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

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

This version: February 2020

Abstract

By integrating frictional sorting into the neoclassical framework with post-match investment by firms, this paper analyzes the social optimum and the equilibrium. A main theoretical result is that post-match investment restores positive assortative matching (PAM) to the social optimum in the long run, but increases the likelihood of negative assortative matching (NAM) at the beginning of a match. The stronger is the complementarity between capital and the skill in production, the *more likely is NAM* to occur at the beginning of a match. I calibrate the model to show that sorting is NAM initially but becomes PAM on average over time. Workers with a higher skill have higher labor productivity and wages throughout the match. They also have steeper time profiles of these variables. There is sizable within-skill inequality in labor productivity and wages. Moreover, an increase in the matching efficiency reduces the labor share and increases the skill premium.

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

* Address: Department of Economics, Pennsylvania State University, 502 Kern Building, University Park, PA 16802, USA. An earlier version of this paper was presented at Rutgers, Virginia, Peking, Queen's, and the Federal Reserve Bank of Philadelphia. I thank the participants for comments. All errors are mine.

Electronic copy available at: https://ssrn.com/abstract=3240652

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). This result is well known when there are no matching frictions (Becker, 1973). When matching frictions exist, neither does PAM necessarily maximize social welfare (Shi, 2001, 2005), nor is PAM necessarily an equilibrium outcome (Shimer and Smith, 2000). Because matching frictions are significant, as evidenced by the existence of persistent unemployment, the failure of PAM casts serious doubt on the neoclassical theory of production.[1]

The potential inefficiency of PAM arises from the tradeoff between the matching rate and productivity. PAM increases labor productivity, but can reduce the matching rate for high skills. The reason is that the vacancy cost increases in the capital stock, as a firm must put some of the capital in place before hiring a worker. PAM increases the vacancy cost for hiring a high skill. If the vacancy cost increases sufficiently in the capital stock, negative assortative matching (NAM) can increase social welfare. By reducing the capital stock for a vacancy targeting a high skill, NAM increases the number of vacancies and, hence, the matching rate for a high skill.

This intuitive explanation applies to the sorting pattern when a match is formed. However, the capital stock a firm allocates to a worker is not a fixed effect. Hiring a worker is only the beginning of a match. After the hiring, a firm can invest to increase the capital stock. The average sorting pattern over time may be opposite to that at the beginning of a match. To understand and identify 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 the source and the dynamics of endogenous heterogeneity in the earnings profile between skills and within

[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).

Electronic copy available at: https://ssrn.com/abstract=3240652

each skill. It is surprising that the voluminous research on sorting has not constructed such an integrated model.[2]

The contribution of this paper is to integrate frictional sorting into the neoclassical framework with post-match investment by firms to analyze the social optimum and the equilibrium. On the theoretical side, I prove that post-match investment restores PAM to the social optimum in the long run, but increases the likelihood of NAM at the beginning of a match. The stronger is the complementarity between capital and the skill in production, the *more likely is NAM* to occur at the beginning of a match. On the quantitative side, I calibrate the model to show that the model provides a unified mechanism for explaining several facts such as the heterogeneous slope of the earnings profile, within-group inequality, and the declining labor share.

Section 2 describes the 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. A worker's productivity is measured by *net output*, defined as output in a match minus the rental cost of capital. Firms rent capital from the rest of the world at a constant rental rate. They can create vacancies with different levels of capital to target different skills, i.e., to direct workers' search. The vacancy cost increases in the initial capital stock for the job. After filling a vacancy, the firm chooses investment to adjust the capital stock, with an adjustment cost 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.[3]

In section 3, I characterize the social optimum that maximizes the sum of expected surpluses with the same matching frictions as in the market. For unemployed workers, the planner chooses the number of vacancies and the capital stock for each vacancy to target each skill. The choice of the number of vacancies is equivalent to the choice of a matching

[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 and sorting.

Electronic copy available at: https://ssrn.com/abstract=3240652

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 employed workers, 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 of a match. Specifically, the option to invest after a match is formed 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: it is higher for a high skill and increases with the complementarity. Thus, stronger complementarity between capital and the skill *increases* the likelihood of NAM in the initial assignment. In contrast, in sorting models without post-match investment, increasing the complementarity increases the likelihood of PAM.[4]

Post-match investment follows dynamics described in a macroeconomic textbook. Given the 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 neoclassical and hence PAM, because it equates the marginal productivity of capital to the rental rate of capital.

The social optimum yields interesting predictions on the time profile and the distribution of labor productivity measured by net output. Suppose that capital and the skill are strongly complementary so that the initial assignment is NAM. Because of the initial NAM, a high skill has a lower capital stock initially than a low skill has. But because the final assignment is PAM, the path of the capital stock for a high skill will eventually overtake 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 initial dispersion in net output between skills, but amplifies 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. This distribution of

[4]In this model, firms delay some investment and NAM can arise not because firms first hire workers and then figure out who their workers are. To the opposite, because search is directed, all firms know exactly the skill they will be matched with. NAM arises precisely because firms offer each skill with the optimal tradeoff between the initial capital stock and the matching rate.

Electronic copy available at: https://ssrn.com/abstract=3240652

capital stocks among workers of the same skill gives rise to two notions of overall sorting in the economy. PAM-D is the first-order stochastic dominance in the distribution and PAM-M is the increasing mean of the distribution. Section 4 provides the conditions for each notion of PAM to hold.

Section 5 characterizes the market equilibrium that implements the planner's allocation. In the market, recruiting firms create vacancies and direct workers' search. For each skill, an offer consists of the expected lifetime utility to a worker and the initial capital stock of a job. After a match is formed, the firm chooses investment, and the two sides bargain over wages to deliver the lifetime utility promised to the worker.

I calibrate the model in section 6 to illustrate several quantitative results. First, the initial assignment is NAM, and the matching rate increases in the skill. Second, the within-skill mean of the capital stock increases in the skill; i.e., PAM-M holds. Convergence toward the neoclassical final state exerts a strong force on dynamic sorting by reverting the pattern from NAM to PAM in the transition. However, the mean of the capital stock for each skill is substantially lower than in the final state. Thus, the model's quantitative behavior differs significantly both from frictional sorting models without post-match investment and from the neoclassical model without matching frictions.

Third, despite the initial NAM, workers with a higher skill have higher net output and wages throughout the match. They also have steeper profiles of these variables over time. In any cohort of workers who become employed at the same time, the variance in net output increases over time. The model provides an explanation for the empirical regularity 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 wages 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).

Fourth, the model generates significant inequality in net output and wages within each skill. Weighted by the employment share of each skill, the average within-skill coefficient of variation is 0.22 in net output and 0.17 in wages. The average mean-min ratio within a

Electronic copy available at: https://ssrn.com/abstract=3240652

skill, a measure of inequality proposed by Hornstein et al. (2011), is 2.3 in net output and 1.75 in wages. 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).

Finally, an increase in the matching efficiency reduces the labor share and increases the skill premium, as explored in section 6.5.

There is a large literature on sorting originated in Becker (1973). As mentioned earlier in this Introduction, the literature has treated the firm type as a fixed effect and focused on sorting at the time of a match. Also, a small number of recent papers (e.g., Lise and Robin, 2017) have examined 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.

Some papers have examined the effect of pre-match investment on sorting, e.g., Burdett and Coles (2001), Peters and Siow (2002), Chade and Lindenlaub (2017). Pre-match investment is intended to increase the probability of matching with the desired type. Because investment ends with the formation of a match, these models do not provide a theory of how sorting or within-group inequality changes over time after a match is formed. In contrast, post-match investment is intended to increase productivity within a match. It tends to make the initial assignment NAM and generates dynamics of sorting over time.[5]

2. The Model

2.1. Model Environment

Time is continuous with an infinite horizon. There is a unit measure of risk-neutral 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. Competitive entry determines the measure of firms. Firms are homogeneous ex ante but become heterogeneous ex post in the capital stock, k. A

[5]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).

Electronic copy available at: https://ssrn.com/abstract=3240652

firm needs one worker, as in most models of labor market search. I refer to h as the skill, to k as a firm's (endogenous) type, and to a match between the capital stock k and the skill h as a (k,h) match. In a (k,h) match, the output flow is $\tilde{f}(k,h)$. The 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, $\tilde{f}_{12}(k,h) > 0$ so that k and h are *strictly complementary* with each other in production.

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 $f(k^*,h) > 0$. This is the optimal capital stock for skill h in the neoclassical theory when the labor market is frictionless. Denote $f^*(h) = f(k^*(h),h)$. For an unemployed worker, home production is $f_u \geq 0$, which is constant over time.[6]

A firm chooses the time path of capital stocks. At the time of hiring, a firm can choose any capital stock k but must rent at least a fraction of capital in advance for the job. The flow cost of a vacancy, denoted as $\psi(k)$, includes this rental cost of capital and other costs of the vacancy. The vacancy cost satisfies $\psi'(k) > 0$ and $\psi''(k) > 0$ for all $k > 0$, with $\psi(0) = \psi'(0) = 0$. The assumptions $\psi' > 0$ and $\psi'' \geq 0$ are easily satisfied because the rental cost of capital for a vacancy is linear in k. Strict convexity of ψ is not critical for the analytical part of the paper other than easing the task of ensuring uniqueness of the social optimum, but it is important for the quantitative analysis. Strict convexity of ψ is realistic. In addition to the rental cost, a firm may need to incur costs to maintain vacant capital rented for a vacancy. The marginal cost of such maintenance can be increasing. More broadly, one can interpret a job with a higher k as a better job, as in Acemoglu (2001), because such a job yields higher output for any given skill. The better is a job, the increasingly more resources a firm needs to devote to recruiting.

Once a vacancy is filled, the match generates output. A firm can adjust the capital

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

Electronic copy available at: https://ssrn.com/abstract=3240652

stock by undertaking investment, denoted as i. The capital stock evolves according to

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

It is costly to put new capital to work together with existing capital. The adjustment cost of investment is $c\left(i\right)$, which has the standard properties: $c\left(\infty\right)=\infty$, $c\left(0\right)=c'\left(0\right)=0$, $c'\left(i\right)>0$ for all $i>0$, and $c''\left(i\right)>0$ for all $i\geq 0$. With the adjustment cost, a firm can adjust the capital stock of an existing job only smoothly. To focus on capital accumulation, instead of decumulation, I assume $c\left(i\right)=c'\left(i\right)=0$ for all $i<0$ so that there is no adjustment cost to reduce the capital stock. This assumption is unimportant for the analysis because $i\geq 0$ in the entire duration of a match.[7] Note that vacancy creation is a unique time in a firm's capital formation – it is the only time when the firm can choose whatever level of capital 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. 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.

Unemployed workers search for jobs. Firms can direct workers' search by offering a matching rate and a capital stock to each skill. The searchers of the same skill and the vacancies targeting them form a submarket. In any submarket where the matching rate for a searcher is p, the number of vacancies per searcher is $\theta\left(p\right)$, and the matching rate for a vacancy is $q\left(p\right)=\frac{p}{\theta(p)}$. The function $\theta\left(p\right)$ summarizes matching frictions and is implied by a matching function. Precisely, if the matching rate function for a searcher is $P\left(\theta\right)$, then

[7] Job separation is the only time when a firm reduces the capital stock. If there is a cost to reduce the capital stock, a firm may choose a lower capital stock in anticipation of job separation.

Electronic copy available at: https://ssrn.com/abstract=3240652

$\theta(p)$ is the inverse function of P. Let $\bar{p} \le \infty$ be the natural upper bound on p. Assume:

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

Denote $\varepsilon(p) = \frac{p\theta'(p)}{\theta(p)}$ (> 1).[8]

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 expected surpluses, 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.[9] Because a worker's skill and home production are constant over time, $V_u(h)$ is constant over time.

For each skill h, the planner chooses a particular submarket (p, ϕ) to direct unemployed workers' search, where p is the matching rate for a worker and ϕ the capital stock of a vacancy. To deliver the matching rate p, the number of vacancies per searcher must be $\theta(p)$. Each vacancy incurs a flow cost $\psi(\phi)$. Because a searcher of skill h finds a match at the rate p and a (ϕ, h) match 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 efficient choices (p, ϕ) maximize the expected social surplus of search and yields:

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

Denote the optimal choices by the policy functions, $(p(h), \phi(h))$. The function ϕ describes the initial assignment of capital to the skill.

[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 $\lim_{\theta \to \infty} \frac{P(\theta)}{\theta} = 0$.

[9]If search is random instead, value functions can be nonstationary because a worker's acceptance set for a match may exhibit cycles (see Shimer and Smith, 2001). This does not occur when search is directed, because firms and workers can choose which type to search for.

Electronic copy available at: https://ssrn.com/abstract=3240652

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 adjustment cost of investment is $c(i(t))$. Thus, for any $t \geq \tau$, investment maximizes the return on investment and yields:

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

subject to the constraint (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.[10] It is useful to express the planner's choice of investment as a choice of the path $\{i(t)\}_{t \geq \tau}$. Integrating (2.3) over time generates:

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

subject to (2.1) for all $t \geq \tau$. The initial capital stock at τ is given as $k(\tau)$.[11]

There is a non-degenerate distribution of capital stocks among employed 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. 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.

[10] An assumption implicit in (2.3) is that a firm incurs the adjustment cost to make room for new capital but rents new capital with an instant delay. The rental cost of new capital appears in $f(k(t_+), h)$. This assumption makes the neoclassical final state attainable. If the rental cost of new capital is incurred without delay, then $c(i(t))$ should be re-interpreted as the sum of the rental cost of new capital and the adjustment cost. In this case, $c'(0_+) = r$ instead of 0, and the final stock of capital should be redefined by $f_1(k^*(h), h) = (r+\delta) c'(0_+)$. The modifications to the analytical and quantitative results are very small.

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

Electronic copy available at: https://ssrn.com/abstract=3240652

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.[12]

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:

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

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

These conditions are intuitive. If 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).[13] 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:

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

The marginal benefit of investment is equal to the present value of the marginal productivity of capital, $\frac{f_1}{r+\delta}$. If this marginal benefit exceeds the marginal cost of investment, $c' (i)$, efficient investment must be high currently and, hence, declining over time. Conversely, if the marginal benefit of investment is lower than the marginal cost, then efficient investment must be low currently and, hence, increasing over time.

Equations (3.3) and (2.1) form a 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,

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

[13] 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.

Electronic copy available at: https://ssrn.com/abstract=3240652

$f_1(k^*, h) = (r+\delta) c'(0) = 0$, as defined earlier. To study the dynamics of sorting, I linearize the dynamic system around the final state:[14]

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

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. *The final state has PAM, 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.

The final state has PAM because the marginal productivity of capital is equal to the rental rate. With complementarity between the two factors of production, a higher skill increases the marginal productivity of capital. The capital stock must increase to reduce the marginal productivity of capital back to the level of the rental rate.

[14]The complementary Appendix E computes the non-linearized model with the parameter values calibrated in section 6.1. The results are very close to those in the linearized model. In fact, the non-linearized results indicate that the linearization under-represents the main messages of the model.

Electronic copy available at: https://ssrn.com/abstract=3240652

Post-match investment is important for the final assignment to be positive assortative. As shown later, if post-match investment is not possible, search frictions can make the assignment "stuck" in a non-PAM or even NAM initial assignment. Post-match investment restores PAM to the efficient assignment 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 can benefit from the increased capital stock.[15] Investment is inversely related to the gap between the current and the final stock (see (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 and the concavity of the output function. If the adjustment cost is more convex in investment, 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 both cases, the capital stock increases more slowly toward the final stock.

The skill of the worker in a match affects $-f_{11}(k^*(h), h)$ 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:* $\tilde{f}(k, h) = A_0 \left[\alpha k^\xi + (1-\alpha) h^\xi\right]^{1/\xi}$ *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* $-f_{11}(k^*(h), h) = A_1/h$, *where* A_1/h *is a decreasing function of* h.

Slower convergence to the final state for a higher skill does not necessarily imply lower investment in such a match. 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

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

Electronic copy available at: https://ssrn.com/abstract=3240652

the gain from investment is greater. Investment is higher for a high-skill worker than for a low-skill worker. The slower convergence of k for a high-skill reflects higher persistence of investment. I will establish this result in Proposition 3.4 later.

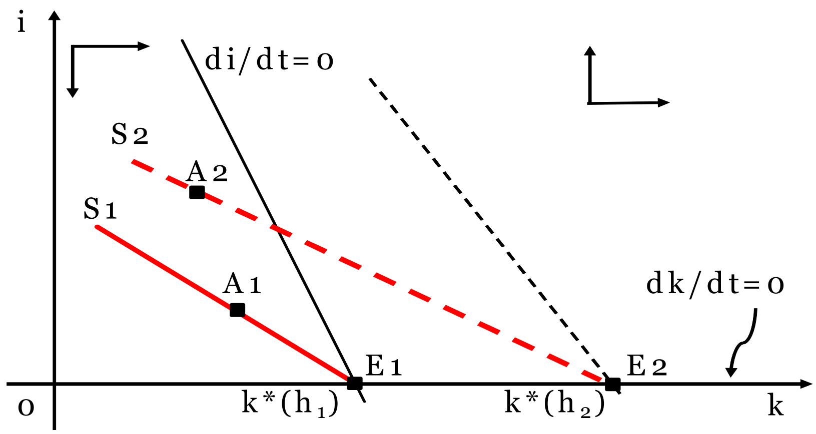

Figure 1. Dynamics of investment and the capital stock

Figure 1 is the phase diagram of the dynamic system (3.4) for two skill levels, 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 $\frac{dk}{dt} = 0$. The schedule $\frac{di}{dt} = 0$ 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.[16] 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 endogenous 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 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.

[16]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.

Electronic copy available at: https://ssrn.com/abstract=3240652

3.2. Efficient Search: Existence, Uniqueness, and $\phi(h) < k^*(h)$

For each skill h of unemployed workers, the planner chooses the initial capital stock, $\phi(h)$, and the job-finding rate, $p(h)$, to solve the search problem in (2.2). The expected social surplus of a (ϕ, h) match is $p(h)\Delta(\phi, h) - \psi(\phi)\theta(p)$, where $\Delta(\phi, h)$ denotes the social gain from a (ϕ, h) match:

$$\Delta(\phi, h) \equiv V(\phi, h) - V_u(h). \tag{3.8}$$

If post-investment is exogenously set to zero, a marginal increase in p from 0 increases the expected social surplus of a (k, h) match by

$$\tilde{S}(k, h) \equiv \frac{f(k, h) - f_u}{r + \delta} - \theta'(0)\psi(k). \tag{3.9}$$

I focus on the nontrivial case where $p(h) > 0$ and $\phi(h) > 0$ for all h. For this focus and uniqueness of the social optimum, I impose:

Assumption 1. *(i) For every h, $\max_{k\in[0,k^*(h)]} \tilde{S}(k, h) > 0$. (ii) Define $\underline{k}(h)$ $(< k^*(h))$ as the solution to:*

$$[\psi(k)V_1(k, h) - \psi'(k)\Delta(k, h)]_{k=\underline{k}(h)} = 0. \tag{3.10}$$

For all h and all $k \in (\underline{k}(h), k^*(h))$, the following inequality holds:

$$\frac{I\,c''(I)}{c'(I)}\frac{k}{k^*(h) - k} + \frac{k\psi''}{\psi'} - \frac{k\psi'}{\psi}\frac{(\varepsilon - 1)^2\theta}{p^2\theta''} > 0 \tag{3.11}$$

where $\psi = \psi(k)$, $I = I(k, h)$, $\varepsilon = \frac{p\theta'(p)}{\theta(p)}$ and $p = \theta'^{-1}(\frac{\Delta(k,h)}{\psi(k)})$.

Part (i) of Assumption 1 ensures $p(h) > 0$ for all h. Because post-match investment increases the social value of a match, a marginal increase in p from 0 increases the expected social surplus of a (k, h) match by more than $\tilde{S}(k, h)$. If $\tilde{S}(k, h) > 0$ for some k, then there exist $\phi > 0$ and $p(h) > 0$ that yield a positive expected social surplus. This assumption also implies $\phi(h) > \underline{k}(h)$, where $\underline{k}(h)$ is defined by (3.10). Part (ii) of Assumption 1 requires the planner's objective function to be strictly quasi-concave at all local maxima.

Electronic copy available at: https://ssrn.com/abstract=3240652

This assumption is needed to ensure the social optimum to be unique for each h, because the planner's objective function contains the term $-\psi(\phi)\theta(p)$ that is not jointly concave in (ϕ, p). Intuitively, the assumption is satisfied if either the adjustment cost $c(i)$ or the vacancy cost $\psi(k)$ is sufficiently convex. Note that even if $\psi'' = 0$, sufficient convexity of $c(i)$ can still make (3.11) satisfied.

Appendix A proves the following proposition:

Proposition 3.3. *If (i) in Assumption 1 holds, then $(p(h), \phi(h))$ are interior and distinct for each h. In particular, $\phi(h) \in (\underline{k}(h), k^*(h))$, and so $I(\phi(h), h) > 0$, for all h. The efficient choices of search solve:*

$$\text{for } \phi\text{: } V_1(\phi, h) = \psi'(\phi)\frac{\theta(p)}{p} \tag{3.12}$$

$$\text{for } p\text{: } \Delta(\phi, h) = \psi(\phi)\theta'(p). \tag{3.13}$$

For any h_1 and $h_2 \neq h_1$, either $\phi(h_1) \neq \phi(h_2)$, or $p(h_1) \neq p(h_2)$, or both. If (ii) in Assumption 1 holds in addition to (i), then $(p(h), \phi(h))$ are unique for each h.

The conditions (3.12) and (3.13) are intuitive. (3.12) equates the social marginal benefit of capital in the initial assignment, $pV_1(\phi, h)$, to the social marginal cost of capital in recruiting, $\psi'(\phi)\theta(p)$. (3.13) equates the social marginal benefit of increasing the matching rate, $\Delta(\phi, h)$, to the marginal cost of vacancies, $\psi(\phi)\theta'(p)$.

The initial assignment of capital to a skill is strictly lower than the final assignment, i.e., $\phi(h) < k^*(h)$ **for every** h. This result is necessary and sufficient for post-match investment to be positive (see (3.7)). The cause of $\phi(h) < k^*(h)$ is $\psi'(k) > 0$ in addition to matching frictions. If the initial capital stock is equal to the final stock, the marginal benefit of capital is zero, because the marginal net output of capital and post-match investments are both zero at the final stock. Since the marginal cost of capital is always positive under $\psi'(k) > 0$, there is a net gain from reducing the capital stock for a vacancy. Thus, with $\psi'(k) > 0$, the social optimum uses both the initial assignment and post-match investment to build capital toward the final stock.

Electronic copy available at: https://ssrn.com/abstract=3240652

The efficient choices of search are distinct for each skill level. This is an implication of the complementarity between capital and skill. Because the marginal productivity of capital increases in the skill, the tradeoff between the initial capital stock and the job-finding rate varies with the skill. This variation makes it socially efficient to separate different skills into different submarkets. A higher skill should either match with a higher capital stock, or match more quickly, or both.

Note that strict convexity $\psi'' > 0$ does not play a role for the results discussed so far: $\phi(h) < k^*(h)$, $I(\phi, h) > 0$, and the distinct choices $(\phi(h), p(h))$ for each h. However, $\psi'' > 0$ will be important quantitatively for the calibration in section 6.1.

At this point, it is useful to ask: what will happen in the model if matching is frictionless but there is exogenous separation? No interesting dynamics will occur. To see this, recall that a worker's matching rate is $p = P(\theta)$. The absence of matching frictions leads to $P(0) = \bar{p}$, where $\bar{p} \leq \infty$ is the natural upper bound on p. In this case, the function $P(\theta)$ is not invertible to generate the inverse $\theta(p)$. The condition (3.13) becomes an inequality "$>$", which implies the corner solution $p = \bar{p}$. The condition (3.12) becomes $V_1(\phi, h) = 0$, which implies $f_1(\phi, h) = 0$ and, hence, $\phi(h) = k^*(h)$. The initial assignment of capital to each skill jumps immediately to the final state.

3.3. Initial Sorting and the Matching Rate

The initial assignment has PAM if and only if $\phi'(h) > 0$. A higher skill has a higher matching rate if and only if $p'(h) > 0$. The following proposition characterizes the dependence of (p, ϕ) on h, and Appendix A gives a proof:

Proposition 3.4. *Maintain Assumption 1. 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,*

$$\phi'(h) > 0 \quad \textit{iff} \quad \frac{(a_2 - I) f_{12}(\phi, h)}{r + \delta + \beta} > f_2(\phi, h) \tag{3.14}$$

$$p'(h) > 0 \quad \textit{iff} \quad \frac{(a_1 - I) f_{12}(\phi, h)}{r + \delta + \beta} < f_2(\phi, h), \tag{3.15}$$

Electronic copy available at: https://ssrn.com/abstract=3240652

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

$$a_1 = \frac{(r+\delta+p)(\varepsilon-1)}{\psi''/\psi' + \beta c''(I)/c'(I)}, \qquad a_2 = \frac{(r+\delta+p)\,p\psi\theta''}{(\varepsilon-1)\,c'(I)}. \tag{3.16}$$

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

In the social optimum, a higher skill has a higher social value, both in employment and in unemployment. However, a higher skill is not always assigned to match with a higher capital stock even though capital and the skill are complementary in production. There are two ways to capture the high social surplus from a more productive worker. One is PAM and the other is a higher matching rate. The social optimum uses either PAM or a higher matching rate for a higher skill, or both. It is never socially efficient to assign both a lower capital stock and a lower matching rate to a higher skill. Given post-match investment, there can be three cases of the initial assignment:

Case 1: $\phi' < 0$ and $p' > 0$. This case occurs when the complementarity between the two factors of production is weak. In this case, the marginal 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 as to create many low-capital vacancies for high-skill workers to increase their matching rate.

Case 2: $\phi' > 0$ and $p' > 0$. This case occurs when the two factors of production 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 occurs when the two factors of production are strongly complementary with each other. In this case, it is socially efficient to capture the strong complementarity through 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.

Post-match investment can affect which of the three cases occurs. To show this effect, Table 1 divides investment into three intervals. In each interval, Table 1 lists the initial assignment and the matching rate according to the complementarity between the

Electronic copy available at: https://ssrn.com/abstract=3240652

two factors, measured by $\frac{f_{12}(\phi,h)}{r+\delta+\beta}$. The two threshold levels of the complementarity are $\frac{f_2}{a_1-I}$ and $\frac{f_2}{a_2-I}$, where $a_2 > a_1 > 0$ are defined by (3.16). 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 above can occur. In the middle 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. Only Cases 1 and 2 can occur. At the bottom of Table 1, investment is high. Because $a_1 - I < a_2 - I \leq 0$, both thresholds become inapplicable. Only Case 1 can occur, which has NAM in the initial assignment.

Post-match investment increases the benefit of a higher matching rate relative 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. Even if this may require the initial assignment to be NAM, post-match investment can restore 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.

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

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

The complementarity between k and h increases the likelihood of NAM in the initial assignment. This result is opposite to sorting models without post-match investment, e.g., Shi (2001, 2005) and Eeckhout and Kircher (2010). The cause of this result is that investment is likely to be higher when capital and the skill are more complementary with each other. With high investment to come after a match is formed, there is little

Electronic copy available at: https://ssrn.com/abstract=3240652

benefit for high-skill workers to wait for a long time just for PAM. 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 with $\phi'(h) > 0$ and $p'(h) < 0$ may not exist at all in the social optimum, because the matching features are inconsistent with high investment.

NAM in the initial assignment has an important implication for the dynamic paths of capital stocks for different skill levels: These paths must cross each other over time. Figure 2 depicts the time paths of capital stocks for skills h_1 and h_2, with $h_2 > h_1$. With the initial NAM, the path for the higher skill h_2 starts below the path for the lower skill h_1. The former overtakes the latter at the time point t_0, because the final state has PAM. Sorting between capital and the skill is negative for $t \in [0, t_0)$ and positive for $t > t_0$. This crossing implies further that matching is mixed in the population at any given point of time. Even though 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. I turn to the distribution of workers next.

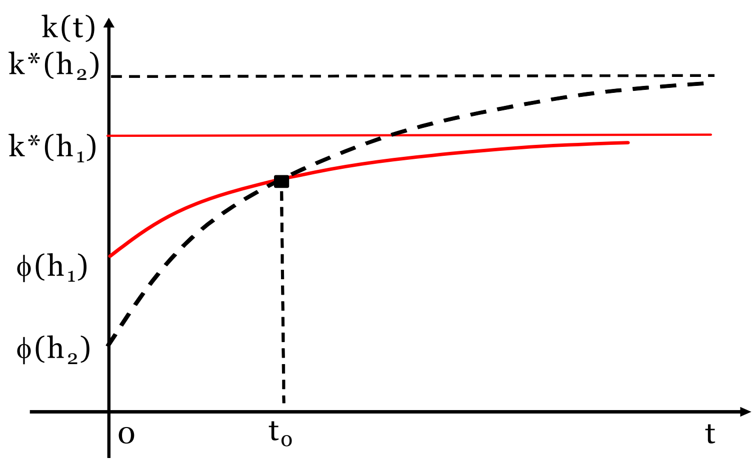

Figure 2. Crossing paths of capital stocks for two types of workers

Electronic copy available at: https://ssrn.com/abstract=3240652

4. Dynamic Sorting and the Stationary Distribution

Unemployed 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 succeed earlier have 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. This process generates a distribution of workers of the same skill over capital stocks. I characterize the ergodic distribution and the overall sorting pattern.

Among the workers of any given h, let the fraction of unemployed be $u(h)$ and the fraction of employed be $n_e(h) = 1 - u(h)$. In the steady state, the flow of unemployed skill-h workers exiting unemployment, $u(h)\,p(h)$, is equal to the flow of employed workers of skill h into unemployment, $\delta n_e(h)$. This equality determines

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

It is clear that $u'(h) < 0$ if and only if $p'(h) > 0$. Among employed workers of skill h, let $G(k|h)$ be the cumulative distribution over k. The conditional mean of capital stocks is:

$$E(k|h) \equiv \int_{\phi(h)}^{k^*(h)} k \;\mathrm{d}G(k|h)\,. \tag{4.2}$$

Definition 4.1. *Capital and skills have PAM-D (positive assortative matching in the distribution) if $G(k|h)$ is decreasing in h for all $k \in (\phi(h), k^*(h))$. Capital and skills have PAM-M (positive assortative matching in the mean) if $E(k|h)$ is increasing in h.*

The following proposition is proven in Appendix B:

Proposition 4.2. *Approximate the dynamic system of (i, k) by (3.4). Then,*

$$G(k|h) = 1 - \left[\frac{k^*(h) - k}{k^*(h) - \phi(h)}\right]^{\frac{\delta}{\beta(h)}} = 1 - [R(k,h)]^{\frac{\delta}{\beta(h)}}\,, \tag{4.3}$$

$$E(k|h) = \frac{\delta\phi(h) + \beta(h)\,k^*(h)}{\beta(h) + \delta} = \phi(h) + \frac{I(\phi(h), h)}{\beta(h) + \delta}, \tag{4.4}$$

Electronic copy available at: https://ssrn.com/abstract=3240652

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*

$$R(k,h)\,\phi'(h) + [1 - R(k,h)]\,k^{*\prime}(h) > \frac{\beta'(h) I(k,h)}{[\beta(h)]^2} \ln R(k,h) \quad \text{for all } k \in (\phi(h), k^*(h)). \tag{4.5}$$

$\phi'(h) \geq 0$ *is necessary and sufficient for (4.5) to hold. However,* $\phi'(h) \geq 0$ *is sufficient, but may not be necessary, for PAM-M.*

The density of capital among workers of the same skill h is decreasing if and only if $\beta(h) < \delta$. This result is intuitive. Among employed workers of skill h, the fraction with capital stocks in the upper tail $[k, k^*(h)]$ is $1 - G(k|h)$. Workers exit this upper 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 interval $[k, k^*(h)]$. When k increases, this width narrows, which reduces the size of both the exit and the entry into the upper tail. 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 must be decreasing.[17] Conversely, if $\beta(h) > \delta$, the density must be increasing.[18]

The condition (4.5) 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 either the initial or the final assignment contributes to PAM-D by increasing the capital stock assigned to a higher skill at one end of the dynamic path. Also contributing to PAM-D is faster convergence to the final state, i.e., $\beta'(h) > 0$, which allows the capital stock for a higher skill to increase more quickly. It is clear that the initial PAM is necessary for PAM-D, but surprising that the initial PAM is

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

[18] Define $z = [k^*(h) - k]^{-1}$. Then, (4.3) 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.

Electronic copy available at: https://ssrn.com/abstract=3240652

also sufficient for PAM-D. The explanation for the sufficiency is that, for any given capital stock, investment increases in the skill, i.e., $I_2(k,h) > 0$ for all $k < k^*(h)$ (see Proposition 3.4). If high-skill workers employment with a higher capital stock in the initial assignment, higher investment will keep capital stocks in their matches higher than those for low-skill workers throughout the transition to the final state.

The three determinants of PAM-D above are also the determinants of PAM-M, but their weights for PAM-M differ from those for PAM-D. For any given skill, the mean of capital stocks is a weighted average of the initial and the final assignment (see (4.4)). The weight is $\frac{\delta}{\beta(h)+\delta}$ for the initial assignment and $\frac{\beta(h)}{\beta(h)+\delta}$ for the final assignment. A higher exit rate δ increases the weight for the initial assignment by terminating a match earlier before the capital stock in the match gets closer to the final stock. A higher convergence speed β does the opposite by increasing the capital stock in a match quickly toward the final stock. The second equality in (4.4) expresses the mean of the increase in the capital stock during employment, $E(k|h) - \phi(h)$, as the present value of the initial investment, where the discount rate is $(\beta + \delta)$.

PAM-D implies PAM-M. Because the initial PAM is sufficient for PAM-D, it is also sufficient for PAM-M. However, the initial PAM is not necessary for PAM-M. Even if the initial assignment is NAM, higher investment for a high skill makes capital grow more quickly than for a low skill and, hence, can generate a higher mean of the capital stock.

The above analysis examines the distribution of capital conditional on the skill. It is also useful to examine the joint distribution between capital and the skill. 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

$$s(h) = \frac{n_e(h)H'(h)}{N_e}. \tag{4.6}$$

Let $G(k,h)$ be the joint cumulative distribution function of employed workers over (k,h).

Electronic copy available at: https://ssrn.com/abstract=3240652

Then, the joint density is $G_{12}(k,h) = s(h)\, G'(k|h)$, and so

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

The employment share of a skill accentuates the matching pattern of that skill by allowing such matches to be present 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 3.4, I explained that the social optimum is more likely to have higher matching rates for 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.

5. Equilibrium Implementation of the Social Optimum

In this section I characterize the market equilibrium that implements the planner's allocation. In the market, recruiting firms create vacancies and make offers to direct workers' search, and firms with filled jobs choose investments. After a match is formed, the worker and the firm bargain over wages.

It is convenient to formulate a recruiting firm's offer as a value x, which is the expected lifetime utility to a worker taking future separation into account. A firm delivers an offer by wages formulated later. For each skill h, the market consists of a continuum of potential submarkets indexed by (x, ϕ, h), where ϕ is the initial capital stock of a job. Unemployed workers choose the submarket to search and firms choose the submarket to locate vacancies. They rationally expect a worker's job-finding rate in the submarket to be $p = \tilde{p}(x, \phi, h)$. In any submarket where the job-finding rate is p, the matching technology determines the tightness as $\theta(p)$, which is described in section 2. Moreover, since matching is bilateral, the following accounting identity holds: $q(p) = \frac{p}{\theta(p)}$.

Denote $V_u^e(h)$ as the value for a skill h unemployed worker in the equilibrium. For a match that has lasted for a length t (tenure) reaching the current capital stock $k(t)$, denote $\Omega(k(t), h, t)$ as the worker value. At the time when the worker is recruited in submarket

Electronic copy available at: https://ssrn.com/abstract=3240652

(x,ϕ,h), $\Omega(k(0),h,0)=x$ and $k(0)=\phi$. When searching for a job, a skill h unemployed worker chooses the submarket (x,ϕ,h) to enter to maximize the expected surplus of search, as in the following Bellman equation for V_u^e:

$$rV_u^e(h) = f_u + \max_{(x,\phi)} \tilde{p}(x,\phi,h)\left[x - V_u^e(h)\right]. \tag{5.1}$$

Denote the solution to the maximization problem as $x^e(h)$ for x and $\phi^e(h)$ for ϕ. The job-finding rate at the optimal choice is $p^e(h) \equiv \tilde{p}(x^e(h),\phi^e(h),h)$.

For a recruiting firm in submarket (x,ϕ,h), denote $J(x,\phi,h,0)$ as the firm value at the moment of succeeding in recruiting a worker, where 0 indicates the beginning of a match. The vacancy cost is $\psi(\phi)$ and the vacancy-filling rate is $q(\tilde{p}(x,\phi,h))$. In every submarket (x,ϕ,h) with $J(x,\phi,h,0) \geq \psi(\phi)$, competitive entry of vacancies drives net profit of recruiting to zero, i.e., $q(\tilde{p}(x,\phi,h))\,J(x,\phi,h,0) = \psi(\phi)$. In every submarket with $J(x,\phi,h,0) < \psi(\phi)$, no firm enters and so the tightness is zero, which implies $p=0$. These requirements determine the function $\tilde{p}(x,\phi,h)$ as

$$\tilde{p}(x,\phi,h) = \begin{cases} q^{-1}\left(\frac{\psi(\phi)}{J(x,\phi,h,0)}\right), & \text{if } J(x,\phi,h,0) \geq \psi(\phi) \\ 0, & \text{otherwise.} \end{cases} \tag{5.2}$$

In a match with a skill h worker at tenure t, denote $V^e(k(t),h)$ as the joint value between a firm and the worker.[19] Then, V^e obeys an equation similar to (2.3) for V in the social planner's problem. That is,

$$(r+\delta)V^e(k(t),h) = f(k(t),h) + \delta V_u^e(h) + \frac{\mathrm{d}}{\mathrm{d}t}V^e(k(t),h) - c(i(t)). \tag{5.3}$$

The firm delivers the worker value $\Omega(k(t),h,t)$ by a wage path, $\{w(k(\tau),h,\tau)\}_{\tau\geq t}$, which will be determined later. The worker value Ω satisfies:

$$(r+\delta)\Omega(k(t),h,t) = w(k(t),h,t) + \delta V_u(h) + \frac{\mathrm{d}}{\mathrm{d}t}\Omega(k(t),h,t). \tag{5.4}$$

Given the worker value Ω, the value of the firm in the match at tenure t is:

$$J(\Omega,k(t),h,t) = V^e(k(t),h) - \Omega. \tag{5.5}$$

[19] V^e depends on t only through $k(t)$. If the worker value depends on t directly in addition to $k(t)$, this dependence is canceled by the opposite dependence of the firm value on t directly.

Electronic copy available at: https://ssrn.com/abstract=3240652

Setting $t = 0$, $k(0) = \phi$ and $\Omega = x$ in (5.5) yields the firm value at the moment of successfully recruiting a worker. At each $t \geq 0$, the firm chooses the path of investment, $\{i(\tau)\}_{\tau \geq t}$, to solve the following problem by taking Ω and $k(t)$ as given:

$$\max_{\{i(\tau)\}_{\tau \geq t}} J(\Omega, k(\tau), h, \tau) \text{ s.t. } \frac{\mathrm{d}}{\mathrm{d}\tau} k(\tau) = i(\tau). \tag{5.6}$$

Denote the path of optimal investments in the equilibrium as $\{i^e(t,h)\}_{t \geq 0}$.

In this economy, workers and firms are risk neutral. At any $t > 0$, there are infinitely many wage paths to deliver the promised value $\Omega(k(t), h, t)$ to the worker and the value $J(\Omega, k(t), h, t)$ to the firm. To select one wage path, I assume that wages are continuously bargained between the worker and the firm for all $t > 0$. Specifically, at any $t > 0$, the match surplus is $[\Omega - V_u^e(h)]$ for the worker and $J(\Omega, k(t), h, t)$ for the firm, where $\Omega = \Omega(k(t), h, t)$. The path $\{w(k(\tau), h, \tau)\}_{\tau \geq t}$ solves the following problem:

$$\max [\Omega - V_u^e(h)]^b [J(\Omega, k(t), h, t)]^{1-b}, \text{ for } t > 0, \tag{5.7}$$

where $b \in (0,1)$ is the worker's bargaining weight and $\Omega = \Omega(k(t), h, t)$.

An equilibrium consists of (i) value functions, $[V_u^e(h), V^e(k,h), J(\Omega,k,h,t), \Omega(k,h,t)]$, that satisfy (5.1), (5.3), (5.4) and (5.5), (ii) search policy functions, $[x^e(h), \phi^e(h)]$, that solve the maximization problem in (5.1), (iii) investments, $i^e(t,h)$, that solve (5.6), (iv) the job-finding rate function, $\tilde{p}(x, \phi, h)$, that satisfies (5.2), and (v) the wage function, $w(k,h,t)$, that solves the bargaining problem in (5.7). The proof of the following proposition is in Appendix C:

Proposition 5.1. *The market equilibrium implements the constrained social optimum. In particular, $[p^e(h), \phi^e(h)] = [p(h), \phi(h)]$ and $i^e(t,h) = I(k(t), h)$ for all t and h. The optimal worker value searched by skilled h unemployed workers is:*

$$x^e(h) = V(\phi(h), h) - \frac{\psi(\phi(h))}{q(p(h))}. \tag{5.8}$$

For any tenure $t > 0$, a skill h worker's wage determined by bargaining in (5.7) is:

$$w(k(t), h, t) = b[f(k(t), h) - c(I(k(t), h))] + (1-b) r V_u(h). \tag{5.9}$$

Electronic copy available at: https://ssrn.com/abstract=3240652

The search decision is socially efficient because of directed search and competitive entry of vacancies, similar to the literature, e.g., Moen (1997), Menzio and Shi (2011). Competitive entry of vacancies generates the job-finding rate function as in (5.2), which links the job-finding rate to the offer terms (x, ϕ). In particular, the job-finding rate is a decreasing function of the offer x in all submarkets with $p > 0$. With directed search, searchers take into account this link between p and x. The optimal tradeoff between the two internalizes search externalities. Adding to the literature, Proposition 5.1 shows that the optimal search choice ϕ is also socially efficient. It is clear from (5.1) that ϕ affects a searcher's expected surplus only through the job-finding rate $\tilde{p}(x, \phi, h)$. (5.2) shows that the effect of ϕ on $\tilde{p}$ is hump-shaped because of $\psi' > 0$. As a result, the optimal choice of ϕ is interior.

Investment is socially efficient in the equilibrium for a simple reason. The bargaining protocol in (5.7) is monotone in the sense that the surpluses of both sides of a match increase in the joint surplus. Under any monotone protocol, it is optimal for a firm to choose investment to maximize the joint surplus, which is also what the planner does.

The wage function (5.9) is intuitive. Net income flow of a firm is $f - c$, and a worker's outside option, in terms of flow income, is rV_u. The wage is a weighted average of these two, where the weight on $(f - c)$ is the worker's bargaining weight. Note that the wage $w(k(t), h, t)$ depends on t only through $k(t)$ and not through t directly.

6. Quantitative Analysis

6.1. Calibration and Simulation

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

$$\begin{aligned}
&\tilde{f}(k, h) = f_0 k^{\alpha} h^{1-\alpha}, \quad f_0 > 0, \alpha \in (0,1); \qquad H(h) = \tfrac{h-h_L}{h_H-h_L}, \ h_H > h_L > 0; \\
&\theta(p) = p_0 \left[p^{-\rho} - 1\right]^{-1/\rho}, \quad p_0, \rho \in (0, \infty); \qquad c(i) = c_1 i^2, \quad c_1 > 0; \\
&\psi(k) = \psi_0 k^{\psi_1}, \quad \psi_0 > 0, \ \psi_1 \geq 1.
\end{aligned} \tag{6.1}$$

Net output is $f(k, h) = \tilde{f}(k, h) - rk$. The inverse of $\theta(p)$ is a Dagum (1975) function. I normalize the adjustment cost of capital, $c(i)$, to be quadratic and identify the curvature of the vacancy cost. It is the relative curvature between these two cost functions that is important for the initial capital stock in a match relative to the final stock. The functional

Electronic copy available at: https://ssrn.com/abstract=3240652

form of the distribution of skills, H, has no consequence on the assignment of capital to workers with each skill, because the skill is contractible and the social optimum features separation by the skill. Without loss of generality, I set H to be uniform, where $h_H = 1$ and $h_L = 0.4$. Denote the average value of h in the population as $h_m = \frac{h_H + h_L}{2} = 0.7$. I interpret the skill level as education attainment, where h_H corresponds to a college degree or higher, h_m to a high school degree only, and h_L to no high school degree. The model is calibrated monthly. Table 2 lists the parameters, their values and the calibration targets (see the Supplementary Appendix D for the procedure).

Some aspects of the calibration are worth noting. First, the parameter $\alpha = 0.53$ in the Cobb-Douglas production function results from targeting the average capital share at the commonly used value, 0.36. The capital share in a (k, h) match is $rk/\tilde{f}(k, h) = \frac{r}{f_0}\left(\frac{k}{h}\right)^{1-\alpha}$. Because the rental rate of capital is fixed, the capital share increases in k/h rather than being a constant. The capital share is equal to α only in the final state. In the transition to the final state, the capital share is lower than α and increases toward α as the capital stock increases. A value of α higher than 0.36 is necessary for the average capital share in the model to match the target 0.36.

Table 2. Identification of the parameters

Parameter	Value	Target
r	4.15×10^{-3}	quarterly interest rate = 0.0125
α	0.53	average capital share = 0.36
f_0	0.0682	normalize $k^*(h_H) = 100$
δ	0.026	monthly EU rate in CPS
ρ	0.5	elasticity $\frac{\mathrm{d}\ln p}{\mathrm{d}\ln\theta} = 0.362$ at average $u = 0.06$
p_0	0.2313	$\theta = 0.72$ at average u (Pissarides 2009)
f_u	0.0824	$\frac{\text{home production}}{\text{average market net output}} = 0.4$
ψ_1 ψ_0 c_1	3.6837 3.2785×10^{-4} 0.0161	(ψ_1, ψ_0, c_1) minimize the distance between $(u(h_L), u(h_m), u(h_H))$ and the targets $(0.10, 0.06, 0.03)$
b	0.5	symmetric bargaining weights

Second, the average unemployment rate is targeted at 6%, and the ratio of home production of an unemployed worker to the average net output in the market is targeted at 0.4. These targets put important disciplines on wage dispersion, as emphasized by Hornstein

Electronic copy available at: https://ssrn.com/abstract=3240652

et al. (2011). The three cost parameters, (ψ_1, ψ_0, c_1), minimize the distance between the model and the data in unemployment rates of three skill levels, (h_L, h_m, h_H). These rates are a good representation of U.S. workers between 25 to 64 years of age with the three education attainment levels (see Bureau of Labor Statistics, 2019).

Third, the identified ψ_1 is significantly higher than one. Since strict convexity of the vacancy cost has not played a critical role in the analytical sections, why is it needed for the calibration? The answer is that the model needs to match the features of unemployment rates in the data. The unemployment rate is low on average and has sizable variations across skills. When the vacancy cost is sufficiently convex, it is optimal to create a large number of low-capital vacancies, which increase the matching rate for searching workers and keep the average unemployment rate low. Moreover, the matching rate must be significantly higher for skill h_H than for h_L in order to match the difference in the unemployment rate between the two skill levels. This requires the vacancies cost to be sufficiently convex in the capital stock.

With the calibrated parameters, the model satisfies Assumption 1, and so the social optimum and the equilibrium are unique. I compute the social optimum (see the Supplementary Appendix D for the procedure). For equilibrium wages, I assume symmetric bargaining weights, $b = 0.5$, as in Hagedorn et al. (2017).

6.2. Initial Assignment and Matching Rate

Figure 3 depicts the initial assignment $\phi(h)$ (divided by 10) 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 3.4. As the skill increases from h_L to h_H, the initial capital stock decreases from 3.76 to 2.44, which translates into 17% reduction in net output if the skill remained at h_L. However, the size of this effect pales in comparison with the increase in the matching rate from 0.194 to 0.539 when the skill increases from h_L to h_H. The large increase in the matching rate reflects the feature in the data that the unemployment rate of skill h_H is less than a half of that of skill h_L.

Electronic copy available at: https://ssrn.com/abstract=3240652

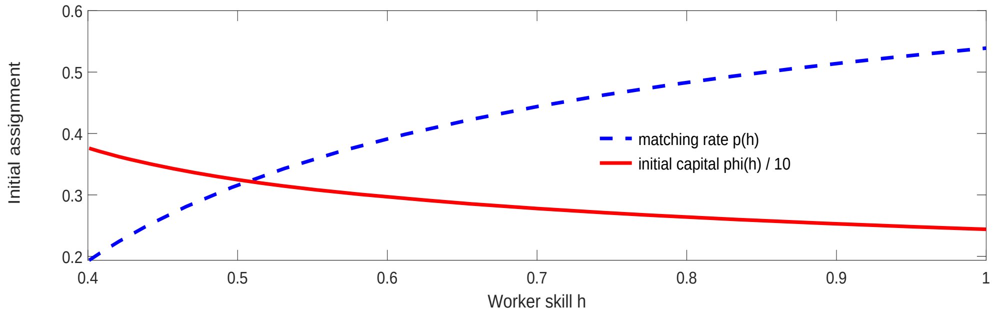

Figure 3. Initial assignment $\phi(h)/10$ and matching rate $p(h)$

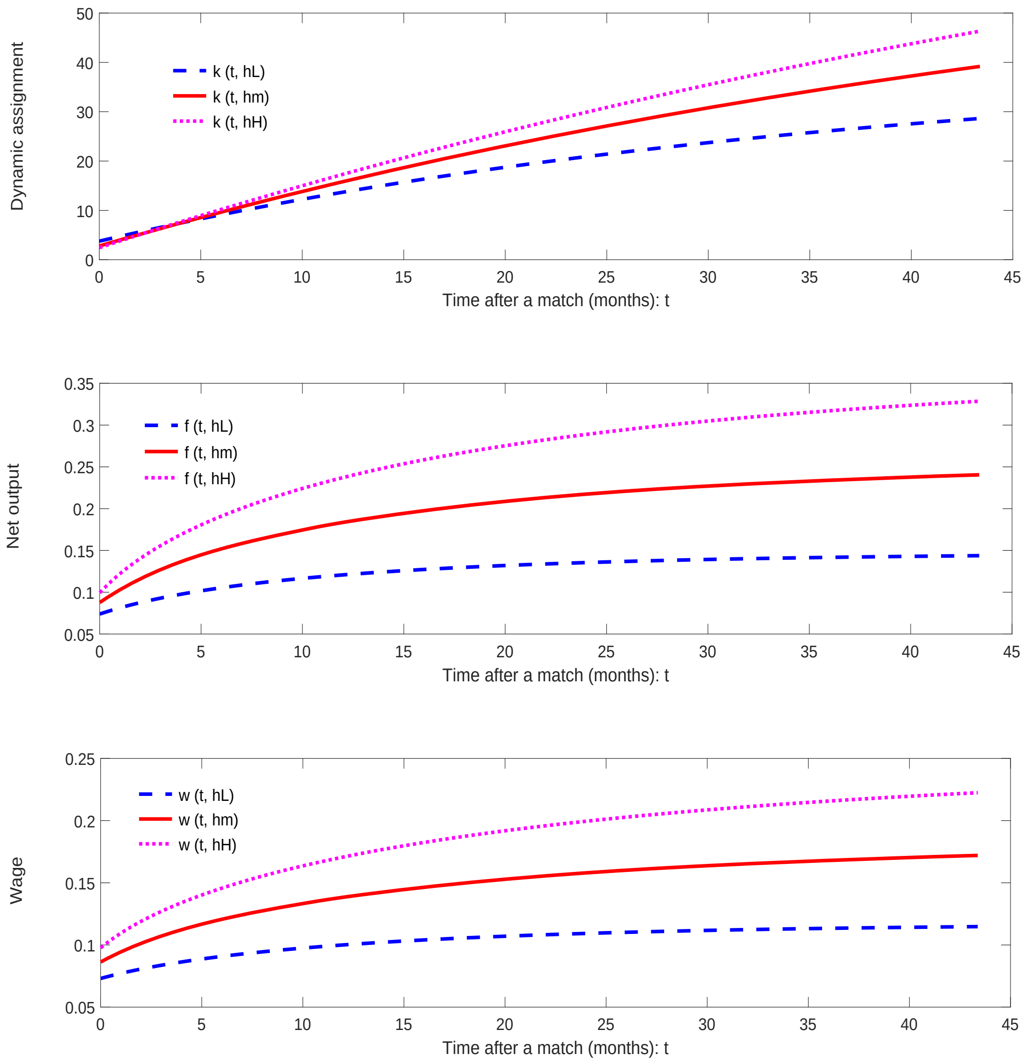

Figure 4. Capital, net output and wages over time for skills (h_L, h_m, h_H)

Electronic copy available at: https://ssrn.com/abstract=3240652

Although the initial NAM appears to be mild, its effects on time profiles of net output and earnings are profound. As discussed earlier (see Figure 2), the initial NAM implies that the time paths of capital stocks for different skill levels necessarily cross each other. With the calibrated parameters, this crossing occurs in less than 5 months, as shown in the top panel in Figure 4. For $h \in \{h_L, h_m, h_H\}$, the middle panel in Figure 4 depicts the time paths of net output. The initial NAM does not eliminate the productivity advantage of the higher skill. That is, a higher skill has higher net output throughout a match. Moreover, the time profile of net output is steeper for a higher skill, and so the difference in net output between skills increases over time.

The time profile of wages mimics that of net output. The wage path, given by (5.9), is depicted in the bottom panel in Figure 4. Throughout a match, workers with a higher skill have higher wages.[20] These workers also have a steeper wage path. This implies that if one traces a cohort of workers who become employed at the same time, the variance in wages in the cohort increases over time. Both the heterogeneity in the slope of the earnings profile and the fanning-out of the wage paths are consistent with the evidence on earnings dynamics, e.g., Deaton and Paxson (1994), Guvenen (2007).

6.3. Distribution of Workers, the Mean, and Sorting

Figure 5 depicts the density of capital stocks among employed workers of the same skill $h \in \{h_L, h_m, h_H\}$. Because of the initial NAM, the lower bound on the support of the distribution is lower for a higher skill. The upper bound on the support is higher for a higher skill, because the final state has PAM. For any given h, the density is a decreasing function of capital, because the calibration yields $\beta(h) < \delta$ for all h. That is, the job separation rate exceeds the speed of convergence to the final state. As explained earlier, in this case the density must become smaller in order to keep the distribution of workers stationary. Also, the density of the wage distribution within each skill is a decreasing function (not depicted).

[20] This result suggests that using wages or earnings to identify sorting, as is common in the empirical literature, would fail to detect the initial NAM.

Electronic copy available at: https://ssrn.com/abstract=3240652

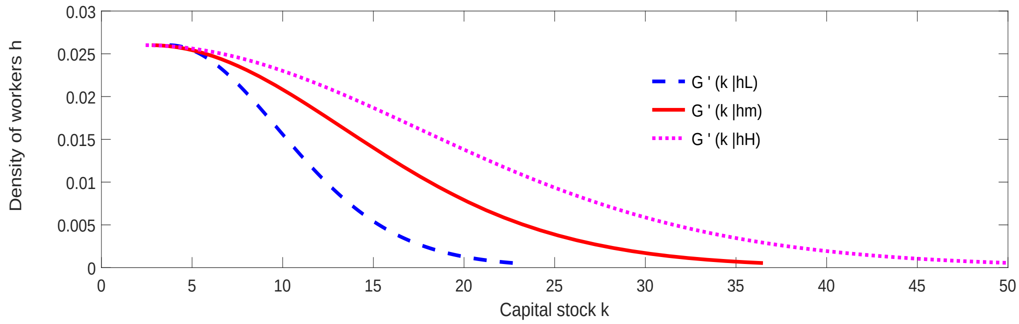

Figure 5. The density of capital among skill h workers: $G'(k|h)$

Although PAM-D fails, PAM-M holds in the calibrated model. That is, the mean of the capital stock among employed workers of the same skill increases in the skill. The left panel in Figure 6 depicts within-skill means of capital, net output and wages, where $E(k|h)$ is divided by 200. All three within-skill means increase in h. Recall that h_H is 2.5 times h_L in the calibration. In comparison, the ratio between the means of these two skills is 2.16 in net output and 1.84 in wages. A significant part of the higher means of net output and wages for a higher skill comes from post-match investment. The average capital stock among skill h_H workers is 1.64 times that among skill h_L. Therefore, the neoclassical theory of investment exerts a strong force on dynamic sorting between capital and the skill, by pushing the match toward the final state of PAM.

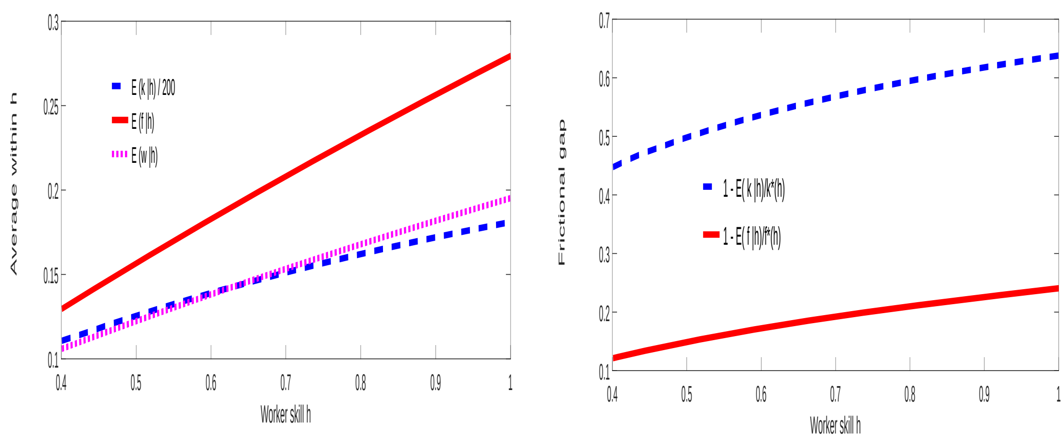

Figure 6. Within-skill mean $E(z|h)$ $(z = k, f, w)$
and the frictional gap $1 - \frac{E(z|h)}{z^*(h)}$ $(z = k, f)$

However, labor market frictions do matter quantitatively. They generate a significant gap between the social optimum and the final state. Let me refer to this gap as a *frictional*

Electronic copy available at: https://ssrn.com/abstract=3240652

gap within a skill and measure it by $(1 - \frac{E(z|h)}{z^*(h)})$ for $z \in \{k, f\}$ and any h. The right panel in Figure 6 depicts these frictional gaps. On average across skill levels, the frictional gap is 18% in net output and 55% in the capital stock. Both gaps increase in the skill. The causes are the adjustment cost in investment and job separation. The higher is the skill, the longer is the distance from the initial state to the final state, and the more likely is that a match is terminated by separation before the capital stock gets close to the final state. For a higher skill, although the average capital stock is higher in the level, the ratio to the final state is lower than for a lower skill. The frictional gap in the capital stock generates the frictional gap in net output.[21]

6.4. Frictional Inequality

Dynamic sorting induces large frictional inequality, which refers to inequality in productivity or earnings among workers of the same skill. The left panel in Figure 7 depicts the within-skill coefficient of variation (cv) in net output and wages. The right panel depicts the within-skill mean-min ratio (Mm) in these variables, a measure of frictional inequality proposed by Hornstein et al. (2007). Frictional inequality is large and increases in the skill. As the skill increases from h_L to h_H, the coefficient of variation in net output increases from 0.15 to 0.253, and the mean-min ratio increases from 1.75 to 2.81. For wages, the coefficient of variation increases from 0.11 to 0.20, and the mean-min ratio increases from 1.45 to 2.0. In matches with a high skill, there is more room for the capital stock to grow, which makes net output and wages fan out by more than in matches with a low skill.[22] Frictional inequality is smaller in wages than in net output because a worker's value of unemployment varies with skills by less than net output does.

Frictional inequality in this model is comparable with the data but much larger than in most search models. Hornstein et al. (2011) estimate the mean-min ratio in wages in the

[21] Looking back at Figure 4, the capital stock for skill h_H at $t = 45$ months (3.75 years) is only 22% of the final stock. However, the fraction of matches that survive to $t = 45$ is only $(1 - \delta)^{45} = 0.3056$.

[22] The empirical evidence is mixed on whether frictional inequality increases in the worker type. 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.

Electronic copy available at: https://ssrn.com/abstract=3240652

U.S. data to be between 1.7 and 2. They find that most search models generate mean-min ratios in wages lower than 1.05. They attribute this failure to two empirical restrictions imposed by the data – the low unemployment rate and the sizable value of home production to market production. The two targets imply that the option value of continuing to search must be low which, in most search models, requires wage dispersion to be small. Since the calibration here uses the two targets, the option value of continuing to search is low in this model. But it is accompanied by large, instead of small, frictional dispersion. The causes of this result are post-match investment and the convex vacancy cost, as first explored in Shi (2018). Briefly, the lower option value means that it is optimal to make workers employed quickly. To do so in the presence of convex vacancy costs, a large number of vacancies must be created, each with a low capital stock. From such low capital stocks, matches can experience large investment which widens dispersion between matches of the same skill that are formed at different points in time. In turn, to benefit from post-match investment, unemployed workers are willing to accept matches with initially low capital stocks, resulting in low unemployment rates.

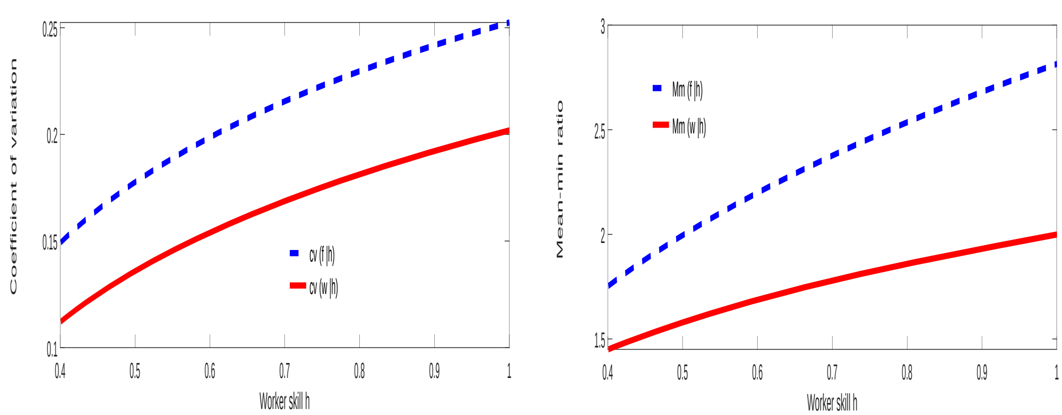

Figure 7. Coefficient of variation and mean-min ratio in net output f and wages w among skill-h workers

Electronic copy available at: https://ssrn.com/abstract=3240652

6.5. The Labor Share and the Skill Premium

Dynamic sorting between capital and the skill can be a potential explanation for the observed decline in the U.S. labor share over the last 3 decades (Elsby, et al., 2013). Recall from the calibration that the capital share in a (k, h) match is $rk/\tilde{f}(k, h)$, which increases in k/h even though the production function is Cobb-Douglas. Since the frictional gap in the capital stock increases in the skill (Figure 6), the capital share decreases in the skill. Define the average labor share within skill h as $1 - E(\frac{rk}{\tilde{f}(k,h)}|h)$. Then, the labor share increases in the skill. If a change in the economy affects dynamic sorting between capital and labor, then it can change the labor share. For example, consider one likely event: an increase in the matching efficiency.[23] A reduction in p_0 in the function $\theta(p)$ represents an increase in the matching efficiency (see (6.1)). Suppose that p_0 falls by 25%. The left panel in Figure 8 depicts the within-skill average labor share which is an increasing function of the skill. The increase in the matching efficiency shifts down this function, thus reducing the labor share for every skill. The reason is simple. When the matching efficiency increases, firms increase the initial capital stock for all vacancies. The time path of capital shifts up for every skill level, resulting in a lower labor share.

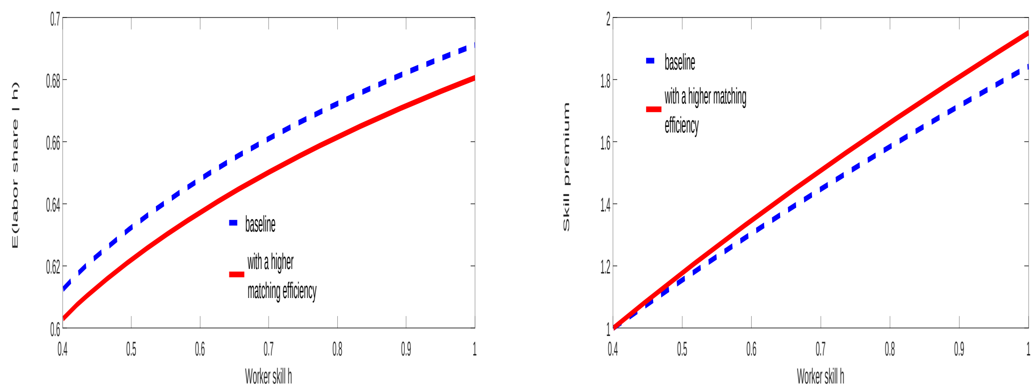

Figure 8. Effects of higher matching efficiency on within-skill average labor share, E(labor share $|h$), and the skill premium $\frac{E(w|h)}{E(w|h_L)}$

[23]Another example is a reduction in the vacancy cost, which has similar effects to an increase in the matching efficiency.

Electronic copy available at: https://ssrn.com/abstract=3240652

The increase in the capital stock is non-uniform across skills, which has an implication for the skill premium. Matching frictions are the cause of NAM in the initial assignment. By reducing these frictions, an increase in the matching efficiency weakens the efficiency consideration for NAM. The capital stock in the initial assignment increases by more for a high skill than for a low skill. This reduces the length of time needed for the sorting pattern to change from NAM to PAM in the transition. Once this change occurs, the gap in the capital stock between a high-skill match and a low-skill match widens quickly. Thus, the skill premium increases. The right panel in Figure 8 depicts the skill premium measured by the ratio in the average wage between any skill h and the lowest skill h_L. As an increasing function of h, this ratio rotates counter-clockwise after an increase in the matching efficiency. Therefore, changes in frictional dynamic sorting can be the common cause for both the increase in the skill premium and the declining labor share.

7. 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 social optimum and the equilibrium. 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 effect in the absence of post-match investment. However, investment makes sorting positive eventually. Because the sorting pattern in a match reverses over time, the time profiles of labor productivity and wages have steeper slopes for a high skill than for a low skill. Between-skill dispersion in labor productivity and wages widens over time. There is also dispersion in labor productivity and wages within each skill. The calibrated model shows that sorting is negative initially but becomes positive on average, with significant within-skill dispersion. Moreover, an increase in the matching efficiency reduces the labor share and increases the skill premium.

Frictional sorting and post-match investment are both important. Post-match investment quickly reverses negative sorting in the initial assignment and widens the gap in net

Electronic copy available at: https://ssrn.com/abstract=3240652

output between skills. Without such investment, a model with frictional matching alone would mistake the initial sorting pattern for the final pattern and, hence, reach the wrong conclusion of negative sorting even if sorting is positive on average. On the other hand, when there are no matching frictions, the initial capital assignment will be equal to the final assignment, which significantly exaggerates the extent of positive sorting. Moreover, a frictionless model does not provide an endogenous mechanism to generate inequality within each skill, the heterogeneity in the slope of the time profile of labor productivity and wages between skills, or the fanning-out of labor productivity and wages over time.[24]

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.

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

Electronic copy available at: https://ssrn.com/abstract=3240652

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 proof establishes results (i)-(v) below, which constitute the proposition:

(i) The expressions for $V(\phi, h)$ and its derivatives: Setting $t = 0$ in (2.3) and substituting $i(0) = I(\phi, h)$, I get:

$$(r + \delta) V(\phi, h) - \delta V_u(h) = f(\phi, h) + [V_1(\phi, h) I - c(I)]_{I = I(\phi, h)}. \tag{A.1}$$

Subtracting (2.2) from (A.1) into (3.8) yields the social gain from a (ϕ, h) match as

$$\Delta(\phi, h) = \frac{f(\phi, h) - f_u + \psi(\phi)\theta(p) + [V_1(\phi, h) I - c(I)]_{I = I(\phi, h)}}{r + \delta + p}, \tag{A.2}$$

where $p = p(h)$. For all $t \geq 0$, (3.1) implies $V_1(k(t), h) = c'(i(t))$. Setting $t = 0$ and using $I_1(\phi, h) = -\beta$ from (3.7), I have:

$$V_1(\phi, h) = c'(I(\phi, h)), \quad V_{11}(\phi, h) = -c''(I(\phi, h))\beta(h) < 0. \tag{A.3}$$

Electronic copy available at: https://ssrn.com/abstract=3240652

(ii) The solution to the search problem in (2.2) is interior: Before proving this result, I characterize the maximum of $\tilde{S}(k,h)$, where $\tilde{S}$ is defined by (3.9). Recall that $\theta'(0) > 0$ by assumption. It is clear that $\tilde{S}_1(0,h) = \infty$, $\tilde{S}_1(k^*(h),h) < 0$ and $\tilde{S}_{11}(k,h) < 0$. These properties imply that, for any given h, there is a unique $\hat{k}(h) \in (0, k^*(h))$ that achieves the maximum of $\tilde{S}(k,h)$. Moreover, $\tilde{S}_1(\hat{k}(h), h) = 0$, i.e.,

$$f_1(\hat{k}(h), h) = (r+\delta)\,\theta'(0)\,\psi'(\hat{k}(h)). \tag{A.4}$$

Note that $\tilde{S}(0,h) < 0$. If $\tilde{S}(\hat{k}(h),h) > 0$, then there exist $k_1(h)$ and $k_2(h)$, with $0 < k_1(h) < \hat{k}(h) < k_2(h) \le k^*(h)$, such that $\tilde{S}(k,h) > 0$ for all $k \in (k_1(h), k_2(h))$.

Now consider the maximization problem in (2.2). Denote the objective function as

$$S(p,\phi,h) \equiv p\Delta(\phi,h) - \psi(\phi)\,\theta(p).$$

Consider the optimal choice of p. When p approaches the natural upper bound $\bar{p} \le \infty$, the assumption, $\lim_{p\to\bar{p}} \frac{\theta(p)}{p} = \infty$, implies $\lim_{p\to\bar{p}} \frac{S(p,\phi,h)}{p} < 0$. Thus, the optimal p satisfies $p(h) < \bar{p}$. At the lower bound $p = 0$, it is clear that $S(0,\phi,h) = 0$. (A.2) implies $\Delta(\phi,h)|_{p=0} \ge \frac{f(\phi,h)-f_u}{r+\delta}$ which, in turn, implies:

$$S_1(0,\phi,h) = \Delta(\phi,h)|_{p=0} - \psi(\phi)\,\theta'(0) \ge \tilde{S}(\phi,h).$$

Recall that $\tilde{S}(\phi,h) > 0$ for all $\phi \in (k_1(h), k_2(h))$, where k_1 and k_2 are characterized after (A.4). For such levels of ϕ, $S_1(0,\phi,h) > 0$, in which case $S(p,\phi,h) > S(0,\phi,h) = 0$ for sufficiently small $p > 0$. Thus, the optimal choice of p satisfies $p(h) > 0$.

To prove that the optimal ϕ is interior, use (A.2) to derive:

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

where $I = I(\phi,h)$ and I have substituted V_1 from (A.3). Note that $I(k^*(h),h) = 0$. Because $f_1(k^*(h),h) = 0$ and $\psi' > 0$, then $S_2(p,k^*(h),h) < 0$ and, hence, the optimal ϕ satisfies $\phi(h) < k^*(h)$. (3.7) implies $I(0,h) = \beta(h)\,k^*(h) < \infty$, and so (A.3) yields $-V_{11}(0,h) = c''(I(0,h))\,\beta(h) < \infty$. With this result, the assumptions $f_1(0,h) = \infty$ and $\psi'(0) = 0$ imply $S_2(p,0,h) = \infty$. Thus, the optimal ϕ satisfies $\phi(h) > 0$. I will develop a tighter lower bound, $\underline{k}(h)$, on $\phi(h)$ later.

Electronic copy available at: https://ssrn.com/abstract=3240652

(iii) Conditions (3.12) and (3.13), and their solutions for different h: Because the solution to the search problem is interior, it satisfies the first-order conditions, which are (3.12) and (3.13). To prove that the solution is distinct for distinct skill levels, consider any two different levels, h_1 and h_2. Suppose $p(h_1) = p(h_2)$ and $\phi(h_1) = \phi(h_2)$, contrary to the distinct solution. Then (3.12) implies $V_1(\phi, h_1) = V_1(\phi, h_2)$, which contradicts the result $V_{12}(\phi, h) > 0$ established by (A.14) later.

(iv) The results $\phi(h) > \underline{k}(h)$ and $I(\phi(h), h) > 0$, where $\underline{k}(h)$ is defined by (3.10): By (3.7), $\phi(h) > \underline{k}(h)$ implies $I(\phi(h), h) > 0$. To prove $\phi(h) > \underline{k}(h)$, divide (3.12) by (3.13) to obtain: $\psi(\phi) V_1(\phi, h) = \frac{1}{\varepsilon}\psi'(\phi)\Delta(\phi, h)$. Because $\varepsilon \equiv \frac{p\theta'(p)}{\theta(p)} > 1$, then

$$\psi(\phi) V_1(\phi, h) - \psi'(\phi)\Delta(\phi, h) < 0.$$

This relationship changes into equality if $\phi = \underline{k}(h)$. Examine the expression on the left-hand side of the above inequality. The expression strictly decreases in ϕ for all $\phi \leq k^*(h)$. Because $f_1(0, h) = \infty$ and $f(0, h) = 0$, then $V_1(0, h) > 0$ and $\Delta(0, h) \leq 0$, and so the expression is positive at $\phi = 0$. Also, since $f_1(k^*(h), h) = 0$ and $0 < f(k^*(h), h) < \infty$, then $V_1(k^*(h), h) = 0$ and $\Delta(k^*(h), h) > 0$. The expression is negative at $\phi = k^*(h)$. Thus, there exists a unique $\underline{k}(h) \in [0, k^*(h)]$ that sets the above expression to 0. Because the expression is negative at ϕ, then $\phi(h) > \underline{k}(h)$.

(v) The solution to the search problem in (2.2) is unique: I show that (3.12) and (3.13) have a unique solution. Use (3.13) to solve:

$$p = \pi(\phi, h) \equiv \theta'^{-1}\left(\frac{\Delta(\phi, h)}{\psi(\phi)}\right). \tag{A.5}$$

Substituting into (3.12) yields $D(\phi, h) = 0$, where

$$D(\phi, h) \equiv \psi'(\phi) \left.\frac{\theta(p)}{p}\right|_{p=\pi(\phi,h)} - V_1(\phi, h).$$

Fix h. If $D_1(\phi, h) > 0$ for all interior ϕ such that $D(\phi, h) = 0$, then there is a unique solution of ϕ to $D(\phi, h) = 0$ for each h, in which case $p(h)$ is also unique for each h. Consider any interior ϕ that solves $D(\phi, h) = 0$. From (A.5), I can compute $\pi_1(\phi, h) = \frac{1}{\psi\theta''}(V_1 - \theta'\psi')$, where I have substituted $\Delta = \theta'(p)\psi$. Differentiate D with respect to ϕ:

$$D_1(\phi, h) = -V_{11} + \frac{\psi''}{\psi'}V_1 - \frac{\psi'}{\psi}\frac{(\varepsilon - 1)^2\theta}{p^2\theta''}V_1. \tag{A.6}$$

Electronic copy available at: https://ssrn.com/abstract=3240652

Substituting (A.3) and (3.7) yields:

$$\frac{\phi}{V_1}D_1(\phi,h)=\frac{c''(I)\,I}{c'(I)}\frac{\phi}{k^*(h)-\phi}+\frac{\psi''\phi}{\psi'}-\frac{\psi'\phi}{\psi}\frac{(\varepsilon-1)^2\theta}{p^2\theta''}.$$

(3.11) ensures $D_1(\phi,h)>0$ for all $\phi\in[\underline{k}(h),k^*(h)]$ and, hence, guarantees uniqueness of the solution to the search problem in (2.2). **QED**

Proof of Proposition 3.4:

I derive $(\phi'(h),p'(h))$ and prove (3.14) and (3.15). The envelope conditions of h for (2.4) and (2.2) yield:

$$V_2(k(t),h)=\frac{\delta}{r+\delta}V_u'(h)+\int_t^\infty f_2(k(\tau),h)\,e^{-(r+\delta)(\tau-t)}\mathrm{d}\tau, \tag{A.7}$$

$$V_u'(h)=\frac{pV_2(\phi,h)}{r+p}, \tag{A.8}$$

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 (3.12) and (3.13) yields:

$$\begin{bmatrix}\phi'(h)\\ p'(h)\end{bmatrix}=\frac{1}{\psi\theta''D_1(\phi,h)}\begin{bmatrix}\psi\theta''\frac{V_{12}}{V_1}-\frac{\Delta_2}{p}(\varepsilon-1)\\ (\frac{\psi''}{\psi'}-\frac{V_{11}}{V_1})\Delta_2-(\varepsilon-1)V_{12}\end{bmatrix} \tag{A.9}$$

where $\varepsilon=\frac{p\theta'}{\theta}\ (>1)$ and (A.6) shows $D_1>0$. The arguments of $(V_1,V_{11},V_{12},\Delta_2)$ are (ϕ,h) and I substituted $\Delta_1(\phi,h)=V_1(\phi,h)$. 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 $\Delta_2(\phi,h)$, I differentiate Δ in (A.2) with respect to h by fixing ϕ but taking the dependence $p(h)$ into account. Substituting (3.1) and (3.13) into the result, I get:

$$\Delta_2(\phi,h)=\frac{f_2(\phi,h)+V_{12}I(\phi,h)}{r+\delta+p}. \tag{A.10}$$

To compute V_{12}, differentiate (A.7) with respect to t:

$$V_{12}(k(t),h)\,i(t)=-f_2(k(t),h)+(r+\delta)\int_t^\infty f_2(k(\tau),h)\,e^{-(r+\delta)(\tau-t)}\mathrm{d}\tau. \tag{A.11}$$

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

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

Electronic copy available at: https://ssrn.com/abstract=3240652

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

$$\int_0^\infty f_2(k(t),h)\,e^{-(r+\delta)t}\mathrm{d}t \approx \frac{1}{r+\delta}\left[f_2(\phi,h)+\frac{f_{12}(\phi,h)}{r+\delta+\beta}I(\phi,h)\right], \tag{A.13}$$

where $I(\phi,h)=\beta(k^*-\phi)$. Set $t=0$ in (A.11). Substituting (A.12) and (A.13) yields:

$$V_{12}(\phi,h)\approx\frac{f_{12}(\phi,h)}{r+\delta+\beta}>0. \tag{A.14}$$

Substitute (A.10), (A.14) and (A.3) into (A.9):

$$\begin{aligned}\phi'(h) &\approx \tfrac{(\varepsilon-1)/p}{(r+\delta+p)\psi\theta''D_1}\left\{(a_2-I)\tfrac{f_{12}(\phi,h)}{r+\delta+\beta}-f_2(\phi,h)\right\}\\ p'(h) &\approx \tfrac{\frac{\psi''}{\psi'}+\beta\frac{c''(I)}{c'(I)}}{(r+\delta+p)\psi\theta''D_1}\left\{f_2(\phi,h)-(a_1-I)\tfrac{f_{12}(\phi,h)}{r+\delta+\beta}\right\},\end{aligned} \tag{A.15}$$

where $I=I(\phi,h)$, and (a_1,a_2) are defined by (3.16). Because $D_1>0$ and $\varepsilon>1$, then (A.15) implies (3.14) and (3.15). Also, $D_1>0$ implies $a_2>a_1$. Clearly, $a_1>0$ and $a_2>0$.

To determine the signs of $I_2(k,h)$ and $\frac{\mathrm{d}}{\mathrm{d}h}I(\phi(h),h)$, differentiate the relationship $V_1(k,h)=c'(I(k,h))$ in (A.3) with respect to h. This 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.14). Since $I_1(\phi,h)=-\beta(h)$, then

$$\frac{\mathrm{d}}{\mathrm{d}h}I(\phi(h),h)=\frac{V_{12}(\phi,h)}{c''(I)}-\beta(h)\,\phi'(h). \tag{A.16}$$

(A.14) shows $V_{12}>0$. If $\phi'(h)\le 0$, then (A.16) shows $\frac{\mathrm{d}}{\mathrm{d}h}I(\phi(h),h)>0$.[25]

Finally, I prove that $\phi'(h)\le 0$ implies $p'(h)>0$. Suppose $\phi'(h)\le 0$. The first relationship in (A.15) implies $f_2(\phi,h)\ge(a_2-I)\frac{f_{12}(\phi,h)}{r+\delta+\beta}$, and so

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

In this case, the second relationship in (A.15) implies $p'(h)>0$. **QED**

[25]Note that I obtained this derivative by differentiating (3.1), instead of (3.7), with respect to h. (3.7) is an approximation. Differentiating (3.7) may yield an ambiguous result.

Electronic copy available at: https://ssrn.com/abstract=3240652

B. Proof of Proposition 4.2

To compute the distribution $G(k|h)$, 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:

$$n_e(h)\lim_{\mathrm{d}t\to0}\frac{G(k|h)-G(k-I(k,h)\,\mathrm{d}t|h)}{\mathrm{d}t}=n_e(h)\,G'(k|h)\,I(k,h)\,.$$

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

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

Approximating the dynamics of (i,k) by (3.4), then $I(k,h)$ satisfies (3.7). Substituting into (B.1), integrating from $\phi(h)$ to k, and using $G(\phi(h)|h)=0$, I obtain the cumulative distribution function in the first equality in (4.3). 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 (4.3). The corresponding density function is

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

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

For any fixed k, differentiating (4.3) with respect to h yields:

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

Thus, $\frac{\partial}{\partial h}G(k|h)<0$ for all $k\in(\phi(h),k^*(h))$ if and only if (4.5) holds. If $\phi'(h)<0$, then (4.5) is violated for k sufficiently close to $\phi(h)$, because $R(k,h)$ is sufficiently close to 1 for such values of k. Thus, $\phi'(h)\geq0$ is necessary for (4.5) to hold. To prove that $\phi'(h)\geq0$ is also sufficient for (3.7) to hold, suppose $\phi'(h)\geq0$ and consider interior $k\in(\phi(h),k(h))$. Note that (3.7) implies $I_2(\phi,h)=\beta'(h)\left[k^*(h)-\phi\right]+\beta(h)\,k^{*\prime}(h)$. Because $I_2(\phi,h)>0$ by Proposition 3.4, 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

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

Electronic copy available at: https://ssrn.com/abstract=3240652

The function $(1 - z + z \ln z)$ is strictly decreasing in z for all $z \in (0, 1)$, and it is equal to 0 at $z = 1$. Thus, the function is strictly positive for all $z \in (0, 1)$. 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 (4.3) into (4.2) and integrate by parts. This yields the first equality in (4.4). The second equality in (4.4) follows from $I(\phi, h) = \beta(h)[k^*(h) - \phi(h)]$. Differentiating the first equality in (4.4) with respect to h, I can verify that $E(k|h)$ increases in h if and only if

$$\frac{\delta\phi'(h) + \beta(h) k^{*\prime}(h)}{k^*(h)} + \frac{\delta\, \beta'(h)}{\beta(h) + \delta} > 0. \tag{B.2}$$

Even if $\phi'(h) < 0$, (B.2) can still hold. Thus, $\phi'(h) \geq 0$ is not necessary for (B.2) 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 $\frac{\mathrm{d}}{\mathrm{d}h} E(k|h) > 0$.[26] **QED**

C. Proof of Proposition 5.1

Assume $V_h^e(h) = V_u(h)$ for all h, which I will verify later. Then (5.3) is the same as (2.3), and so $V^e(h) = V(h)$ for all h. By (5.5), it is clear that a firm's optimal investment under any given Ω maximizes the joint value V^e. As a result, equilibrium investment is socially efficient, i.e., $i^e(t, h) = I(k(t), h)$, provided that the initial $k(0) = \phi$ is socially efficient which I will prove below.

To prove that the optimal search choice is socially efficient, note that it is always optimal for a worker to search in a submarket where the market tightness and, hence, the job-finding rate is strictly positive. In such a submarket (x, ϕ, h), (5.2) implies $J(x, \phi, h, 0) = \frac{\psi(\phi)}{q(p)}$, where $p = \tilde{p}(x, \phi, h)$. With $J(x, \phi, h, 0) = V(\phi, h) - x$, this further implies

$$x = V(\phi, h) - \frac{\psi(\phi)}{q(p)}. \tag{C.1}$$

[26] To directly prove that $\phi'(h) \geq 0$ is sufficient for $\frac{\mathrm{d}}{\mathrm{d}h} E(k|h) > 0$, suppose $\phi'(h) \geq 0$. If $\beta'(h) \geq 0$, then (B.2) shows $\frac{\mathrm{d}}{\mathrm{d}h} E(k|h) > 0$. If $\beta'(h) < 0$, differentiate the second equality in (4.4) to obtain:

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

where $I = I(\phi(h), h)$. Substituting (A.16) yields: $\frac{\mathrm{d}}{\mathrm{d}h} E(k|h) > \frac{1}{\beta(h)+\delta} \left[\frac{V_{12}}{c''(I)} + \delta\phi'(h) \right]$. Since $V_{12} > 0$, then $\phi'(h) \geq 0$ implies $\frac{\mathrm{d}}{\mathrm{d}h} E(k|h) > 0$.

Electronic copy available at: https://ssrn.com/abstract=3240652

Using this equation to substitute x, I can express a worker's search choices as (p, ϕ), instead of (x, ϕ), and rewrite (5.1) as

$$rV_u^e(h) = f_u + \max_{(p,\phi)} \; p\left[V(\phi, h) - \frac{\psi(\phi)}{q(p)} - V_u^e(h)\right]. \tag{C.2}$$

Because $p/q = \theta(p)$, (C.2) is the same as (2.2) in the planner's problem. The two equations solve for the same function. That is, $V_u^e(h) = V_u(h)$ for all h. In addition, optimal search choices (p, ϕ) in the equilibrium are the same as the planner's allocation, because they solve the same maximization problem. In particular, the initial capital in a match, $k(0) = \phi$, is socially efficient. Then, (C.1) implies (5.8).

To characterize the bargaining result, it is simpler to treat the worker value Ω, instead of the wage path, as the object to be bargained. Substitute J from (5.5) into the bargaining problem in (5.7). It is easy to verify that the solution for Ω to the bargaining problem satisfies:

$$\Omega(k(t), h, t) - V_u(h) = b\left[V(k(t), h) - V_u(h)\right], \tag{C.3}$$

where I have substituted the earlier results $V^e = V$ and $V_u^e = V_u$. To recover wages from the bargaining result, substitute Ω from (C.3) into (5.4) to obtain:

$$(r+\delta)\left[V^e(k(t), h) - V_u^e(h)\right] = \frac{1}{b}\left[w(k(t), h, t) - rV_u^e(h)\right] + \frac{\mathrm{d}}{\mathrm{d}t}V^e(k(t), h).$$

Subtracting this equation from (5.3) yields (5.9). **QED**

Electronic copy available at: https://ssrn.com/abstract=3240652

References

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

[2] Acemoglu, D., 2001, "Good Jobs versus Bad Jobs," Journal of Labor Economics 19, 1-21.

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

[4] Bureau of Labor Statistics, 2019, "Table 501.80. Unemployment rates of persons 16 to 64 years old, by age group and highest level of educational attainment: Selected years, 1975 through 2018", U.S. Department of Labor.

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

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

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

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

[9] Chade, H. and I. Lindenlaub, 2017, "Risky Investment and Matching," manuscript, Yale University.

[10] 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.

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

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

[13] Elsby, M.W., Hobijn, B. and A. Sahin, 2013, "The Decline of the U.S. Labor Share," Brookings Papers on Economic Activity 2013 (2), 1-63.

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

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

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

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

Electronic copy available at: https://ssrn.com/abstract=3240652

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

[34] Shimer, R. and L. Smith, 2001, "Matching, Search, and Heterogeneity," B.E. Journal of Macroeconomics 1, 1-18.

Electronic copy available at: https://ssrn.com/abstract=3240652

For Online Publication

Supplementary Appendix for

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

Shouyong Shi

2020

D. 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 parameters (α, f_u) are chosen to allow the model to match the targets on the average capital share and the ratio of home production to the average net output in the market. This procedure requires $\alpha = 0.53$ and $\frac{f_u}{f^*(h_m)} = 0.32$. 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 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 0.06. Let $\hat{h}$ be the skill of workers at this average u; i.e., $u(\hat{h}) = 0.06$. 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 skill, I apply the targets to skill $\hat{h}$ workers whose unemployment rate is at the average level, 0.06. 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 $\frac{1}{\varepsilon} \equiv \frac{\mathrm{d}\ln p_u}{\mathrm{d}\ln\theta(p_u)} = 0.362$. This value lies within the range used in the literature.[27] For skill $\hat{h}$ workers, the job-finding rate the steady state yields $p(\hat{h}) = \delta(\frac{1}{u(\hat{h})} - 1) = 0.407$ 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 value $\rho = 0.5$ delivers the target $\frac{1}{\varepsilon(\hat{h})} = 0.362$. 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 $p_0 = \theta(\hat{h})\left[p^{-\rho}(\hat{h}) - 1\right]^{1/\rho} = 0.2313$.

The remaining parameters, (ψ_1, ψ_0, c_1), minimize the distance between the model and

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

Electronic copy available at: https://ssrn.com/abstract=3240652

three targets: the unemployment rates at h_L, h_m and h_H. The distance measure is:

$$drest\ (\%) \equiv \Big[a_1 \left[u\left(h_L\right) - 0.10\right]^2 + a_2 \left[u\left(h_m\right) - 0.06\right]^2 \\ + a_3 \left[u\left(h_H\right) - 0.03\right]^2\Big]^{1/2} \times 100\%$$

where (a_1, a_2, a_3) are weights on the targets, with $a_1 + a_2 + a_3 = 1$. I use $a_1 = 0.3$, $a_2 = 0.4$ and $a_3 = 0.3$. The target on the unemployment rate at h_m is given relatively higher weights in order for the average unemployment rate to be close to the target 0.06. Table 3 lists the result of this minimization and some other statistics. The unemployment rates are higher than the targets at h_L and h_H, and lower than the target at h_m. 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 and can be a good outcome. The actual capital share in the model can be higher than the definition which excludes the vacancy cost and the adjustment cost of investment.

Table 3. Performance of the calibration

	Model	Targets			Model	Targets
$u\left(h_L\right)$	0.118	0.100		average u	0.061	0.06
$u\left(h_m\right)$	0.056	0.060		$\frac{\text{home production}}{\text{average } f}$	0.397	0.40
$u\left(h_H\right)$	0.046	0.030		average share of k	0.341	0.36

To compute the social optimum, I substitute $V_1(\phi, h) = c'(I(\phi, h))$ from (3.1) into (A.2) to obtain:

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

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 (3.12) and (3.13) as

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

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

These two equations determine (ϕ, p). In fact, p can be solved first as a function of ϕ from (D.2) and then substituted into (D.3) to solve ϕ.[28] Once (ϕ, p) are solved, I can compute

[28] Equations (D.2) and (D.3) can also be used to prove Proposition 3.4. Doing so generates expressions for $\phi'(h)$ and $p'(h)$ that are equivalent to those in (A.9), with V_1 and V_{11} being obtained from differentiating (3.8) with respect to ϕ.

Electronic copy available at: https://ssrn.com/abstract=3240652

$\{k(t), i(t)\}_{t\geq 0}$ by (3.6) and compute the distribution of workers and capital according to section 4.

Wages in the equilibrium are computed according to (5.9).

E. The Non-Linearized Model

The non-linearized dynamic system formed by (3.3) and (2.1) has no analytical solution. To check the accuracy of linearized dynamics, I compute the non-linearized model numerically with the parameter values calibrated in section 6.1. The computation centers around the iteration on the value function. In fact, it is more accurate to iterate on the gain from a match, $\Delta(k, h)$, defined by (3.8). Subtracting (2.2) from (2.3) yields:

$$\begin{aligned}(r+\delta)\,\Delta(k,h) = \;& f(k,h) - f_u + \max\nolimits_i \;\left[\Delta_1(k,h)\, i - c(i)\right] \\ & - \max\nolimits_{(p,\phi)} \left[p\Delta(\phi,h) - \psi(\phi)\,\theta(p)\right].\end{aligned} \tag{E.1}$$

Similarly, rewrite (2.2) as

$$rV_u(h) = f_u + \max_{(p,\phi)} \left[p\Delta(\phi,h) - \psi(\phi)\,\theta(p)\right]. \tag{E.2}$$

The iteration uses Chebyshev projection and follows Steps 1-4 below:

Step 1. Discretize the support of k into $\{k_1, k_2, ..., k_{M_k}\}$, which are the zeros of the Chebyshev polynomials of k. Similarly, discretize the support of h into $\{h_1, h_2, ..., h_{M_h}\}$. Construct the Chebyshev basis for each of the two state variables.

Step 2. Start with initial guesses of the coefficients of the Chebyshev polynomials that approximate $\Delta(k,h)$ and $V_u(h)$. Compute the right-hand sides of (E.1) and (E.2). The derivative $\Delta_1(k,h)$ is computed forward; that is, $\Delta_1(k,h) = \frac{\Delta(k+dk,h)-\Delta(k,h)}{dk}$ for a small $dk > 0$. For any k not in the discretized support, use Chebyshev projection to compute the values of functions. For example, ϕ in $\Delta(\phi,h)$ and $(k+dk)$ in the computation of $\Delta_1(k,h)$ typically do not lie in the discretized support of k. Use grid search to solve the two maximization problems.

Step 3. Update $\Delta(k,h)$ and $V_u(h)$ on the discretized grid. For Δ, use the following rewritten form of (E.1) to update:

$$\Delta(k,h) = \frac{1}{1+(r+\delta)\,\mathrm{d}t}\left[\Delta(k,h) + RHS(\text{E.1})\ \mathrm{d}t\right].$$

Choose $\mathrm{d}t > 0$ to be relatively small. This mimics the discrete-time equation. Update $V_u(h)$ Similarly.

Electronic copy available at: https://ssrn.com/abstract=3240652

Step 4. Use the updated values of Δ and V_u on the grid to compute the updated coefficients of the Chebyshev polynomials for the approximation of Δ and V_u.

Repeat Steps 1-4 until the Chebyshev coefficients converge. The policy functions representing the optimal choices are the results of Step 2 in the final iteration.

With the computed value and policy functions, I compute the time paths of (k, f, w) and the distribution of workers. The time path of k comes from integrating (3.3), with the policy function $i(k, h)$ replacing i. Substitute $k(t)$ into $f(k(t), h)$ and $w(k(t), h)$ in (5.9). For the within-skill distribution, $G(k|h)$, note that it obeys (B.1) where $I(k, h)$ is now the policy function in the non-linearized model. Integrating (B.1) yields $G(k|h)$. Other variables and statistics can be computed accordingly.

The non-linearized model yields results similar to the linearized model. Figures E-1 through E-5 present the results in the non-linearized model. The counterparts in the linearized model are Figures 3 through 7. The main similarities are as follows:
(i) The initial assignment is NAM and the matching rate increases in the skill;
(ii) The time path of the capital stock for a higher skill starts below that for a lower skill and overtakes the latter quickly;
(iii) A higher skill has higher net output throughout the match and the slope of net output is steeper for a higher skill (The time path of wages has similar features and is not depicted);
(iv) Within each skill, the density of capital among worker is a decreasing function;
(v) PAM-M holds, and the within-skill means of net output and wages increase in the skill;
(vi) Frictional gaps in capital and net output are significant;
(v) There is sizable within-skill inequality, measured by the coefficient of variation and the mean-min ratio.

The main difference between the non-linearized model and the linearized model lies in the initial assignment. The capital stock in the initial assignment is lower and the matching rate is higher in the non-linearized model than in the linearized model. A lower capital stock in the initial assignment means that the social optimum delays more capital accumulation to the time after a match is formed. As a result, the non-linearized model has higher investment than in the linearized model (not depicted). Because the linearization under-represents the importance of the matching rate and post-match investment in the social optimum, the bias works in the favor of the model's main messages.

Electronic copy available at: https://ssrn.com/abstract=3240652

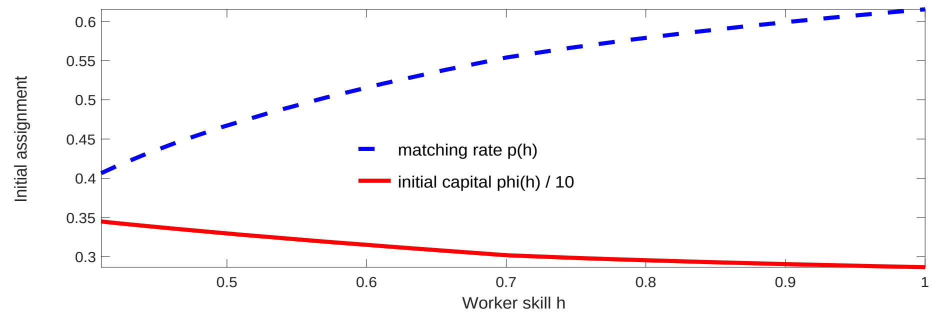

Figure E-1. Initial assignment $\phi(h)/10$ and matching rate $p(h)$

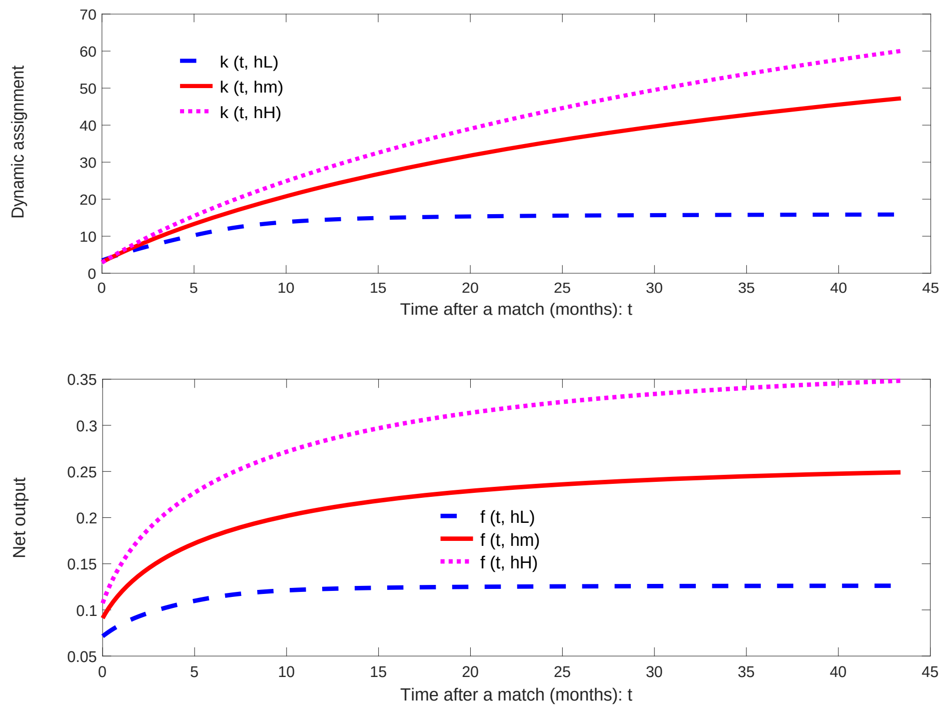

Figure E-2. Capital and net output over time for skills (h_L, h_m, h_H)

The quantitative differences are small between Figures E-2 through E-5 and their counterparts in the linearized model, and these differences trace back to the difference in the initial assignment. Relative to Figure 4, the time paths of capital for different skills cross each other earlier in Figure E-2, and the slope of the time path of net output is steeper. Both differences arise from the fact that investment is higher in the non-linearized model. For the same reason, the within-skill means in Figure E-4 have slightly steeper slopes than in Figure 6. Within-skill inequality measured by the mean-min ratio is larger in Figure E-5 than in Figure 7 because workers start with lower capital stocks in the non-linearized model.

Electronic copy available at: https://ssrn.com/abstract=3240652

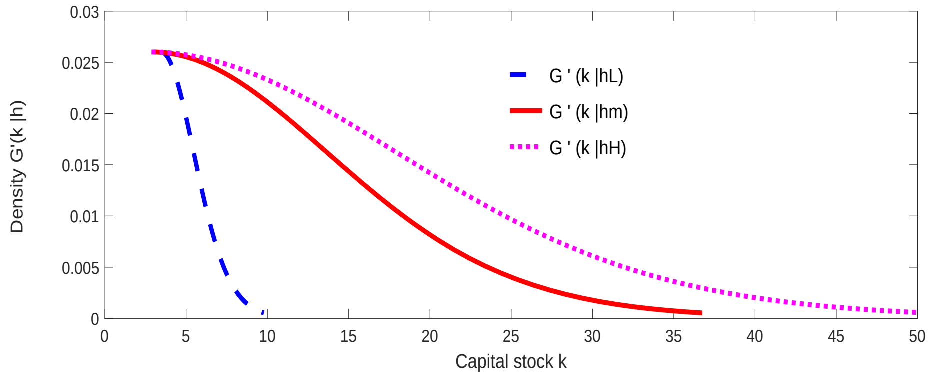

Figure E-3. The density of capital among skill h workers over capital: $G'(k|h)$

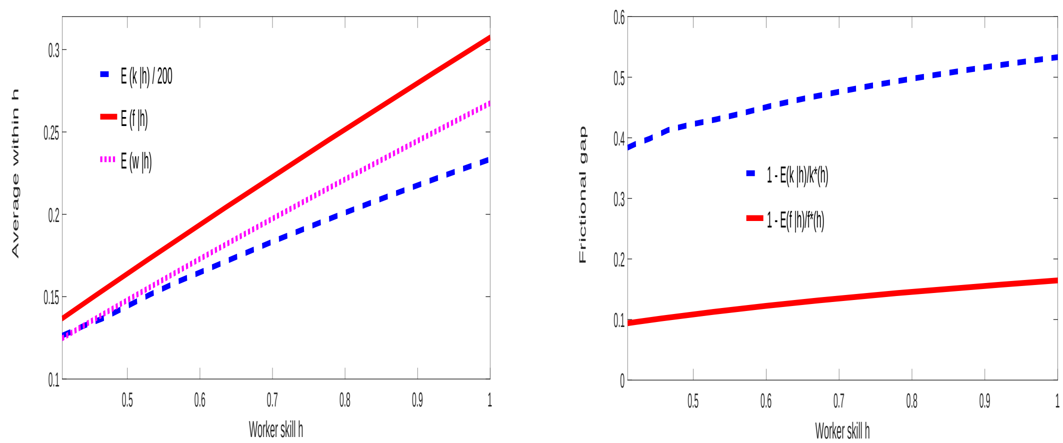

Figure E-4. Within-skill mean $E(z|h)$ $(z = k, f, w)$
and the frictional gap $1 - \frac{E(z|h)}{z^*(h)}$ $(z = k, f)$

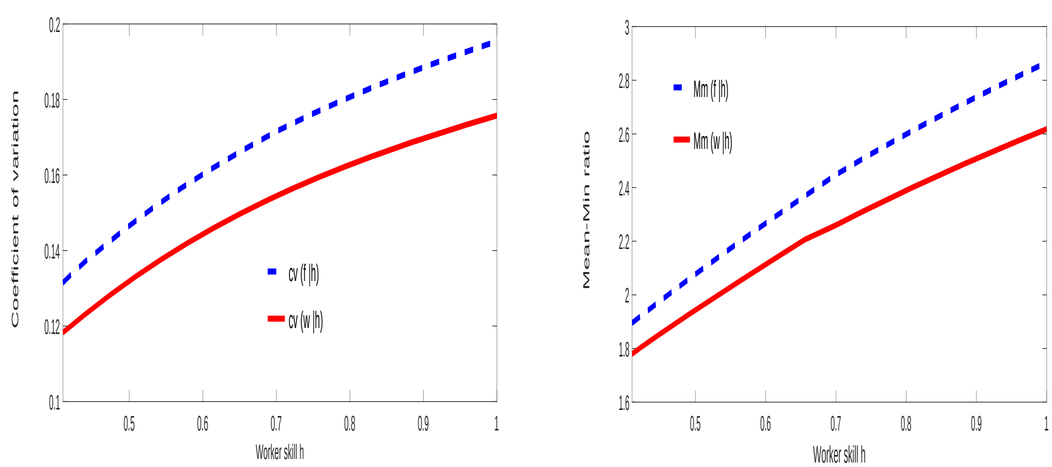

Figure E-5. Coefficient of variation and mean-min ratio in net output f
and wages w among skill-h workers

Electronic copy available at: https://ssrn.com/abstract=3240652