RIETI Discussion Paper Series 11-E-048

Persistent Productivity Differences Between Firms

TAKII Katsuya
Osaka University

RIETI
Research Institute of Economy, Trade & Industry, IAA

The Research Institute of Economy, Trade and Industry
http://www.rieti.go.jp/en/

RIETI Discussion Paper Series 11-E-048

April 2011

Persistent Productivity Differences Between Firms*

Katsuya TAKII† (Osaka University)

Abstract

We construct a dynamic assignment model that explains persistent productivity differences between firms. Large expected organization capital (firm-specific knowledge) attracts skilled workers, who help to accumulate organization capital. Accumulated large organization capital leads to good performances, which, in turn, confirm high expectations. It is shown that the sluggish movement of expected productivity that occurs through this positive feedback can play a role similar to an unobserved fixed effect in the productivity dynamics. Our calibration exercises suggest that the proposed feedback accompanied by amplification mechanisms inherent in the assignment model can explain a major part of the observed persistence and disparity in productivity.

Keywords: organization capital, assignment, productivity, disparity, and persistence.

JEL Classification: J24; L25

* This research is being conducted as a part of the project "Basic Research on Japanese Labor Market" by the Research Institute of Economy, Trade and Industry. I would like to thank Eric Bartelsman, Julen Esteban-Pretel, Masahisa Fujita, Robert Gibbons, Shigeru Iwata, Rasmus Lentz, Makoto Saito, Ryuichi Tanaka, Futoshi Yamauchi, and seminar participants at numerous seminars for helpful comments and encouragement. 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) is gratefully acknowledged.

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

1 Introduction

Why are some firms persistently more productive than others? Evidence reveals substantial and persistent differences in productivity between plants or between firms [e.g., Baily et al. (1992)]. The same evidence can be found in several countries [e.g. Fox and Smeets (2010), Fukao and Kwon (2006)]. This indicates that the persistent differences in productivity are universal. Taking the persistent heterogeneity between firms as given, researchers investigate how the reallocation of resources across heterogeneous firms influences the aggregate economy [e.g., Melitz (2003) and Lentz and Mortensen (2008)]. However, the reasons for the productivity differences remain an open question and an unobserved firm-specific fixed effect explains large portions of the productivity differences [e.g., Fox and Smeets (2010)].

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 the coexistence of persistent differences in several variables. If a quasi-fixed firm-specific resource and workers' skills are complementary to each other, a firm endowed with large resources 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, an assignment model 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.

To answer these questions and explain the observed persistent productivity differences between firms, 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) Skills and organization capital are complementary inputs. 2) Skills are an input in the accumulation of organization capital. 3) Although we cannot directly observe the amount of firm-specific knowledge, we can infer it from the firm's productivity. These assumptions allow us to analyze how not only the assignment mechanism, but also the discrepancy between beliefs and fundamental values influences the persistence of

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

observed productivity.

The main logic can be explained as follows. If a firm's organization capital is believed to be high, this belief attracts skilled workers. On the other hand, because skill is an input in 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 confirms the perception that the firm has a higher level of organization capital. We argue that this positive feedback mechanism causes persistent productivity differences.

We derive the dynamics of relative productivity, where relative productivity is measured by the logarithm of the total factor productivity (TFP) relative to industry and year averages. When assignment between beliefs and skills exists, the dynamics of relative productivity is shown to exhibit a reversion to a firm-specific fixed effect plus the expected relative productivity. As long as the adjustment of the expectation is very slow, the expected relative productivity is not empirically distinguishable from the unobserved fixed effect in the standard productivity dynamics regressions. In fact, when productivity has no predictive power for organization capital, nobody can update their beliefs. Therefore, the expected relative productivity never changes. That is, it is shown that a fixed effect can be derived without any real heterogeneity across firms.

In order to quantify the persistence of relative productivity, we measure it by the correlation between current relative productivity and its lagged values. This measure can be decomposed into two parts: the share of the variance of the long-run average relative productivity in the variance of the relative productivity (the between-to-overall variance ratio), which measures the permanent persistence, and the correlation between the current deviation of the relative productivity from its long-run average and its lagged values (within-correlation), which measures the temporal persistence. The diversity of productivity can be measured by the variance of the relative productivity that is the sum of the variance of the long-run average relative productivity (between-variance) and the variance of the deviation of the relative productivity from its long-run average (within-variance). We derive analytical predictions on each component. Examining these predictions, we analytically and quantitatively analyze the reason behind the persistence and diversity of productivity differences.

Three analytical results are worthy of special attention. First, it is shown that the within-correlation (i.e., the measure of the temporal persistence) is strongly influenced by the correlation between the current expected deviation of relative productivity from its long-run average and its lagged values (the measure for the persistence of the expected deviation, hereafter). Theory suggests that the assignment between the beliefs regarding organization capital and skills induces the persistence of the beliefs that anchor the dynamics of relative productivity. The measure for the persistence of the expected deviation captures the persistence of the beliefs through the proposed feedback mechanism.

Second, we identify three main factors that can influence the persistence of the expected deviation – the variation of skills, the difficulty in inferring organization capital, and the variation of a firm-specific fixed effect. When the variance of skills is high, the top organization has the greatest advantage because it can attract the best workers, who can provide the firm with the best knowledge and promote the accumulation of organization capital, which, in turn, helps maintain beliefs in high organization capital. This effect increases the persistence of the expected deviation. We measure the importance of assignment by the variation of skills because if there is no heterogeneity of skill, the assignment problem does not exist.

There are two other factors. As we discussed before, if it is difficult to infer organization capital, people cannot change their beliefs. Because the assignment occurs between the beliefs and skills, people's beliefs can be confirmed by the assignment of skilled workers who correspond to their beliefs. Therefore, the expected deviation of relative productivity is more persistent. More interestingly, an increase in firm-specific heterogeneity lowers the persistence of the expected deviation. If a top organization already has a big advantage, the benefits from assigned skilled workers can be small relative to what it has already. Hence, the temporal deviation from the long-run average is expected to be short. Therefore, the persistence of the expected deviation is small.

Third, among the three factors that we identify, we find that only an increase in skill variation can unambiguously raise the persistence and diversity of productivity. In other words, an increase in firm-specific heterogeneity may lower the persistence and diversity of productivity. A rise in skill variation increases not only the persistence of the beliefs, but also the impacts of the beliefs on productivity because it makes it possible for a firm believed to have high organization capital to attract more highly skilled workers. Both mechanisms amplify small firm heterogeneity and increase the persistence and diversity of productivity. On the other hand, because a rise in firm-specific heterogeneity lowers the benefits from assignment, the amplification effect becomes smaller. Therefore, the persistence and diversity of productivity might be reduced.

We use the theory to quantify the importance of three factors as a source of the persistence and diversity of productivity. For this purpose, we calibrate parameters so that the derived dynamics and the disparity of relative productivity is consistent with observed dynamics and the diversity of productivities, using firm-level data from Japan (the Basic Survey of Japanese Business Structure and Activities [BSJBSA]) between 1994 to 2004.

The calibrated parameters reveal that the implied variations of a firm-specific fixed effect are very small. This means that the observed variations in productivity differences can be explained without having large variations on the firm side. Based on the calibrated parameters, we simulate our analytical correlation. It is shown that the model can quantitatively capture changes in the correlation of the relative productivity over time.

We conduct several counterfactual experiments. They show that if there was no assignment problem, the correlation of relative productivity diminishes to less than 10 percent after two years and the variance of relative productivity drops by 79 percent. On the contrary, even if there is no variations in the firm-specific factors, the correlation remains about 34 percent after 10 years and the variance drops only by 15 percent. Finally, noisy information has only a minor influence on the correlation and variance of relative productivity.

This result is less likely to be specific to Japan. It is also shown that similar results can be confirmed by the analysis of labor productivity using an industry annual dataset in COMPUSTAT covering 1970 to 2004. In sum, the quantitative exercises repeatedly suggest that the assignment mechanism that causes the sluggish movement of beliefs and a rise in the sensitivity of productivity to the beliefs has quantitatively large impacts on the observed persistence and disparity of productivity; firm-specific heterogeneity and noisy information have only modest impacts.

It has long been recognized that an individual firm possesses firm-specific knowledge. 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) quantified the macroeconomic effects of organization capital. 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 vintage human capital. For any organization, ancestors determine a particular routine, culture, and organizational structure 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 have long-run effects on organization capital.

Positive assortative assignment models also have a long history [e.g., Becker (1973) and Sattinger (1979)]. Kremer (1993) demonstrated that the model of positive assortative matching among workers can explain a variety of evidence. Our model conveys the spirit of Kremer's (1993) idea in a dynamic framework; that is, current skilled workers attract skilled successors.

Most assignment models are static and the distribution of assigned variables is treated as given. Notable exceptions are Acemoglu (1997) and Jovanovic (1998). Both authors examine persistent income inequality. Unlike them, we endogenize the distribution of organization capital and examine persistent differences in productivity.

Learning is another important feature of the model. As Jovanovic (1982) explained, a firm gradually learns its own productive capacity. However, unlike Jovanovic (1982), we assume that a firm's productive capacity itself changes because of investment in organization capital and 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. First, evidence shows that a productive organizational arrangement demands skill [e.g., Caroli and Van Reenen (2001) and Bresnahan et al. (2002)]. This is consistent with our assumption of complementarity between organization capital and skill[2]. Second, the evidence is also consistent with the assumption that skill is important for the accumulation of organization capital. Evidence from Caroli and Van Reenen (2001) suggests that firms need the intangible assets accumulated by skilled workers to make organizational changes productive. In addition, evidence shows that the intangible assets accumulated by skilled workers are an important determinant of technology adoption [e.g.,

[2]Recently, researchers have tried to find more direct evidence using matched employer–employee data. Although controversial evidence has been found by some researchers (e.g., Abowd, Kramarz, and Margolis 1999), other researchers have criticized the identification strategy used by these authors (e.g., Bagger and Lentz (2007), Eeckhout and Kircher (2009)) and developed alternative methods to examine empirically the positive assortative matching hypothesis in a labor market. Among them, Mendes, Van den Berg, and Lindeboom (2007) and Lopes de Melo (2008) provided evidence based on their methods that supports a positive assortative matching hypothesis. At least, there is no evidence against a positive assortative matching hypothesis if observable characteristics are not controlled for, which is more relevant in this paper.

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 explain an intuition regarding how persistent productivity differences occur. In Section 3, we set up a dynamic positive assortative matching model between unobserved organization capital and skill. The existence and stability of a stationary distribution is proved. In Section 4, we derive the dynamics of relative productivity and provide analytical predictions. In Section 5, we calibrate parameters and conduct several quantitative exercises, which produce the quantitative predictions of the model. In Section 6, we summarize the results and conclude by providing three implications for empirical research on productivity dynamics.

2 Intuitive explanation for the mechanism of the persistent productivity difference

The persistent difference in productivity can be derived without any uncertainty. Hence, it is instructive to start with the assumption of perfectly observable organization capital. In the next section, we extend the model to unobserved organization capital in order to emphasize the role of beliefs .

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 there is a set of firms, the total mass of which is also normalized to unity. Suppose that $\ln k_t^o$ is normally distributed with a mean of μ_{ot} and a standard deviation of σ_{ot} at the date t. Assume also that $\ln q_t$ is normally distributed with a mean of μ_q and a standard deviation of σ_q at any date. For simplicity, we assume that firms and workers have reservation values of 0. Because the number of firms is the same as the number of workers, nobody chooses the outside option and every agent can find a partner. Hence, these assumptions make it possible to focus on the assignment problem.

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

$$1 - \Phi\left(\frac{\ln k_t^o - \mu_{ot}}{\sigma_{ot}}\right) = 1 - \Phi\left(\frac{\tilde{\chi}(\ln k_t^o) - \mu_q}{\sigma_q}\right), \ \forall \ln k_t^o. \tag{1}$$

where $\tilde{\chi}(\cdot)$ is a firm's optimal policy function, which maps the space of organization capital to the space of quality of workers. We will explicitly describe the firm's optimization decision in the next section, but whatever the firm's decision problem is, equation (1) states that the policy function must satisfy:

$$\tilde{\chi}(\ln k_t^o) = \frac{\sigma_q}{\sigma_{ot}}\left[\ln k_t^o - \mu_{ot}\right] + \mu_q, \ \forall \ln k_t^o.$$

Note that a one percent increase in $\ln k_t^o$ makes the firm employ $\frac{\sigma_q}{\sigma_{ot}}$ percent more talented workers. That is, because of assignment, the benefits from having an additional large $\ln k_t^o$ depends on the variation of skill relative to the variation of organization capital, $\frac{\sigma_q}{\sigma_{ot}}$.

The dynamics of organization capital is described by the following equation:

$$k_{t+1}^o = B\left(k_t^o\right)^{\phi}\left(q_t\right)^{\gamma} e^{\varepsilon_t}, 0 \leq \phi < 1, \ \gamma > 0, \tag{2}$$

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

Because $\ln q_t = \chi(\ln k_t^o) = \frac{\sigma_q}{\sigma_{ot}}(\ln k_t^o - \mu_{ot}) + \mu_q$, in equilibrium, the dynamics of organization capital can be written as:

$$\ln k_{t+1}^o = \ln B + \phi \ln k_t^o + \gamma\left[\frac{\sigma_q}{\sigma_{ot}}\left(\ln k_t^o - \mu_{ot}\right) + \mu_q\right] + \varepsilon_t. \tag{3}$$

Because $\ln k_t^o$ and ε_t are normally distributed, $\ln k_{t+1}^o$ is also normally distributed. Hence, the dynamics of μ_{ot} and σ_{ot} can be derived from equation (3) as follows:

$$\mu_{ot+1} = \ln B + \phi\mu_{ot} + \gamma\mu_q - \frac{\sigma_\varepsilon^2}{2},\ \sigma_{ot+1} = \sqrt{\lambda_t^2\sigma_{ot}^2 + \sigma_\varepsilon^2},$$

where $\lambda_t = \phi + \frac{\gamma\sigma_q}{\sigma_{ot}}$. These two equations characterize the dynamics of the aggregate state variables. By using the dynamics of μ_{ot}, equation (3) can be rewritten as:

$$\ln k_{t+1}^o - \mu_{ot+1} = \lambda_t\left(\ln k_t^o - \mu_{ot}\right) + \varepsilon_t^*,$$

where $\varepsilon_t^* = \varepsilon_t + \frac{\sigma_\varepsilon^2}{2}$ is normally distributed with a mean of 0 and a standard deviation of σ_ε. This equation implies that when $\ln k_t^o$ is larger than its industry mean, μ_{ot}, the fraction λ_t of this relative advantage is carried over to the next period. Because persistence is only influenced by λ_t, we refer to this as the persistence parameter in what follows. Note that the persistence parameter is composed of technological persistence ϕ and persistence due to the assignment mechanism, $\frac{\gamma\sigma_q}{\sigma_{ot}}$. It shows that the larger the assignment effect is, the larger the persistence.

We can show that σ_{ot} and μ_{ot} globally converge to the stationary points, which are denoted by $\sigma_{o\infty}$ and $\mu_{o\infty}$. Then, the dynamics of organization capital are as follows:

$$D\ln k_{t+1}^o = \lambda_\infty D\ln k_t^o + \varepsilon_t^*, \tag{4}$$

where $D\ln k_t^o = \ln k_t^o - \mu_{o\infty}$, $\lambda_\infty = \phi + \frac{\gamma\sigma_q}{\sigma_{o\infty}}$ and $\frac{\sigma_{o\infty}}{\gamma\sigma_q} = \frac{\phi+\sqrt{1+\left(1-\phi^2\right)\left(\frac{\sigma_\varepsilon}{\gamma\sigma_q}\right)^2}}{1-\phi^2}$. Note that $\sigma_{o\infty}$ is not 0. Because σ_q is always positive, when σ_{ot} is small, the assignment effect, $\frac{\gamma\sigma_q}{\sigma_{ot}}$, is large. Hence, a firm with a relatively large organization capital benefits substantially. This mechanism increases σ_{ot}. Ultimately, $\sigma_{o\infty}$ does not converge to 0.

We can show that the persistence parameter, λ_∞, is increasing in skill variation relative to the standard deviation of random shocks, $\frac{\sigma_q}{\sigma_\varepsilon}$, which measures the importance of assignment.

$$\lambda_\infty \in (\phi, 1),\ \frac{d\lambda_\infty}{d\frac{\sigma_q}{\sigma_\varepsilon}} > 0,\ \lim_{\frac{\sigma_q}{\sigma_\varepsilon}\to\infty}\lambda_\infty = 1. \tag{5}$$

As the persistence parameter is always less than 1, equation (4) is covariance stationary. Hence, the correlation between current organization capital and the jth organization capital, which is equivalent to the jth autocorrelation of $D\ln k_t^o$ in the stationary environment, is entirely determined by the persistence parameter λ_∞, $\rho_{oj} \equiv \frac{E\left[D\ln k_t^o D\ln k_{t-j}^o\right]}{\sigma_{o\infty}^2} = \lambda_\infty^j$.

Note that if there is no shock, $\frac{\sigma_q}{\sigma_\varepsilon} = \infty$ and $\rho_{oj} = 1$ for any j. When there are no idiosyncratic shocks, the top organization always attracts the best workers, who, in turn, equip the firm with the best knowledge. Hence, the firm remains at the top and maintains exactly the same level of organization capital in the long run. When we introduce idiosyncratic shocks into the accumulation of organization capital, $\frac{\sigma_q}{\sigma_\varepsilon} < \infty$, a reversion to the mean occurs. Idiosyncratic shocks make changes in rankings possible. A firm that receives a positive shock climbs the rankings, which enables it to attract higher quality workers. This means that top organizations cannot remain the best and might slip down the rankings.

The next section incorporates uncertainty into the model and formally defines the equilibrium.

3 The Model

In this section, we explicitly describe the firm's optimization decision and define the equilibrium with unobserved organization capital. In addition, we add the time-invariant component in organization capital and physical capital in the production function. These factors are not necessary to describe the main logic of this paper. However, as it is known that a fixed effect explains a large variation of the TFP between firms, this is a requirement for a quantitative theory of the TFP dynamics.

We assume that a firm with the TFP A_t employs physical capital k_t and a unit mass of workers[3] and produces output of y_t according to the production function

[3] A fixed amount of labor is used for production. As Lentz and Mortensen (2008) shows, it is known that there is almost no correlation between labor productivity and labor inputs. Hence, as our main focus is the productivity dynamics, this is a convenient assumption for a simple analysis.

$y_t = A_t k_t^{\delta}$, where $\delta \in (0,1)$. Because the total mass of workers is 1, we can alternatively interpret y_t (k_t) as the output per unit of workers (physical capital per unit of workers).

Assume that A_t is a function of organization capital, k_t^o, and the average quality of workers, q_t:

$$A_t = e^{u_t} k_t^o q_t^{\psi}, \; q_t = \left[\int_0^1 q_{jt}^{\alpha} dj \right]^{\frac{1}{\alpha}}, \tag{6}$$

where $\psi > 0$ and $\alpha \leq 1$ are constant parameters and u_t is a firm-specific productivity shock, which 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. The range of the parameter α allows that each worker can be complementary to each other. This assumption captures Kremer's (1993) idea of team production.

Assume that organization capital has two components, $\ln k_t^o = \ln k^{of} + \ln k_t^{ov}$, where k^{of} is a firm-specific fixed component and k_t^{ov} is a time-varying transitory component. We assume that k^{of} is known and perfectly observable, but that k_t^{ov} cannot be directly observed and must be inferred from the realizations of the TFPs. When employment decisions are made, we assume that the TFP is not realized. Hence, a decision must be based on a conditional expectation given the prior beliefs about the level of organization capital. We assume that the prior distribution of $\ln k_t^{ov}$ is normally distributed with a mean of μ_{ot}^v and a variance of σ_{ot}^2. Then, the prior distribution of $\ln k_t^o$ is normally distributed with a mean of $\mu_{ot} = \ln k^{of} + \mu_{ot}^v$ and a variance of σ_{ot}^2. Hence, the expected output is $E\left[y_t | \mu_{ot}, \sigma_{ot}, \ln q_t, k_t\right] = \exp\left(\mu_{ot} + \frac{\sigma_{ot}^2}{2} + \psi \ln q_t + \delta \ln k_t\right)$. Assume that a firm rents physical capital each period at a rental price r. Then, it is easy to show that $\max_{k_t} \left\{E\left[y_t | \mu_{ot}, \sigma_{ot}, \ln q_t, k_t\right] - r k_t\right\} = (1-\delta) E\left[y_t | \mu_{ot}, \sigma_{ot}, \ln q_t\right]$, where:

$$E\left[y_t | \mu_{ot}, \sigma_{ot}, \ln q_t\right] \equiv \left(\frac{\delta}{r}\right)^{\frac{\delta}{1-\delta}} \exp \frac{1}{1-\delta} \left(\mu_{ot} + \frac{\sigma_{ot}^2}{2} + \psi \ln q_t\right). \tag{7}$$

As we implicitly assume that there is a risk-neutral representative household with a constant discount rate, r is constant in the equilibrium. Hence, we take r as given.

Assume that a firm-specific fixed component of organization capital, $\ln k^{of}$, is normally distributed across firms with the mean $-\frac{\sigma_f^2}{2}$ and the variance σ_f^2. We also assume the same transition equation (2) for the dynamics of the transitory component of organization capital, $\ln k_t^{ov}$. Then, the dynamics of $\ln k_t^o$ would be:

$$\ln k_{t+1}^o = \ln B^* + \phi \ln k_t^o + \gamma \ln q_t + F + \varepsilon_t, \tag{8}$$

where $F = (1-\phi)\left(\ln k^{of} + \frac{\sigma_f^2}{2}\right)$ is normally distributed with the mean 0 and the variance $\sigma_F^2 = (1-\phi)^2 \sigma_f^2$ and $\ln B^* = \ln B - (1-\phi)\frac{\sigma_f^2}{2}$. The parameter F represents a firm-specific fixed factor in this paper.

In order to characterize a firm's problem, we need to describe how a firm's current choice influences the beliefs regarding $\ln k_t^o$. After the firm employs a worker, output is produced. From the realized output, the firm knows the TFP and, therefore, $e^{\left(u_t + \frac{\sigma_u^2}{2}\right)} k_t^o$. Hence, a firm uses a signal, $s_t \equiv \ln k_t^o + u_t^*$, to infer $\ln k_t^o$, where $u_t^* = u_t + \frac{\sigma_u^2}{2}$ is normally distributed with a mean of 0 and a standard deviation of σ_u.

Because $\mu_{ot+1} = E\left[\ln k_{t+1}^o | s_t, \mu_{ot}, \sigma_{ot}\right]$ and $\sigma_{ot+1} = \sqrt{Var\left[\ln k_{t+1}^o | s_t, \mu_{ot}, \sigma_{ot}\right]}$, the dynamics of μ_{ot} and σ_{ot} can be written as follows:

$$\mu_{ot+1} = \ln B^* + \phi E\left[\ln k_t^o | s_t, \mu_{ot}, \sigma_{ot}\right] + \gamma \ln q_t + F - \alpha \frac{\sigma_\varepsilon^2}{2}, \tag{9}$$

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

where:

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

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

Equation (11) shows that $E\left[\ln k_t^o | s_t, \mu_{ot}, \sigma_{ot}\right]$ is a weighted average of the prior belief, μ_{ot}, 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, a small weight is placed on s_t. In this way, the variable h_t measures the reliability of new information.

All firms are assumed to have the same σ_{ot} at date t. However, the beliefs regarding organization capital, μ_{ot}, differ 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 beliefs regarding organization capital, μ_{ot}, characterize a firm's position in the economy. We assume that:

$$(\mu_{ot}, F) \tilde{} N\left(\left(\mu_{ot}^e, 0\right), \begin{bmatrix} \sigma_{\mu t}^2 & cov_{\mu F t} \\ cov_{\mu F t} & \sigma_F^2 \end{bmatrix} \right)$$

where μ_{ot}^e and $\sigma_{\mu t}^2$ are the mean and the variance of μ_{ot}, and $cov_{\mu F t}$ is the covariance between μ_{ot} and F. We consider the assignment between beliefs regarding organization capital and the quality of workers. The wage function depends not only on the quality of workers, but also on the aggregate state variables: μ_{ot}^e, $\sigma_{\mu t}$, $:\sigma_{ot}^a$ and $cov_{\mu F t}$, where σ_{ot}^a is a prevailing standard deviation of $\ln k_t^o$. Let $\mathbf{x}_t = \left(\mu_{ot}^e, \sigma_{\mu t}, \sigma_{ot}^a, cov_{\mu F t}\right)'$. Assume that a firm faces a competitive wage function, $w\left(\ln q_t : \mathbf{x}_t\right)$. The profit maximization problem of a firm is written as follows:

$$\begin{aligned} V\left(\mu_{ot}, \sigma_{ot}, F : \mathbf{x}_t\right) &= \max_{\{\ln q_{jt}\}} \left\{ \begin{array}{l} (1-\delta) E\left[y_t | \mu_{ot}, \sigma_{ot}, \ln q_t\right] - w\left(\ln q_{jt} : \mathbf{x}_t\right) + \\ \beta \int V\left(\mu_{ot+1}, \sigma_{ot+1}, F : \mathbf{x}_{t+1}\right) d\Gamma_s\left(s_t | \mu_{ot}, \sigma_{ot}\right) \end{array} \right\} \quad (13) \\ & \quad s.t.\ equations\ (7),\ (9),\ (10) \\ \mu_{ot+1}^e &= f\left(\mathbf{x}_t\right),\ \sigma_{\mu t+1} = g\left(\mathbf{x}_t\right),\ \sigma_{ot+1}^a = m\left(\mathbf{x}_t\right),\ cov_{\mu F t+1} = n\left(\mathbf{x}_t\right) \end{aligned}$$

where $q_t = \left[\int_0^1 q_{jt}^\alpha dj\right]^{\frac{1}{\alpha}}$ and $\Gamma_s\left(s_t | \mu_{ot}, \sigma_{ot}\right)$ is a conditional distribution function of a signal s_t, given μ_{ot} and σ_{ot}, and functions $f(\cdot)$, $g(\cdot)$, $m(\cdot)$ and $n(\cdot)$ represent firms' expectations about the transition of the aggregate state variables.

We examine a positive assortative matching equilibrium between beliefs regarding organization capital, μ_{ot}, and a skill, $\ln q_{jt}$. We first describe a candidate equilibrium. Later, we prove that the candidate equilibrium exists. Let $\chi\left(\mu_{ot} : \mathbf{x}_t\right)$ denote a policy function of the problem (13). Following the argument in the previous section, the policy function must satisfy the following:

$$\chi\left(\mu_{ot}:\mathbf{x}_t\right)=\frac{\sigma_q}{\sigma_{\mu t}}\left(\mu_{ot}-\mu_{ot}^e\right)+\mu_q,\ \forall\mu_{ot}. \tag{14}$$

That is, a firm chooses the same quality of workers in an equilibrium. Hence, $q_t=\chi\left(\mu_{ot}:\mathbf{x}_t\right)$. Using equation (14), the dynamics of $\ln k_t^o$ and μ_{ot} are described by:

$$\begin{aligned}\ln k_{t+1}^o &= \ln B^*+\phi\ln k_t^o+\gamma\left[\frac{\sigma_q}{\sigma_{\mu t}}\left(\mu_{ot}-\mu_{ot}^e\right)+\mu_q\right]+F+\varepsilon_t\\ \mu_{ot+1} &= \ln B^*+\phi\left[\left(1-h_t\right)\mu_{ot}+h_t\ln k_t^o\right]+\gamma\left[\frac{\sigma_q}{\sigma_{\mu t}}\left(\mu_{ot}-\mu_{ot}^e\right)+\mu_q\right]+F-\frac{\sigma_\varepsilon^2}{2}+\phi h_t u_t^*\end{aligned}$$

Using the two equations, μ_{ot+1}^e, $\sigma_{\mu t+1}$, and $cov_{\mu Ft+1}$ can be derived:

$$\mu_{ot+1}^e = \ln B^*+\phi\mu_{ot}^e+\gamma\mu_q-\frac{\sigma_\varepsilon^2}{2}, \tag{15}$$

$$\sigma_{\mu t+1} = \sqrt{\lambda_{1t}^2\sigma_{\mu t}^2+\phi^2h_t\sigma_{ot}^2+2\lambda_{1t}cov_{\mu Ft}+\sigma_F^2}, \tag{16}$$

$$cov_{\mu Ft+1} = \lambda_{1t}cov_{\mu Ft}+\sigma_F^2, \tag{17}$$

where $\lambda_{1t}=\phi+\frac{\gamma\sigma_q}{\sigma_{\mu t}}$ is the persistence parameter when organization capital is unobservable.

We are ready to define a recursive positive assortative matching equilibrium between unobserved organization capital and skill.

Definition 1 *A recursive positive assortative matching equilibrium between unobserved organization capital and skill comprises values of* $\chi\left(\mu_{ot}:\mathbf{x}_t\right)$, $V\left(\mu_{ot},\sigma_{ot},F:\mathbf{x}_t\right)$, $w\left(\ln q_t:\mathbf{x}_t\right)$, $f\left(\mathbf{x}_t\right)$, $g\left(\mathbf{x}_t\right)$, $m\left(\mathbf{x}_t\right)$ *and* $n\left(\mathbf{x}_t\right)$ *that satisfy the following conditions.*

1. *An individual firm solves its maximization problem (13).*

2. *Equation (14) is satisfied to clear the labor market:*

3. *Expectations are rational:* $f\left(\mathbf{x}_t\right)=\ln B^*+\phi\mu_{ot}^e+\gamma\mu_q-\frac{\sigma_\varepsilon^2}{2}$, $g\left(\mathbf{x}_t\right)=\sqrt{\begin{array}{c}\lambda_{1t}^2\sigma_{\mu t}^2+\phi^2h_t\left(\sigma_{ot}^a\right)^2\\ +2\lambda_{1t}cov_{\mu Ft}+\sigma_F^2\end{array}}$, $m\left(\mathbf{x}_t\right)=\sqrt{\phi^2\left(1-h_t\right)\left(\sigma_{ot}^a\right)^2+\sigma_\varepsilon^2}$, $n\left(\mathbf{x}_t\right)=\lambda_{1t}cov_{\mu Ft}+\sigma_F^2$, *where* $\lambda_{1t}=\phi+\frac{\gamma\sigma_q}{\sigma_{\mu t}}$, $h_t=\frac{\left(\frac{\sigma_{ot}^a}{\sigma_u}\right)^2}{1+\left(\frac{\sigma_{ot}^a}{\sigma_u}\right)^2}$ *and* $\sigma_{ot}^a=\sigma_{ot}$.

Because $\sigma_{ot}^a = \sigma_{ot}$ in equilibrium, σ_{ot} is used subsequently to denote the aggregate state variable. Because the labor-market clearing condition and rational expectations assumption determine the functions $\chi(\mu_{ot} : \mathbf{x}_t)$, $f(\mathbf{x}_t)$, $g(\mathbf{x}_t)$, $m(\mathbf{x}_t)$ and $n(\mathbf{x}_t)$ by construction, we need to find a value function and a wage function that are consistent with the definition of equilibrium. The next theorem derives the value function and the wage function. The proof of the theorem is provided in the Appendix.

Theorem 2 *There exists a unique recursive positive assortative matching equilibrium between unobserved organization capital and skill. At the equilibrium, a wage function and a value function are solved by:*

$$w(\ln q_t : \mathbf{x}_t) = \frac{\frac{\psi\sigma_q}{\sigma_{\mu t}}(1-\delta) E\left[y_t | \chi^{-1}(\ln q_t : \mathbf{x}_t), \sigma_{ot}, F\right]}{1 + \frac{\psi\sigma_q}{\sigma_{\mu t}}} + \frac{\beta\frac{\gamma\sigma_q}{\sigma_{\mu t}}}{\lambda_{1t}} \int V\left(\mu_{ot+1}, \sigma_{ot+1} : \mathbf{x}_{t+1}, F\right) d\Gamma_s\left(s_t | \chi^{-1}(\ln q_t : \mathbf{x}_t), \sigma_{ot}\right) \quad (18)$$

$$V(\mu_{ot}, \sigma_{ot} : \mathbf{x}_t, F) = \sum_{i=0}^{\infty} \Pi_{s=1}^{i} \frac{\beta\phi}{\lambda_{1t+s-1}} \frac{(1-\delta) E\left[y_{t+i} | \mu_{ot}, \sigma_{ot}, F\right]}{1 + \frac{\psi\sigma_q}{\sigma_{\mu t+i}}}, \quad (19)$$

where:

$$E\left[y_{t+i} | \mu_{ot}, \sigma_{ot}, F\right] = \left(\frac{\delta}{r}\right)^{\frac{\delta}{1-\delta}} \exp \frac{1}{1-\delta} \left\{ \begin{array}{c} \psi\mu_q + \mu_{ot+i}^e + \frac{\sigma_{ot+i}^2}{2} \\ + \left(1 + \frac{\psi\sigma_q}{\sigma_{\mu t+i}}\right)\left[E\left[\mu_{ot+i} | \mu_{ot} : F\right] - \mu_{ot+i}^e\right] \\ + \frac{\left(1 + \frac{\psi\sigma_q}{\sigma_{\mu t+i}}\right)^2}{2} Var\left[\mu_{ot+i} | \sigma_{ot}\right] \end{array} \right\}$$

and $\Pi_{s=1}^{0} \frac{\beta\phi}{\lambda_{1t+s-1}} = 1$, $E\left[\mu_{ot+i} | \mu_{ot} : F\right] - \mu_{ot+i}^e = \Pi_{\tau=1}^{i} \lambda_{1t+i-\tau}\left(\mu_{ot} - \mu_{ot}^e\right) + \sum_{x=1}^{i} \Pi_{\tau=1}^{x-1} \lambda_{1t+i-\tau} F$ *and* $Var\left[\mu_{ot+i} | \sigma_{ot}\right] = \sum_{\tau=1}^{i} \Pi_{s=1}^{\tau-1} \lambda_{1t+i-s}^2 \phi^2 h_{t+i-\tau} \sigma_{ot+i-\tau}^2$.

Note that the $\ln q_t$ and $V(\mu_{ot}, \sigma_{ot} : \mathbf{x}_t, F)$ are strictly increasing functions of μ_{ot} and that $w(\ln q_t : \mathbf{x}_t)$ is a strictly increasing function of $\ln q_t$. Moreover, we know that capital stock, $\ln k_t$, is a strictly increasing function of μ_{ot}, which means that the firm believed to have high organization capital is large. Hence, the dynamics for skills, wages, expected profits, and physical capital stock follow the dynamics of μ_{ot}. Although the main focus in this paper is the dynamics of the TFP, it is worth emphasizing that the theoretical predictions for persistence and correlation of these variables are largely consistent with the evidence. For readers who are interested in the persistence of other variables, we refer to Takii (2008).

Because $\ln A_t$ depends on $\ln k_t^o$ and $\chi(\mu_{ot} : \mathbf{x}_t)$, we first discuss the dynamics of $\ln k_t^o$ and μ_{ot}. Using the equation (15), we can rewrite the dynamics of $\ln k_t^o$ and μ_{ot} as follows:

$$\ln k_{t+1}^o - \mu_{ot+1}^e = \phi [\ln k_t^o - \mu_{ot}^e] + \frac{\gamma \sigma_q}{\sigma_{\mu t}} (\mu_{ot} - \mu_{ot}^e) + F + \varepsilon_t^*, \tag{20}$$

$$\begin{aligned} \mu_{ot+1} - \mu_{ot+1}^e &= \phi h_t [\ln k_t^o - \mu_{ot}^e] \\ &\quad + \left[\phi (1 - h_t) + \frac{\gamma \sigma_q}{\sigma_{\mu t}}\right] (\mu_{ot} - \mu_{ot}^e) + F + \phi h_t u_t^*, \end{aligned} \tag{21}$$

where $\varepsilon_t^* = \varepsilon_t + \frac{\sigma_\varepsilon^2}{2}$ is normally distributed with the mean 0 and the variance σ_ε^2.

Equation (20) shows the dynamics of $\ln k_t^o$. The first term of equation (20) is influenced by technological persistence, ϕ. That is, if organization capital is above average, the fraction ϕ of this relative advantage is carried over to the next period. On the other hand, the second term is influenced by positive assignment. If organization capital is believed to be above average, the firm attracts skilled workers who help the firm to accumulate further organization capital. Note that when $\frac{\gamma \sigma_q}{\sigma_{\mu t}}$ is large, the effect of μ_{ot} on $\ln k_{t+1}^o$ is large. The firms with the largest μ_{ot} derive the highest benefits from a 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.

Note that equation (20) shows that the dynamics of $\ln k_t^o - \mu_{ot}^e$ exhibits a reversion to the established beliefs plus a firm fixed effect, $\frac{\gamma \sigma_q}{\sigma_{\mu t}} (\mu_{ot} - \mu_{ot}^e) + F$ 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_{ot}. More importantly, it indicates that if $\mu_{ot} - \mu_{ot}^e$ is persistent, the slow movement of $\mu_{ot} - \mu_{ot}^e$ plays a similar role to an unobserved fixed effect.

The dynamics of $\mu_{ot} - \mu_{ot}^e$ is depicted by equation (21). The first term captures how new information influences the dynamics of beliefs regarding organization capital.

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 the current TFP. The high TFP 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 beliefs. This effect is captured by the third term, $\phi h_t u_t^*$, in equation (21).

The second term of equation (21) 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 that 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 the TFP 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.

We can first prove the existence of a unique stationary distribution. The proof of the proposition is provided in the Appendix.

Proposition 3 *Suppose that* $\phi \in (0,1)$, $\frac{\sigma_u}{\sigma_\varepsilon} \in (0,\infty)$, $\frac{\sigma_q}{\sigma_\varepsilon} \in (0,\infty)$ *and* $\frac{\sigma_F}{\sigma_\varepsilon} \in (0,\infty)$. *There exists the steady state value of* $\mathbf{x}_t$*, which is denoted by* $\mathbf{x}_\infty$*. At the steady state, the dynamics of an individual firm are described by the following vector autoregression (VAR):*

$$\mathbf{O}_{t+1} = \mathbf{M}\mathbf{O}_t + \mathbf{F} + \boldsymbol{\xi}_t, \tag{22}$$

where $\mathbf{M} \equiv \begin{bmatrix} \phi, & \frac{\gamma\sigma_q}{\sigma_{\mu\infty}} \\ \phi h_\infty, & \phi(1-h_\infty)+\frac{\gamma\sigma_q}{\sigma_{\mu\infty}} \end{bmatrix}$, $\mathbf{O}_t \equiv \begin{bmatrix} D\ln k_t^o \\ D\mu_{ot} \end{bmatrix}$, $\mathbf{F} \equiv \begin{bmatrix} F \\ F \end{bmatrix}$, $\boldsymbol{\xi}_t \equiv$

$\begin{bmatrix} \varepsilon_t^* \\ \phi h_\infty u_t^* \end{bmatrix}$, *$D \ln k_t^o \equiv \ln k_t^o - \mu_{o\infty}^e$ and $D\mu_{ot} \equiv \mu_{ot} - \mu_{o\infty}^e$. Moreover, there exist functions $\eta(\cdot)$ and $\Sigma(\cdot,\cdot)$ such that:*

$$h_\infty = \eta\left(\frac{\sigma_u}{\sigma_\varepsilon}\right) \in (0,1), \tag{23}$$

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

$$\frac{\sigma_q}{\sigma_{\mu\infty}} = \Sigma\left(\frac{\sigma_q}{\sigma_\varepsilon}, \frac{\sigma_u}{\sigma_\varepsilon}, \frac{\sigma_F}{\sigma_\varepsilon}\right) \in \left(0, \frac{1-\phi}{\frac{\sigma_F}{\sigma_q} + \gamma}\right), \tag{24}$$

where $\Sigma_1\left(\frac{\sigma_q}{\sigma_\varepsilon}, \frac{\sigma_u}{\sigma_\varepsilon}, \frac{\sigma_F}{\sigma_\varepsilon}\right) > 0$, $\Sigma_2\left(\frac{\sigma_q}{\sigma_\varepsilon}, \frac{\sigma_u}{\sigma_\varepsilon}, \frac{\sigma_F}{\sigma_\varepsilon}\right) > 0$, $\Sigma_3\left(\frac{\sigma_q}{\sigma_\varepsilon}, \frac{\sigma_u}{\sigma_\varepsilon}, \frac{\sigma_F}{\sigma_\varepsilon}\right) < 0$, $\lim_{\frac{\sigma_q}{\sigma_\varepsilon}\to 0} \Sigma\left(\frac{\sigma_q}{\sigma_\varepsilon}, \frac{\sigma_u}{\sigma_\varepsilon}, \frac{\sigma_F}{\sigma_\varepsilon}\right) = 0$, and $\lim_{\frac{\sigma_u}{\sigma_\varepsilon},\to\infty} \Sigma\left(\frac{\sigma_q}{\sigma_\varepsilon}, \frac{\sigma_u}{\sigma_\varepsilon}, \frac{\sigma_F}{\sigma_\varepsilon}\right) = \frac{1-\phi}{\frac{\sigma_F}{\sigma_q}+\gamma}$.

Equation (23) shows that h_∞ is negatively related to $\frac{\sigma_u}{\sigma_\varepsilon}$. If the standard deviation of a noise term is relatively large, firms cannot learn much and h_∞ is small. Because equation (23) shows that h_∞ and $\frac{\sigma_u}{\sigma_\varepsilon}$ have a one-to-one relationship in the steady state, without loss of generality, h_∞ can be treated as an exogenous parameter.

As we discussed before, the assignment effect is captured by $\frac{\gamma\sigma_q}{\sigma_{\mu\infty}}$. Hence, equation (24) shows what influences the persistence through the assignment mechanism. First, it shows that the large $\frac{\sigma_q}{\sigma_\varepsilon}$ induces a large $\frac{\gamma\sigma_q}{\sigma_{\mu\infty}}$. This is the effect discussed in the previous section.

Second, $\frac{\gamma\sigma_q}{\sigma_{\mu\infty}}$ is increasing in $\frac{\sigma_u}{\sigma_\varepsilon}$. When information is noisy, rational agents ignore new information and rely more on their prior beliefs to make inferences about the current level of organization capital. Therefore, rational agents cannot change their beliefs very much. This makes the variance of μ_{ot} small and, therefore, makes $\frac{\gamma\sigma_q}{\sigma_{\mu\infty}}$ large. Hence, more noisy information is likely to increase the persistence.

Finally, when $\frac{\sigma_F}{\sigma_\varepsilon}$ is larger, the assignment effect is smaller. The large variation in a firm-specific fixed factor, σ_F is reflected by the large variation in beliefs regarding organization capital, $\sigma_{\mu\infty}$ in the long run. Therefore, it reduces the relative importance of assignment on the persistence, $\frac{\gamma\sigma_q}{\sigma_{\mu\infty}}$. This indicates that a rise in σ_F may not increase the persistence of organization capital.

With some additional conditions, we can also prove global stability. The proof of the proposition is provided in the Appendix.

Proposition 4 *Suppose that* $\phi \in (0,1)$, $\frac{\sigma_u}{\sigma_\varepsilon} \in (0,\infty)$, $\frac{\sigma_q}{\sigma_\varepsilon} \in (0,\infty)$ *and* $\frac{\sigma_F}{\sigma_\varepsilon} \in (0,\infty)$. *Suppose also that* $\frac{\sigma_q}{\sigma_\varepsilon}\frac{\sigma_F}{\sigma_\varepsilon} \le \frac{\phi^2(1-\phi)h_\infty}{2\gamma\left[1-\phi^2(1-h_\infty)\right]}$. *Then, the stationary distribution is globally stable.*

Owing to proposition 4, we presume that the productivity dynamics of the established firms can be approximated by the productivity dynamics under stationary distribution, which we assume in the next section.

4 Productivity Dynamics

In this section, we analytically examine how skill heterogeneity, noisy information and a firm fixed heterogeneity influence the dynamics of difference in the TFP. Let us define the relative productivity by $D\ln A_t = \ln A_t - E[\ln A]$. The dynamics of the relative productivity and its expectation in the steady state are derived from equation (22), as follows:

$$\mathbf{A}_{t+1} = \mathbf{M}^A \mathbf{A}_t + \mathbf{F}_A + \mathbf{v}_t \tag{25}$$

where $\mathbf{A}_t = \begin{bmatrix} D\ln A_t \\ E[D\ln A_t|\mu_{ot}] \end{bmatrix}$, $\mathbf{M}^A = \mathbf{M} + \frac{\psi\sigma_q}{\sigma_{\mu\infty}}\phi h_\infty \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix}$, $\mathbf{F}_A = \begin{bmatrix} F_A \\ F_A \end{bmatrix}$, $\mathbf{v}_t = \begin{bmatrix} v_t \\ 0 \end{bmatrix}$, $F_A = \left(1 + \frac{\psi\sigma_q}{\sigma_{\mu\infty}}\right) F$ and $v_t = \varepsilon_t^* - \phi u_t^* + u_{t+1}^*$ is normally distributed with the mean 0 and the variance $\left[\frac{1+\phi^2[1-h_\infty]^2}{h_\infty\left[1-\phi^2(1-h_\infty)\right]}\right]\sigma_\varepsilon^2 \left(= \left[\frac{h_\infty}{1-h_\infty} + 1 + \phi^2(1-h_\infty)\right]\sigma_u^2\right)$.

Comparing equation (25) with equation (22), we can point out an important difference: the parameter $\frac{\psi\sigma_q}{\sigma_{\mu\infty}}$ in equation (25) does not appear in equation (22). If ψ is 0, equation (25) is equivalent to equation (22) except for the error term. The parameter $\frac{\psi\sigma_q}{\sigma_{\mu\infty}}$ measures the static assignment effect that increases the sensitivity of productivity to beliefs: if a firm is perceived to have large organization capital, it can attract talented workers, which increases the productivity of the firm.

The sensitivity of productivity to beliefs influences the dynamics of productivity for two reasons. First, it magnifies the impacts of a firm-specific factor on the productivity, $F_A = \left(1 + \frac{\psi\sigma_q}{\sigma_{\mu\infty}}\right) F$. A firm with a large firm-specific factor is productive not only because it directly improves productivity, but also because it attracts excellent workers. In other words, a firm fixed effect in the TFP dynamics can already be magnified by the assignment. Second, it magnifies the impact of a prediction error, $\frac{\psi\sigma_q}{\sigma_{\mu\infty}}\phi h_\infty \left[D \ln A_t - E\left[D \ln A_t | \mu_{ot}\right]\right]$. When the realized TFP exceeds the expected TFP, people update their beliefs. The updated beliefs attract better workers and raise productivity in the next period.

Note that, similarly to the dynamics of organization capital, the dynamics of relative productivity exhibits a reversion to the expected relative productivity plus a firm-specific fixed effect, $D \ln A_{t+1} = b_1 D \ln A_t + \bar{F}_t + v_t$, where $b_1 = \phi + \frac{\psi\sigma_q}{\sigma_{\mu\infty}}\phi h_\infty$, $\bar{F}_t = \left(\frac{\gamma\sigma_q}{\sigma_{\mu\infty}} - \frac{\psi\sigma_q}{\sigma_{\mu\infty}}\phi h_\infty\right) E\left[D \ln A_t | \mu_{ot}\right] + F_A$. Hence, as long as $E\left[D \ln A_t | \mu_{ot}\right]$ moves slowly, it is difficult to distinguish the expectation from a firm-specific fixed factor. We examine the properties of productivity dynamics in detail below.

Extreme Cases : It is instructive to start with the extreme cases in which there is no assignment effect $\frac{\sigma_q}{\sigma_\varepsilon} = 0$ and the information is perfectly noisy $\frac{\sigma_u}{\sigma_\varepsilon} = \infty$. The following proposition can be easily proved.

Proposition 5 *1. Suppose that* $\frac{\sigma_q}{\sigma_\varepsilon} = 0$. *Then:*

$$\begin{aligned} D \ln A_{t+1} &= \phi D \ln A_t + F + v_t \qquad (26) \\ E\left[D \ln A_{t+1} | \mu_{ot+1}\right] &= \phi h_\infty D \ln A_t + \phi\left(1 - h_\infty\right) E\left[D \ln A_t | \mu_{ot}\right] + F \end{aligned}$$

2. Suppose that $\frac{\sigma_u}{\sigma_\varepsilon} = \infty$. *Then:*

$$\begin{aligned} D\ln A_{t+1} &= \phi D\ln A_t + (1-\phi) E\left[D\ln A_t | \mu_{ot}\right] + v_t, \qquad (27) \\ E\left[D\ln A_t|\mu_{ot}\right] &= \frac{\left(1+\frac{\psi}{\gamma}\frac{1-\phi}{\frac{\sigma_F}{\gamma\sigma_q}+1}\right)F}{1-\left(\phi+\frac{1-\phi}{\frac{\sigma_F}{\gamma\sigma_q}+1}\right)}, \; if\ \sigma_F > 0 \\ E\left[D\ln A_t|\mu_{ot}\right] &= E\left[D\ln A_0|\mu_{o0}\right], \; if\ \sigma_F = 0 \end{aligned}$$

Equation (26) shows that if there is no skill variation and, therefore, there is no assignment problem, there is no link between expectation and productivity. Even if people believe the organization capital in a firm is high, when there is no assignment mechanism, there is no way that beliefs can influence productivity. That is, in order for the expectation to influence real productivity, the assignment between beliefs and skill is necessary.

Equation (27) shows that if A_t does not contain any information for predicting the level of organization capital, the expected relative productivity converges to constant values, and the dynamics of relative productivity exhibits a reversion to the fraction of the values. Hence, the expected relative productivity itself plays the role of an unobserved firm-specific fixed effect.

When there is variation in the firm-specific fixed factors ($\sigma_F > 0$), because the factors are observable, the expectation is influenced by these firm-specific factors. Therefore, the observed fixed variation coincides with the amplified firm heterogeneity.

More interestingly, even if there is no variation in a firm-specific fixed factor ($\sigma_F = 0$), beliefs can differ. Because there is no information to update beliefs, beliefs never change and initial prior beliefs influence beliefs even in the long run. In this way, the model can endogenize the fixed effect in the productivity dynamics without any real firm-specific heterogeneity.

General Case: Let us examine a more general case. Suppose that $\phi \in (0,1)$,

$\frac{\sigma_u}{\sigma_\varepsilon} \in (0,\infty)$, $\frac{\sigma_q}{\sigma_\varepsilon} \in (0,\infty)$ and $\frac{\sigma_F}{\sigma_\varepsilon} \in (0,\infty)$. Let us define $\mathbf{A}_t^F = \begin{bmatrix} D\ln A_t^F \\ E\left[D\ln A_t^F | \mu_{ot}\right] \end{bmatrix}$ where $D\ln A_t^F = D\ln A_t - F_A^*$ and $F_A^* = \frac{F_A}{1-\lambda_{1\infty}}$. Note that $E\left[D\ln A_t\right] = F_A^*$. That is, F_A^* measures the long-run average of the relative productivity and $D\ln A_t^F$ measures the deviation of the relative productivity from its long-run average. Then, the following equation can be derived.

$$\mathbf{A}_{t+1}^F = \mathbf{M}^A \mathbf{A}_t^F + \mathbf{v}_t, \tag{28}$$

Let $\lambda_{1\infty}$ and $\lambda_{2\infty}$ denote the eigenvalues of the matrix $\mathbf{M}^A$. Then, it is shown that $\lambda_{1\infty} = \phi + \frac{\gamma\sigma_q}{\sigma_{\mu\infty}} < 1$ and $\lambda_{2\infty} = \phi\left(1 - h_\infty\right) < 1$. Hence, equation (28) is covariance stationary. Note that $\lambda_{1\infty}$ is equivalent to the persistence parameter.

In order to analyze the persistence and diversity of relative productivity, we measure the correlation between the current relative productivity and the jth lagged relative productivity, $\rho_{D\ln Aj}$, and the variance of the relative productivity, $Var\left[D\ln A_t\right]$. It can be shown that $\rho_{D\ln Aj}$, can be decomposed into the between to overall variance ratio, $\frac{Var\left[F_A^*\right]}{Var\left[D\ln A_t\right]}$, which measures the permanent persistence, and the within-correlation, $\rho_{D\ln A^F j}$, which measures the temporal persistence, and that $Var\left[D\ln A_t\right]$ can be decomposed into within-variance, $Var\left[D\ln A_t^F\right]$, and between-variance, $Var\left[F_A^*\right]$, by the following equations.

$$\begin{aligned} \rho_{D\ln Aj} &= \left[1 - \frac{Var\left[F_A^*\right]}{Var\left[D\ln A_t\right]}\right]\rho_{D\ln A^F j} + \frac{Var\left[F_A^*\right]}{Var\left[D\ln A_t\right]}, \qquad (29) \\ Var\left[D\ln A_t\right] &= Var\left[D\ln A_t^F\right] + Var\left[F_A^*\right] \qquad (30) \end{aligned}$$

The following proposition is used to derive our theoretical predictions on the persistence and the diversity of the relative productivity, $D\ln A_t$. The proof of the proposition is provided in the Appendix.

Proposition 6

$$
\begin{aligned}
Var\left[F_A^*\right] &= \frac{\left(1+\frac{\psi\sigma_q}{\sigma_{\mu\infty}}\right)^2}{\left(1-\lambda_{1\infty}\right)^2}\sigma_F^2, \\
Var\left[D\ln A_t^F\right] &= \left[\frac{\left(1+\frac{\psi\sigma_q}{\sigma_{\mu\infty}}\right)^2\left(\phi h_\infty\right)^2}{1-\lambda_{1\infty}^2}+1\right]\frac{\sigma_\varepsilon^2}{h_\infty\left[1-\phi^2\left(1-h_\infty\right)\right]},
\end{aligned}
$$

and

$$
\begin{aligned}
\rho_{D\ln A^F j} &= z\rho_{E[D\ln A^F|\mu]j-1}, \\
\rho_{E[D\ln A^F|\mu]j-1} &= \lambda_{1\infty}^{j-1},
\end{aligned}
$$

where $z=\lambda_{1\infty}\frac{Var\left[E\left[D\ln A_t^F|\mu_{ot}\right]\right]}{Var\left[D\ln A_t^F\right]}+\left(1+\frac{\psi\sigma_q}{\sigma_{\mu\infty}}\right)\phi h_\infty\left[1-\frac{Var\left[E\left[D\ln A_t^F|\mu_{ot}\right]\right]}{Var\left[D\ln A_t^F\right]}\right]$ *and* $\frac{Var\left[E\left[D\ln A_t^F|\mu_{ot}\right]\right]}{Var\left[D\ln A_t^F\right]}=\frac{\frac{\left(1+\frac{\psi\sigma_q}{\sigma_{\mu\infty}}\right)^2(\phi h_\infty)^2}{1-\lambda_{1\infty}^2}}{\frac{\left(1+\frac{\psi\sigma_q}{\sigma_{\mu\infty}}\right)^2(\phi h_\infty)^2}{1-\lambda_{1\infty}^2}+1}$.

Proposition 6 shows that the measure for the persistence of the expected deviation, $\rho_{E[D\ln A^F|\mu]j-1}$, is entirely determined by the persistence parameter, $\lambda_{1\infty}$. It means that $\lambda_{1\infty}$ also measures the persistence of beliefs.

In addition, it shows that the between-variance $Var\left[F_A^*\right]$ and the within-variance $Var\left[D\ln A_t^F\right]$ are influenced by $\lambda_{1\infty}=\phi+\frac{\gamma\sigma_q}{\sigma_{\mu\infty}}$, which measures the persistence of beliefs, and the sensitivity of productivity to beliefs, $\frac{\psi\sigma_q}{\sigma_{\mu\infty}}$. It suggests that the variance of relative productivity can be largely influenced by the assignment mechanism.

Finally, proposition 6 shows that the within-correlation is determined by the persistence measure of the expected deviation $\rho_{E[D\ln A^F|\mu]j-1}$ times a factor z. The influence of $\rho_{E[D\ln A^F|\mu]j-1}$ represents the importance of the persistence of beliefs for the within-correlation. The factor z is constructed by the weighted average of the persistence parameter, $\lambda_{1\infty}$ and the impacts of new information from $\ln A_{t-1}$, $\left(1+\frac{\psi\sigma_q}{\sigma_{\mu\infty}}\right)\phi h_\infty$, which changes the beliefs on organization capital and, therefore, the assigned skilled workers in the next period. The reliability of the expectation,

$\frac{Var\left[E\left[D\ln A_t^F|\mu_{ot}\right]\right]}{Var\left[D\ln A_t^F\right]}$, plays the role of a weight. This shows that, if the expectation is perfectly reliable, $\frac{Var\left[E\left[D\ln A_t^F|\mu_{ot}\right]\right]}{Var\left[D\ln A_t^F\right]} = 1$, the within-correlation is entirely explained by the persistence of beliefs, $\rho_{D\ln A^F j} = \rho_{E[D\ln A^F|\mu]j} = \lambda_{1\infty}^j$. As far as $\frac{Var\left[E\left[D\ln A_t^F|\mu_{ot}\right]\right]}{Var\left[D\ln A_t^F\right]} < 1$, new information has some values.

Knowing that $\lambda_{1\infty} = \phi + \gamma\Sigma\left(\frac{\sigma_q}{\sigma_\varepsilon}, \frac{\sigma_u}{\sigma_\varepsilon}, \frac{\sigma_F}{\sigma_\varepsilon}\right)$, $\frac{\psi\sigma_q}{\sigma_{\mu\infty}} = \psi\Sigma\left(\frac{\sigma_q}{\sigma_\varepsilon}, \frac{\sigma_u}{\sigma_\varepsilon}, \frac{\sigma_F}{\sigma_\varepsilon}\right)$ and $h_\infty = \eta\left(\frac{\sigma_u}{\sigma_\varepsilon}\right)$, we can first derive the analytical predictions on the diversity of the relative productivity. The following proposition can be easily proved.

Proposition 7 *For any j*

$$\frac{dVar\left[F_A^*\right]}{d\frac{\sigma_q}{\sigma_\varepsilon}} > 0,\ \frac{dVar\left[F_A^*\right]}{d\frac{\sigma_u}{\sigma_\varepsilon}} > 0,\ \frac{dVar\left[F_A^*\right]}{d\gamma} > 0,\ \frac{dVar\left[F_A^*\right]}{d\psi} > 0,$$
$$\frac{dVar\left[D\ln A_t^F\right]}{d\frac{\sigma_q}{\sigma_\varepsilon}} > 0,\ \frac{dVar\left[D\ln A_t^F\right]}{d\frac{\sigma_F}{\sigma_\varepsilon}} < 0,\ \frac{dVar\left[D\ln A_t^F\right]}{d\gamma} > 0,\ \frac{dVar\left[D\ln A_t^F\right]}{d\psi} > 0,$$
$$\frac{dVar\left[D\ln A_t\right]}{d\frac{\sigma_q}{\sigma_\varepsilon}} > 0,\ \frac{dVar\left[D\ln A_t\right]}{d\gamma} > 0,\frac{dVar\left[D\ln A_t\right]}{d\psi} > 0.$$

Proposition 7 shows that, among $\frac{\sigma_q}{\sigma_\varepsilon}$, $\frac{\sigma_u}{\sigma_\varepsilon}$ and $\frac{\sigma_F}{\sigma_\varepsilon}$, only an increase in $\frac{\sigma_q}{\sigma_\varepsilon}$, has unambiguous predicted effects on $Var\left[D\ln A_t\right]$. Because the assignment magnifies a small differences in a firm fixed effect, a rise in $\frac{\sigma_q}{\sigma_\varepsilon}$ increases the between-variance and the within-variance, and, therefore, the overall variance.

In order to understand its mechanism, proposition 7 also provides the comparative statics with respect to γ and ψ, where the parameters γ and ψ measure the importance of skill for the accumulation of organization capital and for the current production, respectively. The parameter γ influences the persistence of beliefs, $\lambda_{1\infty} = \phi + \frac{\gamma\sigma_q}{\sigma_{\mu\infty}}$, whereas ψ influences the sensitivity of productivity to expectation, $\frac{\psi\sigma_q}{\sigma_{\mu\infty}}$. Therefore, the comparative statics with respect to γ and ψ are useful to understand the two separate effects that the assignment mechanism has. Proposition 7 shows that, regardless of the role of skill, the more important that skill is in the firm, the more important is assignment, and, therefore, the more diverse is productivity.

Although an increase in $\frac{\sigma_u}{\sigma_\varepsilon}$ raises the between-variance, $Var\left[F_A^*\right]$, its effects on the within-variance are ambiguous. Similarly, an increase in $\frac{\sigma_F}{\sigma_\varepsilon}$ reduces the within-variance, $Var\left[D\ln A_t^F\right]$, but its impacts on the between-variance are ambiguous. One may be surprised by the fact that an increase in firm-specific heterogeneity may reduce between-variance, $Var\left[F_A^*\right]$. This is because an increase in $\frac{\sigma_F}{\sigma_\varepsilon}$ reduces $\frac{\sigma_q}{\sigma_{\mu\infty}}$, and therefore, the dynamic assignment effect, $\frac{\gamma\sigma_q}{\sigma_{\mu\infty}}$ and the static assignment effect, $\frac{\psi\sigma_q}{\sigma_{\mu\infty}}$. Therefore, the amplification through the assignment is smaller, which may offset the direct positive effect on $Var\left[F_A^*\right]$.

Next, we derive the analytical predictions on persistence of the relative productivity. The following proposition can be easily proved.

Proposition 8 *For any j*

$$\frac{d\rho_{E[D\ln A^F|\mu]j-1}}{d\frac{\sigma_q}{\sigma_\varepsilon}} > 0,\ \frac{d\rho_{E[D\ln A^F|\mu]j-1}}{d\frac{\sigma_u}{\sigma_\varepsilon},} > 0,\ \frac{d\rho_{E[D\ln A^F|\mu]j-1}}{d\frac{\sigma_F}{\sigma_\varepsilon}} < 0,\ \frac{d\rho_{E[D\ln A^F|\mu]j-1}}{d\gamma} > 0,$$
$$\frac{d\frac{Var\left[F_A^*\right]}{Var[D\ln A_t]}}{d\frac{\sigma_q}{\sigma_\varepsilon}} > 0\ \frac{d\frac{Var\left[F_A^*\right]}{Var[D\ln A_t]}}{d\gamma} > 0.,\ \frac{d\frac{Var\left[F_A^*\right]}{Var[D\ln A_t]}}{d\psi} > 0.$$

Hence, there exists j^* such that for all $j > j^*$

$$\frac{d\rho_{D\ln A^F j}}{d\frac{\sigma_q}{\sigma_\varepsilon}} > 0,\ \frac{d\rho_{D\ln A^F j}}{d\frac{\sigma_u}{\sigma_\varepsilon}} > 0,\ \frac{d\rho_{D\ln A^F j}}{d\frac{\sigma_F}{\sigma_\varepsilon}} < 0, \frac{d\rho_{D\ln A^F j}}{d\gamma} > 0,$$

and

$$\frac{d\rho_{D\ln Aj}}{d\frac{\sigma_q}{\sigma_\varepsilon}} > 0, \frac{d\rho_{D\ln Aj}}{d\gamma} > 0, \frac{d\rho_{D\ln Aj}}{d\psi} > 0.$$

Similarly to proposition 7, the proposition 8 shows that only an increase in $\frac{\sigma_q}{\sigma_\varepsilon}$ unambiguously increases the persistence of productivity after enough time has passed. An increase in $\frac{\sigma_q}{\sigma_\varepsilon}$ not only increases $\rho_{E[D\ln A^F|\mu]j-1}$, but also increases the between-to-overall variance ratio, $\frac{Var\left[F_A^*\right]}{Var[D\ln A_t]}$. Because actual productivity is subjected to temporal shocks, it can deviate from beliefs temporally. However, as time passes, an increase in the persistence of beliefs dominates the temporal disturbance and increases the persistence of productivity itself. This is what Proposition 8 says. Note also that both γ and ψ can increase the persistence of relative productivity. This indicates that, although the assignment raises both the persistence of beliefs and the sensitivity of productivity to the beliefs, both influence the persistence of productivity in the same direction.

An increase in $\frac{\sigma_F}{\sigma_\varepsilon}$ and a reduction in $\frac{\sigma_u}{\sigma_\varepsilon}$ unambiguously reduce within-correlation, $\rho_{D \ln A^F j}$, because they lower the persistence parameter $\lambda_{1\infty}$. However, the effects on $\rho_{D \ln Aj}$ are ambiguous because the effects of $\frac{\sigma_F}{\sigma_\varepsilon}$ and $\frac{\sigma_u}{\sigma_\varepsilon}$ on the between to overall variance ratio, $\frac{Var\left[F_A^*\right]}{Var[D \ln A_t]}$, are ambiguous.

5 Quantitative Analysis

In this section, we use the theory to quantify the importance of three factors, $\frac{\sigma_q}{\sigma_\varepsilon}$, $\frac{\sigma_F}{\sigma_\varepsilon}$, and $\frac{\sigma_u}{\sigma_\varepsilon}$ as the sources of persistence and diversity of productivity. Owing to proposition 6, the measure of persistence and diversity can be derived as a function of technological persistence, ϕ, a measure of the accuracy of information, h_∞, the heterogeneity of a firm fixed effect, σ_F, technological disturbance σ_ε, the measure of a dynamic assignment effect, $\frac{\gamma\sigma_q}{\sigma_{\mu\infty}}$, which influences the persistence of beliefs, and the measure of a static assignment effect, $\frac{\psi\sigma_q}{\sigma_{\mu\infty}}$, which influences the sensitivity of productivity to the beliefs. We identify these parameters to be consistent with the dynamics and variations of the TFP of a firm and the ratio of the average wage to the average labor productivity.

We constructed a firm level TFP, A_{ft}, using BSJBSA from 1994 to 2004. BSJBSA is based on a survey conducted by the Ministry of Economy, Trade and Industry every year. The BSJBSA covers all the enterprises with 50 employees or more and more than a 30 million yen capitalization that are at least partly engaged in mining, manufacturing, wholesale and retail sales, and restaurant activities in Japan. Hence, the data set contains all relatively established firms in these industries, which are more likely to satisify one of our steady state conditions that $\sigma_{o\infty}^2$ is constant for all firms. Although we can obtain data for the publicly traded firms, many established firms are not traded. Hence, our data set would be more appropriate for the analyisis of the assignment model.

A particular feature of BSJBSA is that it is a survey of firms, not establishments, which are commonly used in the study of productivity dynamics. Because more publicly accessible data is available at the firm level, it would be easier for people to embrace their beliefs regarding the capability of a firm than their beliefs regarding an establishment. In this regard, this data is more desirable for the analysis of our model than is the establishment data available in other countries. More discussions about data can be found in Matsuura and Kiyota (2004), Fukao and Kwon (2006), and Nishimura, Nakajima, and Kiyota (2005).

In order to obtain a reasonable correlation measure, in our study, we retain only the firms for which there are data available in all years. Because we are interested in the assignment within an industry, we use industries that have more than five firms for all years, where industries are classified by the three-digit industry code in the BSJBSA. In addition, we drop the observations with nonpositive values for value added, number of employed workers or capital stock. After these deletions, our balanced panel data set contains 68,838 observations.

We estimate $D\ln A_{ft}$ by $D\ln A_{ft} = \left[\ln y_{ft} - \frac{\sum_j^{m_t} \ln y_{ft}}{m_t}\right] - \delta\left[\ln k_{ft} - \frac{\sum_j^{m_t} \ln k_{ft}}{m_t}\right]$, where y_{ft} and k_{ft} are the value added per worker and the physical capital per worker, respectively, in the fth firm in year t, and m_t is the number of firms in the three-digit industry in year t. The parameter δ is calibrated by the average capital share $\frac{E\left[r_t k_{ft}\right]}{E\left[y_{ft}\right]}$, which is 0.22[4]. Detailed data construction can be found in the Appendix.

In order to calibrate six parameters ϕ, $\frac{\gamma\sigma_q}{\sigma_{\mu\infty}}$, h_∞, σ_F^2, σ_ε^2 and $\frac{\psi\sigma_q}{\sigma_{\mu\infty}}$ from the data, we use several sources of information. The first source of information comes from productivity dynamics:

$$D\ln A_t = b_1 D\ln A_{t-1} + b_2 E\left[D\ln A_{t-1}|\mu_{ot-1}\right] + F_A + v_{t-1}, \qquad (31)$$

$$E\left[D\ln A_t|\mu_{ot}\right] = b_3 D\ln A_{t-1} + b_4 E\left[D\ln A_{t-1}|\mu_{ot-1}\right] + F_A, \qquad (32)$$

where $b_1 = \phi + \frac{\psi\sigma_q}{\sigma_{\mu\infty}}\phi h_\infty, b_2 = \frac{\gamma\sigma_q}{\sigma_{\mu\infty}} - \frac{\psi\sigma_q}{\sigma_{\mu\infty}}\phi h_\infty, b_3 = \phi h_\infty + \frac{\psi\sigma_q}{\sigma_{\mu\infty}}\phi h_\infty$ and $b_4 = b_2 +$

[4] This is between 0.20, calibrated by Atkeson and Kehoe (2005), and 0.25, calibrated by Samaniego (2006).

$(b_1 - b_3)$. Given $\frac{\psi\sigma_q}{\sigma_{\mu\infty}}$, we can choose ϕ, $\frac{\gamma\sigma_q}{\sigma_{\mu\infty}}$ and h_∞, which are matched up to b_1, b_2 and b_3. The second source of information comes from the between-variance, $Var\left(F_A^*\right)$, and the within-variance, $Var\left[D\ln A_t^F\right]$. We estimate these variances, and use them to calibrate σ_F^2 and σ_ε^2 using proposition 6. Finally, in order to calibrate the static assignment effect, $\frac{\psi\sigma_q}{\sigma_{\mu\infty}}$, we take the expectation on both sides of equation (18) in the steady state. Using equation (19), we can obtain the following relationship:

$$\frac{\psi\sigma_q}{\sigma_{\mu\infty}} = \frac{\frac{E[w_t]}{(1-\delta)E[y_t]} - \frac{\beta\left(1-\frac{\phi}{\lambda_{1\infty}}\right)}{1+\frac{\beta\phi}{\lambda_{1\infty}}}}{1 - \frac{E[w_t]}{(1-\delta)E[y_t]}}.$$

where $\lambda_{1\infty} = \phi + \frac{\gamma\sigma_q}{\sigma_{\mu\infty}}$, the ratio of the average wage to the average labor productivity, $\frac{E[w_t]}{E[y_t]}$ is estimated to be 0.65 and we assume that $\beta = 0.95$. We can derive a unique closed-form solution of $\left(\phi, \frac{\gamma\sigma_q}{\sigma_{\mu\infty}}, h_\infty, \sigma_F^2, \sigma_\varepsilon^2, \frac{\psi\sigma_q}{\sigma_{\mu\infty}}\right)'$ from $\left(b_1, b_2, b_3, Var\left(F_A^*\right), Var\left[D\ln A_t^F\right], \frac{E[w_t]}{(1-\delta)E[y_t]}\right)'$ and find a one-to-one relationship between these estimates and the calibrated parameters.

In order to obtain information on b_1, b_2, b_3 and F_A^* from the data, we extract information from the productivity dynamics equations, (31) and (32). First, we estimate b_3 and b_4, and construct $E\left[D\ln A_t|\mu_{ot}\right]$ by the following procedure. Note that the following regression equation can be derived from equation (32):

$$D\ln A_t = b_3 D\ln A_{t-1} + b_4 E\left[D\ln A_{t-1}|\mu_{ot-1}\right] + F_A + \varpi_t, \tag{33}$$

where $\varpi_t = D\ln A_t - E\left[D\ln A_t|\mu_{ot}\right]$. Assume that $E\left[D\ln A_{t-1}|\mu_{ot-1}\right] = D\ln A_{ft-2}$. First, we estimate equation (33) using Allenano–Bover/Blundell–Bond Estimation, although we exclude $D\ln A_{ft-j}$ where $j \geq 3$ from its instruments because, once $E\left[D\ln A_{t-j}|\mu_{ot-j}\right]$ $j \geq 2$ are controlled, $D\ln A_{ft-j}$ $j > 3$ are not correlated with $D\ln A_{ft-1}$. It gives us $\hat{b}_3$ and $\hat{b}_4$, which are the estimated values of b_3 and b_4. Second, then we construct $\hat{F}_A$ by $\hat{F}_A = \frac{\sum_t^T D\ln A_t - \left(\hat{b}_3 D\ln A_{t-1} + \hat{b}_4 E\left[D\ln A_{t-1}|\mu_{ot-1}\right]\right)}{T}$. Third, we construct $E\left[D\ln A_t|\mu_{ot}\right]$ using equation (32) if $E\left[D\ln A_{t-1}|\mu_{ot-1}\right]$ exists and set $\left(\hat{b}_3 + \hat{b}_4\right) D\ln A_{t-1} + \hat{F}_A$ if $E\left[D\ln A_{t-1}|\mu_{ot-1}\right]$ is missing. Fifth, we repeat the same procedure until the estimated $E\left[D\ln A_{t-1}|\mu_{ot-1}\right]$ converges to the assumed

	$D\ln A_{t-1}$	$E\left[D\ln A_{t-1}\vert\mu_{ot-1}\right]$	# of Obs.
$D\ln A_t$	0.471* (0.013)	0.242* (0.026)	56322
	$D\ln A_{t-1}-E\left[D\ln A_{t-1}\vert\mu_{ot-1}\right]$		# of Obs.
$D\ln A_t - E\left[D\ln A_t\vert\mu_{ot}\right]$	0.165* (0.040)		56322

Table 1: Regression Results

WC-Robust standard errors are reported in parentheses. *denotes significance at the one percent level.

$E\left[D\ln A_{t-1}|\mu_{ot-1}\right]$. This process gives us the estimates of $\hat{b}_3$, $\hat{b}_4$ and $\hat{F}_A$. Knowing that $\lambda_{1\infty}=b_3+b_4$, we estimate the long-run average relative productivity F_A^* by $\hat{F}_A^*=\frac{\hat{F}_A}{1-\left(\hat{b}_3+\hat{b}_4\right)}$.

Given the estimated value of $E\left[D\ln A_t|\mu_{ot}\right]$, we estimate b_1 and b_2. Note that the difference between equation (31) and (33) is just their error terms, v_{t-1} and ϖ_t. Although $v_{t-1}=\varepsilon_{t-1}^*-\phi u_{t-1}^*+u_t^*$ is correlated with $D\ln A_{t-1}$, ϖ_t is not. Because a large $\ln A_{t-1}$ not only indicates a large $\ln k_{t-1}^o$ but also indicates a large temporal luck u_{t-1}^*, v_{t-1} is correlated with $D\ln A_{t-1}$. When rational agents predict a future $\ln A_t$, they efficiently utilize this information too. That is why the prediction error, ϖ_t, is orthogonal to $D\ln A_{t-1}$. In other words, b_3 is a biased estimator of b_1. This difference is more accurate when we subtract (32) from (31):

$$D\ln A_t-E\left[D\ln A_t|\mu_{ot}\right]=b_5\left[D\ln A_{t-1}-E\left[D\ln A_{t-1}|\mu_{ot-1}\right]\right]+v_{t-1}, \qquad (34)$$

where $b_5=b_1-b_3=\phi\left(1-h_\infty\right)$. That is, a large difference between b_1 and b_3 indicates a small h_∞. Using this equation, we can identify $\hat{b}_1=\hat{b}_3+\hat{b}_5$. From the theoretical restriction on parameters, we can identify $\hat{b}_2=\hat{b}_4-\hat{b}_5$.

We estimate equation (34) by Allenano–Bover/Blundell–Bond Estimation provided that the error term follows a first-order moving average process because $v_{t-1}=\varepsilon_{t-1}^*-\phi u_{t-1}^*+u_t^*$. As instruments, we use $D\ln A_{t-3}$ for the difference equation and $D\ln A_{t-2}-D\ln A_{t-3}$ for the level equation rather than the standard lagged variables because the standard instruments can be weak. The reasons relate to

ϕ	$\frac{\gamma\sigma_q}{\sigma_{\mu\infty}}$	h_∞	σ_F^2	σ_ε^2	$\frac{\psi\sigma_q}{\sigma_{\mu\infty}}$
0.298	0.415	0.447	0.001	0.019	2.535

Table 2: The Calibrated Parameters

the specific structure of the equation (34). If information is nearly perfect, then $b_1 \approx b_3$ and, therefore, the lagged variables are less likely to be correlated with $D \ln A_{t-1} - E\left[D \ln A_{t-1}|\mu_{ot-1}\right]$.

In fact, the theoretical predictions on the correlations of relative productivity are independent of the choice of $\hat{b}_5$. In order to understand the reason for this, consider proposition 6. It shows that the correlation is decomposed into $\lambda_{1\infty}$, z and $\frac{Var\left[F_A^*\right]}{Var[D \ln A_t]}$, and $\lambda_{1\infty}$ is identified by $\hat{b}_3 + \hat{b}_4$; z is determined by $\left(1 + \frac{\psi\sigma_q}{\sigma_{\mu\infty}}\right)\phi h_\infty = \hat{b}_3$ and $\lambda_{1\infty}$; and $\frac{Var\left[F_A^*\right]}{Var[D \ln A_t]}$ is taken from the data. Hence, it is determined without knowing $\hat{b}_5$. However, the choice of $\hat{b}_5$ may change the results of our counterfactual exercises. It turns out that the most of our quantitative results are not sensitive to the choice of $\hat{b}_5$. We discuss this issue in detail later.

Table 1 reports the results from the regression equation (33) when the estimated $E\left[D \ln A_t|\mu_{ot}\right]$ converges to the assumed $E\left[D \ln A_t|\mu_{ot}\right]$ and the results from the regression equation (34). Table 1 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 $E\left[D \ln A_{t-1}|\mu_{ot-1}\right]$ 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 beliefs that influence future performance. Table 1 also indicates that we cannot perfectly infer the level of organization capital from the TFP.

Using our estimate, we can calibrate our structural parameters. Table 2 reports our calibration results. It shows fairly low values of σ_F^2 and σ_ε^2, 0.001 and 0.019. Note also that the implied values of the variation of a firm-specific factor are smaller than that of transitory shocks $\sigma_F^2 < \sigma_\varepsilon^2$. As shown in Table 3, our estimates of $Var\left(F_A^*\right)$ and $Var\left[D \ln A_t^F\right]$ are 0.171 and 0.064, respectively. It indicates that the relatively

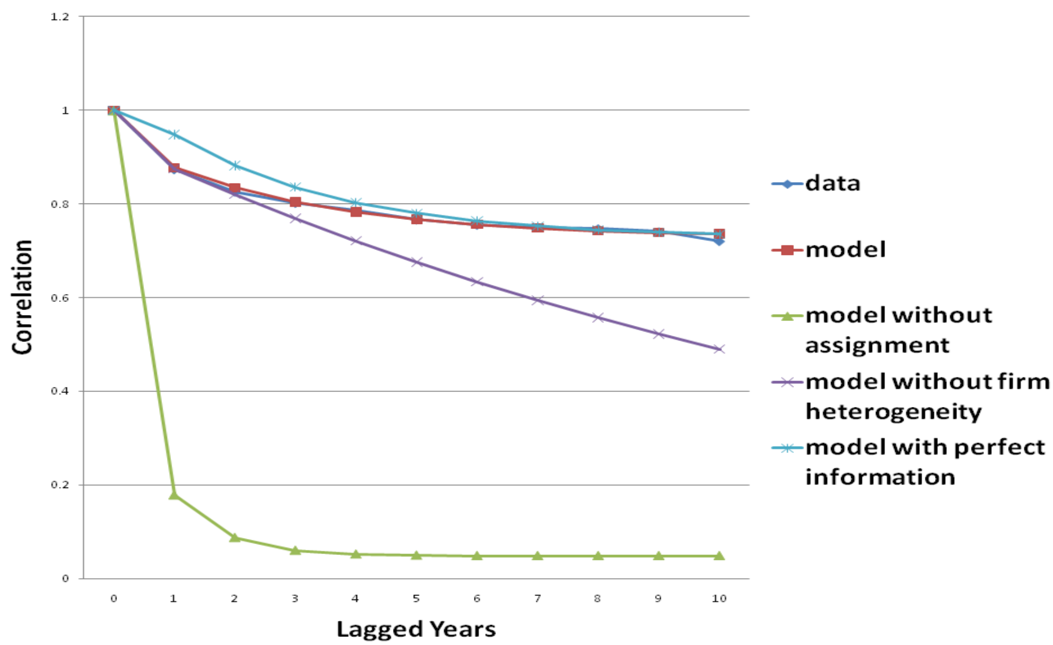


Figure 1: The Persistence of Relative Productivity Between Firms

large between-variance of $\ln A_t$ can mostly be explained by the amplification though assignment. We use these calibrated values to examine the predictions of our model.

Persistence and Disparity of Productivity: We compare the predicted correlations with the correlations observed in the data. Our measures of correlation and variance are constructed by using equations (29) and (30), where $Var\left[F_A^*\right]$, $Var\left[D\ln A_t^F\right]$ and $\rho_{D\ln A_t^F j}$ are constructed by the sample variance of $\hat{F}_A^*$ and $D\ln A_t - \hat{F}_A^*$ and the simple correlation of $D\ln A_t - \hat{F}_A^*$ and $D\ln A_{t+j} - \hat{F}_A^*$, respectively. Our theoretical counterparts are constructed by using proposition 6.

Figure 1 summarizes the results of our simulations. Figure 1 shows that the predicted correlation sequence almost perfectly explains the overall movement of actual correlation. We use this model to conduct several counterfactual experiments.

We first ask "What would happen, if there were no assignments in the economy

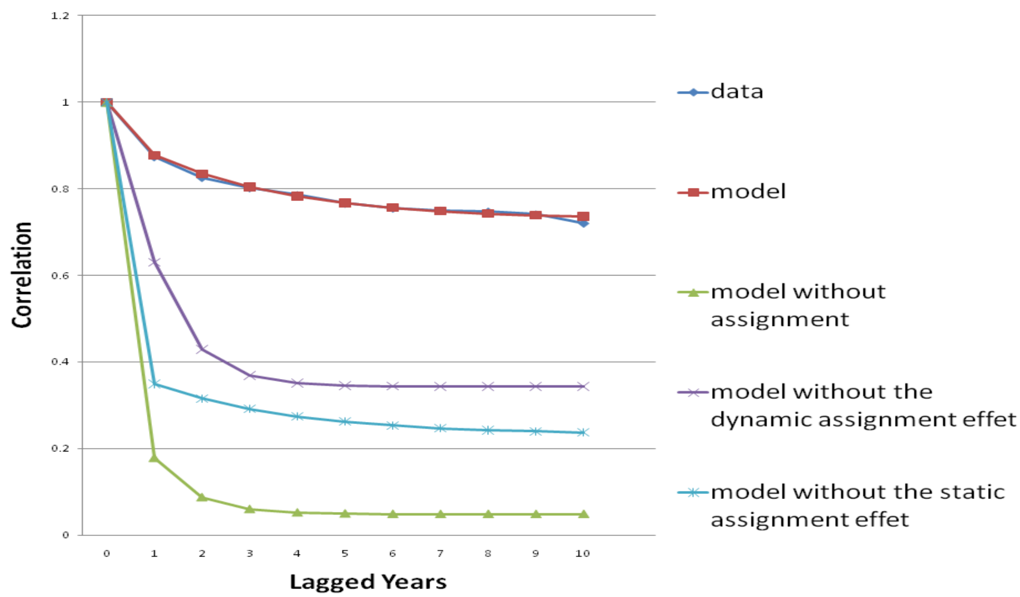


Figure 2: The Decomposition of the Role of Assignment on the Persistence

($\frac{\sigma_q}{\sigma_\varepsilon} = 0$)?" This experiment can be done by assuming that $\frac{\gamma\sigma_q}{\sigma_{\mu\infty}} = \frac{\psi\sigma_q}{\sigma_{\mu\infty}} = 0$ and the other parameters are constant. Figure 1 shows that, if $\frac{\sigma_q}{\sigma_\varepsilon} = 0$, the correlation of relative productivity diminishes to less than 10 percent after 2 years. That is, the relative temporal advantages disappear quickly if there are no benefits from positive assignment.

We also conduct an experiment involving no variation in the firm-specific factor $\frac{\sigma_F}{\sigma_\varepsilon} = 0$. When $\frac{\sigma_F}{\sigma_\varepsilon} = 0$, $\frac{\sigma_q}{\sigma_{\mu\infty}}$ becomes larger. This indirect effect is also computed using the equations in the proof of proposition 3. Compared to the case of $\frac{\sigma_q}{\sigma_\varepsilon} = 0$, the reduction of the correlation is slow. The correlation of the relative productivity is about 34 percent even after 10 years. This indicates that the assignment effect seems strong even without firm heterogeneity.

In addition, we ask "What would happen if $\frac{\sigma_u}{\sigma_\varepsilon} = 0$?". When $\frac{\sigma_u}{\sigma_\varepsilon} = 0$, $h_\infty = 1$ and $\frac{\sigma_q}{\sigma_{\mu\infty}}$ becomes smaller. In the same way as the case of $\frac{\sigma_F}{\sigma_\varepsilon} = 0$, these combined effects are reported in Figure 1. It shows that an improvement in information causes only slight changes in the persistence of productivity.

To understand the impacts of assignment on the persistence, Figure 2 decomposes the effect of assignment into the dynamic effect, $\gamma = 0$ ($\frac{\gamma\sigma_q}{\sigma_{\mu\infty}} = 0$), which influences the persistence of beliefs, and the static effect, $\psi = 0$ ($\frac{\psi\sigma_q}{\sigma_{\mu\infty}} = 0$), which influences the sensitivity of productivity to beliefs. It shows that both the dynamic effect and the static effect have sizable impacts on the persistence. However, the compounded effects are much bigger, which explains large reductions of the persistence found in the model without assignment in Figure 1.

Table 3 shows an alternative decomposition, which assists in understanding our quantitative result[5]. When there is no assignment, not only $\lambda_{1\infty}$ drops by 58 percent, from 0.713 to 0.298, but z also drops by 75 percent from 0.547 to 0.136. Both effects explain a rapid reduction of the correlation in Figure 1. No assignment also causes a large reduction of $\frac{Var\left[F_A^*\right]}{Var[D\ln A_t]}$, from 0.73 to 0.05. This explains why a reduction in correlation in Figure 1 is so large. Table 3 also reveals the reason behind a change in $\frac{Var\left[F_A^*\right]}{Var[D\ln A_t]}$. Although no assignment reduces within-variance, the reduction of between-variance is more drastic than that of within-variance. Table 3 shows that if there is no assignment, the between-variance drops by 99 percent. It indicates that the most of between variations can be explained by the amplification through the assignment mechanism.

Table 3 can also show the impacts of $\gamma = 0$ ($\frac{\gamma\sigma_q}{\sigma_{\mu\infty}} = 0$) and $\psi = 0$ ($\frac{\psi\sigma_q}{\sigma_{\mu\infty}} = 0$) on the persistence. First, the lack of the dynamic assignment effect, $\frac{\gamma\sigma_q}{\sigma_{\mu\infty}} = 0$, reduces $\lambda_{1\infty}$ and z, which causes a rapid reduction of the correlation shown in Figure 2. Second, although the static assignment effect cannot influence $\lambda_{1\infty}$, it can greatly influence z. A reduction in ψ lowers the impact of changes in beliefs on productivity, which reduces not only the importance of new information, $\left(1 + \frac{\psi\sigma_q}{\sigma_{\mu\infty}}\right)\phi h_\infty$, but also the reliability of beliefs, $\frac{Var\left[E\left[D\ln A_t^F|\mu_{ot}\right]\right]}{Var\left[D\ln A_t^F\right]}$. Both reduces z. This explains why a big initial decline exists when $\psi = 0$ in Figure 2. Third, it also shows that although

[5]Note that as σ_F and σ_ε are chosen to meet between- and within-variances, the benchmark values of the variance are the same as the statistics constructed from the data.

	$\lambda_{1\infty}$	z	$\frac{Var[F_A^*]}{Var[D\ln A_t]}$	$Var\left[D\ln A_t\right]$	$Var\left(F_A^*\right)$	$Var\left[D\ln A_t^F\right]$
Benchmark	0.713	0.547	0.729	0.234	0.171	0.064
$\frac{\sigma_q}{\sigma_\varepsilon}=0$	0.298	0.136	0.049	0.047	0.002	0.045
$\gamma=0$	0.298	0.437	0.343	0.083	0.028	0.054
$\psi=0$	0.713	0.153	0.231	0.059	0.014	0.045
$\frac{\sigma_F}{\sigma_\varepsilon}=0$	0.938	0.875	0	0.198	0	0.198
$\frac{\sigma_u}{\sigma_\varepsilon}=0$	0.703	0.810	0.726	0.212	0.154	0.058

Table 3: The Decomposition of Persistence and Variance

both the lack of the dynamic effect ($\frac{\gamma\sigma_q}{\sigma_{\mu\infty}}=0$) and that of the static effect ($\frac{\psi\sigma_q}{\sigma_{\mu\infty}}=0$) have sizable impacts on $\frac{Var[F_A^*]}{Var[D\ln A_t]}$, the compounded impacts are much greater. These observations confirm that the large impacts of assignment on the persistency can be generated by the combinations of both the slow adjustment of the beliefs and a rise in the sensitivity of productivity to the beliefs.

When there is no firm heterogeneity, by definition, $\frac{Var[F_A^*]}{Var[D\ln A_t]}=0$. However, $\lambda_{1\infty}$ and z increase to 0.938 and 0.875, respectively. This is because a reduction in $\frac{\sigma_F}{\sigma_\varepsilon}$ increases the dynamic assignment effect, $\frac{\gamma\sigma_q}{\sigma_{\mu\infty}}$ and the static assignment effect, $\frac{\psi\sigma_q}{\sigma_{\mu\infty}}$. It explains why a reduction in $\frac{\sigma_F}{\sigma_\varepsilon}$ brings about a slow reduction of correlation in Figure 1.

Finally, even if information is perfect, $\frac{\sigma_u}{\sigma_\varepsilon}=0$, $\lambda_{1\infty}$ and $\frac{Var[F_A^*]}{Var[D\ln A_t]}$ do not change much. This explains why an improvement in information does not change the persistence of productivity in Figure 1 very much. Note that because the TFP is realized after the employment decisions, even if the information contained in the TFP is perfect, the assignment must rely on the beliefs. Small changes in $\lambda_{1\infty}$ suggests that, even when the TFP is perfectly informative, the slow adjustment of beliefs can be an important factor for the persistence of productivity dynamics.

Table 3 can also provide information about how each factor influences the overall variance of relative productivity, $Var\left[D\ln A_t\right]$. When $\frac{\sigma_q}{\sigma_\varepsilon}=0$, $Var\left[D\ln A_t\right]$ drops by 79 percent, from 0.234 to 0.047. Table 3 shows that the strong reduction of between-variance is the main reason for this large reduction. Table 3 also suggests that, similarly to the effect on the persistence, although both $\gamma = 0$ and $\psi = 0$ have large impacts on the reduction in overall variance, the compounded effect is bigger.

On the other hand, when $\frac{\sigma_F}{\sigma_\varepsilon} = 0$ or $\frac{\sigma_u}{\sigma_\varepsilon} = 0$, a change in $Var\left[D \ln A_t\right]$ is small. A small reduction of overall variance under $\frac{\sigma_F}{\sigma_\varepsilon} = 0$ is surprising. When $\frac{\sigma_F}{\sigma_\varepsilon} = 0$, $Var\left(F_A^*\right) = 0$ by definition. Because a reduction of $\frac{\sigma_F}{\sigma_\varepsilon}$ increases the persistence parameter through the dynamic assignment effect, $\lambda_{1\infty} = \phi + \frac{\gamma\sigma_q}{\sigma_{\mu\infty}}$ and the static assignment effect, $\frac{\psi\sigma_q}{\sigma_{\mu\infty}}$, it greatly increases $Var\left[D \ln A_t^F\right]$. A rise in the within-variance mostly offsets the reduction of the between-variance. It indicates that if the assignment mechanism is important, reducing the variations in firm-specific factors does not change the disparity of productivity very much.

The Robustness of the Results: Our results are fairly robust. Table 4 provides a summary of our robustness checks. More comprehensive results for the robustness checks are reported in the Appendix 2.

First, we examine how much the results are sensitive to changes in b_5. The theory predicts that the possible range of b_5 is between 0 and $\frac{E[w]b_3}{(1-\delta)E[y_t]} + b_4$. When b_5 is 0, $b_1 = b_3$. In this case, the TFP can perfectly predict organization capital, $\frac{\sigma_u}{\sigma_\varepsilon} = 0$. When $b_5 = \frac{E[w]b_3}{(1-\delta)E[y_t]} + b_4$, $\frac{\gamma\sigma_q}{\sigma_{\mu\infty}}$ must be 0. That is, skill does not have any impacts on the accumulation of organization capital. In order to obtain the robust rsults, we set $\max b_5 = \frac{E[w]b_3}{(1-\delta)E[y_t]} + b_4 - 0.001$. Our benchmark estimate is located between these values.

Table 4 shows that most of the important results are insensitive to changes in b_5. Assignment has large impacts on the persistence and diversity of the relative productivity. The impact of the noisy information on the persistence and the diversity of relative productivity is minor. A change in firm heterogeneity does not change the diversity of relative productivity very much.

The reason is as follows. Note that when b_5 is set at the minimized (maximized) value, b_2 is set to be the maximized (minimized) value, which largely influences our

	$\min b_5$	$\min b_5$	$\max b_5$	$\max b_5$	COMPUSTAT	COMPUSTAT
Statistics	$\rho_{D\ln A5}$	$Var\left[D\ln A_t\right]$	$\rho_{D\ln A5}$	$Var\left[D\ln A_t\right]$	$\rho_{D\ln y5}$	$Var\left[D\ln y_t\right]$
Benchmark	0.767	0.234	0.767	0.234	0.619	0.423
$\frac{\sigma_q}{\sigma_\varepsilon}=0$	0.067	0.049	0.121	0.049	0.027	0.128
$\gamma=0$	0.282	0.075	0.766	0.233	0.168	0.188
$\psi=0$	0.396	0.073	0.122	0.049	0.209	0.154
$\frac{\sigma_F}{\sigma_\varepsilon}=0$	0.692	0.191	0.222	0.193	0.510	0.372
$\frac{\sigma_u}{\sigma_\varepsilon}=1$	0.767	0.234	0.710	0.270	0.628	0.400

Table 4: Robustness of the Results

The benchmark value of $Var\left[D\ln A_t\right]$ is the same as the estimates from the data by construction. The benchmark of $\rho_{D\ln A5}$ is almost the same as the estimates from the data ($\rho_{D\ln A5}=0.767$ by BSJBSA, $\rho_{D\ln y5}=0.637$ by COMPUSTAT).

calibrations of $\frac{\gamma\sigma_q}{\sigma_{\mu\infty}}$. As workers can contribute to the firm by increasing output and/or increasing organization capital for the future, if there are no investment effects of skill, $\gamma=0$, skilled workers should be extremely productive for current production in order to be consistent with the data. Otherwise, firms would not pay observed high wages for workers ($\frac{E[w_t]}{(1-\delta)E[y_t]}=\frac{E[w_t]}{E[y_t]-rE[k]}=0.83$). This effect offsets the lack of the dynamic assignment effect. That is why a change in b_5 causes sizable changes in the impacts of $\gamma=0$ and $\psi=0$ on the persistence and diversity of relative productivity. in Table 4. This argument indicates that, even if one has an extreme view that skilled workers are not important for the accumulation of knowledge in a firm, the strong impacts of assignment cannot be rejected.

There is one exception. Table 4 shows that the impacts of $\frac{\sigma_F}{\sigma_\varepsilon}=0$ on the persistence of relative productivity can be sensitive to the choice of b_5. When b_5 is set at its maximum value, $\frac{\sigma_F}{\sigma_\varepsilon}=0$ results in $\rho_{D\ln A5}=0.222$, which is significantly smaller than in other cases. In this case, because $\frac{\gamma\sigma_q}{\sigma_{\mu\infty}}$ is assumed to be almost 0, $\lambda_{1\infty}\approx\phi$. This means that a reduction of firm heterogeneity does not increase $\lambda_{1\infty}$. Because $\frac{\sigma_F}{\sigma_\varepsilon}=0$ means $\frac{Var\left[F_A^*\right]}{Var[D\ln A_t]}=0$ by definition, the correlation of relative productivity should drop. However, note that, even in this extreme case, after five years, the lack of assignment still has larger impacts on the reduction of the correlation, $\rho_{D\ln A5} = 0.121$. That is, we cannot change the conclusion that a reduction of assignment effect has more drastic impacts on the persistence than that of firm heterogeneity, at least in the short run.

We also conduct robustness checks using an alternative dataset. It is well-known that the Japanese economy suffered from a large productivity slowdown during the 1990s [e.g., Hayashi and Prescott (2002)]. Although the previous literature reveals that, similarly to other countries, the productivity differences in Japanese firms were large and persistent in this period [e.g., Fukao and Kwon (2006)], the literature also found that, during 1997, relatively productive firms exited from the market more than did unproductive firms, which is unusual [e.g., Nishimura, Nakajima and Kiyota (2005)]. One may wonder whether this influences the results of this paper.

In order to examine this possibility, we use a COMPUSTAT dataset containing data between 1970 and 2004. Because most of the companies in COMPUSTAT do not report labor costs, we are not able to construct the TFP. Therefore, as a proxy, we measure the labor productivity by sales per employee. For this exercise, we use a three-digit industry code to obtain enough firms in an industry and retain firms with more than 10 observations and industries that have more than five firms for all years. We also delete observations with nonpositive values for sales, number of employees or expenses (= Data41+Data189 in COMPUSTAT). This leaves 84,686 observations for our analysis. We set $\delta = 0.25$ and $\frac{E(w)}{E(y)} = 0.63$ [6]. Using the same

[6] The number $\delta = 0.25$ is taken from Samaniego (2006) and $\frac{E(w)}{E(y)} = 0.63$ is estimated by applying the method used in Takii (2008). It turns out that 0.63 is the value for the labor share used in Samaniego (2006).

Note that $\frac{E(w)}{E(y)}$ can be still used for our calibration although we use sales per employee as a proxiy of labor productivity. In order to apply our method to the dynamics of sales per worker, we implicitly assume that a firm can employ physical capital stock and intermediate goods after the TFP is realized. Assume that the production function is $y_t = A_t k_t^{\delta} m_t^{\delta_m}$, where m_t is intermediate goods and the price for the intermediate goods is p_m. After optimally choosing the intermediate goods, procedure as before, we obtained parameters. Table 4 reports the summaries of our counterfactual experiments using COMPUSTAT. As it shows, the overall result is the same. It indicates that our results are less likely to be subject to the Japanese-specific environment.

In summary, these exercises consistently suggest that positive assortative matching accounts for a large component of the observed persistence and disparity of a firm's relative productivity, and the large impacts of positive assortative matching can be explained by the compounded impacts of the slow adjustment of beliefs and an increase in the sensitivity of productivity to beliefs.

6 Conclusion

In this paper, we developed a dynamic assignment model between the skills of workers and unobserved firm-specific knowledge, which we term a firm's organization capital, to account for observed large and persistent productivity differences between firms. We can analytically show that when the assignment between the beliefs regarding organization capital and skill exists, the slow adjustment of unobserved heterogeneous expected relative productivity can play a role similar to an unobserved fixed effect in the productivity dynamics. Our quantitative exercises suggest that the assignment mechanism that causes the sluggish movement of beliefs and a rise in the sensitivity of productivity to the beliefs have quantitatively large impacts on the observed persistence and disparity of productivity. In contrast, firm-specific heterogeneity and noisy information have only modest impacts.

At this point, it is appropriate to discuss how friction may influence the results in this paper. Frictions such as search costs and training cost cannot themselves

---

the production function is expressed as $y_t = xe^{\tilde{u}_t}\tilde{k}_t^o q_t^{\psi^*} k_t^{\delta^*}$, where $x = \left(\frac{\delta_m}{p_m}\right)^{\frac{\delta_m}{1-\delta_m}}$, $\tilde{k}_t^o = (k_t^o)^{\frac{1}{1-\delta_m}}$, $\tilde{u}_t = \frac{u_t}{1-\delta_m}$, $\psi^* = \frac{\psi}{1-\delta_m}$ and $\delta^* = \frac{\delta}{1-\delta_m}$. Hence, what we need to calibrate is $\frac{\psi^*\sigma_q}{\sigma_\varepsilon}$, but not $\frac{\psi\sigma_q}{\sigma_\varepsilon}$. Because the value added is proportional to sales under the Cobb–Douglas production function, our estimate provides the moment we need.

explain large productivity differences, but they can certainly act as an additional source of persistence in our model. However, they may induce the opposite effect, too. As suggested by Acemoglu (1997), a frictional economy increases mismatches[7], which reduces the persistence of productivity. Hence, the overall effect is uncertain. More interestingly, these frictions are likely to interact with the positive feedback mechanism proposed in this paper. Although we cannot predict the consequences of all possible interactions, the previous literature suggests that there may be multiple equilibria [see the survey by Burdett and Coles (1999)]. This means that small changes in skill variations might have more drastic impacts on the persistence and diversity of productivity. This is an interesting topic for future research.

As a final point, we emphasize that the implications of the obtained results offer quite valuable lessons for the empirical studies of productivity dynamics. We provide three implications. First, large persistent productivity differences cannot be taken as evidence of large firm-specific time-invariant heterogeneity. There may not be a positive link between the persistent productivity differences and large firm heterogeneity and, even if the link exists, it can be quantitatively small. Second, empirical research on the persistent productivity differences must seriously take into account the effect of assignment. Our analytical and quantitative results suggest that the mechanism of assignment influencing the persistence and diversity of productivity is really complex, but all effects move in the same direction and the compounded effects are huge. Therefore, the cost of ignoring the assignment mechanism cannot be negligible. Finally, if one could not suitably control the expectations on the firm's capability in the productivity dynamics, researchers might encounter difficulties in finding causal reasons for persistent productivity differences. In order to construct suitable expectations, our derived productivity dynamics can be useful.

Hopefully, these suggestions can assist in advancing research on productivity differences.

[7] Training cost also causes mismatches when the results of training are uncertain.

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] Abowd John M., Francis Kramarz and David N. Margolis (1999), "High Wage Workers and High Wage Firms," Econometrica, Vol. 67, No. 2, March, 251–333.

[3] Acemoglu, Daron (1997), "Matching, Heterogeneity and the Evolution of Income Distribution," Journal of Economic Growth, Vol. 2, No. 1, 61–92.

[4] Atkeson, Andrew and Patrick J. Kehoe (2005), "Modeling and Measuring Organization Capital," Journal of Political Economy, Vol. 113, No. 5, 1026–1053.

[5] Bagger, Jesper, and Rasmus, Lentz (2007), "An Empirical Model of Wage Dispersion with Sorting," mimeo.

[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, Vol. 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] Burdett, Kenneth and Melvyn G. Coles (1999), "Long-Term Partnership Formation: Marriage and Employment, " Economic Journal, Vol.109, June, F307–F334.

[12] 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.

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

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

[15] Eeckhout, Jan and Philipp Kircher (2009), "Identifying Sorting – In Theory," PIER Working Paper, 09–007.

[16] Fox, Jeremy T., and Valérie Smeets (2010), "Does Input Quality Drive Measured Differences in Firm Productivity," forthcoming in International Economic Review.

[17] Fukao, Kyoji, and Hyeog Ug Kwon (2006), "Why Did Japan's TFP Growth Slow Down in the Lost Decade? An Empirical Analysis Based on Firm-Level Data of Manufacturing Firms," Japanese Economic Review, Vol. 57, No. 2, June, 195–228.

[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, Vol. 14, 575–602.

[21] Hayashi, Fumio and Edward C. Prescott (2002), "The 1990s in Japan: A Lost Decade," Review of Economic Dynamics, Vol. 5, No. 1, January, 206–235.

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

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

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

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

[26] Lentz, Rasmus and Dale T. Mortensen (2008), "An Empirical Model of Growth Through Product Innovation," Econometrica, Vol. 76, No. 6, November, 1317–1373.

[27] Lopes de Melo, Rafael (2008), "Sorting in the Labor Market: Theory and Measurement," mimeo

[28] Matsuura, Toshiyuki and Kozo Kiyota (2004), "The Guide for the Use of Basic Survey on Business Activities by Enterprises: The Application to Economic Analysis and the Issues for Data Maintenance (Kigyo Katsudo Kihon Chosa Panel Data no Sakusei Riyou ni tsuite: Keizai Bunseki heno Ouyou to Deta Seibi no Kadai)," RIETI Policy Discussion Series 04-P-004.

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

[30] Melitz, Marc J. (2003), "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity," Econometrica, Vol. 71, No. 6, November, 1695–1725.

[31] Mendes, Rute, Gerard van den Berg and Maarten Lindeboom (2007), "An Empirical Assessment of Assortative Matching in the Labor Market," IZA Discussion Paper, No. 3053.98.

[32] Nishimura, Kiyohiko G., Takenobu Nakajima and Kozo Kiyota (2005), "Does the Natural Selection Mechanism Still Work in Severe Recessions? Examination of the Japanese Economy in the 1990s," Journal of Economic Behavior and Organization, Vol. 58, 53–78.

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

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

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

[36] Takii, Katsuya (2008) "The Persistence of Differences in Productivity, Wages, Skill Mixes and Profits Between Firms in a Rapidly Changing Environment," OSIPP Discussion Paper: DP-2008-E-003, Osaka University, February.

7 Appendix

The Proof of Theorem 2: Consider a mapping, T

$$
\begin{aligned}
TV &= \max_{\ln q_t} \left\{ \begin{array}{c} (1-\delta) E\left[y_t|\mu_{ot}, \sigma_{ot}, \ln q_t\right] - w\left(\ln q_t : \mathbf{x}_t\right) \\ +\beta \int V\left(\mu_{ot+1}, \sigma_{ot+1} : \mathbf{x}_{t+1}, F\right) d\Gamma_s\left(s_t|\mu_{ot}, \sigma_{ot}\right) \end{array} \right\}, \\
E\left[y_t|\mu_{ot}, \sigma_{ot}^2, \ln q_t\right] &= \left(\frac{\delta}{r}\right)^{\frac{\delta}{1-\delta}} \exp \frac{1}{1-\delta}\left(\mu_{ot} + \frac{\sigma_{ot}^2}{2} + \psi \ln q_t\right), \\
\mu_{ot+1} &= \ln B^* + \phi\left[(1-h_t)\mu_{ot} + h_t s_t\right] + \gamma \ln q_t + F - \frac{\sigma_\varepsilon^2}{2}
\end{aligned}
$$

Suppose that the value function and the wage function are represented by equations (19) and (18). Define:

$$
MPQ\left(\mu_{ot}, \sigma_{ot}, \ln q_t\right) = \frac{d(1-\delta)E\left[y_t|\mu_{ot}, \sigma_{ot}^2, \ln q_t\right]}{d\ln q_t} + \beta \int \frac{dV\left(\mu_{ot+1}, \sigma_{ot+1} : \mathbf{x}_{t+1}, F\right)}{d\ln q_t} d\Gamma_s\left(s_t|\mu_{ot}, \sigma_{ot}\right).
$$

Then, we can derive:

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

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

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

Hence, the objective function of the firm endowed with $\mu_{ot} = \frac{\sigma_{\mu t}}{\sigma_q}\left(\ln q_t - \mu_q\right) + \mu_{ot}^e$ is strictly concave in $\ln q_t$ and $\chi^*\left(\mu_{ot}, \sigma_{ot} : \mathbf{x}_t\right)$ is a unique optimal decision. Using this policy function, TV can be rewritten as:

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

where $E\left[y_t|\mu_{ot}, \sigma_{ot}, F\right]$ is defined in Theorem 2. Hence, $TV = V$.

Finally, we show that $\sum_{j=0}^{\infty} \Pi_{u=1}^{j} \frac{\beta\phi}{\phi+\frac{\gamma\sigma_q}{\sigma_{\mu t+u-1}}} \frac{(1-\delta)E[y_{t+j}|\mu_{ot},\sigma_{ot},F]}{1+\frac{\psi\sigma_q}{\sigma_{\mu t+j}}}$ is bounded. Because $\frac{\beta\phi}{\phi+\frac{\gamma\sigma_q}{\sigma_{\mu t+u-1}}} < 1$, it is enough to show that $\lim_{j\to\infty} \sigma_{\mu t+j} > 0$, and $E\left[\mu_{ot+i}|\mu_{ot}, \sigma_{ot}, F\right]$ and $Var\left[\mu_{ot+i}|\mu_{ot}, \sigma_{ot}, F\right]$ are bounded. We first show that if $\frac{\sigma_u}{\sigma_\varepsilon}$ is finite or $\sigma_F > 0$, $\lim_{j\to\infty} \sigma_{\mu t+j} > 0$ and $E\left[\mu_{ot+i}|\mu_{ot}, \sigma_{ot}, F\right]$ and $Var\left[\mu_{ot+i}|\mu_{ot}, \sigma_{ot} : F\right]$ are bounded. For this purpose, the following lemma is useful.

Lemma 9 ***Lemma 10*** *For any $\sigma_{\mu 0}$, there is some $t > 0$ such that:*

$$\sigma_{\mu t} \geq \frac{\gamma\sigma_q + \sigma_F}{1-\phi}.$$

Proof. *Suppose that the above lemma does not apply. Then, for all $t > 0$, $\sigma_{\mu t} < \frac{\gamma\sigma_q+\sigma_F}{1-\phi}$. This means that:*

$$cov_{\mu F t+1} = \left(\phi + \frac{\gamma\sigma_q}{\sigma_{\mu t}}\right) cov_{\mu F t} + \sigma_F^2 > \left(\phi + \frac{1-\phi}{1+\frac{\sigma_F}{\gamma\sigma_q}}\right) cov_{\mu F t} + \sigma_F^2.$$

Hence, there is some t such that:

$$\frac{\sigma_F^2}{1-\left[\phi + \frac{1-\phi}{1+\frac{\sigma_F}{\gamma\sigma_q}}\right]} < cov_{\mu F t} \leq \sigma_{\mu t}\sigma_F.$$

This means that:

$$\sigma_{\mu t} > \frac{\sigma_F}{1 - \left[\phi + \frac{1-\phi}{1+\frac{\sigma_F}{\gamma\sigma_q}}\right]} = \frac{\gamma\sigma_q + \sigma_F}{1-\phi}.$$

Contradiction. ■

Lemma 11 *Suppose $\frac{\sigma_u}{\sigma_\varepsilon}$ is finite or $\sigma_F > 0$:*

$$\sigma_{\mu t} \geq \frac{\gamma\sigma_q}{1-\phi}.$$

Then:

$$\sigma_{\mu t+1} > \frac{\gamma\sigma_q}{1-\phi}.$$

Proof. *Because $\frac{\sigma_u}{\sigma_\varepsilon}$ is finite, it means that $\sigma_{ot} > 0$ and $h_t > 0$, and $\sigma_F > 0$ means that $cov_{\mu Ft} > 0$. Suppose that $\sigma_{\mu t} \geq \frac{\gamma\sigma_q}{1-\phi}$. Then*

$$\begin{aligned}\sigma_{\mu t+1} &= \sqrt{\left(\phi + \frac{\gamma\sigma_q}{\sigma_{\mu t}}\right)^2 \sigma_{\mu t}^2 + \phi^2 h_t \sigma_{ot}^2 + 2\left(\phi + \frac{\gamma\sigma_q}{\sigma_{\mu t}}\right) cov_{\mu Ft} + \sigma_F^2} \\ &> \phi\sigma_{\mu t} + \gamma\sigma_q \geq \phi\frac{\gamma\sigma_q}{1-\phi} + \gamma\sigma_q = \frac{\gamma\sigma_q}{1-\phi}.\end{aligned}$$

■

As lemmas 10 and 11 show that if $\frac{\sigma_u}{\sigma_\varepsilon}$ is finite or $\sigma_F > 0$, there is a τ such that $\sigma_{\mu t} > \frac{\gamma\sigma_q}{1-\phi}$ and, therefore, $\phi + \frac{\gamma\sigma_q}{\sigma_{\mu t}} < 1$ for $t > \tau$. Hence, $\lim_{j\to\infty} \sigma_{\mu t+j} > 0$, and $E\left[\mu_{ot+i}|\mu_{ot}, \sigma_{ot}, F\right]$ and $Var\left[\mu_{ot+i}|\mu_{ot}, \sigma_{ot}, F\right]$ converge to finite values. Next, suppose that $\frac{\sigma_u}{\sigma_\varepsilon} = \infty$ and $\sigma_F = 0$. Then, $Var\left[\mu_{ot+i}|\mu_{ot}, \sigma_{ot}, F\right] = 0$ and $\sigma_{\mu t+j} = \phi\sigma_{\mu t+j-1} + \sigma_q\gamma$. This means that $\sigma_{\mu t+j} = \frac{1-\phi^j}{1-\phi}\sigma_q\gamma + \phi^j\sigma_{\mu t}$ and $\lim_{j\to\infty}\sigma_{\mu t+j} = \frac{\sigma_q\gamma}{1-\phi} > 0$. Hence, it is enough to show that $E\left[\mu_{ot+i}|\mu_{ot}, \sigma_{ot}, F\right]$ is bounded. We define

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

Note that if D_i is bounded, $E\left[\mu_{ot+i}|\mu_{ot}, \sigma_{ot}, F\right]$ is bounded. Hence, we need to prove that D_i is bounded.

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

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

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

Hence, it can be shown that:

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

This means that:

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

Hence, D_i is bounded. **Q.E.D.**

The Proof of Proposition 3: In order to prove the existence of a unique stationary distribution, it is enough to show that there exists a unique $\mathbf{x}_\infty=(\mu^e_{o\infty},\sigma_{\mu\infty},\sigma_{o\infty},cov_{\mu F\infty})$, where $\mathbf{x}_\infty$ is the steady state value of $\mathbf{x}_t$. The steady state values of two variables $\mu^e_{o\infty}$ and $\sigma_{o\infty}$ are directly solved as follows.

$$\mu^e_{o\infty} = \frac{\ln B^*+\gamma\mu_q-\frac{\sigma_\varepsilon^2}{2}}{1-\phi}, \tag{35}$$

$$\sigma^2_{o\infty} = \frac{\sigma_\varepsilon^2}{1-\phi^2(1-h_\infty)}, \tag{36}$$

where:

$$h_\infty = \eta\left(\frac{\sigma_u}{\sigma_\varepsilon}\right) = \frac{-\left(\left(1-\phi^2\right)\left(\frac{\sigma_u}{\sigma_\varepsilon}\right)^2+1\right)+\sqrt{\left(\left(1-\phi^2\right)\left(\frac{\sigma_u}{\sigma_\varepsilon}\right)^2+1\right)^2+4\phi^2\left(\frac{\sigma_u}{\sigma_\varepsilon}\right)^2}}{2\phi^2\left(\frac{\sigma_u}{\sigma_\varepsilon}\right)^2}. \tag{37}$$

Moreover, using (37), we can show that $\eta'\left(\frac{\sigma_u}{\sigma_\varepsilon}\right) < 0$, $\lim_{\frac{\sigma_u}{\sigma_\varepsilon}\to 0}\eta\left(\frac{\sigma_u}{\sigma_\varepsilon}\right) = 1$ and $\lim_{\frac{\sigma_u}{\sigma_\varepsilon}\to\infty}\eta\left(\frac{\sigma_u}{\sigma_\varepsilon}\right) = 0$. Suppose that $\lambda_{1\infty} = \phi + \frac{\gamma\sigma_q}{\sigma_{\mu\infty}} > 1$. Then, $cov_{\mu F\infty}$ diverges and there is no stationary distribution. Hence, $\lambda_{1\infty} = \phi + \frac{\gamma\sigma_q}{\sigma_{\mu\infty}}$ must be less than one in order to guarantee the existence. Suppose that $\lambda_{1\infty} = \phi + \frac{\gamma\sigma_q}{\sigma_{\mu\infty}} < 1$. Then, the steady state value of $cov_{\mu F\infty}$ is also directly solved:

$$cov_{\mu F\infty} = \frac{\sigma_F^2}{1-\lambda_{1\infty}}. \tag{38}$$

Moreover, using equations (36) and (38), $\sigma^2_{\mu\infty}$ must satisfy the following:

$$\begin{aligned}
\sigma^2_{\mu\infty} &= \lambda^2_{1\infty}\sigma^2_{\mu\infty} + \phi^2 h_\infty \sigma^2_{o\infty} + 2\lambda_{1\infty} cov_{\mu F\infty} + \sigma_F^2, \\
&= \frac{1}{1-\lambda^2_{1\infty}}\left[\frac{\phi^2 h_\infty \sigma_\varepsilon^2}{1-\phi^2\left(1-h_\infty\right)} + \frac{2\lambda_{1\infty}\sigma_F^2}{1-\lambda_{1\infty}} + \sigma_F^2\right].
\end{aligned}$$

Rearranging this equation, we can derive:

$$\left(\frac{\sigma_q}{\sigma_\varepsilon}\right)^2 = \left[\frac{1}{1-\left(\phi+\frac{\gamma\sigma_q}{\sigma_{\mu\infty}}\right)^2}\frac{\phi^2 h_\infty}{1-\phi^2\left(1-h_\infty\right)} + \frac{1}{\left[1-\left(\phi+\frac{\gamma\sigma_q}{\sigma_{\mu\infty}}\right)\right]^2}\left(\frac{\sigma_F}{\sigma_\varepsilon}\right)^2\right]\left(\frac{\sigma_q}{\sigma_{\mu\infty}}\right)^2.$$

Define $D\left(\frac{\sigma_q}{\sigma_{\mu\infty}}, \frac{\sigma_u}{\sigma_\varepsilon}, \frac{\sigma_q}{\sigma_\varepsilon}, \frac{\sigma_F}{\alpha\sigma_\varepsilon}\right) = \frac{\left(\frac{\sigma_q}{\sigma_{\mu\infty}}\right)^2}{1-\left(\phi+\frac{\gamma\sigma_q}{\sigma_{\mu\infty}}\right)^2}\frac{\phi^2\eta\left(\frac{\sigma_u}{\sigma_\varepsilon}\right)}{1-\phi^2\left(1-\eta\left(\frac{\sigma_u}{\sigma_\varepsilon}\right)\right)} + \frac{\left(\frac{\sigma_q}{\sigma_{\mu\infty}}\right)^2}{\left[1-\left(\phi+\frac{\gamma\sigma_q}{\sigma_{\mu\infty}}\right)\right]^2}\left(\frac{\sigma_F}{\sigma_\varepsilon}\right)^2 - \left(\frac{\sigma_q}{\sigma_\varepsilon}\right)^2$. We can show that $D_1\left(\frac{\sigma_q}{\sigma_{\mu\infty}}, \frac{\sigma_u}{\sigma_\varepsilon}, \frac{\sigma_q}{\sigma_\varepsilon}, \frac{\sigma_F}{\sigma_\varepsilon}\right) > 0$, $\lim_{\frac{\sigma_q}{\sigma_{\mu\infty}}\to 0} D\left(\frac{\sigma_q}{\sigma_{\mu\infty}}, \frac{\sigma_u}{\sigma_\varepsilon}, \frac{\sigma_q}{\sigma_\varepsilon}, \frac{\sigma_F}{\sigma_\varepsilon}\right) < 0$ and $\lim_{\frac{\sigma_q}{\sigma_{\mu\infty}}\to\frac{1-\phi}{\frac{\sigma_F}{\sigma_q}+\gamma}<1-\phi} D\left(\frac{\sigma_q}{\sigma_{\mu\infty}}, \frac{\sigma_u}{\sigma_\varepsilon}, \frac{\sigma_q}{\sigma_\varepsilon}, \frac{\sigma_F}{\sigma_\varepsilon}\right) > 0$. Hence, there exists $\frac{\sigma_q}{\sigma_{\mu\infty}} = \Sigma\left(\frac{\sigma_q}{\sigma_\varepsilon}, \frac{\sigma_u}{\sigma_\varepsilon}, \frac{\sigma_F}{\sigma_\varepsilon}\right)$ in $\left(0, \frac{1-\phi}{\frac{\sigma_F}{\sigma_q}+\gamma}\right)$ and, therefore, $\lambda_{1\infty} = \phi + \frac{\gamma\sigma_q}{\sigma_{\mu\infty}} < 1$. Moreover, as $D_2\left(\frac{\sigma_q}{\sigma_{\mu\infty}}, \frac{\sigma_u}{\sigma_\varepsilon}, \frac{\sigma_q}{\sigma_\varepsilon}, \frac{\sigma_F}{\sigma_\varepsilon}\right) < 0$, $D_3\left(\frac{\sigma_q}{\sigma_{\mu\infty}}, \frac{\sigma_u}{\sigma_\varepsilon}, \frac{\sigma_q}{\sigma_\varepsilon}, \frac{\sigma_F}{\sigma_\varepsilon}\right) < 0$ and $D_4\left(\frac{\sigma_q}{\sigma_{\mu\infty}}, \frac{\sigma_u}{\sigma_\varepsilon}, \frac{\sigma_q}{\sigma_\varepsilon}, \frac{\sigma_F}{\sigma_\varepsilon}\right) > 0$, $\Sigma_1\left(\frac{\sigma_q}{\sigma_\varepsilon}, \frac{\sigma_u}{\sigma_\varepsilon}, \frac{\sigma_F}{\sigma_\varepsilon}\right) > 0$, $\Sigma_2\left(\frac{\sigma_q}{\sigma_\varepsilon}, \frac{\sigma_u}{\sigma_\varepsilon}, \frac{\sigma_F}{\sigma_\varepsilon}\right) > 0$ and $\Sigma_3\left(\frac{\sigma_q}{\sigma_\varepsilon}, \frac{\sigma_u}{\sigma_\varepsilon}, \frac{\sigma_F}{\sigma_\varepsilon}\right) < 0$. Finally, using the results that $\lim_{\frac{\sigma_u}{\sigma_\varepsilon},\to\infty} D\left(\frac{\sigma_q}{\sigma_{\mu\infty}}, \frac{\sigma_u}{\sigma_\varepsilon}, \frac{\gamma\sigma_q}{\sigma_\varepsilon}, \frac{\sigma_F}{\sigma_\varepsilon}\right) =$

$\frac{\left(\frac{\sigma_q}{\sigma_{\mu\infty}}\right)^2}{\left[1-\left(\phi+\frac{\gamma\sigma_q}{\sigma_{\mu\infty}}\right)\right]^2}\left(\frac{\sigma_F}{\sigma_\varepsilon}\right)^2-\left(\frac{\sigma_q}{\sigma_\varepsilon}\right)^2$ and $\lim_{\frac{\gamma\sigma_q}{\sigma_\varepsilon}\to 0} D\left(\frac{\sigma_q}{\sigma_{\mu\infty}},\frac{\sigma_u}{\sigma_\varepsilon},\frac{\sigma_q}{\sigma_\varepsilon},\frac{\sigma_F}{\sigma_\varepsilon}\right)=\frac{\sigma_q}{\sigma_{\mu\infty}}$, we can show that $\lim_{\frac{\sigma_q}{\sigma_\varepsilon}\to 0}\Sigma\left(\frac{\sigma_q}{\sigma_\varepsilon},\frac{\sigma_u}{\sigma_\varepsilon},\frac{\sigma_F}{\sigma_\varepsilon}\right)=0$, and $\lim_{\frac{\sigma_u}{\sigma_\varepsilon},\to\infty}\Sigma\left(\frac{\sigma_q}{\sigma_\varepsilon},\frac{\sigma_u}{\sigma_\varepsilon},\frac{\sigma_F}{\sigma_\varepsilon}\right)=\frac{1-\phi}{\frac{\sigma_F}{\sigma_q}+\gamma}$. **Q.E.D.**

The Proof of Proposition 4: Because:

$$\frac{d\mu^e_{ot+1}}{d\mu^e_{ot}}=\phi\in(0,1)\,,\ \frac{d\sigma^2_{ot+1}}{d\sigma^2_{ot}}=\phi^2\left(1-h_t\right)^2\in(0,1)\,,$$

$\mu^e_{o\infty}$ and $\sigma_{o\infty}$ are globally stable. Define $S_{\mu t}=\frac{\sigma_{\mu t}}{\gamma\sigma_q}$, $S_F=\frac{\sigma_F}{\gamma\sigma_q}$, $S_{ot}=\frac{\sigma_{ot}}{\gamma\sigma_q}$, $S_\varepsilon=\frac{\sigma_\varepsilon}{\gamma\sigma_q}$, and $C_t=\frac{cov_{\mu Ft}}{(\gamma\sigma_q)^2}$. Then:

$$\begin{aligned}S_{\mu t+1}&=\sqrt{\left(\phi S_{\mu t}+1\right)^2+\phi^2h_tS^2_{ot}+2\left(\phi+\frac{1}{S_{\mu t}}\right)C_t+S^2_F},\\C_{t+1}&=\Sigma^2_F+\left(\phi+\frac{1}{\Sigma_{\mu t}}\right)C_t.\end{aligned}$$

We will prove that if $\frac{\gamma\sigma_q}{\sigma_\varepsilon}\frac{\sigma_F}{\alpha\sigma_\varepsilon}\leq\frac{\phi^2(1-\phi)\eta\left(\frac{\sigma_u}{\sigma_\varepsilon}\right)}{2\left[1-\phi^2\left(1-\eta\left(\frac{\sigma_u}{\sigma_\varepsilon}\right)\right)\right]}$, or equivalently, if $S_F\leq\frac{\phi^2(1-\phi)\eta\left(\frac{\sigma_u}{\sigma_\varepsilon}\right)S^2_\varepsilon}{2\left[1-\phi^2\left(1-\eta\left(\frac{\sigma_u}{\sigma_\varepsilon}\right)\right)\right]}$, $S_{\mu\infty}$ and C_∞ are globally stable.

Lemma 12 *Suppose that for all S_μ, C and t:*

$$\left|\frac{C}{S^2_\mu}\right|+\phi+\frac{1}{S_\mu}<1$$

Then, a mapping of T such that:

$$T\left(S_\mu,C,t\right)=\begin{bmatrix}\sqrt{\left(\phi S_\mu+1\right)^2+\phi^2h_tS^2_{ot}+2\left(\phi+\frac{1}{S_\mu}\right)C+S^2_F}\\S^2_F+\left(\phi+\frac{1}{S_\mu}\right)C\end{bmatrix}$$

is a contraction mapping.

Proof. *Define $f(\cdot)$ and $g(\cdot)$:*

$$\begin{aligned}f\left(S_\mu,C,t\right)&=\sqrt{\left(\phi S_\mu+1\right)^2+\phi^2h_tS^2_{ot}+2\left(\phi+\frac{1}{S_\mu}\right)C+S^2_F}\\g\left(S_\mu,C\right)&=\Sigma^2_F+\left(\phi+\frac{1}{S_\mu}\right)C\end{aligned}$$

Take any (S_μ, C) *and* $\left(S'_\mu, C'\right)$. *Then, the mean value theorem implies that there exists* $\left(\bar{S}_\mu, \bar{C}^*\right)$, *where* $\bar{S}_\mu \in \left(S_\mu, S'_\mu\right)$ *and* $C \in (C, C')$ *such that:*

$$\begin{aligned} f(S_\mu, C, t) - f\left(S'_\mu, C', t\right) &= \frac{\left(\phi \bar{S}_\mu + 1\right)\phi - \frac{\bar{C}}{\bar{\Sigma}_\mu^2}}{f\left(\bar{S}_\mu, \bar{C}, t\right)}\left(S_\mu - S'_\mu\right) + \frac{\phi + \frac{1}{\bar{S}_\mu}}{f\left(\bar{S}_\mu, \bar{C}, t\right)}(C - C'), \\ g(S_\mu, C^*) - g\left(S'_\mu, C'\right) &= -\frac{\bar{C}^*}{\bar{S}_\mu^2}\left(S_\mu - S'_\mu\right) + \left(\phi + \frac{1}{\bar{S}_\mu}\right)(C - C'). \end{aligned}$$

Because we can show that:

$$\begin{aligned} & \max\left\{\left|f(S_\mu, C, t) - f\left(S'_\mu, C', t\right)\right|, \left|g(S_\mu, C) - g\left(S'_\mu, C'\right)\right|\right\} \\ \leq\ & \max\left\{\begin{array}{c} \frac{\left|\left(\phi \bar{S}_\mu + 1\right)\phi - \frac{\bar{C}}{\bar{S}_\mu^2}\right| + \left(\phi + \frac{1}{\bar{S}_\mu}\right)}{f\left(\bar{S}_\mu, \bar{C}^*, t\right)}, \\ \left|\frac{\bar{C}}{\bar{S}_\mu^2}\right| + \phi + \frac{1}{\bar{S}_\mu} \end{array}\right\} \max\left\{\left|S_\mu - S'_\mu\right|, |C - C'|\right\} \end{aligned}$$

if, for all S_μ, C *and* t,*:*

$$\max\left\{\frac{\left|(\phi S_\mu + 1)\phi - \frac{C}{\bar{\Sigma}_\mu^2}\right| + \left(\phi + \frac{1}{S_\mu}\right)}{\sqrt{(\phi S_\mu + 1)^2 + \phi^2 h_t S_{ot}^2 + 2\left(\phi + \frac{1}{S_\mu}\right)C + S_F^2}},\ \left|\frac{C}{S_\mu^2}\right| + \phi + \frac{1}{S_\mu}\right\} < 1,$$

the mapping of T *is a contraction. Next, we need to show that:*

$$\frac{\left|(\phi S_\mu + 1)\phi - \frac{C}{S_\mu^2}\right| + \left(\phi + \frac{1}{S_\mu}\right)}{\sqrt{(\phi S_\mu + 1)^2 + \phi^2 h_t S_{ot}^2 + 2\left(\phi + \frac{1}{S_\mu}\right)C + S_F^2}} < \left|\frac{C}{S_\mu^2}\right| + \phi + \frac{1}{S_\mu}$$

Suppose that $(\phi S_\mu + 1)\phi \geq \frac{C}{S_\mu^2}$.

$$\begin{aligned} & \frac{(\phi S_\mu + 1)\phi - \frac{C}{S_\mu^2} + \left(\phi + \frac{1}{S_\mu}\right)}{\sqrt{(\phi S_\mu + 1)^2 + \phi^2 h_t S_{ot}^2 + 2\left(\phi + \frac{1}{S_\mu}\right)C + S_F^2}} \\ =\ & \frac{S_\mu\left(\phi + \frac{1}{S_\mu}\right)^2 - \frac{C}{S_\mu^2}}{\sqrt{\left(\phi + \frac{1}{S_\mu}\right)^2 S_\mu^2 + \phi^2 h_t S_{ot}^2 + 2\left(\phi + \frac{1}{S_\mu}\right)C + S_F^2}} \\ \leq\ & \phi + \frac{1}{S_\mu} \leq \left|\frac{C}{S_\mu^2}\right| + \phi + \frac{1}{S_\mu} \end{aligned}$$

Suppose that $(\phi S_\mu + 1)\,\phi < \frac{C}{S_\mu^2}$.

$$\begin{aligned} & \frac{\frac{C}{S_\mu^2} - (\phi S_\mu + 1)\,\phi + \left(\phi + \frac{1}{S_\mu}\right)}{\sqrt{(\phi S_\mu + 1)^2 + \phi^2 h_t S_{ot}^2 + 2\left(\phi + \frac{1}{S_\mu}\right) C + S_F^2}} \\ < \; & \frac{\frac{C}{S_\mu^2} + \left(\phi + \frac{1}{S_\mu}\right)}{\sqrt{(\phi S_\mu + 1)^2 + \phi^2 h_t S_{ot}^2 + 2\left(\phi + \frac{1}{S_\mu}\right) C + S_F^2}} \\ < \; & \frac{C}{S_\mu^2} + \left(\phi + \frac{1}{S_\mu}\right). \end{aligned}$$

Hence, if for all S_μ, C *and* t $\left|\frac{C^*}{S_\mu^2}\right| + \phi + \frac{1}{S_\mu} < 1$, *the mapping* T *is a contraction.* ■

Lemma 13 *Suppose that* $S_\mu > \frac{1+S_F}{1-\phi}$.

$$\left|\frac{C}{S_\mu^2}\right| + \left|\phi + \frac{1}{S_\mu}\right| < 1$$

Proof. *Suppose that* $S_\mu > \frac{1+S_F}{1-\phi}$. *Because* $S_F S_\mu \geq |C|$,

$$1 > \phi + \frac{1}{S_\mu} + \frac{S_F}{S_\mu} \geq \phi + \frac{1}{S_\mu} + \left|\frac{C}{S_\mu^2}\right|.$$

■

Lemma 14 *Suppose that:*

$$S_F \leq \frac{\phi^2 (1-\phi)\,\eta\left(\frac{\sigma_u}{\sigma_\varepsilon}\right) S_\varepsilon^2}{2\left[1 - \phi^2\left(1 - \eta\left(\frac{\sigma_u}{\sigma_\varepsilon}\right)\right)\right]}.$$

There exists a τ *such that, for all* $t > \tau$ *if:*

$$S_{\mu t} \geq \frac{1+S_F}{1-\phi},$$

then:

$$S_{\mu t+1} > \frac{1+S_F}{1-\phi}.$$

Proof. *Because* $C_{t+1} = S_F^2 + \left(\phi + \frac{1}{S_{\mu t}}\right) C_t > S_F^2 + \phi C_t$ *and* $h_t S_{ot}^2$ *converges to* $\frac{\eta\left(\frac{\sigma_u}{\sigma_\varepsilon}\right) S_\varepsilon^2}{1-\phi^2\left(1-\eta\left(\frac{\sigma_u}{\sigma_\varepsilon}\right)\right)}$*, there exists a* τ *such that, for all* $t > \tau$*:*

$$S_{\mu t+1} > \sqrt{(\phi S_{\mu t} + 1)^2 + \frac{\phi^2 \eta\left(\frac{\sigma_u}{\sigma_\varepsilon}\right) S_\varepsilon^2}{1-\phi^2\left(1-\eta\left(\frac{\sigma_u}{\sigma_\varepsilon}\right)\right)} + 2\phi \frac{S_F^2}{1-\phi} + S_F^2}.$$

Suppose that $S_{\mu t} \geq \frac{1+S_F}{1-\phi}$*. Then, for all* $t > \tau$*:*

$$S_{\mu t+1} > \sqrt{\left(\phi \frac{1+S_F}{1-\phi} + 1\right)^2 + \frac{\phi^2 \eta\left(\frac{\sigma_u}{\sigma_\varepsilon}\right) S_\varepsilon^2}{1-\phi^2\left(1-\eta\left(\frac{\sigma_u}{\sigma_\varepsilon}\right)\right)} + 2\phi \frac{S_F^2}{1-\phi} + S_F^2}.$$

Rearranging the equation, we can obtain:

$$S_{\mu t+1} > \sqrt{\left(\frac{1+S_F}{1-\phi}\right)^2 + \frac{\phi^2 \eta\left(\frac{\sigma_u}{\sigma_\varepsilon}\right) S_\varepsilon^2}{1-\phi^2\left(1-\eta\left(\frac{\sigma_u}{\sigma_\varepsilon}\right)\right)} - \frac{2S_F}{(1-\phi)}}.$$

Because $S_F \leq \frac{\phi^2(1-\phi)\eta\left(\frac{\sigma_u}{\sigma_\varepsilon}\right) S_\varepsilon^2}{2\left[1-\phi^2\left(1-\eta\left(\frac{\sigma_u}{\sigma_\varepsilon}\right)\right)\right]}$*,* $S_{\mu t+1} > \frac{1+S_F}{1-\phi}$*.* ■

Because S_{kt} and, therefore, h_t are globally stable, using lemma 10, lemma 12, lemma 13 and lemma 14, the desired result follows. **Q.E.D.**

The Proof of Proposition 6: First, we derive a closed form solution of $\rho_{D \ln A_t^F j}$ and later we discuss how we can derive a closed form solution of $Var\left[F^*\right]$ and $Var\left[D \ln A_t^F\right]$. Because $\mathbf{A}_t^F = \mathbf{M}^A \mathbf{A}_{t-1}^F + \mathbf{v}_{t-1}$,

$$E\left[\mathbf{A}_t^F \left(\mathbf{A}_{t-j}^F\right)'\right] = \left(\mathbf{M}^A\right)^{j-1} E\left[\mathbf{A}_{t-(j-1)}^F \left(\mathbf{A}_{t-j}^F\right)'\right] = \mathbf{T_A} \mathbf{\Lambda}^{j-1} \mathbf{T_A}^{-1} E\left[\mathbf{A}_{t+1}^F \left(\mathbf{A}_t^F\right)'\right],$$

where:

$$\mathbf{T_A} \mathbf{\Lambda}^j \mathbf{T_A}^{-1} = \frac{1}{\phi h_\infty + \frac{\gamma\sigma_q}{\sigma_{\mu\infty}}} \left\{ \begin{array}{c} \begin{bmatrix} \phi h_\infty \lambda_{1\infty}^j + \frac{\gamma\sigma_q}{\sigma_{\mu\infty}} \lambda_{2\infty}^j & \frac{\gamma\sigma_q}{\sigma_{\mu\infty}} \left(\lambda_{1\infty}^j - \lambda_{2\infty}^j\right) \\ \phi h_\infty \left(\lambda_{1\infty}^j - \lambda_{2\infty}^j\right) & \frac{\gamma\sigma_q}{\sigma_{\mu\infty}} \lambda_{1\infty}^j + \phi h_\infty \lambda_{2\infty}^j \end{bmatrix} \\ +\phi \frac{\psi\sigma_q}{\sigma_{\mu\infty}} h_\infty \left(\lambda_{1\infty}^j - \lambda_{2\infty}^j\right) \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix} \end{array} \right\}.$$

Note that $E\left[\mathbf{A}_{t+1}^F\left(\mathbf{A}_t^F\right)'\right] = \mathbf{M}^A E\left(\mathbf{A}_t^F\left(\mathbf{A}_t^F\right)'\right) + E\left(\mathbf{v}_t\mathbf{A}_t^F\right)$. As we can show that $E\left(\mathbf{A}_t^F\left(\mathbf{A}_t^F\right)'\right) = Var\left[E\left[\ln A_t^F|\mu_{ot}\right]\right]\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} + Var\left[\ln A_t^F|\mu_{ot}\right]\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ and $E\left[\mathbf{v}_t\left(\mathbf{A}_t^F\right)'\right] = \begin{bmatrix} -\phi\sigma_u^2 & 0 \\ 0 & 0 \end{bmatrix}$, it is shown that:

$$E\left[\mathbf{A}_{t+1}^F\left(\mathbf{A}_t^F\right)'\right] = \lambda_{1\infty} Var\left[E\left[D\ln A_t^F|\mu_{ot}\right]\right]\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} + \phi h_\infty\left(1+\frac{\psi\sigma_q}{\sigma_{\mu\infty}}\right) Var\left[D\ln A_t^F|\mu_{ot}\right]\begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix},$$

where we use a relationship $\frac{\sigma_u^2}{Var\left[D\ln A_t^F|\mu_{ot}\right]} = 1 - h_\infty$ for this derivation. The relationship, $\frac{\sigma_u^2}{Var\left[D\ln A_t^F|\mu_{ot}\right]} = 1 - h_\infty$, is easily proved from $Var\left[D\ln A_t^F|\mu_{ot}\right] = \frac{\sigma_{o\infty}^2}{h_\infty}$, the derivation of which is discussed later. Now, we can derive the following:

$$E\left[\mathbf{A}_t^F\left(\mathbf{A}_{t-j}^F\right)'\right] = \lambda_{1\infty}^{j-1}\left\{\begin{array}{c} \lambda_{1\infty} Var\left[E\left[D\ln A_t^F|\mu_{ot}\right]\right]\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \\ +\phi h_\infty\left(1+\frac{\psi\sigma_q}{\sigma_{\mu\infty}}\right) Var\left[D\ln A_t^F|\mu_{ot}\right]\begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} \end{array}\right\}.$$

Hence, we can show that:

$$\begin{aligned} \rho_{E[D\ln A^F|\mu]j} &= \lambda_{1\infty}^j, \\ \rho_{D\ln A^F j} &= \left\{\begin{array}{c} \lambda_{1\infty}\frac{Var\left[E\left[D\ln A_t^F|\mu_{ot}\right]\right]}{Var\left[D\ln A_t^F\right]} \\ +\phi h_\infty\left(1+\frac{\psi\sigma_q}{\sigma_{\mu\infty}}\right)\left[1-\frac{Var\left[E\left[D\ln A_t^F|\mu_{ot}\right]\right]}{Var\left[D\ln A_t^F\right]}\right] \end{array}\right\}\lambda_{1\infty}^{j-1}. \end{aligned}$$

This shows that we need to know a closed form solution of $\frac{Var\left[E\left[D\ln A_t^F|\mu_{ot}\right]\right]}{Var\left[D\ln A_t^F\right]}$ for $\rho_{D\ln A^F j}$. Hence, the remaining tasks are the derivations of $Var\left[E\left[D\ln A_t^F|\mu_{ot}\right]\right]$, $Var\left[D\ln A_t^F\right]$ and $Var\left[F_A^*\right]$. As $D\ln A_t^F = \left[(\ln k_t^o - \mu_{o\infty}^e) - \frac{F}{1-\lambda_{1\infty}}\right] + \frac{\psi\sigma_q}{\sigma_{\mu\infty}}\left[(\mu_{ot} - \mu_{o\infty}^e) - \frac{F}{1-\lambda_{1\infty}}\right] + u_t^*$, we can show that $Var\left[E\left[D\ln A_t^F|\mu_{ot}\right]\right] = \left(1+\frac{\psi\sigma_q}{\sigma_{\mu\infty}}\right)^2\frac{\phi^2 h_\infty\sigma_{o\infty}^2}{1-\lambda_{1\infty}^2}$ and $Var\left[D\ln A_t^F|\mu_{ot}\right] = E\left[Var\left[D\ln A_t^F|\mu_{ot}\right]\right] = \frac{\sigma_{o\infty}^2}{h_\infty}$. Using equation (36), $Var\left[E\left[D\ln A_t^F|\mu_{ot}\right]\right] = \frac{\left(1+\frac{\psi\sigma_q}{\sigma_{\mu\infty}}\right)^2\phi^2 h_\infty\sigma_\varepsilon^2}{\left(1-\lambda_{1\infty}^2\right)\left[1-\phi^2(1-h_\infty)\right]}$ and $E\left[Var\left[D\ln A_t^F|\mu_{ot}\right]\right] = \frac{\sigma_\varepsilon^2}{h_\infty\left[1-\phi^2(1-h_\infty)\right]}$. Because $Var\left[D\ln A_t^F\right] = Var\left[E\left[D\ln A_t^F|\mu_{ot}\right]\right] +$

$E\left[Var\left[D\ln A_t^F|\mu_{ot}\right]\right]$, $\frac{Var\left[E\left[D\ln A_t^F|\mu_{ot}\right]\right]}{Var\left[D\ln A_t^F\right]} = \frac{\frac{\left(1+\frac{\psi\sigma_q}{\sigma_{\mu\infty}}\right)^2\phi^2 h_\infty^2}{1-\lambda_{1\infty}^2}}{\frac{\left(1+\frac{\psi\sigma_q}{\sigma_{\mu\infty}}\right)^2\phi^2 h_\infty^2}{1-\lambda_{1\infty}^2}+1}$. Because $Var\left[D\ln A_t^F\right] = Var\left[E\left[D\ln A_t^F|\mu_{ot}\right]\right] + E\left[Var\left[D\ln A_t^F|\mu_{ot}\right]\right]$, it is derived that $Var\left[D\ln A_t^F\right] = \left[\frac{\left(1+\frac{\psi\sigma_q}{\sigma_{\mu\infty}}\right)^2\phi^2 h_\infty^2}{1-\lambda_1^2}+1\right]\frac{(\sigma_\varepsilon)^2}{h_\infty\left[1-\phi^2(1-h_\infty)\right]}$. Finally, because $F^* = \frac{1+\frac{\psi\sigma_q}{\sigma_{\mu\infty}}}{1-\lambda_{1\infty}}F$, $Var\left[F^*\right] = \frac{\left(1+\frac{\psi\sigma_q}{\sigma_{\mu\infty}}\right)^2}{(1-\lambda_{1\infty})^2}\sigma_F^2$ and the desired results are derived . **Q.E.D.**

Data Construction: The construction of the value added and capital stock follows from Matsuura and Kiyota (2004).

- y_{ft}: The value added is defined by sales + an increase or decrease in inventory - operating expense + labor expense + depreciation. The variable y_{ft} is estimated by the value added divided by the number of workers.

- k_{ft}: According to Matsuura and Kiyota (2004), there are many missing values. They suggested the following method.

$$\begin{aligned} K_{ft} &= K_{ft-1} + \left(\tilde{K}_{ft} - \tilde{K}_{ft-1}\right)/P_{It},\ if\ \tilde{K}_{ft} > \tilde{K}_{ft-1} \\ &= K_{ft-1} + \left(\tilde{K}_{ft} - \tilde{K}_{ft-1}\right),\ if\ \tilde{K}_{ft} \leq \tilde{K}_{ft-1}, \end{aligned}$$

  where K_{ft} is teal physical capital stock,, $\tilde{K}_{ft}$ is the book value of tangible fixed assets at year t and P_{It} is an investment goods deflator. The equation assumes that if $\tilde{K}_{ft} > \tilde{K}_{ft-1}$, we have net investment, but if $\tilde{K}_{ft} \leq \tilde{K}_{ft-1}$, there is no net investment.

- r_t: The user cost of capital, r_t, is estimated by the following equation:

$$r_t = P_{It} \times (i_t + dep - g_{pit}),$$

  where i_t is the 10-year-bond yield, which is taken from the Bank of Japan, dep is the depreciation rate of capital and $g_{pit} = \frac{P_{It}-P_{It-1}}{P_{It-1}}$. The depreciation rate is estimated by the sample average of $\frac{Dep_{ft}}{K_{ft}}$, where Dep_{ft} is the depreciation of the fth firm in year t.

Summary Statistics: (The number of firms is 6265 for all years)

	y_{ft} (nominal)		k_{ft}		# of workers		w_{ft} (nominal)	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
1994	7.297	11.85	9.829	15.40	621.4	2398	4.640	1.683
1995	7.664	5.249	9.978	17.06	611.9	2316	5.024	1.676
1996	7.978	6.791	10.24	17.04	610.0	2311	5.179	1.713
1997	7.824	8.389	10.54	17.01	613.0	2367	5.103	1.659
1998	7.483	4.917	10.92	17.48	608.3	2356	5.082	1.643
1999	7.660	5.978	11.18	18.59	604.8	2391	5.041	1.642
2000	8.144	7.624	11.41	17.25	605.4	2509	5.244	1.876
2001	7.775	7.150	11.80	17.87	588.2	2317	5.229	1.913
2002	7.860	7.185	11.88	17.81	586.2	2384	5.154	1.914
2003	8.036	11.18	11.83	16.50	590.6	2402	5.101	2.090
2004	8.337	15.76	11.84	15.91	604.8	2724	5.085	2.011

8 Appendix 2: Robustness

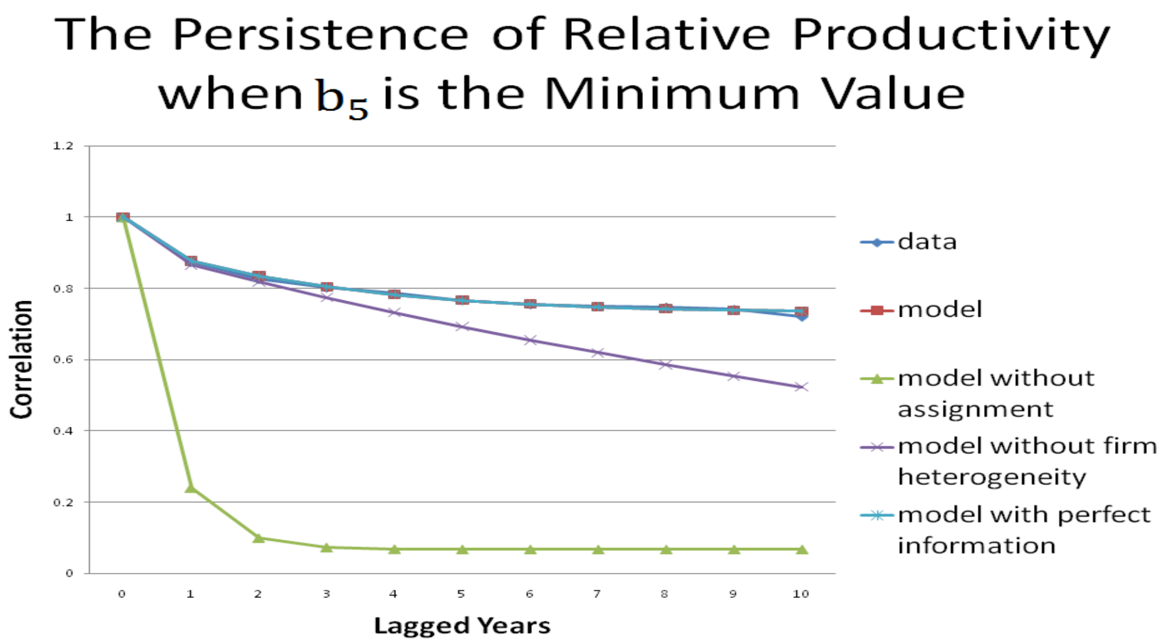
8.1 Robustness Checks using BSJBSA

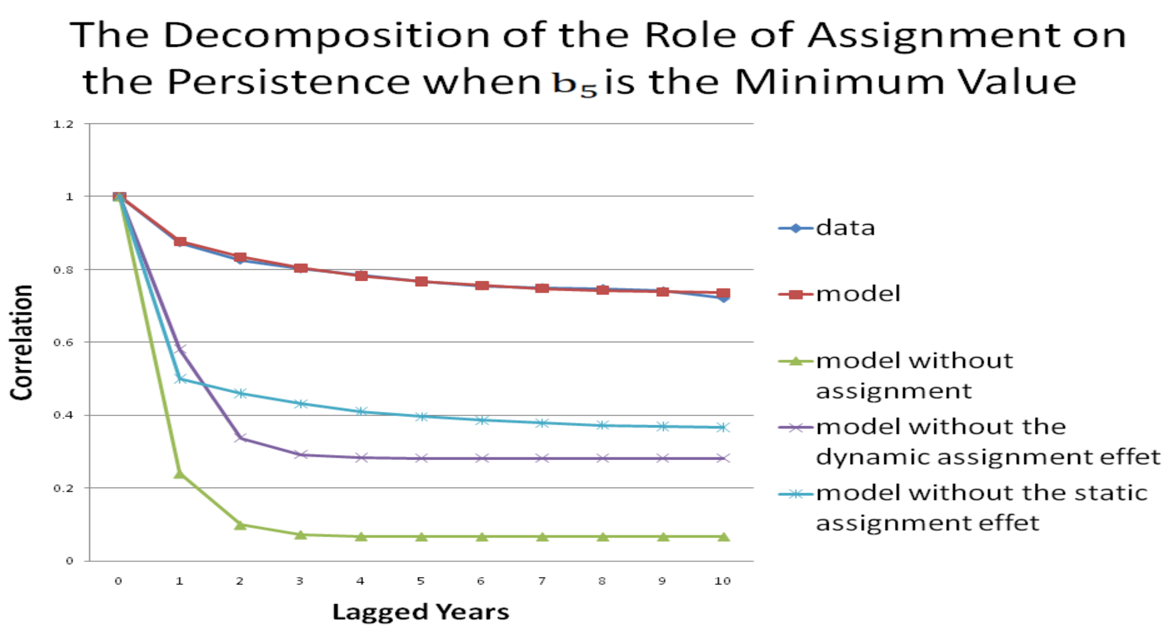
Calibrated Parameters using BSJBSA

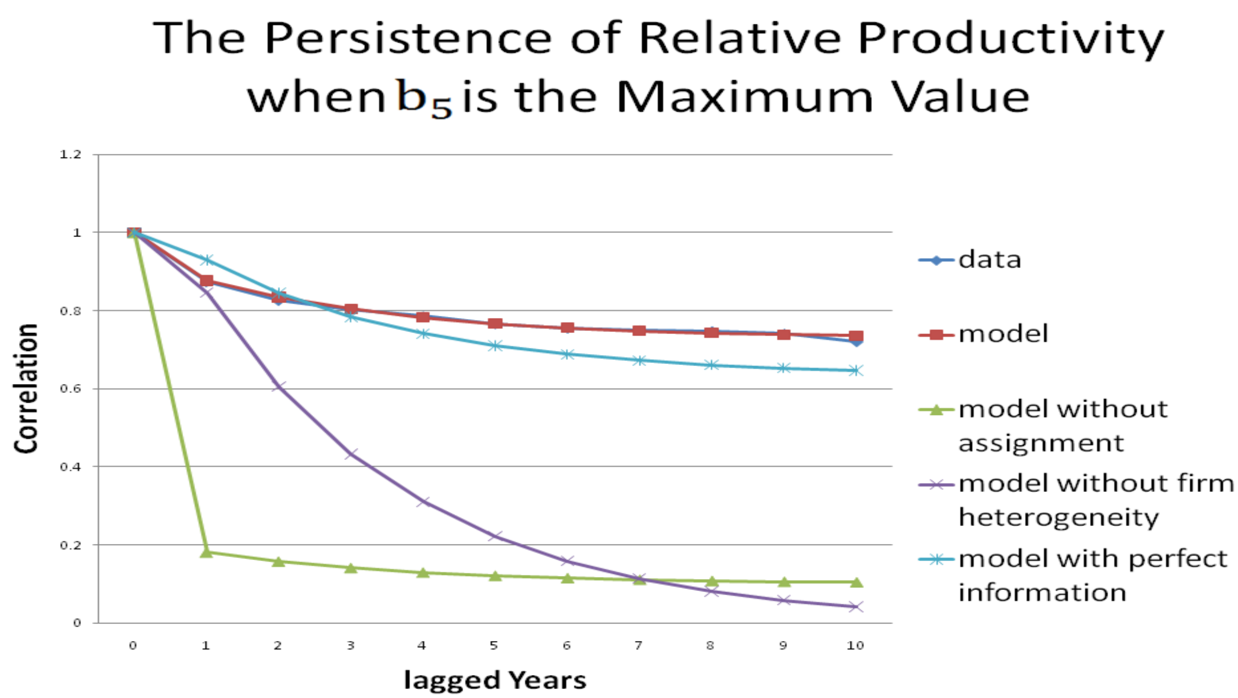
	ϕ	$\frac{\gamma\sigma_q}{\sigma_{\mu\infty}}$	h_∞	σ_F^2	σ_ε^2	$\frac{\psi\sigma_q}{\sigma_{\mu\infty}}$
$\min b_5$	0.185	0.528	1	0.002	0.044	1.547
$\max b_5$	0.713	0.001	0.113	0.0004	0.003	4.846

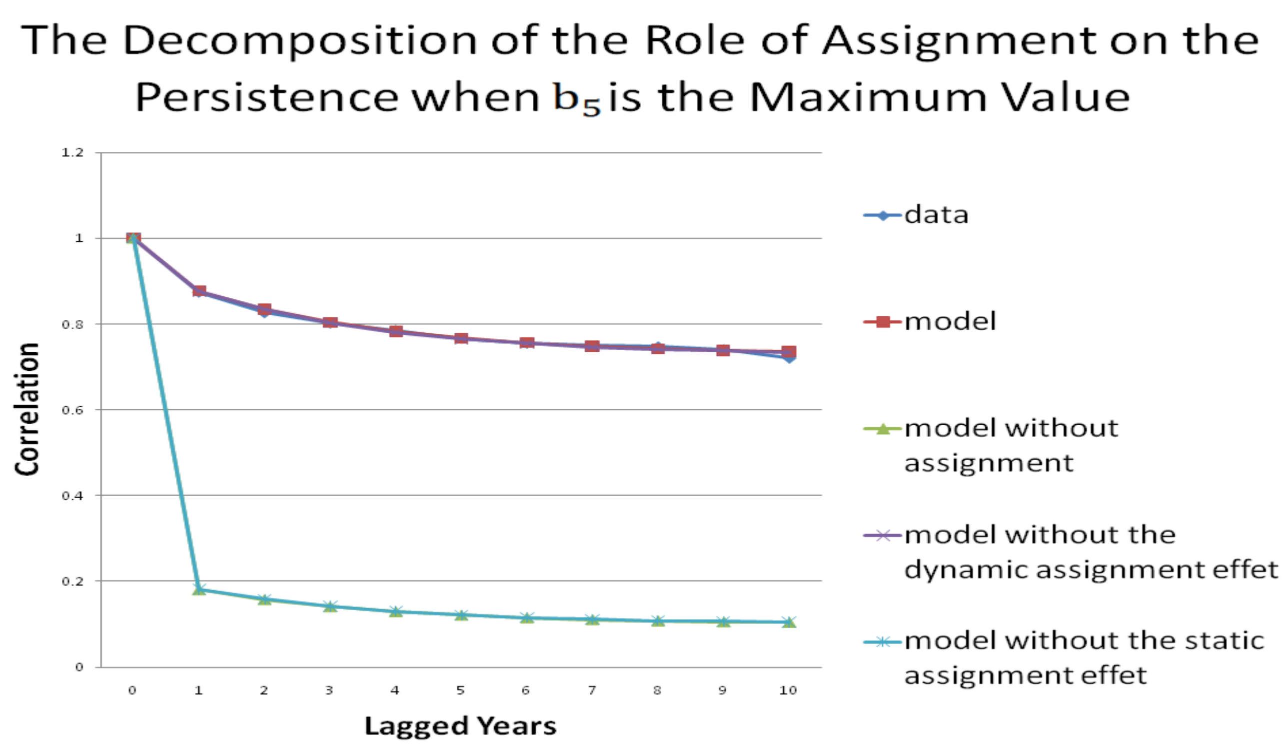
The Decomposition of Persistence and Variance using BSJBSA

		$\lambda_{1\infty}$	z	$\frac{Var[F_A^*]}{Var[D\ln A_t]}$	$Var\left[D\ln A_t\right]$	$Var\left(F_A^*\right)$	$Var\left[D\ln A_t^F\right]$
	Benchmark	0.713	0.547	0.729	0.234	0.171	0.064
$\min b_5$	$\frac{\sigma_q}{\sigma_\varepsilon}=0$	0.185	0.185	0.067	0.049	0.003	0.045
	$\gamma=0$	0.185	0.418	0.282	0.075	0.021	0.054
	$\psi=0$	0.713	0.219	0.360	0.073	0.026	0.047
	$\frac{\sigma_F}{\sigma_\varepsilon}=0$	0.946	0.866	0	0.191	0	0.191
	$\frac{\sigma_u}{\sigma_\varepsilon}=0$	0.713	0.547	0.729	0.234	0.171	0.064
$\max b_5$	$\frac{\sigma_q}{\sigma_\varepsilon}=0$	0.713	0.089	0.101	0.049	0.005	0.044
	$\gamma=0$	0.713	0.547	0.727	0.233	0.170	0.064
	$\psi=0$	0.713	0.089	0.101	0.049	0.005	0.044
	$\frac{\sigma_F}{\sigma_\varepsilon}=0$	0.715	0.846	0	0.193	0	0.193
	$\frac{\sigma_u}{\sigma_\varepsilon}=0$	0.713	0.808	0.633	0.270	0.171	0.099


The Persistence of Relative Productivity when b_5 is the Minimum Value
Correlation
Lagged Years
data
model
model without assignment
model without firm heterogeneity
model with perfect information
0
0.2
0.4
0.6
0.8
1
1.2
0 1 2 3 4 5 6 7 8 9 10


The Decomposition of the Role of Assignment on the Persistence when b_5 is the Minimum Value
Correlation
Lagged Years
data
model
model without assignment
model without the dynamic assignment effet
model without the static assignment effet
0
0.2
0.4
0.6
0.8
1
1.2
0 1 2 3 4 5 6 7 8 9 10


The Persistence of Relative Productivity when b_5 is the Maximum Value
Correlation
lagged Years
0
1
2
3
4
5
6
7
8
9
10
0.2
0.4
0.6
0.8
1
1.2
data
model
model without assignment
model without firm heterogeneity
model with perfect information


The Decomposition of the Role of Assignment on the Persistence when b_5 is the Maximum Value
Correlation
Lagged Years
0
1
2
3
4
5
6
7
8
9
10
0.2
0.4
0.6
0.8
1
1.2
data
model
model without assignment
model without the dynamic assignment effet
model without the static assignment effet

8.2 Robustness Checks using COMPUSTAT

Regression Results using COMPUSTAT

	$D\ln A_{t-1}$	$E\left[D\ln A_{t-1}\vert\mu_{ot-1}\right]$	# of Obs.
$D\ln A_t$	0.531* (0.205)	0.182* (0.021)	74020

	$D\ln A_{t-1} - E\left[D\ln A_{t-1}\vert\mu_{ot-1}\right]$		# of Obs.
$D\ln A_t - E\left[D\ln A_t\vert\mu_{ot}\right]$	0.044 (0.049)		74020

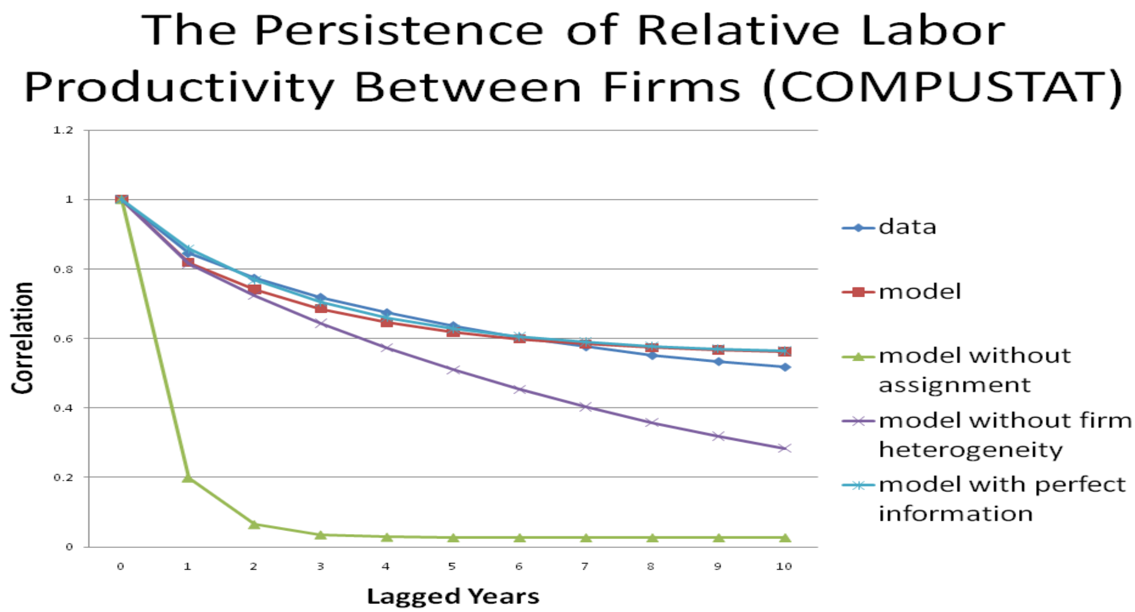
WC-Robust standard errors are reported in parentheses. *denotes significance at the one percent level.

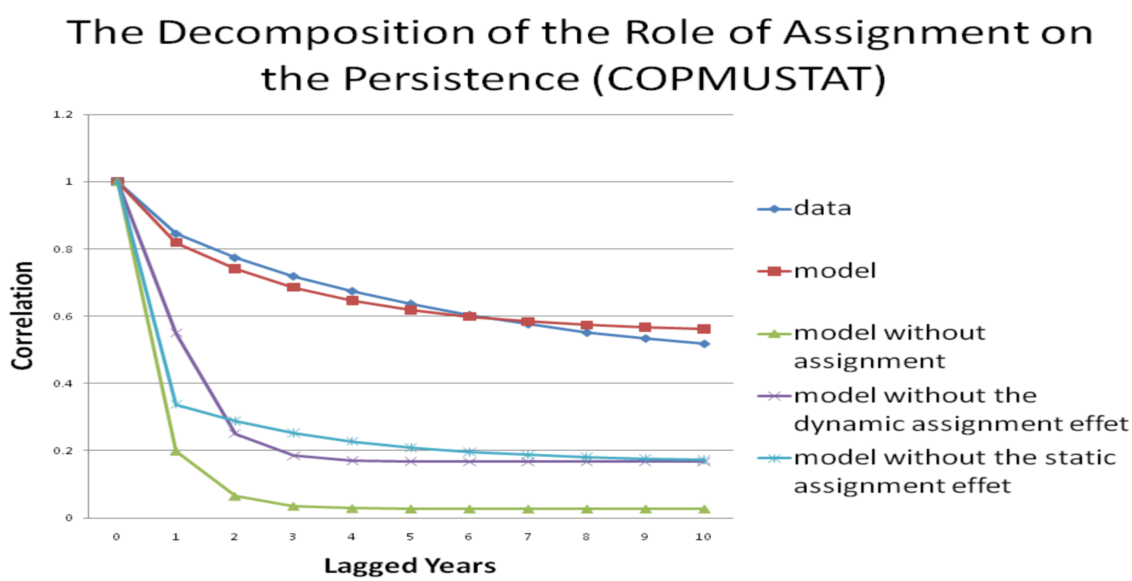
Calibrated Parameters using COMPUSTAT

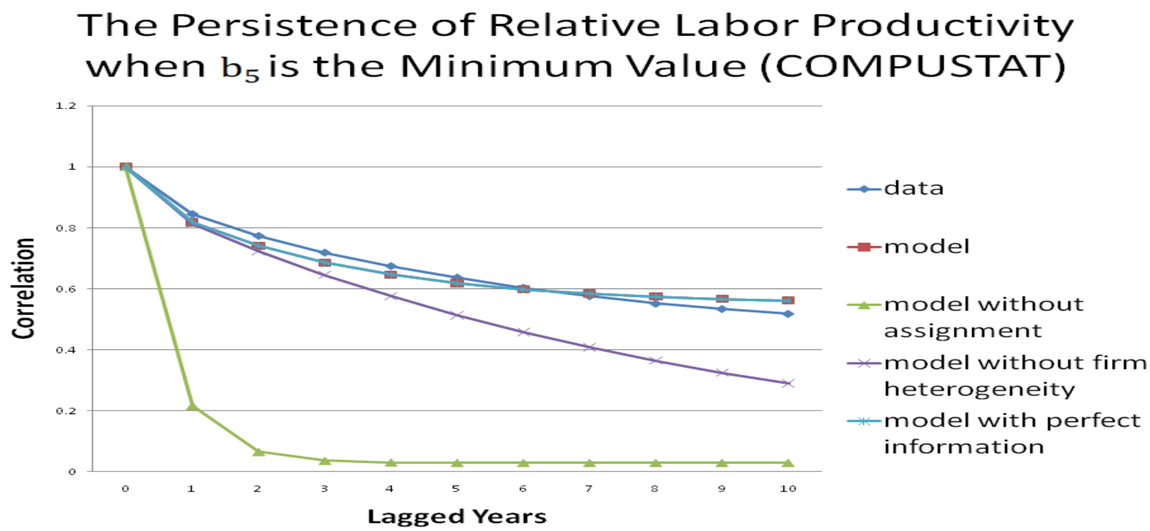
	ϕ	$\frac{\gamma\sigma_q}{\sigma_{\mu\infty}}$	h_∞	σ_F^2	σ_ε^2	$\frac{\psi\sigma_q}{\sigma_{\mu\infty}}$
Benchmark b_5	0.220	0.493	0.798	0.002	0.096	2.025
$\min b_5$	0.193	0.520	1	0.003	0.121	1.757
$\max b_5$	0.712	0.001	0.121	0.001	0.008	5.146

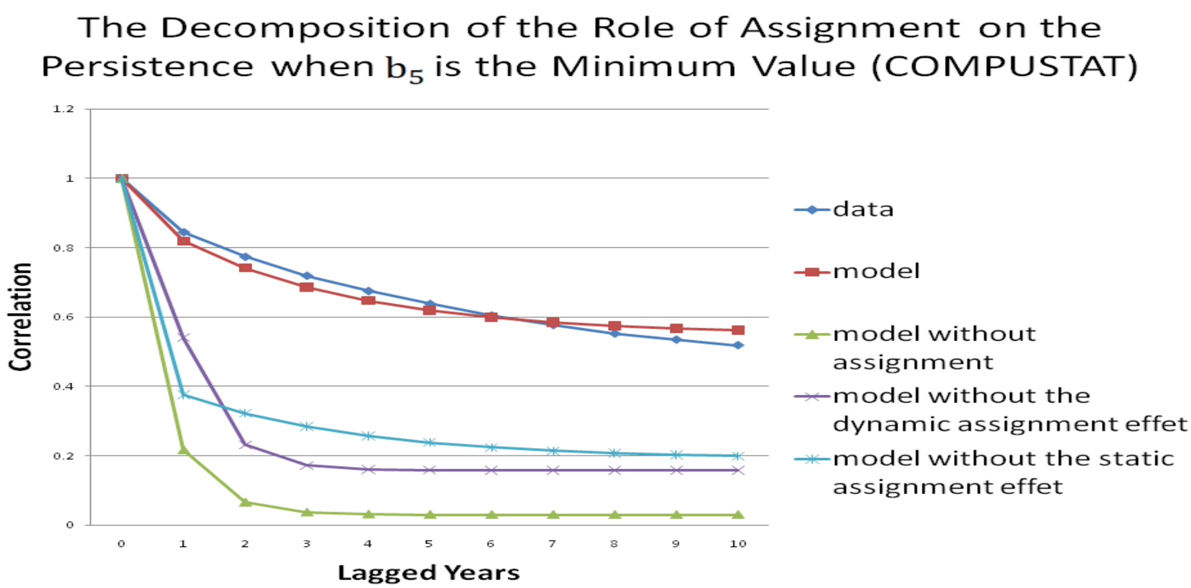
The Decomposition of Persistence and Variance using COMPUSTAT

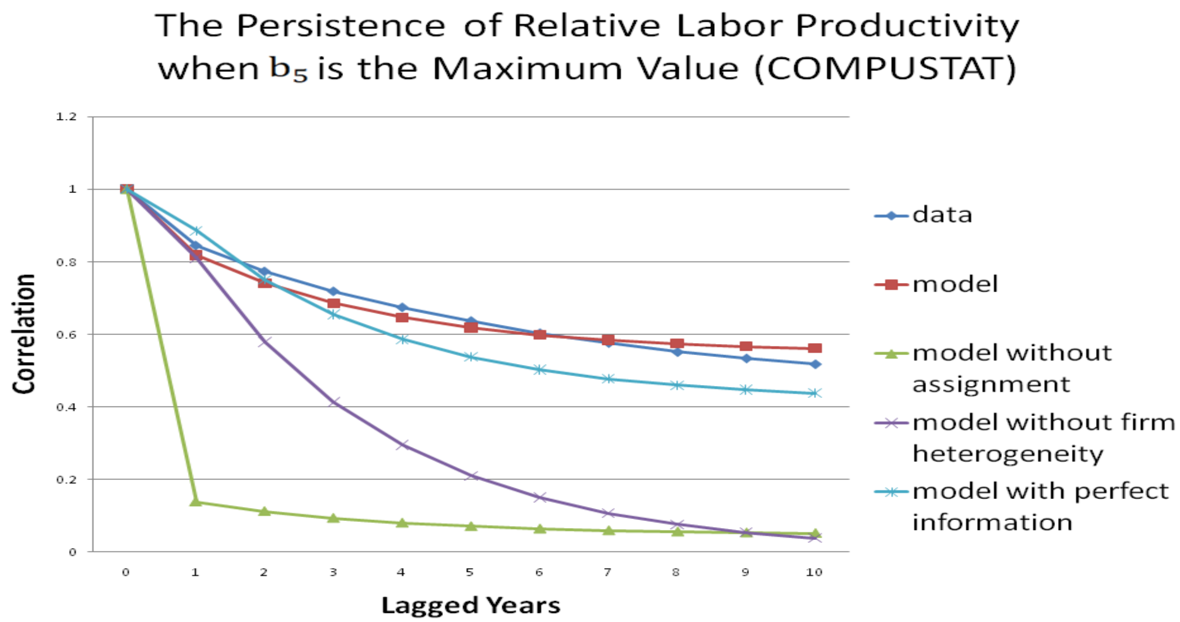
		$\lambda_{1\infty}$	z	$\frac{Var\left[F_A^*\right]}{Var[D\ln A_t]}$	$Var\left[D\ln A_t\right]$	$Var\left(F_A^*\right)$	$Var\left[D\ln A_t^F\right]$
	Benchmark	0.713	0.597	0.549	0.423	0.232	0.191
Benchmark b_5	$\frac{\sigma_q}{\sigma_\varepsilon}=0$	0.220	0.177	0.027	0.128	0.003	0.125
	$\gamma=0$	0.220	0.460	0.167	0.188	0.031	0.157
	$\psi=0$	0.713	0.207	0.165	0.154	0.025	0.129
	$\frac{\sigma_F}{\sigma_\varepsilon}=0$	0.889	0.814	0	0.373	0	0.373
	$\frac{\sigma_u}{\sigma_\varepsilon}=0$	0.708	0.682	0.551	0.400	0.220	0.179
$\min b_5$	$\frac{\sigma_q}{\sigma_\varepsilon}=0$	0.193	0.193	0.030	0.130	0.004	0.126
	$\gamma=0$	0.193	0.454	0.158	0.186	0.029	0.157
	$\psi=0$	0.713	0.230	0.190	0.161	0.031	0.130
	$\frac{\sigma_F}{\sigma_\varepsilon}=0$	0.892	0.812	0	0.350	0	0.350
	$\frac{\sigma_u}{\sigma_\varepsilon}=0$	0.713	0.597	0.549	0.423	0.232	0.191
$\max b_5$	$\frac{\sigma_q}{\sigma_\varepsilon}=0$	0.712	0.096	0.047	0.129	0.006	0.123
	$\gamma=0$	0.712	0.597	0.548	0.421	0.231	0.190
	$\psi=0$	0.713	0.096	0.048	0.129	0.006	0.123
	$\frac{\sigma_F}{\sigma_\varepsilon}=0$	0.714	0.810	0	0.374	0	0.374
	$\frac{\sigma_u}{\sigma_\varepsilon}=0$	0.713	0.805	0.416	0.558	0.232	0.326

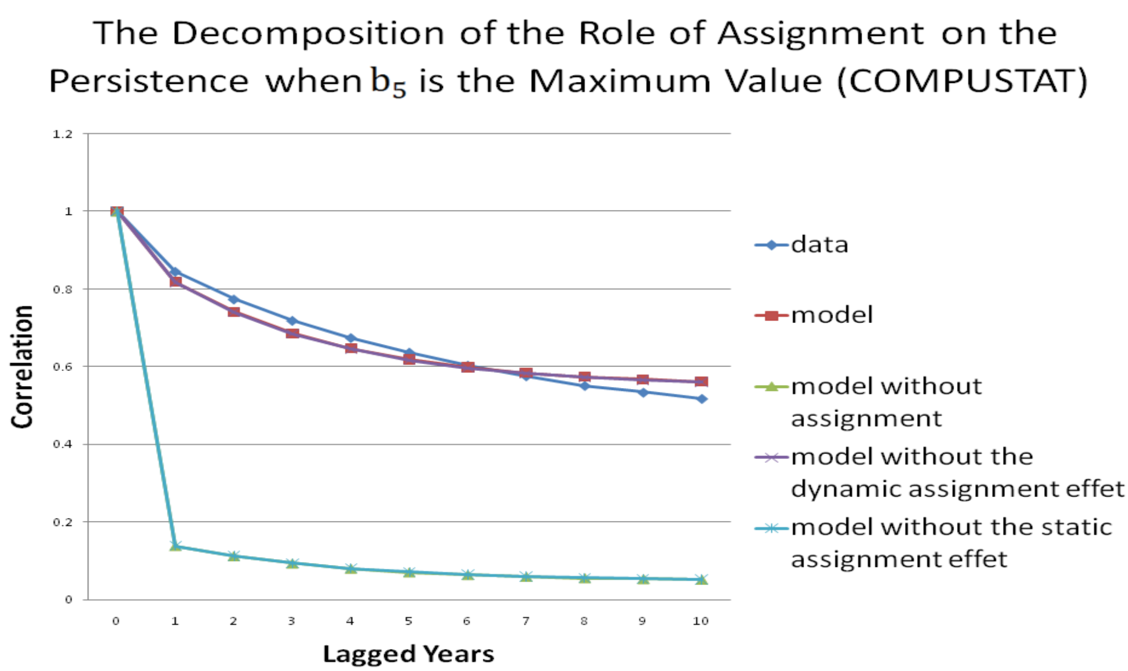

The Persistence of Relative Labor Productivity Between Firms (COMPUSTAT)
Correlation
Lagged Years
1.2
1
0.8
0.6
0.4
0.2
0
0
1
2
3
4
5
6
7
8
9
10
data
model
model without assignment
model without firm heterogeneity
model with perfect information


The Decomposition of the Role of Assignment on the Persistence (COPMUSTAT)
Correlation
Lagged Years
1.2
1
0.8
0.6
0.4
0.2
0
0
1
2
3
4
5
6
7
8
9
10
data
model
model without assignment
model without the dynamic assignment effet
model without the static assignment effet


The Persistence of Relative Labor Productivity when b_5 is the Minimum Value (COMPUSTAT)
Correlation
Lagged Years
data
model
model without assignment
model without firm heterogeneity
model with perfect information


The Decomposition of the Role of Assignment on the Persistence when b_5 is the Minimum Value (COMPUSTAT)
Correlation
Lagged Years
data
model
model without assignment
model without the dynamic assignment effet
model without the static assignment effet


The Persistence of Relative Labor Productivity when b_5 is the Maximum Value (COMPUSTAT)
Correlation
Lagged Years
1.2
1
0.8
0.6
0.4
0.2
0
0
1
2
3
4
5
6
7
8
9
10
data
model
model without assignment
model without firm heterogeneity
model with perfect information


The Decomposition of the Role of Assignment on the Persistence when b_5 is the Maximum Value (COMPUSTAT)
Correlation
Lagged Years
1.2
1
0.8
0.6
0.4
0.2
0
0
1
2
3
4
5
6
7
8
9
10
data
model
model without assignment
model without the dynamic assignment effet
model without the static assignment effet