chi^{-1}(\ln q_t), \ln q_t\right], \ \forall \ln q_t.$$

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:

$$w(\ln q_t) = \frac{\frac{\psi\sigma_q}{\alpha\sigma_{\mu t}} E\left[y_t | \chi^{-1}(\ln q_t), \ln q_t\right]}{1 + \frac{\psi\sigma_q}{\alpha\sigma_{\mu t}}}. \tag{5}$$

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 μ_{kt}. Hence, a firm that has high perceived organization capital pays high wages.

The firm's expected profits are strictly increasing in μ_{kt}.

$$\pi^{e}\left(\mu_{kt}\right)=\int_{0}^{1}\left[E\left[y_{t}|\mu_{kt},\chi\left(\mu_{kt}\right)\right]-w\left(\chi_{i}\left(\mu_{kt}\right)\right)\right]di=\frac{E\left[y_{t}|\mu_{kt},\chi\left(\mu_{kt}\right)\right]}{1+\frac{\psi\sigma_{q}}{\alpha\sigma_{\mu t}}}. \tag{6}$$

In sum, 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: 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:[4]

$$k_{t+1}^{o} = B\left(k_{t}^{o}\right)^{\phi}\left(q_{t}^{e}\right)^{\gamma} e^{\varepsilon_{t}}, 0 \leq \phi < 1,\ \gamma > 0, \tag{7}$$

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. Because of rapid changes in technology or demand, there is uncertainty about the productive usefulness of the accumulated knowledge. The random variable, ε_t, summarizes the shifts in the productivity of the accumulated knowledge because of changes in the environment. 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 (7).

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[5]. Hence, most of our analysis avoids the technical diffi-

[4] Given that every job is filled by workers of the same quality, equation (7) generates dynamics that are the same as those from the transition equation,

$$k_{t+1}^{o} = B_1\left(k_{t}^{o}\right)^{\phi_1}\left(y_{t}\right)^{\gamma_1} e^{\varepsilon_{t}},$$

where $y_t = A\left(k_t^o\right)^{\alpha} q_t^{\psi}$, 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.

[5] In Takii (2007b), we assume that a firm solves a dynamic optimization problem by taking

culties 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 learning-by-doing assumption later.

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

$$\ln k_{t+1}^o = \ln B + \phi \ln k_t^o + \gamma \left[\frac{\sigma_q}{\sigma_{\mu t}} (\mu_{kt} - \mu_{kt}^e) + \mu_q \right] + \varepsilon_t. \tag{8}$$

To derive the dynamics of μ_{kt}, we must describe the information structure of the model. After the job 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 $\mu_{kt+1} = E \left[\ln k_{t+1}^o | s_t, \mu_{kt}, \sigma_{kt} \right]$ and $\sigma_{kt+1} = \sqrt{Var \left[\ln k_{t+1}^o | s_t, \mu_{kt}, \sigma_{kt} \right]}$, the dynamics of μ_{kt} and σ_{kt} can be written as follows:

$$\mu_{kt+1} = \ln B + \phi E \left[\ln k_t^o | s_t, \mu_{kt}, \sigma_{kt} \right] + \gamma \left[\frac{\sigma_q}{\sigma_{\mu t}} (\mu_{kt} - \mu_{kt}^e) + \mu_q \right] - \frac{\sigma_\varepsilon^2}{2}, \tag{9}$$

$$\sigma_{kt+1} = \sqrt{\phi^2 (1 - h_t) \sigma_{kt}^2 + \sigma_\varepsilon^2}, \tag{10}$$

where

$$E \left[\ln k_t^o | s_t, \mu_{kt}, \sigma_{kt} \right] = (1 - h_t) \mu_{kt} + h_t s_t = (1 - h_t) \mu_{kt} + h_t (\ln k_t^o + u_t^*), \tag{11}$$

$$h_t = \frac{\left(\frac{\alpha \sigma_{kt}}{\sigma_u} \right)^2}{1 + \left(\frac{\alpha \sigma_{kt}}{\sigma_u} \right)^2}. \tag{12}$$

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

Equation (11) shows that $E\left[\ln k_t^o | s_t, \mu_{kt}, \sigma_{kt}\right]$ 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 (12), h_t is negatively related to σ_u. If the variance of u_t 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.[6]

Because σ_{kt} is the same in all firms, equation (10) shows that σ_{kt+1} is also the same in all firms. Similarly, because μ_{kt} and s_t are normally distributed, equation (9) shows that μ_{kt+1} is also normally distributed. Hence, the normality of the distribution is preserved. The following mean and standard deviation of the belief in the next period can be derived:

$$\mu_{kt+1}^e = \ln B + \phi \mu_{kt}^e + \gamma \mu_q - \frac{\sigma_\varepsilon^2}{2}, \tag{13}$$

$$\sigma_{\mu t+1} = \sqrt{\left(\phi + \frac{\gamma \sigma_q}{\sigma_{\mu t}}\right)^2 \sigma_{\mu t}^2 + \phi^2 h_t \sigma_{kt}^2}. \tag{14}$$

Furthermore, by substituting equations (11) and (13) into equations (8) and (9), we can also rewrite the dynamics of $\ln k_t^o$ and μ_{kt} as follows:

$$\ln k_{t+1}^o - \mu_{kt+1}^e = \phi\left(\ln k_t^o - \mu_{kt}^e\right) + \frac{\gamma \sigma_q}{\sigma_{\mu t}}\left(\mu_{kt} - \mu_{kt}^e\right) + \varepsilon_t^*, \tag{15}$$

$$\mu_{kt+1} - \mu_{kt+1}^e = \phi h_t\left(\ln k_t^o - \mu_{kt}^e\right) + \left[\phi\left(1 - h_t\right) + \frac{\gamma \sigma_q}{\sigma_{\mu t}}\right]\left(\mu_{kt} - \mu_{kt}^e\right) + \phi h_t u_t^*, \tag{16}$$

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 (15) shows the dynamics of $\ln k_t^o$. The first term of equation (15) is influenced by technological persistence, ϕ. That is, if organization capital is above

[6] In fact, h_t can be also rewritten as follows:

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

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, 2007a), as a tractable measure of prediction ability.

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. Note that when the ratio of the standard deviation of skills to that of perceived organization capital, $\frac{\gamma\sigma_q}{\sigma_{\mu t}}$, is large, the effect of μ_{kt} on $\ln k_{t+1}^o$ is large. The firms with large μ_{kt} derive the high benefits from large $\frac{\gamma\sigma_q}{\sigma_{\mu t}}$ because these leading firms attract the most talented workers, who provide the firms with the best knowledge. Therefore, relative advantages persist longer.

Equation (16) 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, u_t, 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 (16).

The second term of equation (16) 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 (15) and (16) provide some intuition about the dynamics of $\ln k_t^o$ and μ_{kt}. First, the interpretation of equation (15) 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 (16), 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. This means that a firm with large μ_{kt} can persistently attract high $\ln q_t$ and maintain a large μ_{kt}. That is, the noisier is the information, the more persistent is the belief.

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 1 *The aggregate economy converges to a unique stationary distribution. Moreover, the dynamics of an individual firm in the stationary distribution are described by the following vector autoregression (VAR):*

$$\mathbf{k}_{t+1}=\mathbf{M}\mathbf{k}_t+\boldsymbol{\xi}_t, \tag{17}$$

where

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

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. When $\sigma_{\mu t}$ is small, $\frac{\gamma\sigma_q}{\sigma_{\mu t}}$ is large and, therefore, firms with high perceived levels of organization capital have large relative advantages. This increases $\sigma_{\mu t}$. Hence, provided that σ_q is positive, the distribution is not degenerate. We investigate the properties of equation (17) and discuss what influences the persistence of organization capital.

Extreme Cases : It is instructive to start with the extreme cases in which $\sigma_\varepsilon = 0$ and $\sigma_u = \infty$. When $\sigma_\varepsilon = 0$, there are no changes in the environment. Hence, firms that have relatively high levels of organization capital can maintain their relative advantages. On the other hand, when $\sigma_u = \infty$, information is too noisy and the firm can learn nothing about the level of organization capital. Therefore, its belief never changes. The following proposition shows that the level of organization capital persists in both cases.

Proposition 2 *1. Suppose $\sigma_\varepsilon^2 = 0$. Then*

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

2. Suppose that $\sigma_u^2 = \infty$. Then

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

The first part of the proposition shows that if $\sigma_\varepsilon = 0$, the belief does not change and real organization capital eventually converges to the level implied by the constant belief. Because there are no changes in the environment, the belief about organization capital is accurate and there is no additional information from output. Hence, the belief never changes. Because the constant belief determines the quality of workers, real organization capital eventually converges to the level at which it is believed to be.

Similarly to the case of $\sigma_\varepsilon = 0$, if $\sigma_u = \infty$, the belief never changes. However, the level of organization capital fluctuates around this constant belief. Because the firm cannot learn about its own organization capital, the firm never changes its own belief. Hence, the belief is constant. 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.

General Case: Let us examine a more general case. Suppose that $\phi \in (0,1)$, $\frac{\sigma_u}{\alpha\sigma_\varepsilon} \in (0,\infty)$ and $\frac{\gamma\sigma_q}{\sigma_\varepsilon} \in (0,\infty)$. First, we analyze the stability of equation (17). Then, we analyze what influences persistence.

Let λ_1 and λ_2 denote the eigenvalues of the matrix $\mathbf{M}$. Then, equation (17) 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 $\lambda_1 < 1$. Because λ_1 incorporates important information about persistence, it is termed the persistence parameter. It consists of assumed persistence, ϕ, and the assignment effect, $\frac{\gamma\sigma_q}{\sigma_{\mu\infty}}$.

Clearly, $\frac{\gamma\sigma_q}{\sigma_{\mu\infty}}$ and h_∞ are endogenous variables. Thus, there are more fundamental conditions for stability. In order to find out the conditions, we need to understand 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 $\frac{\gamma\sigma_q}{\sigma_\varepsilon}$. It is shown that there exist functions $\eta(\cdot)$ and $\Sigma(\cdot,\cdot)$[7] such that

$$h_\infty = \eta\left(\frac{\sigma_u}{\alpha\sigma_\varepsilon}\right) \in (0,1), \tag{18}$$

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

$$\frac{\gamma\sigma_q}{\sigma_{\mu\infty}} = \Sigma\left(\frac{\gamma\sigma_q}{\sigma_\varepsilon}, h_\infty\right) \in (0, 1-\phi), \tag{19}$$

where $\Sigma_1\left(\frac{\gamma\sigma_q}{\sigma_\varepsilon}, h_\infty\right) > 0$, $\Sigma_2\left(\frac{\gamma\sigma_q}{\sigma_\varepsilon}, h_\infty\right) < 0$, $\lim_{\frac{\gamma\sigma_q}{\sigma_\varepsilon}\to 0} \Sigma\left(\frac{\gamma\sigma_q}{\sigma_\varepsilon}, h_\infty\right) = 0$, $\lim_{\frac{\gamma\sigma_q}{\sigma_\varepsilon}\to \infty} \Sigma\left(\frac{\gamma\sigma_q}{\sigma_\varepsilon}, h_\infty\right) = 1-\phi$, $\lim_{h_\infty\to 1} \Sigma\left(\frac{\gamma\sigma_q}{\sigma_\varepsilon}, h_\infty\right) \in (0, 1-\phi)$ and $\lim_{h_\infty\to 0} \Sigma\left(\frac{\gamma\sigma_q}{\sigma_\varepsilon}, h_\infty\right) = 1-\phi$[8].

Equation (18) shows that h_∞ and $\frac{\sigma_u}{\alpha\sigma_\varepsilon}$ have a one-to-one relationship. Hence,

[7] Explicit solutions for $\eta(\cdot)$ and $\Sigma(\cdot,\cdot)$ can be found in the Takii (2007b).

[8] Note that the properties of the function $\Sigma(\cdot,\cdot)$ imply that when h_∞ converges to 0 or $\frac{\gamma\sigma_q}{\sigma_\varepsilon}$ converges to infinite, the persistence parameter, λ_1, converges to 1. This means that the previous extreme case can be seen as the limit of this general case.

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.

Equation (19) 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. This shows that not only the large variance in skills but also the small variance of shocks to the accumulation of organization capital induces large $\frac{\gamma\sigma_q}{\sigma_{\mu\infty}}$. More interestingly, $\frac{\gamma\sigma_q}{\sigma_{\mu\infty}}$ is decreasing in h_∞. When information is more accurate, rational agents rely more on new information to make inferences about the current level of organization capital. Therefore, rational agents can change their posterior beliefs based on reliable information. This makes the variance of μ_{kt} large and, therefore, makes $\frac{\gamma\sigma_q}{\sigma_{\mu\infty}}$ small.

Note also that equation (19) implies that $\lambda_1 = \phi + \frac{\gamma\sigma_q}{\sigma_{\mu\infty}} < 1$. Hence, the following proposition can be stated.

Proposition 3 *Suppose that $\phi \in (0,1)$, and that $\frac{\sigma_u}{\alpha\sigma_\varepsilon}$ and $\frac{\gamma\sigma_q}{\sigma_\varepsilon}$ are finite. Equation (17) is covariance stationary.*

The two eigenvalues, λ_1 and λ_2, are important determinants of the persistence of the stochastic process, too. 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: $\rho_{\ln kj} \equiv \frac{E\left[D\ln k_t^o D\ln k_{t-j}^o\right]}{Var(D\ln k_t^o)}$, $\rho_{\mu j} \equiv \frac{E\left[D\mu_{kt} D\mu_{kt-j}\right]}{\sigma_{\mu\infty}^2}$. We can derive the autocorrelations of $\ln k_t^o$ and μ_{kt} are functions of λ_1 and λ_2

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

The above equation 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 $\lambda_1 > \lambda_2$. It means that the autocorrelation of the belief about a firm's organization capital exceeds that of its actual organization capital: $\rho_{\mu j} > \rho_{\ln kj}$, $\forall j$. 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.

Next, we show how the parameters $\frac{\gamma\sigma_q}{\sigma_\varepsilon}$ and h_∞(or $\frac{\sigma_u}{\alpha\sigma_\varepsilon}$) affect the autocorrelations.

Proposition 4 *1) There exist j^* and j^{**} such that for all $j \geq j^*$, $\frac{d\rho_{\ln kj}}{d\frac{\gamma\sigma_q}{\sigma_\varepsilon}} > 0$ and for all $j \geq j^{**}$, $\frac{d\rho_{\ln kj}}{dh_\infty} < 0$. 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.

3 Regression Analysis

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 (17), as follows:

$$\begin{aligned} D\ln y_{t+1} &= b_1 D\ln y_t + b_2 E\left[D\ln y_t|\mu_{kt}\right] + v_t, && (20)\\ E\left[D\ln y_{t+1}|\mu_{kt+1}\right] &= b_3 D\ln y_t + b_4 E\left[D\ln y_t|\mu_{kt}\right], && (21) \end{aligned}$$

where $D\ln y_t = \ln y_t - E\left[\ln y\right]$, $b_1 = \phi + \frac{\psi\sigma_q}{\alpha\sigma_{\mu\infty}}\phi h_\infty$, $b_2 = \frac{\gamma\sigma_q}{\sigma_{\mu\infty}} - \frac{\psi\sigma_q}{\alpha\sigma_{\mu\infty}}\phi h_\infty$, $b_3 = \phi h_\infty + \frac{\psi\sigma_q}{\alpha\sigma_{\mu\infty}}\phi h_\infty$, $b_4 = b_1 + b_2 - b_3$ and $v_t = \alpha\left(\varepsilon_t^* - \phi u_t^* + u_{t+1}^*\right)$.

In order to estimate equations (20) and (21), we must estimate $E\left[D\ln y_t|\mu_{kt}\right]$ 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.

Estimation Method 1: The first method is relatively simple and provides evidence that is consistent with the assumption that skilled workers help firms to accumulate assets and raise future productivity. It applies the following proposition, which is proven by equations (5) and (6). Although the dynamics of organization capital are not influenced by the assumption that organization capital is accumulated through learning by doing, equations (5) and (6) are affected by this assumption. Therefore, the following useful proposition represents a benefit of assuming that there is learning by doing.

Proposition 5 *Perceived relative productivity is equal to relative wages and expected* *relative profits per worker.*

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

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

Proposition 5 states that perceived relative productivity can be estimated by using relative wages. Hence, the following testable equation is derived from equations (20) and (21):

$$D\ln y_t = \theta_1 D\ln y_{t-1} + \theta_2 D\ln y_{t-2} + \theta_3 D\ln w_{t-2} + v_{t-1}, \tag{22}$$

where $\theta_1 = b_1$, $\theta_2 = b_2 b_3$ and $\theta_3 = b_2\left(b_1 + b_2 - b_3\right) > 0$. By using the estimated value of θ_1, θ_2 and θ_3, we can identify b_1, b_2 and b_3.

Equation (22) 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 5 provides a suitable instrument. Because the firm makes employment decisions without observing realized output, relative wages, $D\ln w_{t-1} \equiv \ln w_{t-1} - E\left[\ln w\right]$, 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 $E\left[D\ln y_t|\mu_{kt}\right]$ from the data, we derive the following regression equation from equation (21):

$$D\ln y_t = b_3 \sum_{i=0}^{t-1} (b_4)^i \, Dy_{t-1-i} + b_4^t E\left[D\ln y_0|\mu_{k0}\right] + \varpi_t, \tag{23}$$

where $\varpi_t = D\ln y_t - E\left[D\ln y_t|\mu_k\right]$. Note that ϖ_t is not correlated with $D\ln y_{t-1-i}$ for all $i \geq 0$ and $E\left[D\ln y_0|\mu_{k0}\right]$. This contrasts with $E\left[v_t|D\ln y_t\right] \neq 0$ in equation (20).

Note that the parameters b_1 and b_3 differ only if h_∞ is less than 1 or, equivalently, if σ_u is positive. In this case, $E\left[u_t^*|D\ln y_t\right] \neq 0$ and, therefore, $E\left[v_t|D\ln y_t\right] \neq 0$. 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. That is, the bias itself contains useful information on h_∞.

To separate b_1 from b_3, we apply ordinary least squares (OLS) to equation (23) and use an instrumental variables (IV) approach to estimate equation (20). 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. This helps to identify h_∞. We use this technique in applying the second estimation procedure discussed below.

Assume that there is a proxy for $E\left[D\ln y_0|\mu_{k0}\right]$. First, we choose an arbitrary value of b_4, and construct $\sum_{i=0}^{t-1} (b_4)^i \, Dy_{t-1-i}$ and $(b_4)^t \, E\left[D\ln y_0|\mu_{k0}\right]$ from the data. Second, equation (23) is estimated under the constraint that the coefficient of

$(b_4)^t\, E\left[D\ln y_0|\mu_{k0}\right]$ 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 $E\left[D\ln y_t|\mu_{kt}\right]$ by $\hat{b}_3\sum_{i=0}^{t-1}(b_4)^i\, Dy_{t-1-i}+b_4^t E\left[D\ln y_0|\mu_{k0}\right]$. Fourth, using the estimated value of $E\left[D\ln y_t|\mu_{kt}\right]$, we estimate equation (20) 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 $E\left[D\ln y_t|\mu_{kt}\right]$. We need an additional instrument for $E\left[D\ln y_t|\mu_{kt}\right]$ 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 US. It contains information from balance sheets, and information on incomes, cash flows and financial variables. The value added per worker and the average wage rate are constructed for each firm and each year. Details of our data construction procedure are given in Appendix.

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 fth 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, $E\left[D\ln y_0|\mu_{k0}\right]$, 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 1, 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

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 1: 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.

estimation method is described in Appendix.) 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 1 includes only companies that report labor and related expenses. The "Large Sample" includes companies whose labor costs we have estimated.

All coefficients in Table 1 are significant and positive, which is consistent with our 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 positive 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 fth 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 2 and 3. 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 2 reports the results from the regression equation (23). Table 3 reports the results from the regression equation (20).

Table 2 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 3 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 future performance.

Table 3 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, $h_\infty < 1$ 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

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)
$b_4^t E [D \ln y_0 \vert \mu_{k0}]$	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 2: 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)
$E [D \ln y_t \vert \mu_{kt}]$	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 3: 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 2 reports OLS results, and Table 3 reports IV results. Standard errors are reported in parentheses. * denotes significance at the 5 percent level. ** denotes significance at the 0.5 percent level.

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.

In summary, the predictions of our model are broadly supported by the data. In particular, the evidence is consistent with two important predictions: 1) skilled workers help firms to accumulate assets and raise future productivity; and 2) people learn about a firm's capacity from its past performance and form beliefs that influence its future performance.

4 Numerical Exercises

By using the estimated parameters of the previous section, we report our estimates of ϕ, h and $\frac{\gamma\sigma_q}{\sigma_{\mu\infty}}$. Using these structural 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.

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:

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

As discussed in the previous section, b_3 contains useful information on h_∞. Given a knowledge of h_∞, ϕ and $\frac{\gamma\sigma_q}{\sigma_{\mu\infty}}$ are primarily related to b_1 and b_2, respectively. Because the fraction ϕ of current organization capital is transformed into next period's organization capital, b_1 contains information about ϕ. On the other hand, because belief in a high level of organization capital attracts skilled workers, b_2 contains information about the effect of assignment.[9]

[9]Note that we can separately identify ϕ and $\frac{\gamma\sigma_q}{\sigma_{\mu\infty}}$ even if $h_\infty = 1$. 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.

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

Table 4: Labor Share

$\frac{E[w_t]}{E[y_t]} = \frac{\sum_i^I \sum_t^T \left[\frac{\sum_f^{m_{ii}} w_{fit}}{\sum_f^{m_{ii}} y_{fit}}\right]}{IT}$, where w_{fit} and y_{fit} are the wage payments and labor productivity of the fth firm in the ith industry in year t, m_{it} is the number of firms operating in the ith 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.

To identify the parameters, we need to know the value of $\frac{\psi\sigma_q}{\alpha\sigma_{\mu\infty}}$. Hence, we calibrate it from data. The wage function (5) implies that $\frac{\psi\sigma_q}{\alpha\sigma_{\mu\infty}}$ can be calibrated 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 fth firm in the ith industry in year t, m_{it} is the number of firms operating in the ith industry in year t, I is the number of industries and T is the number of years.

Table 4 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 et al. (2002)

	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 5: The Estimated Structural Parameters

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

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 5 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.[10] 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 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. 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 $\rho_{\ln yj} \equiv \frac{E[D\ln y_t D\ln y_{t-j}]}{Var(\ln y_t)}$ and $\rho_{E[\ln y|\mu]j} \equiv \frac{E\left[E[D\ln y_t|\mu_t]E\left[D\ln y_{t-j}|\mu_{t-j}\right]\right]}{Var(E[\ln y_t|\mu_t])}$. The autocorrelations of relative productivity and expected relative productivity are derived as follows:

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

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

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

structural parameters, $\frac{\psi\sigma_q}{\alpha\sigma_{\mu\infty}}$, ϕ, h_∞ and $\frac{\gamma\sigma_q}{\sigma_{\mu\infty}}$. These equations show 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 5 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[11].

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

[11] 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

$$\rho_{\ln xj} = \frac{\sum_t^T \sum_i^I \sum_f^{m_{it}} \left[\ln x_{fit} - \frac{\sum_f^{m_{it}} \ln x_{fit}}{m_{it}}\right]\left[\ln x_{fi(t-j)} - \frac{\sum_f^{m_{it}} \ln x_{fi(t-j)}}{m_{it}}\right]}{\sqrt{\sum_t^T \sum_i^I \sum_f^{m_{it}} \left[\ln x_{fit} - \frac{\sum_f^{m_{it}} \ln x_{fit}}{m_{it}}\right]^2}\sqrt{\sum_t^T \sum_i^I \sum_f^{m_{it}} \left[\ln x_{fi(t-j)} - \frac{\sum_f^{m_{it}} \ln x_{fi(t-j)}}{m_{it}}\right]^2}}, \tag{24}$$

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

persistent differences in wage payments.

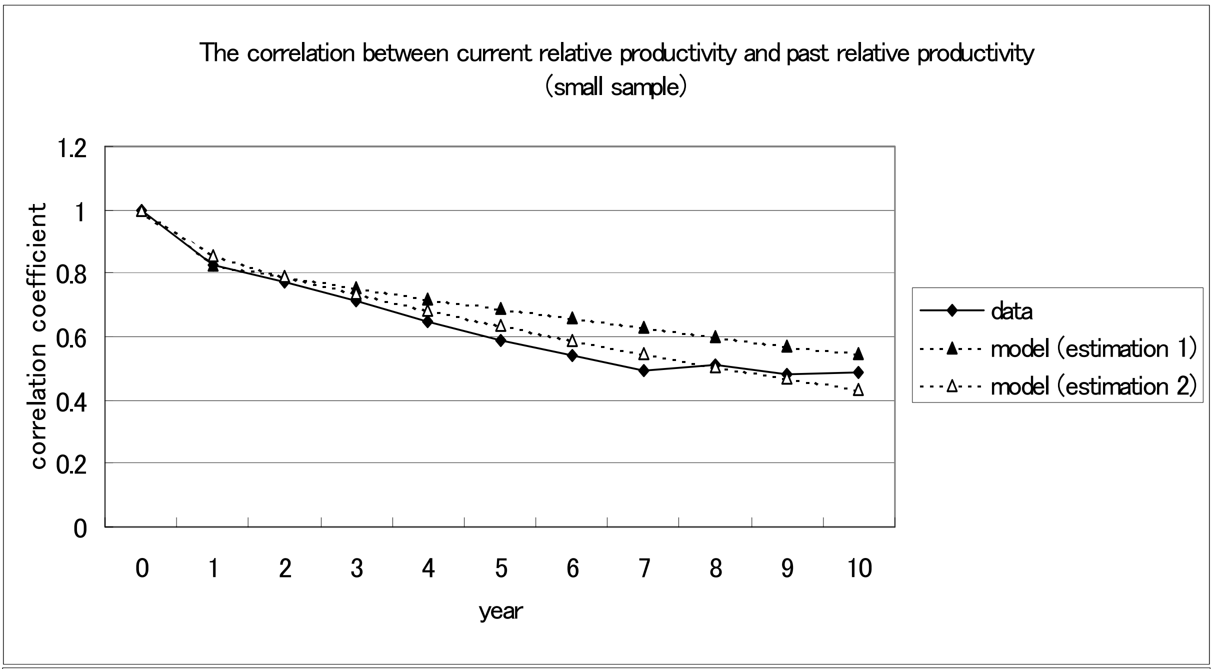
Figure 3 compares the predicted correlation for expected relative profits per worker with the correlation of relative profits per worker. Although the predicted correlation is much larger than the observed correlation, this result is expected. Our theory explains the correlation for expected relative profits per worker, but not the one for real relative profits 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.

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 wages[12] 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.

[12] There is a caution for the interpretation of Figure 5. Equation (5) 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.


The correlation between current relative productivity and past relative productivity
(small sample)
correlation coefficient
1.2
1
0.8
0.6
0.4
0.2
0
0
1
2
3
4
5
6
7
8
9
10
year
data
model (estimation 1)
model (estimation 2)

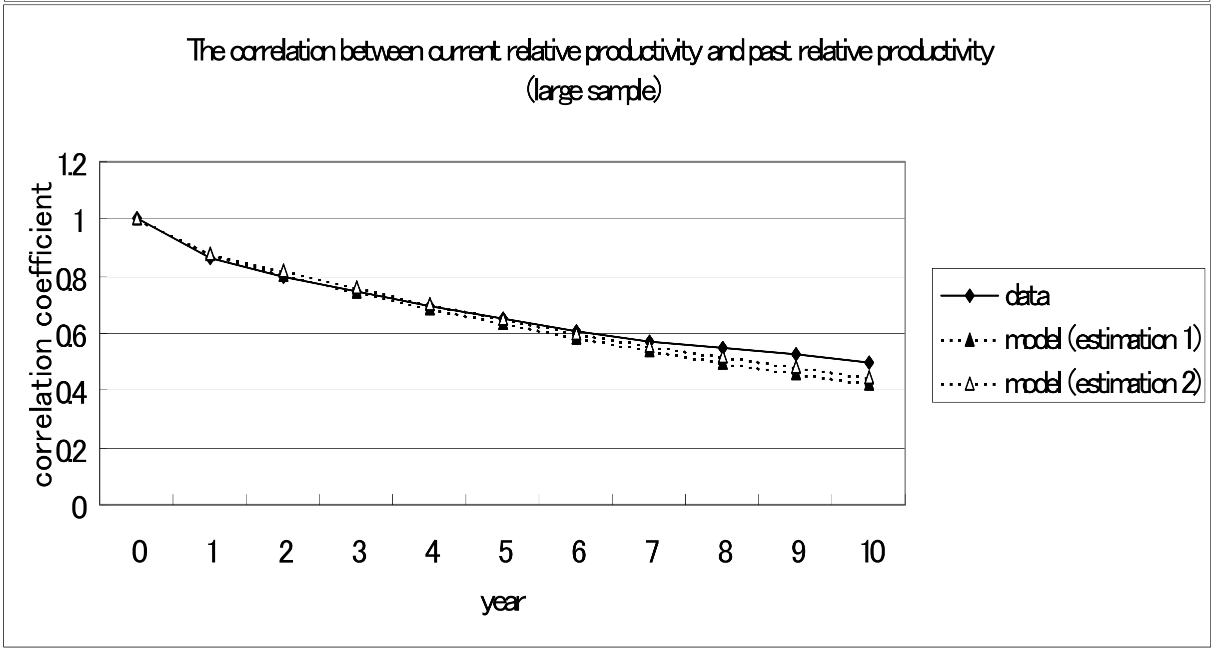
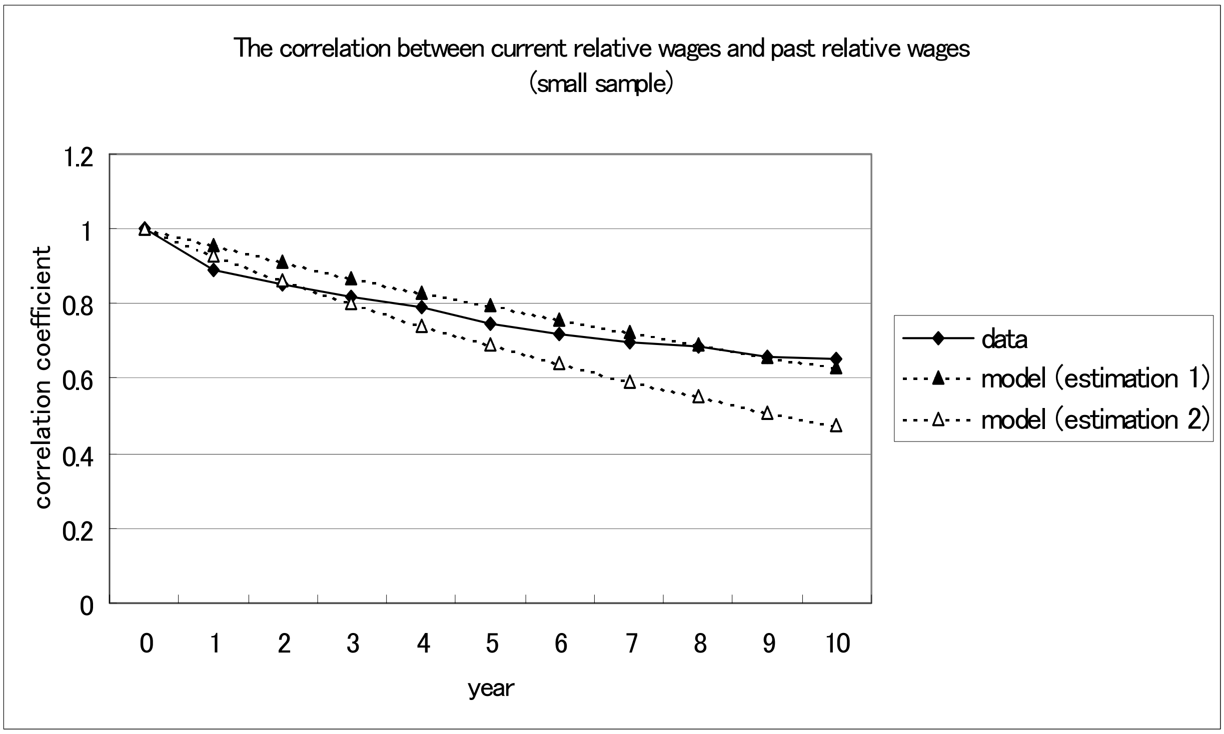

The correlation between current relative productivity and past relative productivity
(large sample)
correlation coefficient
1.2
1
0.8
0.6
0.4
0.2
0
0
1
2
3
4
5
6
7
8
9
10
year
data
model (estimation 1)
model (estimation 2)

Figure 1:


The correlation between current relative wages and past relative wages
(small sample)
1.2
1
0.8
0.6
0.4
0.2
0
correlation coefficient
0
1
2
3
4
5
6
7
8
9
10
year
data
model (estimation 1)
model (estimation 2)

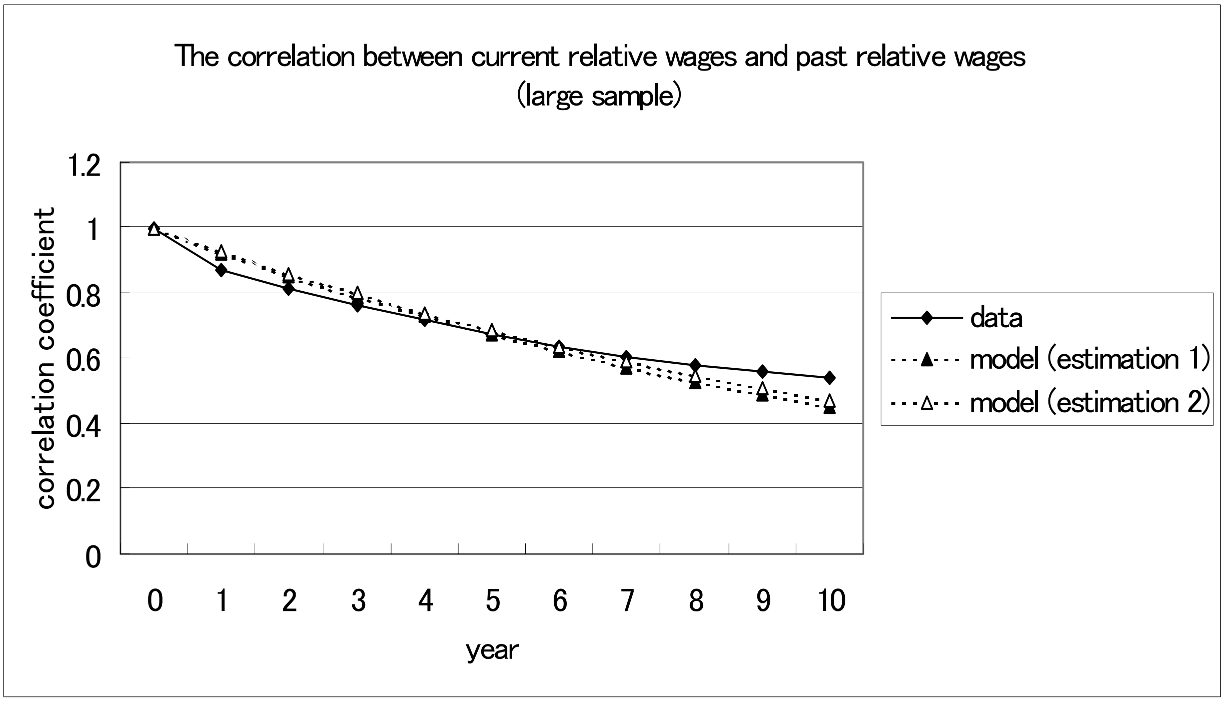
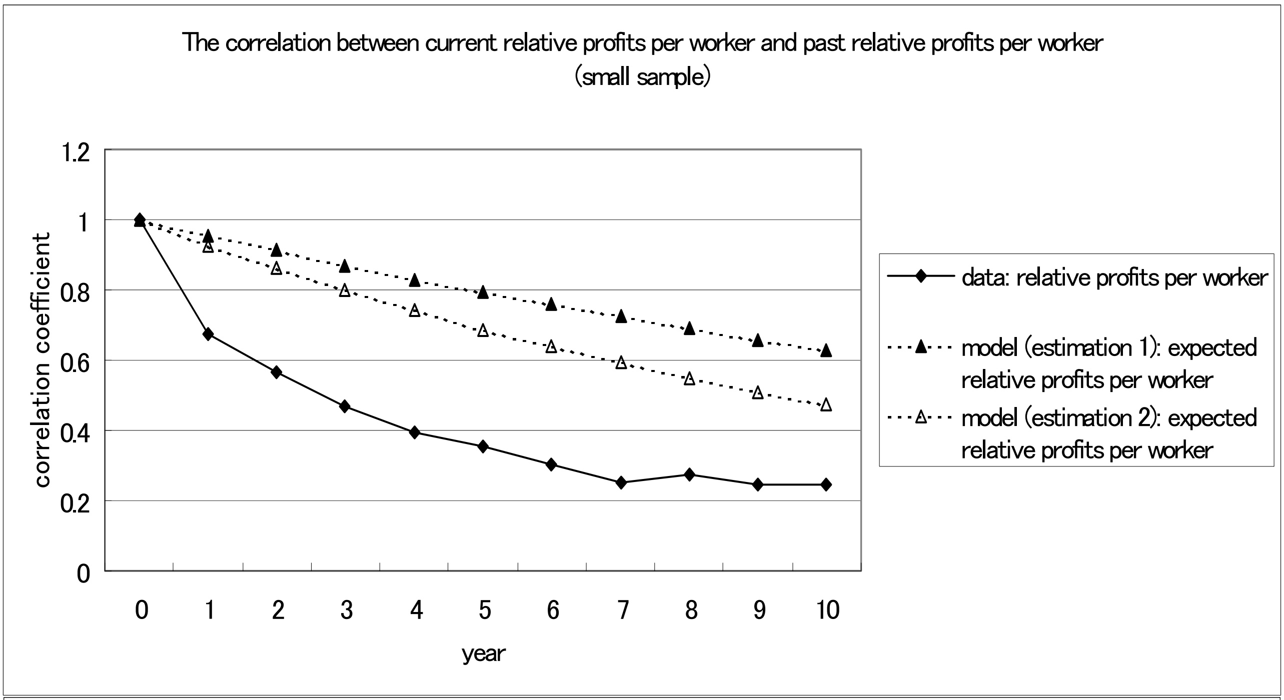

The correlation between current relative wages and past relative wages
(large sample)
1.2
1
0.8
0.6
0.4
0.2
0
correlation coefficient
0
1
2
3
4
5
6
7
8
9
10
year
data
model (estimation 1)
model (estimation 2)

Figure 2:


The correlation between current relative profits per worker and past relative profits per worker
(small sample)
correlation coefficient
year
data: relative profits per worker
model (estimation 1): expected relative profits per worker
model (estimation 2): expected relative profits per worker

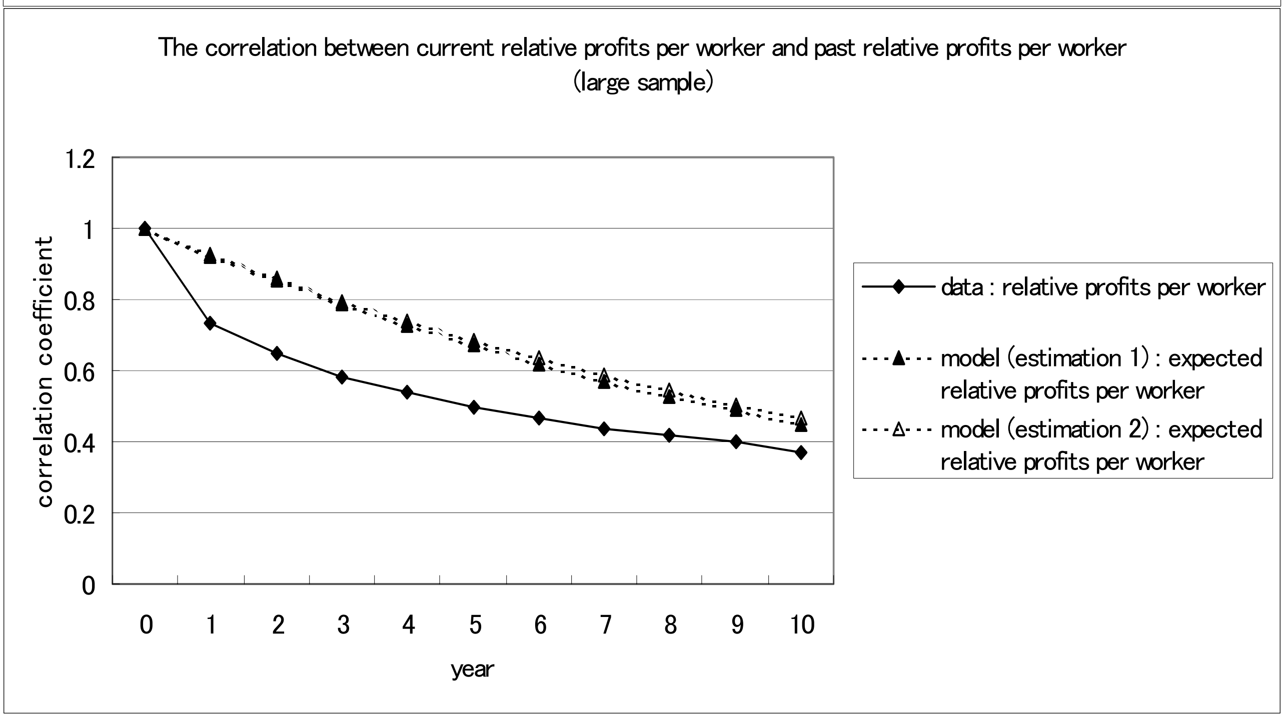

The correlation between current relative profits per worker and past relative profits per worker
(large sample)
correlation coefficient
year
data : relative profits per worker
model (estimation 1) : expected relative profits per worker
model (estimation 2) : expected relative profits per worker

Figure 3:

We also ask "What would happen if $h_\infty = 1$?". Table 5 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.

When h_∞ changes, $\frac{\gamma\sigma_q}{\sigma_{\mu\infty}}$ changes through equation (19), 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, wages and expected profits per worker even in a small sample.

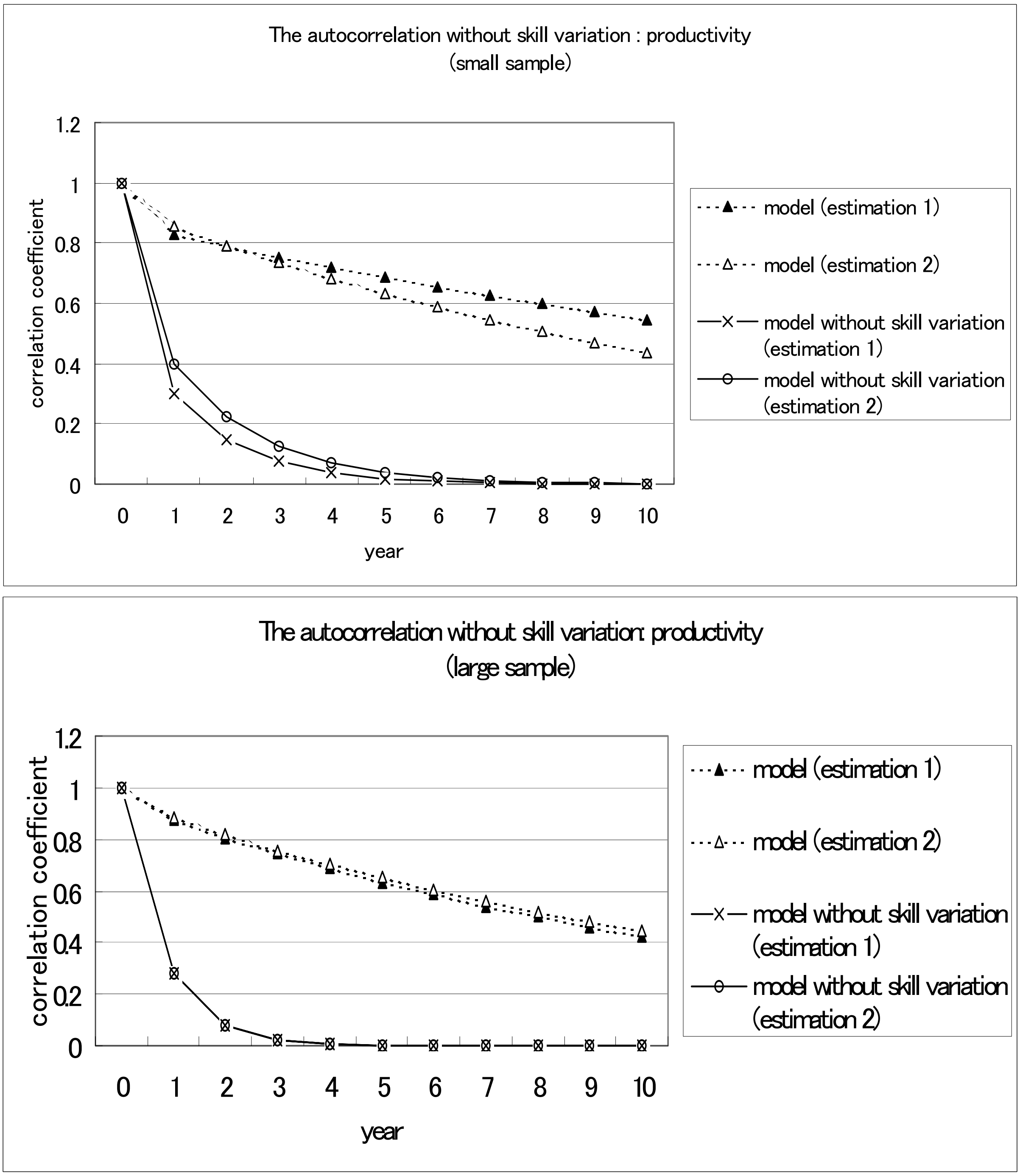
In summary, these exercises consistently suggest that positive assortative assignment accounts for much of the observed persistence of a firm's relative advantages (disadvantages), whereas the noisiness of information plays a relatively minor role.

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 $\rho_{\ln y \ln w} \equiv \frac{E[D\ln y_t D\ln w_t]}{\sqrt{Var(D\ln y_t)Var(D\ln w_t)}}$. The correlation between $D\ln y_t$ and $D\ln w_t$ can be simulated by:

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

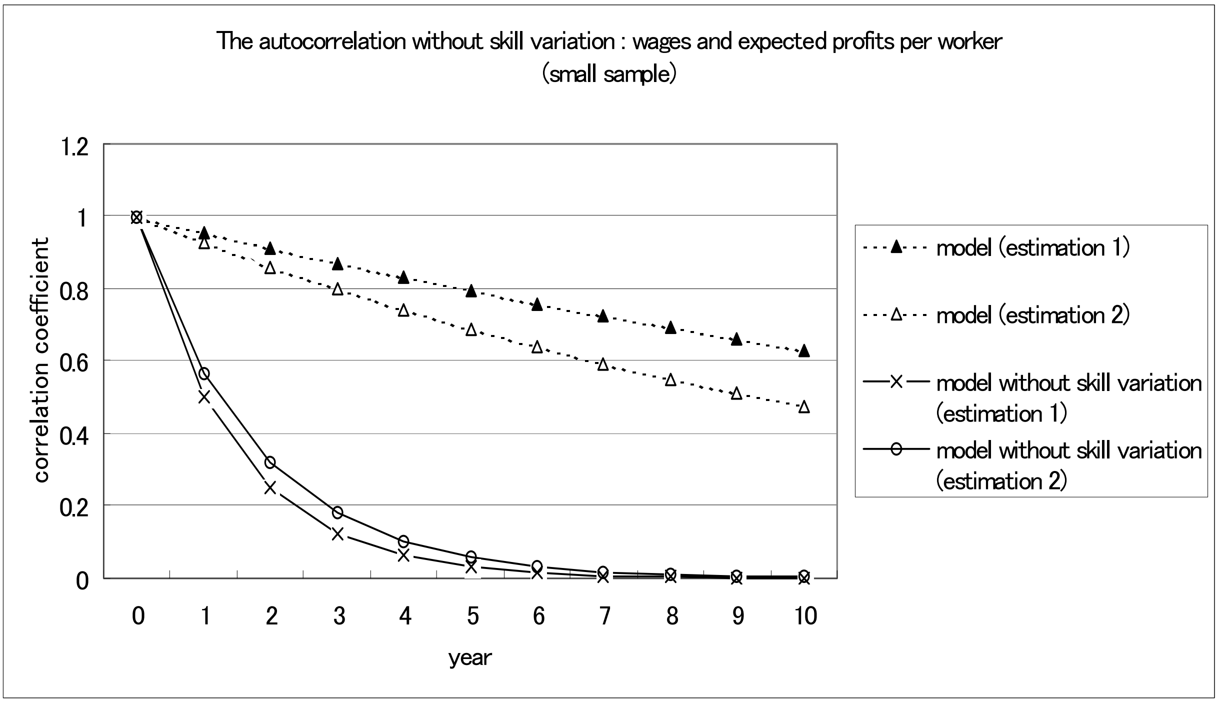
This equation states that the correlation can be predicted by using our estimated parameters. We compare the simulated correlations with the observed ones. The observed correlation between relative productivity and relative wages is estimated by using a method similar to that used to estimate equation (24).

Table 6 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



The autocorrelation without skill variation : productivity
(small sample)
1.2
1
0.8
0.6
0.4
0.2
0
correlation coefficient
0 1 2 3 4 5 6 7 8 9 10
year
model (estimation 1)
model (estimation 2)
model without skill variation (estimation 1)
model without skill variation (estimation 2)
The autocorrelation without skill variation: productivity
(large sample)
1.2
1
0.8
0.6
0.4
0.2
0
correlation coefficient
0 1 2 3 4 5 6 7 8 9 10
year
model (estimation 1)
model (estimation 2)
model without skill variation (estimation 1)
model without skill variation (estimation 2)

Figure 4:


The autocorrelation without skill variation : wages and expected profits per worker
(small sample)
1.2
1
0.8
0.6
0.4
0.2
0
correlation coefficient
0
1
2
3
4
5
6
7
8
9
10
year
model (estimation 1)
model (estimation 2)
model without skill variation (estimation 1)
model without skill variation (estimation 2)

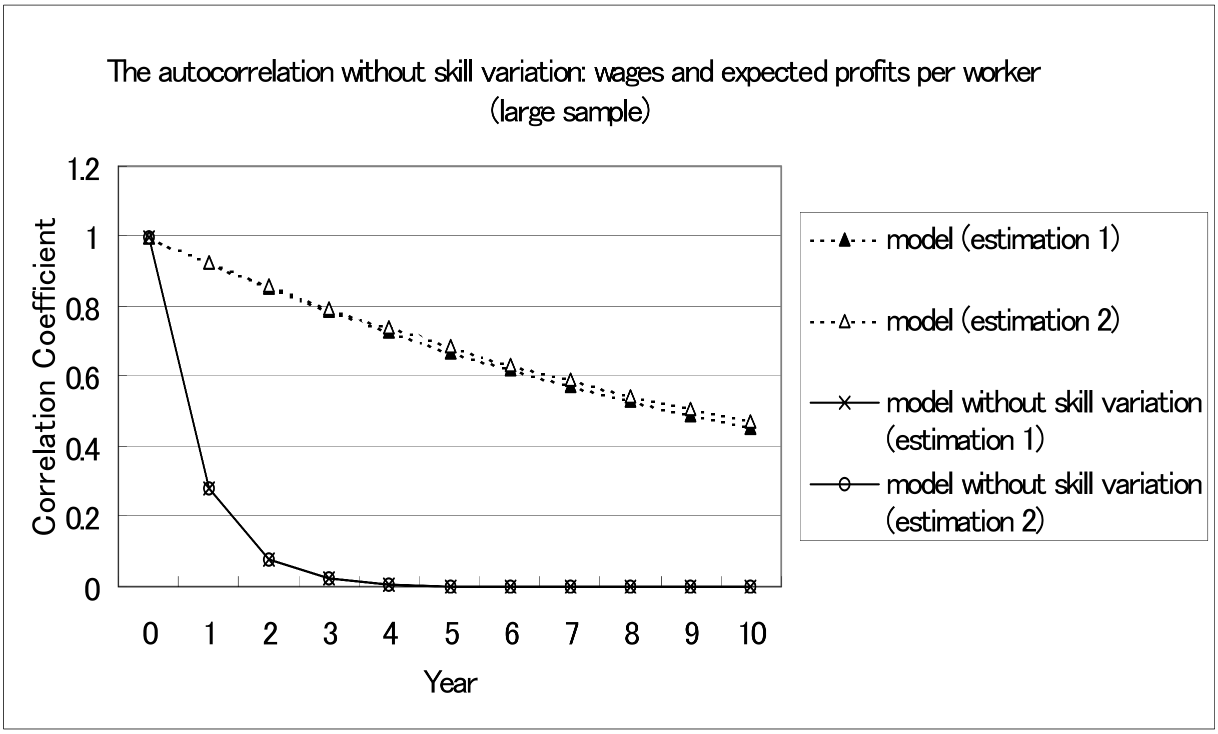
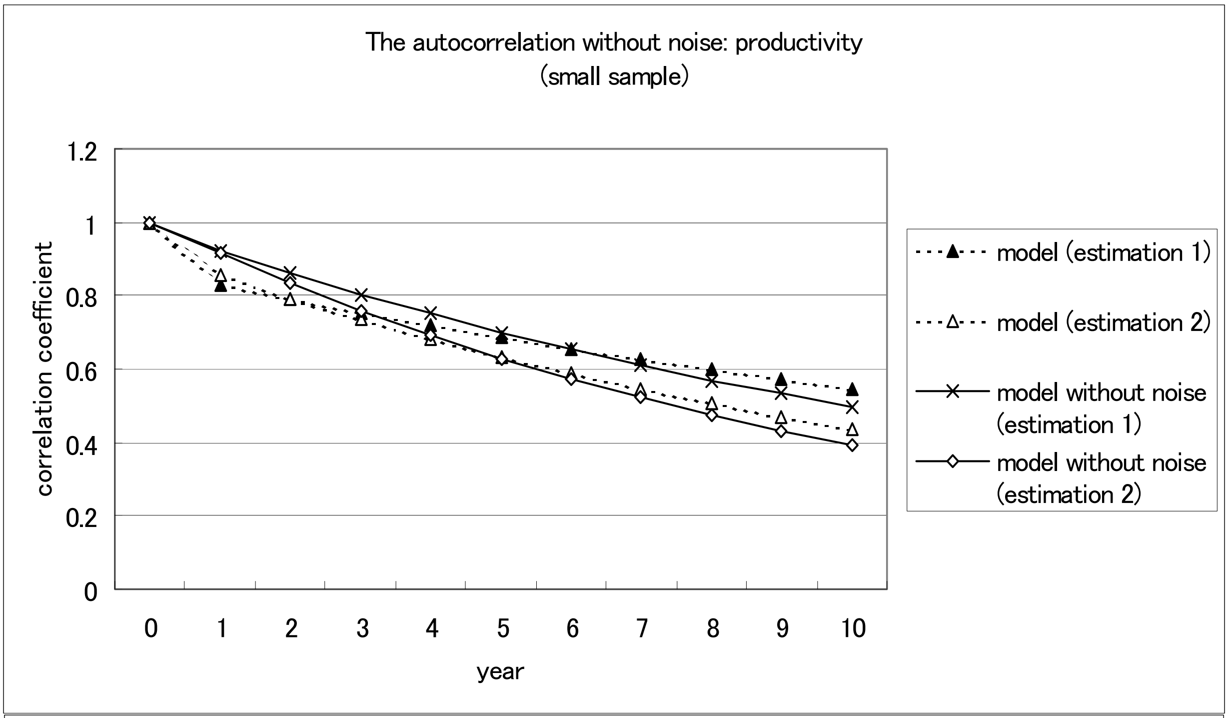

The autocorrelation without skill variation: wages and expected profits per worker
(large sample)
1.2
1
0.8
0.6
0.4
0.2
0
Correlation Coefficient
0
1
2
3
4
5
6
7
8
9
10
Year
model (estimation 1)
model (estimation 2)
model without skill variation (estimation 1)
model without skill variation (estimation 2)

Figure 5:


The autocorrelation without noise: productivity
(small sample)
1.2
1
0.8
0.6
0.4
0.2
0
correlation coefficient
0
1
2
3
4
5
6
7
8
9
10
year
model (estimation 1)
model (estimation 2)
model without noise
(estimation 1)
model without noise
(estimation 2)

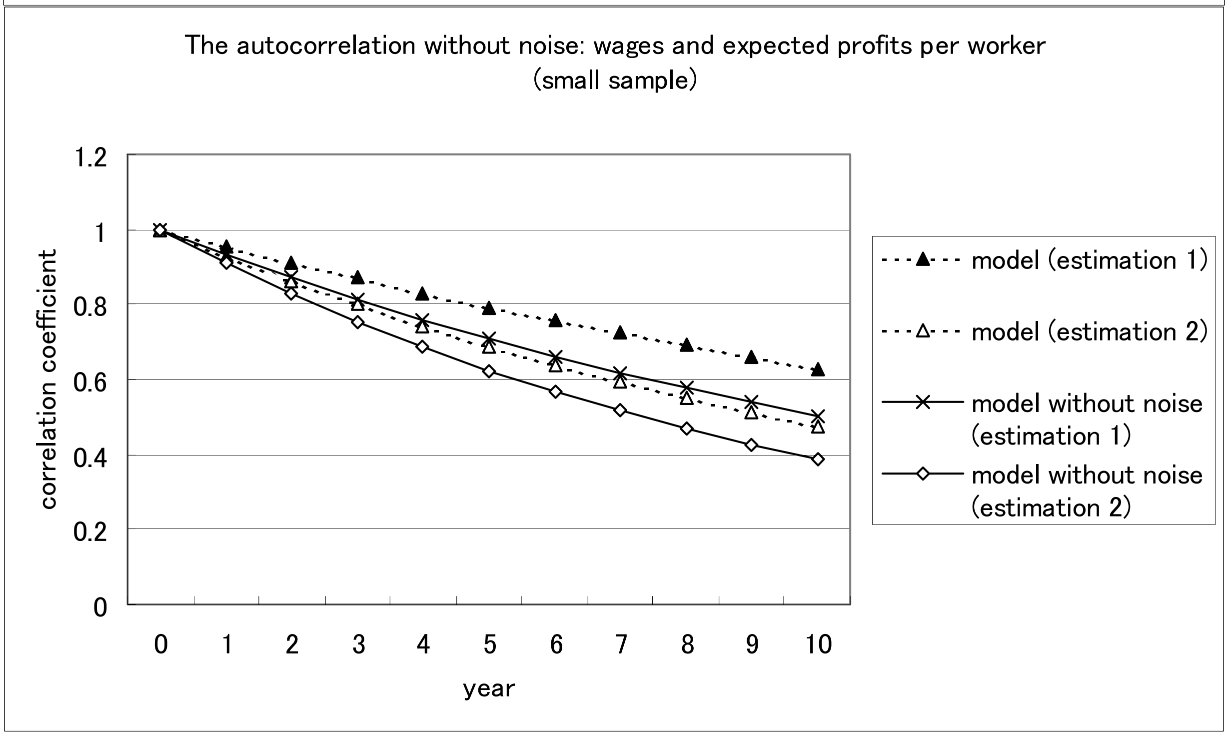

The autocorrelation without noise: wages and expected profits per worker
(small sample)
1.2
1
0.8
0.6
0.4
0.2
0
correlation coefficient
0
1
2
3
4
5
6
7
8
9
10
year
model (estimation 1)
model (estimation 2)
model without noise
(estimation 1)
model without noise
(estimation 2)

Figure 6:

	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 6: 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.

still high. Hence, the high predicted correlation reasonably captures the feature of the actual correlation.

Similarly to the previous argument, if we assume that $\sigma_q = 0$ the correlation is between 0.28 and 0.4.[13]. This means that the observed positive correlation between 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.

5 Conclusion

In this paper, we developed a dynamic assignment model of the relationship between the skills of workers and unobserved firm-specific knowledge, which we term a firm's

[13] When $\sigma_q = 0$, the wage is 0. Hence, the result obtained from the model that does not incorporate skill variations, in Table 6, can be interpreted as the correlation when $\sigma_q \approx 0$.

organization capital, to provide a unified explanation of observed persistence in a changing and uncertain economy. We posed two specific questions. 1) What is the mechanism that enables productive firms to maintain their core resources, and what prevents unproductive firms from investing these resources in a changing environment? 2) To what extent is the observed persistence influenced by the discrepancy between beliefs and fundamental values? The answers based on our analysis are that two feedback mechanisms induced by assignment between unobserved organization capital and skills prevent unproductive firms from investing in firm-specific knowledge and that the inaccuracy of beliefs about fundamental values explains a minor proportion of the observed persistence.

Some points are worth discussing. We defined organization capital as the intangible assets embodied in an organization and modeled this capital as a variant of vintage human capital. To develop the simplest possible model, we assumed that the labor market is perfectly competitive and that a firm receives all the benefits of its organization capital. This is the standard approach used in modeling organization capital in a macroeconomic framework [see, e.g., Atkeson and Kehoe (2005) and Samaniego (2006)]. However, if all workers were to leave a firm at the same time, one would not expect that firm to be able to maintain intangible assets. Hence, it is implicitly assumed that knowledge transferred from senior workers to younger workers maintains organization capital.

It would be interesting to discuss how our results might change if we extended our model to incorporate a long-term relationship in order to analyze explicit interaction between senior workers and junior workers. As explained by Prescott and Visscher (1980), a source of organization capital is firm-specific or relation-specific human capital. Because firm-specific human capital is valuable only to a particular firm, this discussion is particularly important if firm-specific human capital is the main component of organization capital.

We would not expect the qualitative effects of assignment on persistence to be affected by explicitly incorporating a long-term relationship. One could consider a model in which senior workers develop the organization capital of the firm and in which there is assignment between organization capital and the skills of newly employed workers. Profits and wages would continue to depend on organization capital, and the dynamics of organization capital would continue to be influenced by assignment. Of course, there are some differences. We expect that it would alter the wage function and magnify the quantitative impacts of assignment on persistence. Although these considerations would raise several interesting separate issues, because incorporating an internal labor market would complicate the model, it would represent an interesting extension of our model.

In general, assignment models are not suitable for addressing questions about the dynamics of firm size. This is because they require that the number of workers be fixed. However, it would be possible to extend our 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. 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 (4) 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.

6 Data Appendix[14]

- 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

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

  where $(\#29)_{ft}$ is the number of workers in the fth 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_{Yt} is the number of firms in set Y_t and n_{Xt} is the number of firms in set X_t. This is an estimate of labor and related

[14] $(\#X)_{ft}$ implies COMPUSTAT number X of fth firm in year t and $(\#X)_{fit}$ implies COMPUSTAT number X of fth firm in ith industry in year t. Summary statistics on the variables used for estimation are in Takii (2007b).

expenses that includes employee benefits. Second, if a firm does not report labor and related expenses, we estimate these expenses by

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

where Z_{it} is the set of firms that report labor and related expenses in the ith industry in year t and n_{Zit} 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.

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] Bartelsman, Eric J. and Doms, Mark (2000), "Understanding Productivity: Lessons from Longitudinal Microdata," Journal of Economic Literature, 569–594.

[8] Becker, Gary S. (1973), "A Theory of Marriage: Part 1," Journal of Political Economy, 81, 813–846.

[9] Bond, Stephen R. and Jason G. Cummins (2000), "The Stock Market and Investment in the New Economy: Some Tangible Facts and Intangible Fictions," Brookings Papers on Economic Activity, No.1, 61-108

[10] 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.

[11] Chandler, A. (1977), *The Visible Hand, the Managerial Revolution in American Business*, The Belknap Press of Harvard University Press, Cambridge, Massachusetts.

[12] 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.

[13] Caroli, Eve and John, Van Reenen (2001), "Skill-Biased Organizational Change? Evidence from a Panel of British and French Establishments," Quarterly Journal of Economics, Vol.116, No.4., November, 1449-1492.

[14] Davis, Steven J. and Haltiwanger, John (1999), "Gross Job Flows," *In Handbook of Labor Economics, Volume 3*, Ashenfelter, Orley and Card, David (eds), 2711–2805.

[15] 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.

[16] Doms, Mark, Timothy, Dunne and Kenneth R.,Troske (1997), "Workers, Wages, and Technology," Quarterly Journal of Economics, Vol.112, No.1, February, 253-290.

[17] Faria André L. (2003), "Mergers and the Market for Organization Capital," Mimeo.

[18] Hall, Robert D. (2001), "The Stock Market and Capital Accumulation," American Economic Review, Vol. 91, No. 5, December, 1185–1202.

[19] 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.

[20] Haltiwanger, John C., Julia I. Lane and James R. Spletzer (2007), "Wages, Productivity, and the Dynamic Interaction of Businesses and Workers," Labor Economics, 14, 575-602.

[21] Jovanovic, Boyan (1982), "Selection and the Evolution of Industry," Econometrica, Vol. 50, No. 3, May, 649–670.

[22] Jovanovic, Boyan (1998), "Vintage Capital and Inequality," Review of Economic Dynamics, Vol. 1, 497–530.

[23] Jovanovic, Boyan and Peter L. Rousseau (2001), "Vintage Organization Capital," NBER Working Paper 8166.

[24] Kaldor, Nicholas (1934), "The Equilibrium of the Firm," Economic Journal, Vol. 44, 60–76.

[25] Kremer, Michael (1993), "The O-Ring Theory of Economic Development," Quarterly Journal of Economics, Vol. 108, August, 551–575.

[26] Lucas, Robert E. (1978),"On the Size Distribution of Business Firms," Bell Journal of Economics, Vol. 9, 508–523.

[27] McGahan, Anita M. (1999), "The Performance of US Corporations: 1981–1994," Journal of Industrial Economics, Vol. 47, No. 4, December, 373–398.

[28] Nelson, Richard R. and Sidney G. Winter (1982), *An Evolutionary Theory of Economic Change*, The Belknap Press of Harvard University Press, Cambridge, Massachusetts.

[29] Penrose, E. (1959), *The Theory of the Growth of the Firm*, Oxford University Press, Oxford, U.K.

[30] Prescott, E. and M. Visscher (1980), "Organization Capital," Journal of Political Economy, Vol. 88, 446–461.

[31] Robinson, Austin (1934), "The Problem of Management and the Size of Firms," Economic Journal, Vol. 44, 242–257.

[32] Rosen, Sherwin (1972), "Learning by Experience as Joint Production," Quarterly Journal of Economics, Vol. 86, August, 366–382.

[33] Samaniego, Roberto M. (2006), "Organizational Capital, Technology Adoption and the Productivity Slowdown," Journal of Monetary Economics, Vol. 53, October, 1555–1569.

[34] Sattinger, Michael (1979), "Differential Rents and the Distribution of Earnings," Oxford Economic Papers, Vol. 31, March, 60–71.

[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 (2007a), "The Value of Adaptability - Through the Analysis of a Firm's Prediction Ability," Journal of Economics and Business, March-April 2007, vol.59, 144-162.

[38] Takii, Katsuya (2007b) "The Persistence of Differences in Productivity, Wages, Skill Mixes and Profits Between Firms," OSIPP Discussion Paper: DP-2007-E-002 (http://econpapers.repec.org/paper/ospwpaper/07e002.htm) / Discussion Papers in Economics and Business 07-10 (http://ideas.repec.org/p/osk/wpaper/0710.html)