Post-Match Investment and Dynamic Sorting between Capital and Labor*

Shouyong Shi
Pennsylvania State University
(sus67@psu.edu)

This version: August 2018

Abstract

By integrating frictional matching into a neoclassical framework of investment by firms, this paper analyzes dynamic sorting between capital and worker skills in the constrained social optimum. With post-match investment, strong complementarity between capital and skills in production induces the socially efficient pattern of sorting to be negative at the time of a match, but positive eventually. Because the sorting pattern in a match reverses over time, the time profile of labor productivity is steeper for a high skill than for a low skill, and the difference in productivity between skills increases over time. Among any cohort of workers who become employed at the same time, the variance in labor productivity increases over time. There is also dispersion in labor productivity within each skill. The calibrated model shows that sorting is positive on average and that dispersion in labor productivity is significant both between skills and within each skill.

JEL classification: E22; J60; C78.
Keywords: Dynamic sorting; Investment; Inequality.

* Address: Department of Economics, Pennsylvania State University, 502 Kern Building, University Park, PA 16802, USA.

1. Introduction

A cornerstone of the neoclassical theory is a production function in which capital and labor are complementary with each other. With such complementarity, social efficiency requires more capital to be allocated to workers with higher skills. In the terminology of matching or assignment, capital and worker skills should have positive assortative matching (PAM) with each other. This is the result when there are no matching frictions (Becker, 1973). In the presence of matching frictions, neither is PAM necessarily an equilibrium outcome (Shimer and Smith, 2000), nor does PAM necessarily maximize social welfare (Shi, 2001, 2005). Because matching frictions are significant in the labor market, as evidenced by the existence of persistent unemployment, the failure of PAM casts serious doubt on the neoclassical theory of production.[1] The objective of this paper is to examine how post-match investment by firms can correct the failure of PAM between capital and labor. I formulate the dynamic sorting problem and analyze the social optimum constrained by matching frictions. Moreover, I calibrate the model to analyze the quantitative implications of dynamic sorting on the time profile of worker productivity and inequality.

The potential inefficiency of PAM in the presence of matching frictions arises from a tradeoff between the utilization and the productivity of workers. PAM increases productivity, but reduces the utilization of high skills. The reason is that the vacancy cost increases in the capital stock, as a firm has to put some of the capital in place before hiring a worker. When each vacancy has a higher capital stock, only a smaller number of vacancies can be created. If the vacancy cost increases sufficiently in the capital stock, PAM can imply a lower matching rate for high-skill workers. In this case, negative assortative matching (NAM) can increase social welfare by creating a large number of vacancies, each with a low capital stock, to match with high-skill workers.

This intuitive explanation applies to the sorting pattern at the time of a match. If the firm type is permanent or a fixed effect, as assumed in the literature, then this sorting

[1] A growing empirical literature has been investigating the extent of PAM in the data. While some have found little evidence for PAM between workers and firms (e.g., Abowd et al., 1999), others have found strong evidence for PAM (e.g., Hagedorn et al., 2017).

pattern will be preserved in the entire duration of a match.[2] Although some dimensions of the firm type can be permanent, the capital stock a firm allocates to a worker is unlikely to be one of them. Hiring a worker is only the beginning of the production process. After the hiring, a firm can invest to increase the capital stock to change the sorting pattern. The average sorting pattern over time may be opposite to the assignment at the time of a match. Therefore, to understand dynamic sorting between capital and labor, it is necessary to integrate post-match investment by firms into a model of frictional matching. Moreover, such an integrated model can be useful for understanding heterogeneity in the earnings profile between skills and inequality within each skill.

Section 2 describes such an integrated model. Workers are heterogeneous in a permanent and observable type, called the skill. All firms have the same production technology to combine a worker with capital to produce output. Capital and skills are complementary in production. Productivity of a worker is measured by *net output*, defined as output in a match minus the rental cost of capital. The rental market for capital is competitive and frictionless, but the labor market has matching frictions. Firms can create vacancies with different capital stocks to target different skills, i.e., to direct workers' search. The vacancy cost increases in the initial capital stock for the job. Once a vacancy is filled, the firm chooses investment to adjust the capital stock. The adjustment cost is increasing and convex in investment, as in the neoclassical theory (e.g., Lucas, 1967, Gould, 1968). An employed worker separates from a match into unemployment at an exogenous rate. To focus on dynamic sorting inside a firm, I abstract from on-the-job search that generates dynamic sorting of workers across firms.[3]

In section 3, I characterize the socially efficient allocation chosen by the planner constrained by the same matching frictions as in the market. The social welfare function is the sum of expected surpluses in the economy. For unemployed workers, the planner chooses the number of vacancies and the capital stock for each vacancy to target each skill. The

[2]For a survey on sorting models, see Chade et al. (2017). For a list of papers on identifying sorting in the data, see Hagedorn et al. (2017).

[3]In another paper (Shi, 2018), I have incorporated on-the-job search into a model with post-match investment, but abstracted from worker heterogeneity.

choice of the number of vacancies is equivalent to the choice of a matching rate for each skill level of unemployed workers. The capital stock chosen for a vacancy determines the *initial assignment* of capital to the skill. For each employed worker, the planner chooses investment and, hence, the dynamic path of the capital stock in the match.

Post-match investment affects the assignment in all stages, including the initial assignment. Specifically, post-match investment increases the importance of the matching rate relative to PAM in the initial assignment. When investment is possible, the social optimum can increase the capital stock for a high skill in the future to seize the benefit of the complementarity between capital and the skill. This option reduces the need for PAM at the time of hiring. The higher is investment expected to be, the larger is the gain from employing a high skill quickly, and the lower is the initial capital stock assigned to match with a high skill. Of course, investment is endogenous. Post-match investment is likely to be higher if capital and the skill are more complementary with each other. Thus, strong complementarity increases the likelihood for the initial assignment to be NAM, which is opposite to the result in sorting models without post-match investment.

Post-match investment follows dynamics described in a macroeconomic textbook. Given the worker skill in the match and the initial assignment, there is a unique stable saddle path of the capital stock that converges asymptotically to the *final assignment*. The final assignment is PAM, because it equates the marginal productivity of capital to the rental rate of capital. Convergence to the final assignment is slower if the marginal cost of adjustment increases more sharply in investment or if the marginal productivity of capital diminishes less sharply in capital.

The social optimum yields interesting predictions on the time profile and the distribution of labor productivity. Suppose that capital and the skill are strongly complementary so that the initial assignment is NAM. Relative to a low skill, a high skill has a lower capital stock initially. Because the final assignment is PAM, the path of the capital stock for a high skill will eventually surpass that for a low skill. Thus, the time profile of net output is steeper for a high skill than for a low skill. The assignment of capital to labor reduces dispersion in net output between skills initially, but amplifies such dispersion eventually. There is also dispersion in net output within each skill. For the same skill, the longer has a worker been employed, the higher is the capital stock, and the higher is net output of the worker. The extent of sorting between capital and skill depends on this distribution. I examine two notions of PAM: positive assortative matching according to the distribution (PAM-D) and according to the mean of capital stocks (PAM-M) assigned to each skill. Section 4 provides the conditions for each notion of PAM to hold.

To analyze the quantitative implications of the model, I calibrate the model in section 5. With the calibrated parameters, the initial assignment has NAM, and the matching rate increases in the skill. The time paths of capital stocks assigned to different skills cross each other over time quickly, in less than 9 months. Despite NAM at the beginning of a match, net output is higher for a high skill than for a low skill throughout the match. Moreover, the difference in net output between skills increases over time as a result of investment. Thus, for any cohort of workers who become employed at the same time, the variance in net output increases over time. The result that a higher skill has a steeper profile of net output over time is consistent with the empirical evidence that the heterogeneous slope of the earnings profile is an important source of earnings inequality over the life cycle (see Guvenen, 2007). The fanning-out of net output over time between skills is consistent with the evidence that the variance in earnings among workers increases over the life cycle (e.g., Deaton and Paxson, 1994, Guvenen, 2007).

In the calibrated model, the mean of the capital stock increases in the skill, although the initial assignment is NAM. This result of PAM-M illustrates concretely that ignoring post-match investment or treating the firm type as a fixed effect leads to a wrong conclusion about the sorting pattern in a dynamic economy. In this regard, it is useful to note that the mean of the capital stock for each skill is substantially lower than the level in the final assignment, because exogenous separation terminates a match before the capital stock gets close to the final level. This means that the model's behavior is quantitatively different from that of a neoclassical model without matching frictions. Nevertheless, convergence toward the final state predicted by the neoclassical model exerts a strong force on the mean of the capital stock assigned to each skill.

The model generates significant inequality in net output within each skill. Weighted by the employment share of each skill, the average of the coefficient of variation in net output within a skill is 0.226. The mean-min ratio, a measure of inequality proposed by Hornstein et al. (2011), is 1.769 in net output within a skill. Such within-group inequality is comparable to the one observed in the U.S. data but much higher than in most search models (see Hornstein et al., 2011).

There is a large literature on sorting originated in Becker (1973). As mentioned earlier in this Introduction, this literature treats the firm type as a fixed effect and focuses on sorting at the time of a match. Also, a small number of recent papers (e.g., Lise and Robin, 2017) have started to examine dynamic sorting of workers *across firms* through on-the-job search. They, too, assume that the firm type is a fixed effect. In contrast, in this paper, the capital stock is an endogenous type of a firm. The main contribution of this paper is to integrate frictional sorting into the neoclassical framework with investment by firms to examine how post-match investment affects dynamic sorting between capital and the worker skill. As summarized above, this integration yields novel results on the pattern of dynamic sorting, the time profile of labor productivity, and dispersion in labor productivity both between skills and within each skill.

In the literature, some papers have examined the effect of pre-match investment on sorting, e.g., Burdett and Coles (2001), Peters and Siow (2002). Pre-match investment is intended to increase the probability of matching with the desired type. As such, pre-match investment tends to increase the likelihood of PAM at the time of a match. In contrast, post-match investment is intended to increase productivity within a match, and it tends to make the initial assignment NAM. Moreover, the literature on pre-match investment focuses on the externalities that may or may not be internalized by the equilibrium. Instead, I analyze the constrained social optimum.[4]

[4]This paper is related to the literature on directed search, e.g., Peters (1991), Montgomery (1991), Moen (1997), Julien et al. (2000), and Burdett, Shi, Wright (2001). For the use of directed search in sorting models, see Shi (2001, 2002, 2005) and Eeckhout and Kircher (2010).

2. The Model

2.1. Model Environment

Time is continuous. There is a unit measure of workers whose time discount rate is $r \in (0, \infty)$. Workers differ in an exogenous type h, which lies in a compact set $\mathcal{H}$ with $\min \mathcal{H} = h_L > 0$ and $\max \mathcal{H} = h_H < \infty$. The cumulative distribution function of h is H. The measure of firms is determined by competitive entry of vacancies. Firms are homogeneous ex ante but become heterogeneous ex post in the capital stock, k. A firm needs one worker, as in most models of labor market search. I refer to h as the worker skill, to k as a firm's (endogenous) type, and to a match between a job with the capital stock k and a skill-h worker as a (k, h) match. In a (k, h) match, the output flow is $\tilde{f}(k, h)$. The output function $\tilde{f}$ is twice differentiable, strictly increasing and concave in k and h, with $\tilde{f}(0, h) = f(k, 0) = 0$, $\tilde{f}_1(0, h) = \tilde{f}_2(k, 0) = \infty$, and $\tilde{f}_1(\infty, h) = \tilde{f}_2(k, \infty) = 0$, where the subscripts indicate partial derivatives. In addition, assume that k and h are *strictly complementary* with each other in $\tilde{f}$; that is, $\tilde{f}_{12}(k, h) > 0$.

A firm rents capital from the rest of the world at a constant rental rate r. Net output in a (k, h) match is defined as $f(k, h) = \tilde{f}(k, h) - rk$, and is referred to as labor productivity in the match. For each h, assume that there is a unique *final stock* of capital, $k^*(h) \in (0, \infty)$, such that $f_1(k^*, h) = 0$ and $f(k^*, h) > 0$. Denote $f^*(h) = f(k^*(h), h)$. The final stock is the optimal capital stock in the neoclassical theory with a frictionless labor market. For an unemployed worker, home production is $f_u \geq 0$, which is constant over time.[5]

A firm chooses the capital stock dynamically. At the time of creating a job, a firm can choose any capital stock k for a vacancy but must rent capital in advance for the vacancy. The flow cost of a vacancy, denoted as $\psi(k)$, includes both the rental cost of capital and other costs of maintaining the vacancy. The vacancy cost satisfies $\psi'(k) > 0$ and $\psi''(k) > 0$ for all $k > 0$, and $\psi(0) = 0$. The assumption $\psi'(k) > 0$ is easily satisfied since the vacancy cost includes the rental cost of capital, which increases linearly in k. The assumption $\psi''(k) > 0$ is critical for the analysis. It is justified by the intuitive reasoning

[5]The qualitative results are similar if home production depends on h, provided that $f_2(k, h) - f_u'(h) > 0$.

that a firm is unlikely to give a vacancy the same amount of capital as the long-run level of capital for a filled job. Instead, a firm is likely to fill a vacancy at a relatively low capital stock and then invest to build up the capital stock. The assumption $\psi''(k) > 0$ delivers this intuitive outcome naturally.

Once a vacancy is filled, the match generates output. A firm can adjust the capital stock by undertaking investment, denoted as i. The capital stock evolves according to

$$\frac{\mathrm{d}k(t)}{\mathrm{d}t} = i. \tag{2.1}$$

In addition to renting the required addition of capital, a firm must incur an adjustment cost for investment. In terms of the final good, the adjustment cost is $c(i)$, which satisfies $c(\infty) = \infty$, $c(i) = c'(i) = 0$ for all $i \leq 0$, $c'(i) > 0$ for all $i > 0$, and $c''(i) > 0$ for all $i \geq 0$. These assumptions are standard in the neoclassical theory of investment, except $c(i) = c'(i) = 0$ for all $i < 0$. These exceptions are imposed to focus on capital accumulation instead of decumulation, because they imply that reducing the capital stock bears no adjustment cost.[6] With the adjustment cost, a firm can adjust the capital stock of an existing job only smoothly, not because additional capital is not available but because it is costly to put new capital into an existing job. This process contrasts to vacancy creation, where the capital stock can be changed to any amount instantaneously.

Investment is perfectly reversible, because a firm can reduce the capital stock to 0 at no cost and return all capital to the owner at the full value. However, investment is only partially transferable between jobs. To move capital from one job to another filled job, a firm must incur the adjustment cost at the destination. To move capital from one job to a vacancy, a firm must incur the vacancy cost. In both cases, moving capital between two jobs directly is equivalent to moving capital indirectly through the lender of capital.[7]

Unemployed workers search for jobs. Firms can direct workers' search by offering a specific capital stock and matching rate to searchers of each skill. The searchers of the same skill and the vacancies targeting them form a submarket. In any submarket where

[6]The assumption is not activated in the social optimum, where the capital stock increases over time.

[7]Under such partial transferability of capital between jobs, a firm can be interpreted alternatively as a collection of independent jobs, as in most search models of labor.

the matching rate for a searcher is p, the number of vacancies per searcher is $\theta(p)$, and the matching rate for a vacancy is $q(p) = \frac{p}{\theta(p)}$. The function $\theta(p)$ summarizes matching frictions and is implied by a matching function. Precisely, if the matching rate function for a searcher is $P(\theta)$, then $\theta(p)$ is the inverse function of P. Assume:[8]

$$\theta'(p) > \frac{\theta(p)}{p} > 0 \text{ and } \theta''(p) > 0 \text{ for all } p > 0,$$
$$\theta(0) = 0,\ \theta'(0) \in (0, \infty),\ \text{and } \lim_{p\to\infty} \frac{\theta(p)}{p} = \infty.$$

Denote $\varepsilon = \frac{p\theta'(p)}{\theta(p)}$ (> 1).

A match separates exogenously at a rate $\delta \in (0, \infty)$, which makes the worker unemployed. A firm can recover the capital at a vacant job fully and return it to the owner.

2.2. Planner's Problem

The social planner chooses job search for unemployed workers and post-match investment for firms to maximize the sum of social values, subject to search frictions and the process (2.1). To describe the planner's problem, let $V(k, h)$ be the social value of a (k, h) match, and $V_u(h)$ the social value of an unemployed worker of skill h. Focus on the steady state where the value functions change over time only when k changes. Because a worker's skill and home production are constant over time, $V_u(h)$ is constant over time.

For unemployed workers, the planner directs each skill to search in a particular submarket where the vacancies have a particular capital stock. Consider the submarket for skill-h searchers. Let ϕ be the capital stock for each vacancy and p the the matching rate for a searcher. Given the matching function, the choice of p translates into a number of vacancies per searcher, $\theta(p)$. Each vacancy incurs a flow cost $\psi(\phi)$. Because a searcher of skill h finds a match at the rate p and each match (ϕ, h) yields the social value $V(\phi, h)$, the expected surplus of the match is $p[V(\phi, h) - V_u(h)]$. Deducting the vacancy cost from this expected surplus yields the expected social surplus of search. The socially effect choices (p, ϕ) maximize the expected social surplus of search and yields the following Bellman

[8] Since $P(\theta(p)) = p$ for all p, the assumptions on $\theta(p)$ are equivalent to the following assumptions on $P(\theta)$: $0 < P'(\theta) < \frac{P(\theta)}{\theta}$ and $P''(\theta) < 0$ for all $\theta > 0$, $P(0) = 0$, $P'(0) \in (0, \infty)$, and $\lim_{\theta\to\infty} \frac{P(\theta)}{\theta} = 0$.

equation for $V_u(h)$:

$$rV_u(h) = f_u + \max_{(p,\phi)} \{p[V(\phi, h) - V_u(h)] - \psi(\phi)\theta(p)\}. \tag{2.2}$$

The optimal choices are given by the policy functions, $(p(h), \phi(h))$. The function ϕ describes the initial assignment of capital to skills at the time of a match.

In a new match between a firm and a skill-h worker, the capital stock is $k(0) = \phi(h)$. Suppose that the match has lasted for a length of time τ, resulting in a capital stock $k(\tau)$. At τ, the planner chooses the path of investment, $\{i(t)\}_{t \geq \tau}$, to maximize the social value of the match. At any $t \geq \tau$, the benefit of investment $i(t)$ is the increase in the social value $V(k(t), h)$ resulting from a higher capital stock. The cost of investment is $c(i(t))$. Thus, for any $t \geq \tau$, investment maximizes the return on investment and yields the following Bellman equation for $V(k(t), h)$:

$$(r + \delta)V(k(t), h) = f(k(t), h) + \delta V_u(h) + \max_{i(t)} \left[\frac{\mathrm{d}V(k(t), h)}{\mathrm{d}t} - c(i(t))\right], \tag{2.3}$$

where the constraint in the maximization problem is (2.1). The effective discount rate on a match is $(r + \delta)$. The social return to the match, given by the right-hand side of (2.3), is the sum of net output in the match, the expected value of exogenous separation, and the maximized social return on investment. It is useful to express the planner's choice of investment as a choice of the sequence $\{i(t)\}_{t \geq \tau}$. Integrating (2.3) over time generates:

$$V(k(\tau), h) = \max_{\{i(t)\}_{t \geq \tau}} \int_\tau^\infty [f(k(t), h) + \delta V_u(h) - c(i(t))] e^{-(r+\delta)(t-\tau)}\mathrm{d}t, \tag{2.4}$$

subject to (2.1) for all $t \geq \tau$. The initial capital stock at τ is given as $k(\tau)$.[9]

There is a non-degenerate distribution of capital stocks among workers of the same skill and across different skill levels. This distribution evolves endogenously and will be analyzed in section 4. However, the value function and the efficient allocation are independent of this distribution, as formulated above. That is, the allocation is block recursive, as defined by Shi (2009) and Menzio and Shi (2010). Directed search is critical for block recursivity, because it enables the planner to separate different skills to search in different submarkets.

[9]The above step assumes the transversality condition: $\lim_{t\to\infty} V(k(t), h)\, e^{-(r+\delta)t} = 0$.

3. Socially Efficient Search and Investment Dynamics

I analyze the investment problem in (2.4) first and then the search problem in (2.2). For any given h, I restrict the domain of k to $[0, k^*(h)]$ without loss of generality.[10]

3.1. Efficient Investment and Sorting

Let $\lambda(t)$ be the current-value multiplier of (2.1) in the maximization problem in (2.4). The Hamiltonian is: $f(k,h) + \delta V_u(h) - c(i) + \lambda i$. The optimality conditions are:

$$\text{for } i(t) \geq 0\text{: } c'(i(t)) = \lambda(t) = V_1(k(t), h). \tag{3.1}$$

$$\text{for } k(t)\text{:} \quad \frac{\mathrm{d}\lambda(t)}{\mathrm{d}t} = (r+\delta)\lambda(t) - f_1(k(t), h). \tag{3.2}$$

These conditions are intuitive. If the optimal investment is non-negative, then it equates the marginal cost of investment to the shadow value of capital which, in turn, is equal to the marginal social value of capital. This is (3.1).[11] Condition (3.2) requires that the "permanent" income of a marginal unit of capital, $(r+\delta)\lambda$, should be equal to the sum of the marginal productivity of capital and the appreciation in the value of capital.

From (3.1) and (3.2), I obtain:

$$\frac{\mathrm{d}i(t)}{\mathrm{d}t} = \frac{1}{c''}\left[(r+\delta)c'(i) - f_1(k,h)\right], \text{ if } i(t) \geq 0. \tag{3.3}$$

The dynamics of investment are driven by the difference between the marginal cost and the marginal benefit of investment. The marginal benefit is equal to the present value of the marginal productivity of capital, $\frac{f_1}{r+\delta}$. If the marginal benefit of investment exceeds the marginal cost, efficient investment must be high, and so the capital stock must be increasing over time. Because the increase in the capital stock reduces the marginal productivity of capital, this implies that efficient investment is declining over time. Conversely, if the

[10] For all $k > k^*(h)$, net output is decreasing in k. Thus, it is never socially efficient to have $k > k^*(h)$.

[11] The same procedure as in Appendix A in Shi (2018) proves that $V_1(k,h)$ exists. The result $V_1(k,h) = c'(i)$ for $i \geq 0$ can be derived alternatively for the maximization problem in (2.3). Also, if $i(t) < 0$ is optimal, then $k > k^*(h)$. In this case, the optimal investment is $i(t) = -\infty$, which allows the capital stock to fall by a discrete amount to reach $k^*(h)$ immediately.

marginal benefit of investment is lower than the marginal cost, then efficient investment must be low and, hence, must be increasing over time.

Equations (3.3) and (2.1) form the dynamic system of (i, k), with the initial condition $k(0) = \phi(h)$. The *final state* of the system is (i^*, k^*) such that $i^* = 0$ and $\frac{\mathrm{d}i}{\mathrm{d}t} = 0$. Then, $f_1(k^*, h) = c'(0) = 0$, which is consistent with the early definition of k^*. To study the dynamics of sorting, I linearize the dynamic system around the final state:

$$\begin{bmatrix} \mathrm{d}i/\mathrm{d}t \\ \mathrm{d}k/\mathrm{d}t \end{bmatrix} = \begin{bmatrix} r+\delta, & -\frac{f_{11}}{c''} \\ 1, & 0 \end{bmatrix} \begin{bmatrix} i \\ k-k^* \end{bmatrix}. \tag{3.4}$$

Let J denote the (Jacobian) coefficient matrix in (3.4). The elements in J are evaluated in the final state. The matrix J has one stable (negative) eigenvalue and one unstable (positive) eigenvalue. The stable eigenvalue is $-\beta$, where

$$\beta(h) \equiv \frac{1}{2}\left\{\left[(r+\delta)^2 - \frac{4f_{11}(k^*(h), h)}{c''(0)}\right]^{1/2} - (r+\delta)\right\} > 0. \tag{3.5}$$

The following proposition describes the sorting pattern in the final state and the stable saddle path converging to the final state (see Appendix A for a proof):

Proposition 3.1. *In the final state, the matching between the worker skill and the capital stock is positive assortative; i.e., $k^{*\prime}(h) > 0$. For any given h and the initial condition, $k(0) = \phi(h)$, the unique stable saddle path of (3.4) is*

$$\begin{bmatrix} i(t) \\ k(t) - k^*(h) \end{bmatrix} = \begin{bmatrix} \beta(h) \\ -1 \end{bmatrix} [k^*(h) - \phi(h)]\, e^{-\beta(h)t}. \tag{3.6}$$

This implies:

$$i(t) = I(k(t), h) \equiv \beta(h)[k^*(h) - k(t)]. \tag{3.7}$$

Thus, $i(t) > 0$ and $\frac{di(t)}{dt} < 0$ for all $t \in (0, \infty)$ if $\phi(h) < k^*(h)$. The speed of convergence of the capital stock to the final state, measured by β, is lower if $c''(0)$ is higher or if $-f_{11}$ is lower. Moreover, $\beta'(h) < 0$ if and only if $-f_{11}(k^*(h), h)$ is decreasing in h.

In the final state, matching between capital and the worker skill is positive assortative (PAM). This is because the capital stock in the final state equates the marginal productivity of capital to the rental rate. With complementarity between the two factors, the marginal productivity of capital is an increasing function of the worker skill. Thus, in the final state, a higher worker skill will work with a higher capital stock.

Post-match investment is important for the final assignment to be positive assortative. If post-match investment were not possible, search frictions would make the assignment "stuck" at the initial assignment. As shown later, the initial assignment can be non-PAM or even NAM, in contrast to the assignment under frictionless matching (e.g., Becker, 1973). Even if the initial assignment is NAM, post-match investment restores the efficient assignment to be PAM in the long run.

Investment is positive and declines over the duration of a match, provided that the initial capital stock is below the final stock. It is socially efficient to front-load investment. The earlier is investment undertaken, the longer is the time in which the match yields higher output from the increased capital stock.[12] Investment is inversely related to the gap between the current capital stock and the final stock, as shown by the stable saddle path (3.7). As investment increases the capital stock toward the final stock asymptotically, investment declines toward zero. The convergence speed, β, depends on the convexity of the adjustment cost function and the concavity of the output function in the capital stock. If the adjustment cost is more convex, the marginal cost of adjustment increases more quickly with investment. If the output function is less concave in the capital stock, the marginal productivity of capital diminishes less quickly. In either case, the capital stock increases more slowly toward the final stock.

The skill of the worker in a match affects $-f_{11}\left(k^*\left(h\right), h\right)$ directly, and indirectly through the final stock. In general, it is unclear whether a higher worker skill has faster or slower convergence to the final state. However, the following example shows that for a well-known class of output functions, the convergence is slower for a higher skill.

Example 3.2. *Consider the CES output function:* $\tilde{f}\left(k, h\right) = A_0\left[\alpha k^{\xi} + \left(1-\alpha\right)h^{\xi}\right]^{1/\xi}$ *for* $\alpha \in (0,1)$, $\xi \in (-\infty, 1)$ *and* $A_0 \in (0, \infty)$. *Then, there exists a constant* $A_1 > 0$ *such that*

[12] If on-the-job search is introduced, investment can be partially delayed (see Shi, 2018).

$-f_{11}(k^*(h), h) = A_1/h$, *which is a decreasing function of* h.

It is important to clarify that slower convergence does not imply lower investment in a match with a higher h. As shown by (3.7), investment depends on both $\beta(h)$ and the gap $[k^*(h) - k(t)]$. Relative to a low-skill worker, a high-skill worker's capital stock in the final state is higher. For any given current stock, the gap from the final stock is larger, and the gain from investment is greater. Thus, even if a high-skill worker has a lower speed of convergence, investment is higher for such a worker than for a low-skill worker. I will establish this result in Proposition 3.3 later.

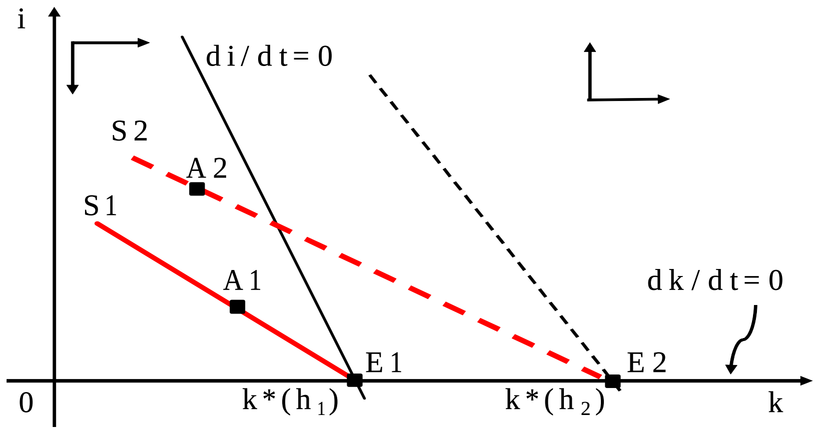


Figure 1. Dynamics of investment and the capital stock

Figure 1 is the phase diagram of the dynamic system (3.4) for two worker types, h_1 and h_2, with $h_2 > h_1$. The horizontal axis is the capital stock, which also represents all combinations of (i, k) that yield the schedule $\frac{\mathrm{d}k}{\mathrm{d}t} = 0$. The schedule $\frac{\mathrm{d}i}{\mathrm{d}t} = 0$ is negatively sloped. The intersection between the two schedules is the final state, which is marked as point E1 for skill h_1 and E2 for skill h_2. The stable saddle path, (3.7), is the line S1-E1 for skill h_1 and the line S2-E2 for skill h_2.[13] The depicted case has $\beta(h_1) > \beta(h_2)$, because the stable path is steeper for the lower skill h_1 than for the higher skill h_2. For the dynamic system to converge to the final state asymptotically, the initial state must be on the stable saddle path. In Figure 1, the initial state is at point A1 for skill h_1 and A2 for skill h_2. The initial capital stock is lower for skill h_2 than for skill h_1, which is possible as analyzed

[13] The phase diagram depicts only $i \geq 0$. Below the horizontal axis, $i = -\infty$. In this case, the capital stock falls in a discrete amount to reach the final state instantaneously.

in the next subsection. Investment is higher for h_2 than for skill h_1, as discussed above. From the initial state, investment declines and the capital stock increases toward the final state. Because investment declines over time, it is front-loaded.

3.2. Efficient Search and the Initial Assignment

For each skill h, the efficient choices of search, $(\phi(h), p(h))$, solve the problem in (2.2). The policy function $\phi(h)$ describes the initial assignment of capital to the worker skill. Denote the social gain from a (ϕ, h) match as $\Delta(\phi, h) = V(\phi, h) - V_u(h)$. The optimality conditions of the efficient choices of search are:

$$\text{for } \phi\text{: } V_1(\phi, h) = \psi'(\phi)\frac{\theta(p)}{p} \tag{3.8}$$

$$\text{for } p\text{: } \Delta(\phi, h) = \psi(\phi)\,\theta'(p). \tag{3.9}$$

The social marginal benefit of capital in the initial assignment is the expected marginal value of capital from a match, $pV_1(\phi, h)$. To generate the matching rate p, the measure of vacancies must be $\theta(p)$, and so the marginal cost of capital in these vacancies is $\psi'(\phi)\,\theta(p)$. The initial assignment of capital equates the social marginal benefit of capital to this marginal cost of vacancies, as in (3.8). In addition, the measure of vacancies targeting each skill must be efficient. By increasing the measure of vacancies marginally, the planner increases workers' matching rate. A match creates the social surplus, $\Delta(\phi, h)$. The marginal cost of increasing p is $\psi(\phi)\,\theta'(p)$. For each h, the efficient allocation equates the social surplus of a match to the marginal cost of increasing p, as shown by (3.9).

The initial assignment has PAM if and only if $\phi'(h) > 0$. To describe the properties of efficient search choices, recall that $I(k, h)$ is the investment function on the stable saddle path, given by (3.7), and $\varepsilon = \frac{p\theta'(p)}{\theta(p)} > 1$. Appendix A proves the following proposition:

Proposition 3.3. *For each given k, $V_2(k, h) > 0$ and $V_u'(h) > 0$. Approximate the derivatives (V_{11}, V_{12}) along the stable path (3.7). Then,*

$$\phi'(h) > 0 \quad \textit{iff} \quad \frac{(a_2 - I)\, f_{12}(\phi, h)}{r + \delta + \beta} > f_2(\phi, h) \tag{3.10}$$

$$p'(h) > 0 \quad \textit{iff} \quad \frac{(a_1 - I)\, f_{12}(\phi, h)}{r + \delta + \beta} < f_2(\phi, h)\,, \tag{3.11}$$

where $I = I(\phi, h)$, and $a_2 > a_1 > 0$ are defined by

$$a_1 = \frac{(r + \delta + p)(\varepsilon - 1)}{\psi''/\psi' + \beta c''(I)/c'(I)}, \qquad a_2 = \frac{(r + \delta + p)\, p\psi\theta''}{(\varepsilon - 1)\, c'(I)}. \tag{3.12}$$

Moreover, $I_2(k, h) > 0$ for all $k < k^*(h)$. If $\phi'(h) \leq 0$, then $p'(h) > 0$ and $\frac{dI(\phi(h),h)}{dh} > 0$.

In the social optimum, a higher skill increases the social value, both in employment and in unemployment. However, a higher skill is not always assigned to match with a higher capital stock. This result holds despite the maintained assumption that capital and the worker skill are complementary in production. To explain the result, note that there are two ways to capture the high social surplus from a more productive worker. One is PAM. The other way is to increase the utilization of the high skill by matching such a worker with a job more quickly. Increasing the matching rate requires an increase in the number of vacancies per searcher, which is optimal only if the capital stock for each vacancy is lower. The social optimum uses either PAM or a higher matching rate for a higher skill, or both. Given post-match investment, there can be three cases:

Case 1: $\phi' < 0$ and $p' > 0$. This case occurs when the two factors are weakly complementary with each other. In this case, the cost of creating high-capital vacancies for PAM exceeds the gain in productivity from PAM. It is socially efficient to make the initial assignment NAM so that many low-capital vacancies can be created for high-skill workers to increase the utilization of such workers.

Case 2: $\phi' > 0$ and $p' > 0$. This case occurs when the two factors are moderately complementary with each other. In this case, the social optimum uses both PAM and a higher matching rate for high-skill workers.

Case 3: $\phi' > 0$ and $p' < 0$. This case may occur when the two factors are strongly complementary with each other. In this case, it is socially efficient to capture such complementarity through strong PAM. Because vacancies with a higher capital stock are more costly to create, the number of vacancies for high-skill workers is small, and so the matching rate for these workers is lower than for low-skill workers.

In all cases, it is never socially efficient to assign both a lower capital stock and a lower matching rate to a higher skill. If a higher skill has a lower matching rate, then the initial assignment must be PAM. If the initial assignment fails to be PAM, then the social optimum matches a higher skill more quickly than a lower skill. In this case, investment immediately after a match is formed is also higher for a higher skill than for a lower skill. More generally, $I_2(k,h) > 0$ for all $k < k^*(h)$, as stated in Proposition 3.3. Thus, for any given capital stock, investment increases in the worker skill.[14]

Table 1. Features of the initial assignment and workers' matching rate

$I < a_1$:	ϕ': <0 >0 >0 |——— |——— |———> $\frac{f_{12}}{r+\delta+\beta}$ 0 (Case 1) $\frac{f_2}{a_2-I}$ (Case 2) $\frac{f_2}{a_1-I}$ (Case 3) p': >0 >0 <0
$a_1 \leq I < a_2$:	ϕ': <0 >0 |——— |———> $\frac{f_{12}}{r+\delta+\beta}$ 0 (Case 1) $\frac{f_2}{a_2-I}$ (Case 2) p': >0 >0
$I \geq a_2$:	ϕ': <0 |———> $\frac{f_{12}}{r+\delta+\beta}$ 0 (Case 1) p': >0

The above explanation for the possible failure of PAM in the initial assignment is similar to that in Shi (2001, 2005) and Eeckhout and Kircher (2010), who abstract from post-match investment. The key difference from those models is that post-match investment affects the threshold levels of complementarity that divide the three cases. To help explain this effect, Table 1 divides the efficient allocation into three regions of investment. It lists the initial assignment and the matching rate according to the measure of the complementarity between the two factors, $\frac{f_{12}(\phi,h)}{r+\delta+\beta}$. The two threshold levels are $\frac{f_2}{a_1-I}$ and $\frac{f_2}{a_2-I}$, where $a_2 > a_1 > 0$ are defined by (3.12). The top part of Table 1 is the region of low investment: $I < a_1$. For such investment, both thresholds of the complementarity are strictly positive, and all three cases explained above exist. In the middle part of Table 1, investment is moderate: $a_1 \leq I < a_2$. Because $a_1 - I \leq 0$, the threshold $\frac{f_2}{a_1-I}$ becomes inapplicable.

[14]To prove this result, I differentiate (3.1), instead of (3.7), with respect to h. (3.7) is an approximation. Differentiating (3.7) may yield an ambiguous result.

Only Cases 1 and 2 exist. In the bottom part of Table 1, investment is high. Because $a_1 - I < a_2 - I \leq 0$, both thresholds become inapplicable. Only Case 1 exists.

This analysis reveals that post-match investment increases the relative benefit of a higher matching rate to PAM in the initial assignment. By making a high skill employed more quickly, the social optimum can capture the benefit of the skill more quickly. This may be socially efficient even if it requires the initial assignment to be NAM, because post-match investment can restore the assignment to be PAM. In Table 1, there are more cases for $p'(h) > 0$ than for $\phi'(h) > 0$. Also, an increase in the marginal productivity of the skill, f_2, widens the region of NAM in the initial assignment in Table 1.

Moreover, strong complementarity between capital and skills may induce the initial assignment to be NAM, which is opposite to the outcome in sorting models without post-match investment. The reason is that investment is likely to be higher when capital and skills are more complementary with each other. With high investment to come after a match is formed, there is little benefit for high-skill workers to wait for a long time just for a PAM match. Instead, such workers should match quickly through NAM. This is the case $I \geq a_2$ in Table 1, where the only outcome is $\phi'(h) < 0$ and $p'(h) > 0$. In contrast, Case 3 discussed earlier requires strong complementarity but has the matching features $\phi'(h) > 0$ and $p'(h) < 0$. This case may not exist at all in the social optimum, because the matching features are inconsistent with high investment.

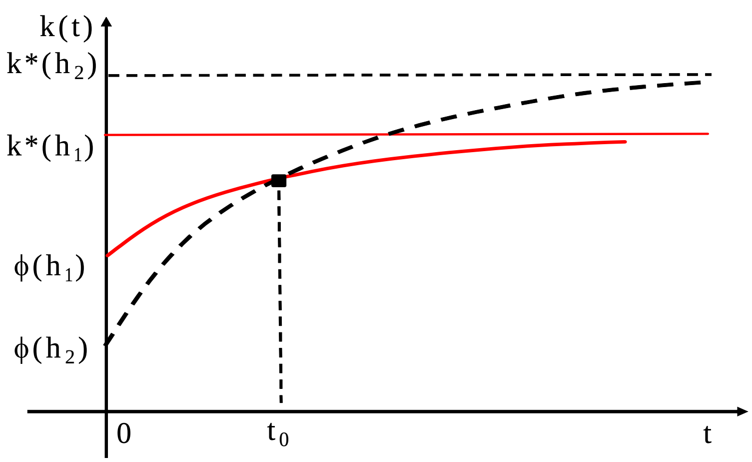


Figure 2. Crossing paths of capital stocks for two types of workers

Negative sorting in the initial assignment has an important implication for the dynamic paths of capital stocks for different skill levels: They must cross each other over time. Figure 2 depicts the time paths of capital stocks for workers of skills h_1 and h_2, with $h_2 > h_1$. With negative sorting in the initial assignment, the path for the high-skill worker starts below the path for the low-skill worker. The relative position between the two skills levels is reversed in the final state, because the final state has PAM. Thus, the two paths must cross each other at some time point, which is depicted as time t_0 in Figure 2. Sorting between capital and the worker skill is negative for $t \in [0, t_0)$ and positive for $t > t_0$. This implies further that matching is mixed in the population at any given point of time. Even if the initial assignment is unique for each worker skill, two workers of the same skill are working with different capital stocks if their matches have lasted for different lengths of time. Similarly, workers with different skill levels can be working with the same capital stock if their matches differ in the duration in a particular way.

4. Dynamic Sorting and the Stationary Distribution

Workers of the same skill are assigned to match with the same capital stock, but they may not succeed in matching at the same time. Those who succeeded earlier have had a longer time of investment and, hence, a higher capital stock. The capital stock in a match continues to increase toward the final state until exogenous separation terminates the match. There is an ergodic distribution of workers of the same skill over capital stocks. To examine the overall sorting pattern, I characterize this distribution.

Among the workers of any given h, let the fraction of unemployed be $u(h)$ and the fraction of employed be $n_e(h) = 1 - u(h)$. The flow of unemployed skill-h workers exiting unemployment is $u(h)\, p(h)$. The flow of employed workers of skill h into unemployment is $\delta n_e(h)$. In the steady state, since the two flows must be equal to each other, then

$$u(h) = \frac{\delta}{\delta + p(h)}. \tag{4.1}$$

Among employed workers of skill h, let $G(k|h)$ be the cumulative distribution function over k. To compute this distribution, consider the group of workers of skill h who are employed with capital stocks greater than k. The measure of this group is $n_e(h)[1-G(k|h)]$, with $G(\phi(h)|h)=0$ and $G(k^*(h)|h)=1$. Consider any $k\in(\phi(h),k^*(h))$. Let $\mathrm{d}t$ be an arbitrarily small interval of time such that $k-I(k,h)\mathrm{d}t>\phi(h)$. In this time interval, the workers who are employed with capital stocks in the interval $(k-I(k,h)\mathrm{d}t,k]$ flow into the described group as a result of investment. This inflow is:

$$n_e(h)\lim_{\mathrm{d}t\to 0}\frac{G(k|h)-G(k-I(k,h)\,\mathrm{d}t|h)}{\mathrm{d}t}=n_e(h)\,G'(k|h)\,I(k,h)\,.$$

Exogenous separation generates the outflow, $n_e(h)\,\delta\,[1-G(k|h)]$. Equating the inflow and outflow yields:

$$\frac{G'(k|h)}{1-G(k|h)}=\frac{\delta}{I(k,h)}. \tag{4.2}$$

Among skill-h employed workers, the mean of capital stocks is

$$Ek(h)\equiv\int_{\phi(h)}^{k^*(h)}k\ \mathrm{d}G(k|h)\,. \tag{4.3}$$

Definition 4.1. *Capital and skills have PAM-D (positive assortative matching according to the distribution) if $G(k|h)$ has the first-order stochastic dominance with respect to h, i.e., if $G(k|h)$ is decreasing in h for all $k\in(\phi(h),k^*(h))$. Capital and workers have PAM-M (positive assortative matching according to the mean) if $Ek(h)$ is increasing in h.*

The following proposition is proven in Appendix B:

Proposition 4.2. *$u'(h)<0$ if and only if $p'(h)>0$. Approximate the dynamic system of (i,k) by (3.4). Then,*

$$G(k|h)=1-\left[\frac{k^*(h)-k}{k^*(h)-\phi(h)}\right]^{\frac{\delta}{\beta(h)}}=1-[R(k,h)]^{\frac{\delta}{\beta(h)}}\ , \tag{4.4}$$

$$Ek(h)=\frac{\delta\phi(h)+\beta(h)\,k^*(h)}{\beta(h)+\delta}=\phi(h)+\frac{I(\phi(h),h)}{\beta(h)+\delta}, \tag{4.5}$$

where $R(k,h)\equiv\frac{I(k,h)}{I(\phi(h),h)}$ and $\beta(h)$ is defined by (3.5). $G''(k|h)<0$ if and only if $\beta(h)<\delta$. Moreover, PAM-D occurs if and only if

$$\begin{gathered}R(k,h)\,\phi'(h)+[1-R(k,h)]\,k^{*\prime}(h)>\frac{\beta'(h)I(k,h)}{[\beta(h)]^2}\ln R(k,h)\\ \text{for all } k\in(\phi(h),k^*(h))\,.\end{gathered} \tag{4.6}$$

$\phi'(h)\geq 0$ is necessary and sufficient for (4.6) to hold. However, $\phi'(h)\geq 0$ is sufficient, but may not be necessary, for PAM-M.

A higher skill has a lower unemployment rate if and only if the matching rate is higher. This occurs in the social optimum except when the two productive factors are strongly complementary with each other as in Case 3 (see Table 1). For any given skill h, the density function of the distribution of workers over the capital stock is decreasing if and only if $\beta(h) < \delta$. This result is intuitive. The upper tail of the distribution of employed workers of any given skill h is $1-G(k|h)$. Workers exit this tail if they are hit by exogenous separation, and workers outside this tail enter it as firms invest to increase the capital stock. Investment is a ratio $\beta(h)$ to the width of the support of the upper tail, $[k, k^*(h)]$. When k increases, the width narrows, which reduces the size of both the exit and the entry into the upper tail. Then, the difference between the two flows depends only on the difference between the rates of exit and entry. If $\beta(h) < \delta$, the exit rate is higher than the entry rate. To keep the distribution stationary, the base for the exit must become increasingly smaller when k increases; i.e., the density of the distribution must be decreasing.[15] Conversely, if $\beta(h) > \delta$, the density of the distribution must be increasing.[16]

The condition (4.6) is necessary and sufficient for PAM-D. It reveals three determinants of PAM-D: the initial assignment $\phi(h)$, the final assignment $k^*(h)$, and the speed of convergence to the final state, $\beta(h)$. PAM of the initial and PAM of the final assignment contribute to PAM-D because they increase capital stocks assigned to high-skill workers at the two ends of the dynamic path. If the capital stock converges to the final state more quickly for high-skill workers, i.e., if $\beta'(h) > 0$, it also contributes to PAM-D by increasing capital for high-skill workers more quickly. PAM in the initial assignment is necessary for PAM-D. If $\phi(h_2) < \phi(h_1)$ for some $h_2 > h_1$, then the lower bound on the support of $G(k|h_2)$ is lower than that of $G(k|h_1)$. In this case, for relative short tenure, capital stocks are lower for skill h_2 workers than for skill h_1 workers (see Figure 2).

PAM in the initial assignment is also sufficient for PAM-D. This result may not be

[15]This contrasts with undirected search models, e.g., Burdett and Mortensen (1998), where the density is increasing and strictly convex.

[16]Define $z = [k^*(h) - k]^{-1}$. Then, (4.4) implies that z is distributed according to $G(z|h) = 1 - \{[k^*(h) - \phi(h)]\, z\}^{-\delta/\beta(h)}$, which is the Pareto distribution with the parameter $\beta(h)/\delta$. The larger is δ relative to $\beta(h)$, the smaller is $\beta(h)/\delta$, and so the fatter is the tail of the distribution of z. Note that the density function of z is always decreasing.

obvious, because the speed of convergence to the final state can be lower for high-skill workers than for low-skill workers. That is, $\beta'(h) < 0$ is possible (see Example 3.2). Even if the initial assignment is PAM, one might think that the slower convergence can reverse the ranking in the capital stock between two skill levels in the transition to the final state. This does not happen in the social optimum. The reason is that, for any given capital stock, investment increases in the worker skill, i.e., $I_2(k,h) > 0$ for all $k < k^*(h)$ (see Proposition 3.3). If high-skill workers start employment with a higher capital stock in the initial assignment, higher investment will keep their capital stocks higher than those for low-skill workers throughout the transition to the final state.

The three determinants for PAM-D are also the determinants for PAM-M. The mean of capital stocks for any given worker skill is a weighted average of the initial and the final assignment, as shown by the first equality in (4.5). In contrast to the distribution where the weights are associated with investment, the weights in the mean depend only on the two rates, β and δ. The weight is $\frac{\delta}{\beta(h)+\delta}$ for the initial assignment and $\frac{\beta(h)}{\beta(h)+\delta}$ for the final assignment. If the exit rate δ is higher, the initial assignment has a higher weight in the mean because the capital stock in the match for a worker remains close to the initial assignment before the worker separates into unemployment. If the convergence speed β is higher, the final assignment has a higher weight in the mean because the capital stock in the match increases quickly toward the final state.

The second equality in (4.5) is also illuminating. It states that the mean of the increase in the capital stock during employment is equal to the present value of the initial investment, where the discount rate is the sum of the convergence speed, β, and the exogenous separation rate, δ. The separation rate increases the discounting because separation terminates a match. The convergence speed β also increases the discounting because investment declines more quickly if β is higher.

PAM-D implies PAM-M. Because PAM in the initial assignment is sufficient for PAM-D, it is also sufficient for PAM-M. However, PAM in the initial assignment is not necessary for PAM-M. Even if high-skill workers are assigned to a lower capital stock at the beginning of employment, investment and the final capital stock are higher for such workers than for low-skill workers. These factors in the transition can dominate negative sorting in the initial assignment to make the mean of the capital stock increase in the worker skill.

The above analysis compares the conditional distribution between different types of workers. It is also useful to examine the joint distribution between capital and the worker skill. For example, the correlation between the two factors in employment is indicative of sorting, but computing the correlation involves the joint distribution. Note that the density of skill-h workers in employment is $n_e(h) H'(h)$, where H' is the density function of types among all workers. Denote the measure of all employed workers as $N_e = \int_{h_L}^{h_H} n_e(h) \mathrm{d}H(h)$. The employment share of skill-h workers is

$$s(h) = \frac{n_e(h) H'(h)}{N_e}.$$

Let $G(k,h)$ be the joint cumulative distribution function of employed workers over (k,h). Then, the joint density is $G_{12}(k,h) = s(h) G'(k|h)$, and so[17]

$$G(k,h) = \int_{h_L}^{h} s(h) G\left(k|\tilde{h}\right) \mathrm{d}\tilde{h}. \tag{4.7}$$

Since the derivative of $G_{12}(k,h)$ with respect to k has the same sign as $G''(k|h)$, Proposition 4.2 implies that the joint density is decreasing in h if and only if $\beta(h) < \delta$.

The joint distribution reveals that employment shares can be an additional factor affecting the extent of sorting. The employment share of a worker skill accentuates the matching pattern of that skill by allowing such matches to be sampled more often in the data. If a worker skill has a high employment share and such workers have PAM, then the skill will contribute a large share to PAM in the employed sample. In the exposition for Proposition 3.3, I explained that the social optimum is more likely to have higher utilization of high-skill workers than PAM in the initial assignment. Thus, $s'(h) > 0$ is more likely than $\phi'(h) > 0$. If the initial assignment fails to be PAM, the high share of high-skill workers in employment may reduce the extent of sorting identified from the data.

[17] The marginal density of h is equal to $s(h) = G_2(k^*(h), h)$ and the marginal density of k is $G_1(k, h_H)$. The density of h conditional on k is $G'(h|k) = \frac{G_{12}(k,h)}{G_1(k,h_H)}$, and the conditional cumulative distribution is $G(h|k) = \frac{G_1(k,h)}{G_1(k,h_H)}$.

5. Quantitative Analysis

5.1. Calibration and Simulation

To calibrate the model, I use the following functional forms:

$$\begin{aligned}
&\tilde{f}(k,h) = f_0 k^{\alpha} h^{1-\alpha}, \quad f_0 > 0,\, \alpha \in (0,1)\,; && H(h) = \tfrac{h-h_L}{h_H-h_L},\, h_H > h_L > 0; \\
&\theta(p) = p_0 \left[p^{-\rho} - 1\right]^{-1/\rho}, \quad p_0, \rho \in (0,\infty)\,; && c(i) = c_1 i^2, \quad c_1 > 0; \\
&\psi(k) = \psi_0 k^{\psi_1}, \quad \psi_0 > 0,\, \psi_1 \geq 1.
\end{aligned}$$

Net output is $f(k,h) = \tilde{f}(k,h) - rk$. The inverse of $\theta(p)$ is a Dagum (1975) function. The model is calibrated monthly. The distribution of worker types is uniform, where $h_H = 1$ and $h_L = 0.25$. This specification of the skill distribution has no consequence on the initial assignment or the dynamics of the capital stock among workers of the same skill, because the social optimum features separation between any two types of workers. Denote the average value of h in the population as $h_a = \frac{h_H+h_L}{2} = 0.625$. Table 2 lists the parameters, their values and the calibration targets (see the Supplementary Appendix C for the procedure).

Table 2. Identification of the parameters

Parameter	Value	Target
r	4.15×10^{-3}	quarterly interest rate $= 0.0125$
α	0.55	capital share at $k^*(h)$ is 0.55
f_0	0.0599	normalize $k^*(h_H) = 100$
δ	0.026	monthly EU rate in CPS
ρ	0.5	elasticity $\frac{\mathrm{d}\ln p}{\mathrm{d}\ln\theta} = 0.39$ at average $u = 0.065$
p_0	0.29	market tightness $\theta = 0.72$ at average u
f_u	0.062	$\frac{f_u}{f^*(h_a)} = 0.29$ ($\frac{\text{home production}}{\text{average market net output}}$=0.4)
ψ_1	4.462	(ψ_1, ψ_0, c_1) minimize the distance
ψ_0	1.817×10^{-5}	between $(u(h_L), u(h_a), u(h_H))$
c_1	0.029	and the targets $(0.15, 0.06, 0.03)$

Several aspects of the calibration are worth noting. First, the average unemployment rate is targeted at 6.5%, and the ratio of home production of an unemployed worker to the average net output in the market is targeted at 40%. These targets are the same as in Hornstein et al. (2011) who emphasize the two targets as constraints on a model's ability to generate frictional wage dispersion. To reduce the computation time, I target, instead, the ratio of home production to the maximum net output of workers of skill h_a to 0.29, and then check the model's prediction on the ratio of home production to average market production. Similarly, rather than targeting the average unemployment rate directly, I minimize the distance between the model and the data in unemployment rates of workers of three types (h_L, h_a, h_H).

Second, the value $\alpha = 0.55$ is chosen with an intended target on the capital share of output, 0.36. The capital share in a (k, h) match is $\frac{rk}{\tilde{f}(k,h)} = \frac{r}{f_0}\left(\frac{k}{h}\right)^{1-\alpha}$. This share varies with k and h, despite the Cobb-Douglas production function. For any given h, the capital shock increases over time as the capital stock increases. Only in the final state is the capital share equal to α. The average capital stock for any given skill h is significantly below the final state. The value of α is set to be greater than 0.36 in order for the average capital share in the model to be close to this realistic value.

Third, I identify (ρ, p_0) using the market tightness, θ, and the elasticity of the job-finding rate with respect to the market tightness, $\frac{1}{\varepsilon}$. Because both statistics depend on the worker skill, I apply the targets to the skill of workers whose unemployment rate is at the average level, 0.065. The market tightness comes from Pissarides (2009) who derives the value from the Job Openings and Labor Turnover Survey (JOLTS) and the Help-Wanted Index. The parameter value $\rho = 0.5$ implies that in the submarket searched by such workers, the job-finding rate varies with the tightness with an elasticity $\frac{1}{\varepsilon} \equiv \frac{\mathrm{d}\ln p_u}{\mathrm{d}\ln\theta(p_u)} = 0.39$. This value lies within the range used in the literature.[18]

Table 3. Performance of the calibration

	Model	Targets
$u(h_L)$	0.150	0.15
$u(h_a)$	0.053	0.06
$u(h_H)$	0.046	0.03
average u	0.060	0.065
$\frac{\text{home production}}{\text{average } f}$	0.377	0.40
average share of k	0.342	0.36

The parameters in the cost functions, (ψ_1, ψ_0, c_1), minimize the distance between the model and the data in unemployment rates of the three types of workers, (h_L, h_a, h_H).

[18]Shimer (2005) estimates this elasticity as 0.27, but Pissarides (2009) uses the estimate 0.5.

Table 3 lists the result of this minimization and some other statistics. The unemployment rate at h_L is close to the target, which unemployment rates at h_a and h_H are lower than the targets. The average unemployment rate is close to the target. So is the ratio of home production to average market output. The average capital share is lower than the target. This discrepancy is small, especially in consideration that the capital share has excluded the costs of vacancies and capital adjustments. With the calibrated parameters, I compute the social optimum (see the Supplementary Appendix C for the procedure).

5.2. Initial Assignment and Matching Rate

Figure 3 depicts the initial assignment $\phi(h)$ and the matching rate $p(h)$. The initial assignment is NAM, and the matching rate increases in the skill, as in Case 1 listed after Proposition 3.3.[19] As the skill increases from h_L to h_H, the initial assignment decreases from $\phi(h_L) = 5.758$ to $\phi(h_H) = 4.128$. This reduction in the capital stock would reduce net output by 12% if the skill remained at h_L. From h_L to h_H, the matching rate increases from $p(h_L) = 0.147$ to $p(h_H) = 0.545$, an increase of 270%. As analyzed in section 3.2, the higher matching rate for a higher skill enables the social optimum to utilize the higher skill more effectively than using PAM in the initial assignment.

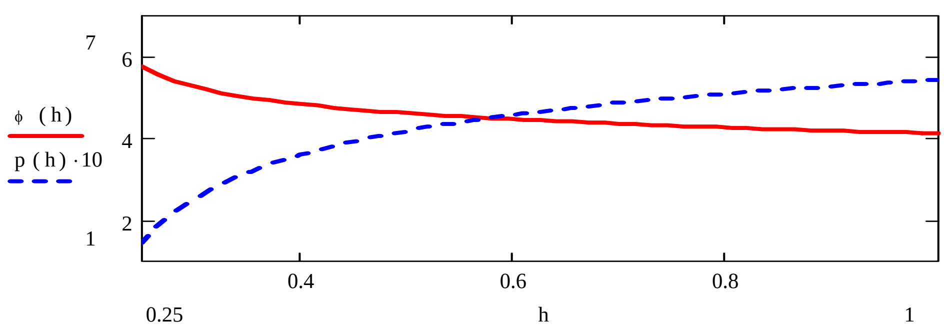
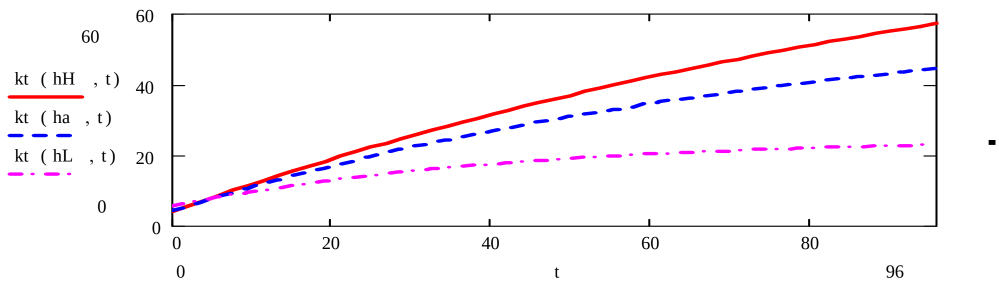


Figure 3. Initial assignment $\phi(h)$ and matching rate $p(h)$

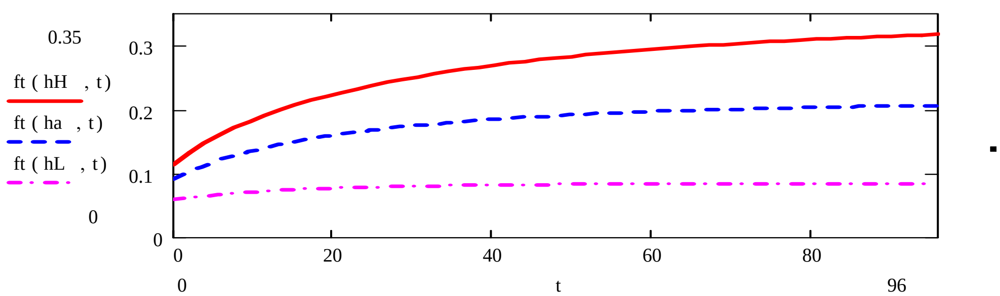
Because the initial assignment of capital to workers is NAM, the time paths of capital stocks for different types necessarily cross each other, as illustrated in Figure 2. With the calibrated parameters, this crossing occurs in less than 9 months, as shown in the upper

[19]Referring to Table 1, the calibrated model has $a_1(h) < I_0(h) < a_2(h)$ and $\frac{f_{12}(\phi(h),h)}{r+\delta+\beta(h)} < \frac{f_2(\phi(h),h)}{a_2(h)-I_0(h)}$ for all h, where $I_0(h) = I(\phi(h), h)$.

panel in Figure 4. The lower panel in Figure 4 depicts the time paths of net output for the three types of workers, (h_L, h_a, h_H). Although workers of a higher skill have a lower capital stock at the time of a match, this negative assignment does not eliminate the productivity advantage of the higher skill. That is, net output at the time of a match increases in the worker skill. At all time points after the match, net output is higher for a high skill than for a low skill. Moreover, the curve $f(k(t, h_2), h_2)$ is steeper than the curve $f(k(t, h_1), h_1)$ for $h_2 > h_1$, and so the difference in net output between skills increases over time.



$kt(h, t) = k(t, h)$: Capital stock for a skill-h worker at t (months)



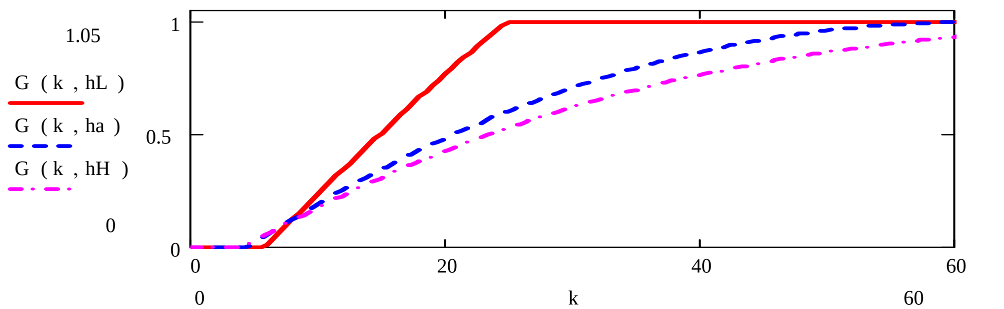
$ft(h, t) = f(k(t, h), h)$: Net output of a skill-h worker at t (months)

Figure 4. Net output over time for (h_L, h_a, h_H)

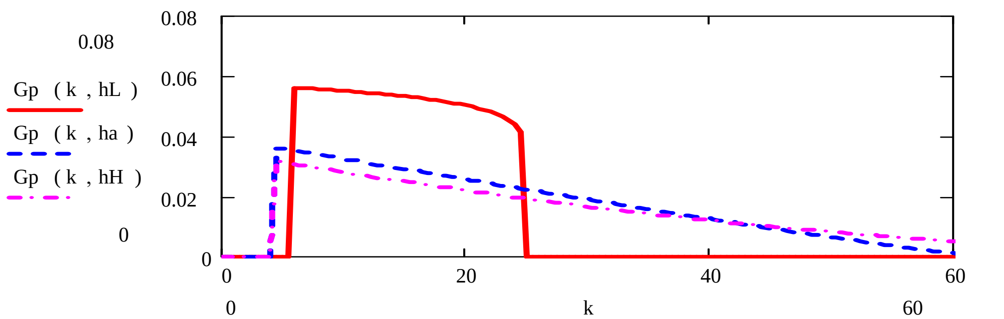
If workers' earnings are proportional to net output, the above results imply that workers of a higher skill have a steeper time profile of earnings than workers of a lower skill. Also, if one traces a cohort of workers who become employed at the same time, the variance in earnings in the cohort increases over time. Both the heterogeneity in the slope of the earnings profile and the fanning-out of earnings are consistent with the evidence on inequality over the life cycle (e.g., Deaton and Paxson, 1994, Guvenen, 2007).

5.3. Distribution of Workers, the Mean, and Sorting

Figure 5 depicts the distribution of workers for h_L, h_a, and h_H. The upper panel is the cumulative distribution $G(k|h)$, and the lower panel is the density function $G'(k|h)$. Because the paths of capital for different skills cross over time, $G(k|h)$ fails to have the first-order stochastic dominance with respect to h. However, this failure occurs for relative low levels of k. After k passes a critical level, $G(k|h_2) < G(k|h_1)$ for any $h_2 > h_1$.



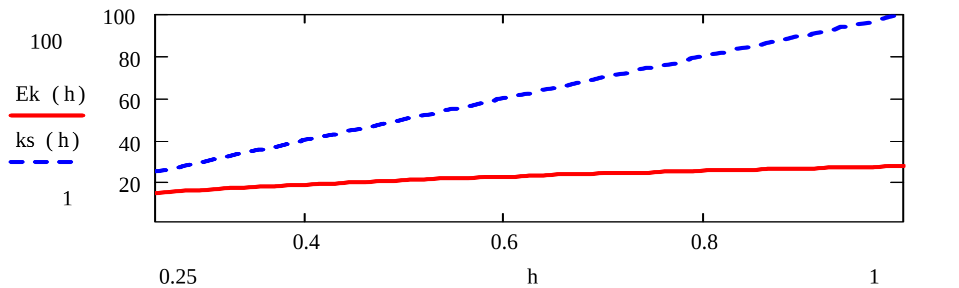
$G(k,h)$: Cumulative distribution of k among skill-h workers



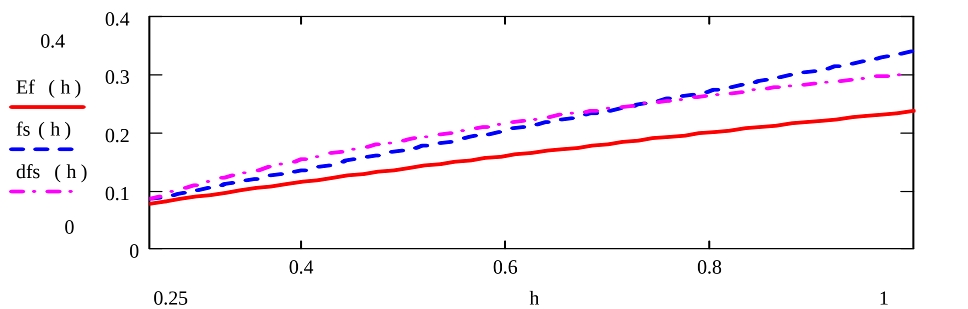
$Gp(k,h) = G_1(k,h)$: Density function of k among skill-h workers

Figure 5. The cumulative distribution $G(k|h)$ and density $G'(k|h)$ for $h = h_L, h_a, h_H$.

The density functions in the lower panel in Figure 5 illustrate the features of the distribution more clearly. Because the initial assignment for skill h_H is lower than for skill h_L, the density function $G(k|h_H)$ becomes positive first at a lower capital stock than $G(k|h_L)$. However, the density is smaller for skill h_H than for skill h_L at low capital stocks if both densities are positive. As the capital stock increases, the density of workers decreases, because $\beta(h) < \delta$ for all h. This decrease is less rapid for a high skill than for a low skill. When the capital stock becomes sufficiently high, the density functions of capital stocks for different skills cross each other. Furthermore, when k increases above $k^*(h_L)$, no workers of skill h_L are employed with such capital stocks, but there are workers of skill h_H employed with such capital stocks.



$Ek(h)$: Expected value of k among skill-h workers; $ks(h) = k^*(h)$



$Ef(h)$: Expected value of f among skill-h workers;
$fs(h) = f^*(h)$; $dfs(h) = 1 - \frac{Ef(h)}{f^*(h)}$

Figure 6. $Ek(h)$, $Ef(h)$ and related values in the final state

The calibrated social optimum features PAM-M. The mean of the capital stock is an increasing function of the worker skill, as depicted in the upper panel in Figure 6. As the worker skill increases from h_L to h_H, the mean of the capital stock increases from 14.99 to 27.58. In addition to the higher average stock of capital, a higher worker skill itself increases net output. Thus, the mean of net output, $Ef(h)$, increases in the worker skill, as depicted in the lower panel in Figure 6. PAM-M, illustrated in Figure 6, and the profile of the assignment, illustrated in Figure 4, indicate that dynamic sorting between capital and the worker skill is strong.[20]

Figure 6 also shows that the capital stock and net output are substantially lower than in the final state predicted by the neoclassical theory. Moreover, this gap between the two increases in the worker skill. In addition to Ef and $f^*(h)$, the lower panel in Figure 6 shows the percentage gap between the two, $dfs\,(h)$. As the worker skill increases from h_L to h_H, the gap between the mean of net output and the final state increases from 8.5% to 30.2%. The explanation is as follows. Because of search frictions, the capital stock in the initial assignment is substantially lower than in the final state. Although investment increases the capital stock toward the final state, the process is slowed down by the adjustment cost in investment and terminated by separation. Among any given skill of employed workers, the average length of time with the job is limited. Moreover, the higher is the worker skill, the larger is the gap between the initial and the final stock, and the longer is the time needed to invest for the capital stock to reach the final state. In the upper panel in Figure 4, at $t = 96$ months (8 years), the capital stock for a skill h_H worker is only 57.3% of the final stock. However, at the separation rate $\delta = 0.026$ per month, the fraction of matches that survive to $t = 96$ is only $(1-\delta)^{96} = 0.0797$.

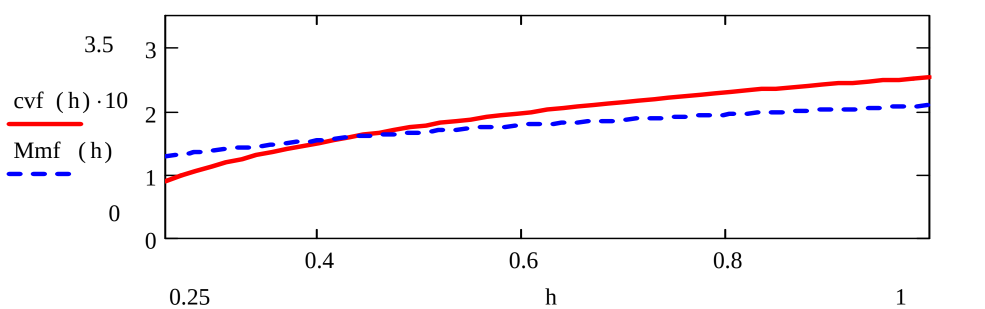
5.4. Frictional Inequality

Frictional inequality refers to inequality in earnings among workers of the same skill. Hornstein et al. (2007) propose to measure frictional inequality by the ratio of the mean to the minimum of wages among workers who are identical in observable types. Another measure is the coefficient of variation. Because the current focuses on the social optimum instead of the equilibrium, I calculate the measures of frictional inequality in net output instead of wages. However, if wages are a constant fraction of net output, the measures of inequality in net output are equal to those in wages.

Figure 7 depicts the coefficient of variation, $cvf\,(h)$, and the mean-min ratio, $Mmf\,(h)$,

[20] However, a linear regression of the capital stock on the worker skill in the model would under-estimate sorting. As a proxy for the regression coefficient, the correlation coefficient between capital and skill among employed workers is 0.26. Non-linearity of output in the capital stock may be a cause of this relatively lower correlation.

in net output as functions of the worker skill. The coefficient of variation ranges from 0.09 among skill h_L workers to 0.253 among skill h_H workers. The mean-min ratio ranges from 1.29 among skill h_L workers to 2.085 among skill h_H workers.[21] Weighted by the employment shares, the average among all employed workers is 0.226 for the coefficient of variation and 1.769 for the mean-min ratio. Such frictional inequality is large and realistic. Hornstein et al. (2011) estimate the mean-min ratio in wages in the U.S. data to be between 1.7 and 2, but find the counterpart in most search models to be lower than 1.05. Note that I have kept the two calibration targets that Hornstein et al. (2011) have explained to be the critical constraints on frictional dispersion: the average unemployment rate and the ratio of home production to market output. Despite these calibration targets, the current model is able to generate large frictional dispersion. The reasons for this success are that post-match investment is available and that the vacancy cost is convex in the capital stock, as explained in detail in another paper (Shi, 2018).



$cvf\,(h)$: Coefficient of variation in f among skill-h workers
$Mmf\,(h)$: Mean-min ratio if f among skill-h workers

Figure 7. Coefficient of variation and mean-min ratio in net output among skill-h workers

Higher skill workers also have lower unemployment rates. This is remarkable for two reasons. First, if different types of workers are weighted by their employment shares to compute the average frictional inequality, then high-skill workers have higher weights than low-skill workers. This explains why the mean of overall frictional inequality is closer

[21] Friction inequality in the capital stock is larger than in net output. The second equality in (4.5) shows that the mean-min ratio in capital among type h workers is $Ek(h)/\phi\,(h) = 1 + \frac{I(\phi(h),h)/\phi(h)}{\beta(h)+\delta}$. This ranges from 2.41 for h_L to 5.56 for h_H.

to that among skill h_H workers than to that among skill h_L workers, as reported above. Second, a low unemployment rate is not a large obstacle to generating high frictional inequality. Hornstein et al. (2011) have intuitively explained that a low unemployment rate implies a low option value of search and, hence, small frictional dispersion. This link between frictional dispersion and the unemployment rate breaks down and is reversed when post-match investment is available. To reduce the unemployment rate for high-skill workers, the social optimum lowers the initial assignment of capital to such workers, which increases frictional dispersion among such workers. Across worker types, the lower the unemployment rate, the larger the frictional dispersion.[22]

The quantitative results from subsections 5.2 to the current subsection demonstrate that both frictional sorting and post-match investment are important. Investment quickly reverses negative sorting in the initial assignment and widens the gap in net output between skills. Without such investment, as in the literature on frictional sorting, a model would reach a wrong conclusion about the relationship between strong complementarity and sorting in the initial assignment. On the other hand, without matching frictions, the neoclassical theory of investment by firms would significantly exaggerate the extent of positive sorting, as illustrated in subsection 5.3 by the gap between the average capital stock for each skill in the current model and in the final state. Moreover, a frictionless model would not provide an endogenous mechanism to generate inequality within each skill, the heterogeneity in the slope of the time profile of labor productivity between skills, or the fanning-out of labor productivity over time.[23]

6. Conclusion

This paper has integrated frictional matching into a neoclassical framework of investment by firms to analyze dynamic sorting between capital and worker skills in the constrained

[22]Figure 7 shows that frictional inequality increases in the worker type. The empirical evidence is mixed on this. Gottschalk and Moffitt (2009) report that, in the late 1980s and early 1990s, the transitory variance in log earnings grew more quickly for the more-educated group than the less-educated group (high school or less). However, the pattern was the opposite in the 1970s and early 1980s.

[23]To account for these facts, the literature on earnings profiles introduces shocks to earnings (e.g., Deaton and Paxson, 1994).

social optimum. The model predicts that strong complementarity between capital and worker skills in production makes the socially efficient pattern of sorting negative at the time of a match, which is opposite to the result in the absence of post-match investment. However, investment makes sorting positive eventually. Because the sorting pattern in a match reverses over time, labor productivity increases more sharply for a high skill than for a low skill, and this between-skill difference increases over time. There is also dispersion in labor productivity within each skill. The calibrated model shows that sorting is positive on average and that dispersion in labor productivity is significant both between skills and within each skill. These results suggest that dynamic sorting with matching frictions and post-match investment can be an important mechanism for generating heterogeneity in earnings profiles, the increasing variance in earnings over the life cycle, and within-group inequality. It is time to incorporate this mechanism as a standard ingredient of macroeconomic models.

There are at least two interesting directions for future research. One is to incorporate on-the-job search to examine how sorting of workers between firms interacts with sorting within each firm. With homogeneous workers, on-the-job search can delay socially efficient investment (see Shi, 2018). With heterogeneous workers, this delay can prolong the negative sorting pattern at the beginning of a match and induce more realistic patterns of job-to-job transition. The other direction of future research is to examine how sorting responds to a skill-biased technological progress. If such a progress increases the gain from employing high skills quickly, then it increases the extent of negative sorting at the beginning of a match. Because post-match investment makes sorting positive eventually, inequality is likely to increase both between skills and within each skill.

Appendix

A. Proofs for Section 3

Proof of Proposition 3.1:

The final capital stock, $k^*(h)$, satisfies $f_1(k^*(h), h) = 0$. Differentiating yields:

$$k^{*\prime}(h) = \frac{f_{12}(k^*, h)}{-f_{11}(k^*, h)} > 0,$$

where the inequality comes from $f_{11} < 0$ and $f_{12} > 0$. It is easy to verify that the eigenvector of the matrix J in (3.4) corresponding to the stable eigenvalue $-\beta$ is $[\beta, -1]^T z$, where z is a constant to be determined. The unique stable saddle path of (3.4) is

$$\begin{bmatrix} i(t) \\ k(t) - k^*(h) \end{bmatrix} = \begin{bmatrix} \beta(h) \\ -1 \end{bmatrix} z\, e^{-\beta(h)t}.$$

Because $k(0) = \phi(h)$, setting $t = 0$ in the second equation yields $z = k^*(h) - \phi(h)$. Thus, the unique stable path is (3.6). Dividing the two equations in (3.6) yields (3.7).

It is evident from (3.6) that if $\phi(h) < k^*(h)$, then $\frac{\mathrm{d}k(t)}{\mathrm{d}t} = i(t) > 0$ and $\frac{\mathrm{d}i(t)}{\mathrm{d}t} < 0$. The speed of convergence to the final state is β as defined in (3.5). Clearly, β is lower if $c''(0)$ is higher or $-f_{11}$ is lower. Moreover, $\beta'(h) < 0$ if and only if $-f_{11}(k^*(h), h)$ is decreasing in h. This completes the proof of Proposition 3.1. **QED**

Proof of Proposition 3.3:

The envelope conditions of h for (2.4) and (2.2) yield:

$$V_2(k(t), h) = \frac{\delta}{r+\delta} V_u'(h) + \int_t^\infty f_2(k(\tau), h)\, e^{-(r+\delta)(\tau - t)} \mathrm{d}\tau, \tag{A.1}$$

$$V_u'(h) = \frac{p V_2(\phi, h)}{r+p}, \tag{A.2}$$

where I switched the indices t and τ. Because $f_2 > 0$, then $V_2(k, h) > 0$ and $V_u'(h) > 0$ for any given k. Differentiating (3.8) and (3.9) yields:

$$\begin{bmatrix} \phi'(h) \\ p'(h) \end{bmatrix} = \frac{1}{\det} \begin{bmatrix} \psi \theta'' \frac{V_{12}}{V_1} - \frac{\Delta_2}{p}(\varepsilon - 1) \\ (\frac{\psi''}{\psi'} - \frac{V_{11}}{V_1})\Delta_2 - (\varepsilon - 1) V_{12} \end{bmatrix} \tag{A.3}$$

where $\varepsilon = \frac{p\theta'}{\theta}$ (> 1) and

$$\det = \psi\theta''(\frac{\psi''}{\psi'} - \frac{V_{11}}{V_1}) - (\varepsilon - 1)^2 \frac{V_1}{p}. \tag{A.4}$$

The arguments of $(V_1, V_{11}, V_{12}, \Delta_2)$ are (ϕ, h) and I substituted $\Delta_1(\phi, h) = V_1(\phi, h)$. Because the optimal search choice maximizes the expected social return on search, the objective function in (2.2) must be concave at the optimal (ϕ, p). This implies $\det > 0$. To determine the signs of $\phi'(h)$ and $p'(h)$, I need to compute $\Delta_2(\phi, h)$ and $V_{12}(\phi, h)$.

To compute $\Delta_2(\phi, h)$, set $t = 0$ in (2.3) and substitute $i(0) = I(\phi, h)$. I get:

$$(r + \delta) V(\phi, h) - \delta V_u(h) = f(\phi, h) + [V_1(\phi, h) I - c(I)]_{I=I(\phi,h)}. \tag{A.5}$$

Subtracting (2.2) from the above equation yields:

$$\Delta(\phi, h) = \frac{f(\phi, h) - f_u + \psi(\phi)\theta(p) + [V_1(\phi, h) I - c(I)]_{I=I(\phi,h)}}{r + \delta + p}, \tag{A.6}$$

where $p = p(h)$. Fixing ϕ but taking the dependence $p(h)$ into account, I differentiate Δ with respect to h. Substituting (3.1) and (3.9) into the result, I get:

$$\Delta_2(\phi, h) = \frac{f_2(\phi, h) + V_{12} I(\phi, h)}{r + \delta + p}. \tag{A.7}$$

To compute V_{12}, differentiate (A.1) with respect to t:

$$V_{12}(k(t), h) i(t) = -f_2(k(t), h) + (r + \delta) \int_t^\infty f_2(k(\tau), h) e^{-(r+\delta)(\tau - t)} d\tau. \tag{A.8}$$

This holds for all $t \geq 0$. Set $t = 0$. Using (3.6), I approximate $f_2(k, h)$ near $k = \phi$ as

$$f_2(k, h) \approx f_2(\phi, h) + f_{12}(\phi, h)(k^* - \phi)\left(1 - e^{-\beta t}\right). \tag{A.9}$$

The last term comes from substituting k from (3.6). With this, I compute:

$$\int_0^\infty f_2(k(t), h) e^{-(r+\delta)t} dt \approx \frac{1}{r + \delta}\left[f_2(\phi, h) + \frac{f_{12}(\phi, h)}{r + \delta + \beta} I(\phi, h)\right], \tag{A.10}$$

where $I(\phi, h) = \beta(k^* - \phi)$. Set $t = 0$ in (A.8). Substituting (A.9) and (A.10) yields:

$$V_{12}(\phi, h) \approx \frac{f_{12}(\phi, h)}{r + \delta + \beta} > 0. \tag{A.11}$$

In addition, because $V_1(k(t), h) = c'(i(t))$ for all $t \geq 0$ by (3.1), then

$$V_1(\phi, h) = c'(I(\phi, h)), \quad V_{11}(\phi, h) = -c''(I(\phi, h))\beta(h), \tag{A.12}$$

where I used $I_1(\phi, h) = -\beta$ from (3.7). Substitute (A.7), (A.11) and (A.12) into (A.3):

$$\begin{aligned} \phi'(h) &\approx \tfrac{(\varepsilon-1)/p}{(r+\delta+p)\det} \left\{ (a_2 - I) \tfrac{f_{12}(\phi,h)}{r+\delta+\beta} - f_2(\phi, h) \right\} \\ p'(h) &\approx \tfrac{\frac{\psi''}{\psi'} + \beta \frac{c''(I)}{c'(I)}}{(r+\delta+p)\det} \left\{ f_2(\phi, h) - (a_1 - I) \tfrac{f_{12}(\phi,h)}{r+\delta+\beta} \right\}, \end{aligned} \tag{A.13}$$

where $I = I(\phi, h)$, and (a_1, a_2) are defined by (3.12). Because $\det > 0$ and $\varepsilon > 1$, then (A.13) implies (3.10) and (3.11). Also, $\det > 0$ implies $a_2 > a_1$. Clearly, $a_1 > 0$ and $a_2 > 0$.

Differentiating $V_1(k, h) = c'(I(k, h))$ with respect to h yields:

$$I_2(k, h) = \frac{V_{12}(k, h)}{c''(I(k, h))} > 0 \text{ for all } k < k^*(h),$$

where the inequality follows from a result similar to (A.11). Since $I_1(\phi, h) = -\beta(h)$, then

$$\frac{\mathrm{d}}{\mathrm{d}h} I(\phi(h), h) = \frac{V_{12}(\phi, h)}{c''(I)} - \beta(h)\phi'(h). \tag{A.14}$$

Because $V_{12} > 0$, then $\phi'(h) \leq 0$ implies $\frac{\mathrm{d}}{\mathrm{d}h} I(\phi(h), h) > 0$. Also, by (A.13), $\phi'(h) \leq 0$ implies $f_2(\phi, h) \geq (a_2 - I) \frac{f_{12}(\phi,h)}{r+\delta+\beta}$ and so

$$f_2(\phi, h) - (a_1 - I) \frac{f_{12}(\phi, h)}{r + \delta + \beta} \geq (a_2 - a_1) \frac{f_{12}(\phi, h)}{r + \delta + \beta} > 0.$$

This implies $p'(h) > 0$. **QED**

B. Proof of Proposition 4.2

It is clear from (4.1) that $u'(h) < 0$ if and only if $p'(h) > 0$ and, hence, if and only if (3.11) is satisfied. The text preceding the proposition has derived (4.2) for the distribution of skill h employed workers over k. Approximating the dynamics of (i, k) by (3.4), then $I(k, h)$ satisfies (3.7). Substituting into (4.2), integrating from $\phi(h)$ to k, and using $G(\phi(h)|h) = 0$, I obtain the cumulative distribution function in the first equality in (4.4). From (3.7), $k^*(h) - k = I(k, h)/\beta(h)$ for all $k \in [\phi(h), k^*(h)]$. Substituting this relationship yields the second equality in (4.4). The corresponding density function is

$$G'(k|h) = \frac{\delta}{\beta(h)} [k^*(h) - k]^{\frac{\delta}{\beta(h)} - 1} [k^*(h) - \phi(h)]^{\frac{\delta}{\beta(h)}}.$$

It is easy to verify that $G''(k|h) < 0$ if and only if $\beta(h) < \delta$.

For any fixed k, differentiating (4.4) with respect to h yields:

$$-\tfrac{I(k,h)}{\delta(1-G)}\tfrac{\partial}{\partial h}G = \quad R(k,h)\,\phi'(h) + [1 - R(k,h)]\,k^{*\prime}(h) - \tfrac{\beta'(h)I(k,h)}{\beta^2(h)}\ln R(k,h)$$

Thus, $\frac{\partial}{\partial h}G(k|h) < 0$ for all $k \in (\phi(h), k^*(h))$ if and only if (4.6) holds. If $\phi'(h) < 0$, then (4.6) is violated for k sufficiently close to $\phi(h)$, because $R(k,h)$ is sufficiently close to 1 for such values of k. Thus, $\phi'(h) \geq 0$ is necessary for (4.6) to hold. To prove that $\phi'(h) \geq 0$ is also sufficient for (3.7) to hold, suppose $\phi'(h) \geq 0$ and consider interior $k \in (\phi(h), k(h))$. Note that (3.7) implies $I_2(\phi,h) = \beta'(h)[k^*(h) - \phi] + \beta(h)\,k^{*\prime}(h)$. Because $I_2(\phi,h) > 0$ by Proposition 3.3, then $\frac{\beta'(h)}{\beta^2(h)} > \frac{-k^{*\prime}(h)}{I(\phi(h),h)}$. For all interior k, $R(k,h) \in (0,1)$, and so

$$\begin{aligned} -\tfrac{I(k,h)}{\delta(1-G)}\tfrac{\partial}{\partial h}G \quad &\geq [1 - R(k,h)]\,k^{*\prime}(h) - \tfrac{\beta'(h)I(k,h)}{\beta^2(h)}\ln R(k,h)\,. \\ &> [1 - R(k,h) + R(k,h)\ln R(k,h)]\,k^{*\prime}(h)\,. \end{aligned}$$

The function $(1 - x + x\ln x)$ is strictly decreasing in x for all $x \in (0,1)$, and it is equal to 0 at $x = 1$. Thus, the function is strictly positive for all $x \in (0,1)$. This shows $-\frac{\partial}{\partial h}G > 0$ for all interior k.

To compute the expected value of k for any given h, I substitute $G(k|h)$ from (4.4) into (4.3) and integrate by parts. This yields the first equality in (4.5). The second equality in (4.5) follows from $I(\phi,h) = \beta(h)[k^*(h) - \phi(h)]$. Differentiating the first equality in (4.5) with respect to h, I can verify that $Ek(h)$ increases in h if and only if

$$\frac{\delta\phi'(h) + \beta(h)\,k^{*\prime}(h)}{k^*(h)} + \frac{\delta\ \beta'(h)}{\beta(h) + \delta} > 0. \tag{B.1}$$

Even if $\phi'(h) < 0$, (B.1) can still hold. Thus, $\phi'(h) \geq 0$ is not necessary for (B.1) to hold. For sufficiency, note that PAM-D implies PAM-M. Because $\phi'(h) \geq 0$ is sufficient for PAM-D, then it is also sufficient for $Ek'(h) > 0$.[24] **QED**

[24] To directly prove that $\phi'(h) \geq 0$ is sufficient for $Ek'(h) > 0$, suppose $\phi'(h) \geq 0$. If $\beta'(h) \geq 0$, then (B.1) shows $Ek'(h) > 0$. If $\beta'(h) < 0$, differentiate the second equality in (4.5) to obtain:

$$Ek'(h) = \phi'(h) + \frac{\mathrm{d}I/\mathrm{d}h}{\beta(h) + \delta} - \frac{I\ \beta'(h)}{[\beta(h) + \delta]^2} > \phi'(h) + \frac{\mathrm{d}I/\mathrm{d}h}{\beta(h) + \delta},$$

where $I = I(\phi(h), h)$. Substituting (A.14) yields: $Ek'(h) > \frac{1}{\beta(h)+\delta}\left[\frac{V_{12}}{c''(I)} + \delta\phi'(h)\right]$. Since $V_{12} > 0$, then $\phi'(h) \geq 0$ implies $Ek'(h) > 0$.

References

[1] Abowd, J.M., Kramarz, F. and D.N. Margolis, 1999, "High Wage Workers and High Wage Firms," Econometrica 67, 251-333.

[2] Becker, G.S., 1973, "A Theory of Marriage: Part I," Journal of Political Economy 81, 813-846.

[3] Burdett, K. and M.G. Coles, 2001, "Transplants and Implants: The Economics of Self-Improvement," International Economic Review 42, 597-616.

[4] Burdett, K. and D. Mortensen, 1998, "Wage Differentials, Employer Size, and Unemployment," International Economic Review 39, 257-273.

[5] Burdett, K., Shi, S. and R. Wright, 2001, "Pricing and Matching with Frictions," Journal of Political Economy 109, 1060-1085.

[6] Chade, H., Eeckhout, J. and L. Smith, 2017, "Sorting through Search and Matching Models in Economics," Journal of Economic Literature 55, 493-544.

[7] Dagum, C., 1975, "A Model of Income Distribution and the Conditions of Existence of Moments of Finite Order, " Bulletin of the International Statistical Institute 46, 199-205.

[8] Deaton, A. and C. Paxson, 1994, "Intertemporal Choice and Inequality," Journal of Political Economy 102, 437-467.

[9] Eeckhout, J. and P. Kircher, 2010, "Sorting and Decentralized Price Competition," Econometrica 78, 539-574.

[10] Gottschalk, P. and R. Moffitt, 2009, "The Rising Instability of U.S. Earnings," Journal of Economic Perspectives 23, 3-24.

[11] Gould, J.P., 1968, "Adjustment Costs in the Theory of Investment of the Firm," Review of Economic Studies 35, 47-55.

[12] Guvenen, F., 2007, "Learning Your Earning: Are Labor Income Shocks Really Very Persistent?" American Economic Review 97, 687-712.

[13] Hagedorn, M., Law, T.H. and I. Manovskii, 2017, "Identifying Equilibrium Models of Labor Market Sorting," Econometrica 85, 29-65.

[14] Hornstein, A., Krusell, P. and G.L. Violante, 2011, "Frictional Wage Dispersion in Search Models: A Quantitative Assessment," American Economic Review 101, 2873-2898.

[15] Julien, B., Kennes, J. and I. King, 2000, "Bidding for Labor," Review of Economic Dynamics 3, 619-649.

[16] Lise, J. and J.-M. Robin, 2017, "The Macro-Dynamics of Sorting between Workers and Firms," American Economic Review 107, 1104-1135.

[17] Lucas, R.E., Jr., 1967, "Adjustment Costs and the Theory of Supply," Journal of Political Economy 75, 321-334.

[18] Menzio, G. and S. Shi, 2011, "Efficient Search on the Job and the Business Cycle," Journal of Political Economy 119, 468-510.

[19] Moen, E.R., 1997, "Competitive Search Equilibrium," Journal of Political Economy 105, 385-411.

[20] Montgomery, J.D., 1991, "Equilibrium Wage Dispersion and Interindustry Wage Differentials," Quarterly Journal of Economics 106, 163-179.

[21] Peters, M., 1991, "Ex Ante Price Offers in Matching Games: Non-Steady State," Econometrica 59, 1425-1454.

[22] Peters, M. and A. Siow, 2002, "Competing Premarital Investments," Journal of Political Economy 110, 592-609.

[23] Pissarides, C.A., 2009, "The Unemployment Volatility Puzzle: Is Wage Stickiness the Answer?" Econometrica 77, 1339-1369.

[24] Shi, S., 2001, "Frictional Assignment I: Efficiency," Journal of Economic Theory 98, 232-260.

[25] Shi, S., 2002, "A Directed-Search model of Inequality with Heterogeneous Skills and Skill-Biased Technology," Review of Economic Studies 69, 467-491.

[26] Shi, S., 2005, "Frictional Assignment, Part II: Infinite Horizon and Inequality," Review of Economic Dynamics 8, 106-137.

[27] Shi, S., 2018, "Efficient Investments, Job Dynamics and Frictional Dispersion," SSRN working paper (http://ssrn.com/abstract=3172147).

[28] Shimer, R., 2005, "The Cyclical Behavior of Equilibrium Unemployment and Vacancies," American Economic Review 95, 25-49.

[29] Shimer, R. and L. Smith, L., 2000, "Assortative Matching and Search," Econometrica 68, 343-370.

Supplementary Appendix for

"Post-Match Investment and Dynamic Sorting between Capital and Labor"

Shouyong Shi

2018

C. Procedures of Calibration and Computation

The model is calibrated at the monthly frequency. The discount rate r is determined by the target, $(1+r)^3 = 1.0125$. The target $\alpha = 0.55$ implies $\frac{rk^*(h)}{\tilde{f}(k^*(h),h)} = \alpha$ for all h. The condition, $f_1(k^*(h), h) = 0$, implies $f_0 = \frac{r}{\alpha}\left(\frac{k^*(h)}{h}\right)^{1-\alpha}$, which determines f_0 under the normalization $k^*(h_H) = 100$. The calibration target on home production yields $f_u = 0.29 f^*(h_a)$.

The monthly transition rate from employment to unemployment is $\delta = 0.026$, which is taken from the Current Population Survey (CPS). The average unemployment rate is targeted to 0.065. Let $\hat{h}$ be the skill of workers at this average u; i.e., $u(\hat{h}) = 0.065$. For such workers, the job-finding rate the steady state yields $p(\hat{h}) = \delta(\frac{1}{u(\hat{h})} - 1) = 0.374$ and $\frac{1}{\varepsilon(\hat{h})} \equiv \frac{\mathrm{d}\ln p(\hat{h})}{\mathrm{d}\ln \theta(p(\hat{h}))} = 1 - p^{\rho}(\hat{h})$. The targets $\frac{1}{\varepsilon(\hat{h})} = 0.39$ yields $\rho = 0.5$. Since this step has used two targets, $\left(u(\hat{h}), \varepsilon(\hat{h})\right)$, and determined only one parameter, ρ, it leaves the target on u to be used later. With the target $\theta(\hat{h}) = 0.72$, the matching function yields $p_0 = \theta(\hat{h})\left[p^{-\rho}(\hat{h}) - 1\right]^{1/\rho} = 0.29$.

The remaining parameters, (ψ_1, ψ_0, c_1), minimize the distance between the model and three targets: the unemployment rates at h_L, h_a and h_H. The distance measure is:

$$\begin{aligned} drest\ \ (\%) \equiv\ & \Big[a_1\left[u\left(h_L\right) - 0.15\right]^2 + a_2\left[u\left(h_a\right) - 1\right]^2 \\ & + a_3\left[u\left(h_H\right) - 1\right]^2\Big]^{1/2} \times 100\% \end{aligned}$$

where (a_1, a_2, a_3) are weights on the targets, with $a_1 + a_2 + a_3 = 1$. I use $a_1 = 0.5$, $a_2 = 0.2$ and $a_3 = 0.3$. The target on the unemployment rate at h_a is given relatively higher weights in order for the average unemployment rate to be close to the target 0.065.

To compute the social optimum, I substitute $V_1(\phi, h) = c'(I(\phi, h))$ from (3.1) into (A.6) to obtain:

$$\Delta(\phi, h) = \frac{f(\phi, h) - f_u + \psi(\phi)\,\theta(p) + \left[c'(I)\,I - c(I)\right]}{r + \delta + p}, \tag{C.1}$$

where $I = I(\phi, h)$ is given by (3.7). Substituting this expression for Δ and $V_1(\phi, h) = c'(I(\phi, h))$, I rewrite (3.8) and (3.9) as

$$c'(I(\phi, h)) = \psi'(\phi) \frac{\theta(p)}{p} \tag{C.2}$$

$$\frac{f(\phi, h) - f_u + \psi(\phi)\theta(p) + [c'(I) I - c(I)]}{r + \delta + p} = \psi(\phi)\theta'(p). \tag{C.3}$$

These two equations determine (ϕ, p). In fact, p can be solved first as a function of ϕ from (C.2) and then substituted into (C.3) to solve ϕ.[25] Once (ϕ, p) are solved, I can compute $\{k(t), i(t)\}_{t \geq 0}$ by (3.6) and compute the distribution of workers and capital according to section 4.

[25]Equations (C.2) and (C.3) can also be used to prove Proposition 3.3. Doing so generates expressions for $\phi'(h)$ and $p'(h)$ that are equivalent to those in (A.3), with V_1 and V_{11} being obtained from differentiating (A.6) with respect to ϕ.