Search, Matching and Training

Christopher Flinn

New York University

Collegio Carlo Alberto

Ahu Gemici

Royal Holloway, U. of London

Steven Laufer

Board of Governors/FRB

February 21, 2016

Search, Matching and Training

Abstract

We estimate a partial and general equilibrium search model in which firms and workers choose how much time to invest in both general and match-specific human capital. To help identify the model parameters, we use NLSY data on worker training and we match moments that relate the incidence and timing of observed training episodes to outcomes such as wage growth and job-to-job transitions. We use our model to offer a novel interpretation of standard Mincer wage regressions in terms of search frictions and returns to training. Finally, we show how a minimum wage can reduce training opportunities and decrease the amount of human capital in the economy.

1 Introduction

There is a long history of interest in human capital investment, both before and after entry into the labor market. In the latter case, it is common to speak of general and specific human capital, which are differentiated in terms of their productivity-enhancing effects across jobs (which may be defined by occupations, industries, or firms). The classic analysis of Becker (1964) considered these types of investments in competitive markets and concluded that workers should pay the full costs of general training, with the costs of specific training (that increases productivity only at the current employer) being shared in some way. Analysis of these investments in a noncompetitive setting is more recent. Acemoglu and Pischke (1999) consider how the predictions of the amount and type of human capital investment in a competitive labor market are altered when there exist market imperfections in the form of search frictions. Frictions create an imperfect "lock in" between a worker and the firm, so that increases in general or specific human capital are generally borne by both the worker and the firm.

We utilize training data in the estimation of what otherwise is a reasonably standard search model with general and specific human capital. These data, described briefly below, indicate that formal training is reported by a not insignificant share of workers, and that the likelihood of receiving training is a function of worker characteristics, in particular, education. Workers do not receive training only at the beginning of job spells, although the likelihood of receiving training is typically a declining function of tenure. Since training influences the likelihood of termination of the job and wages, it is important to examine training decisions in a relatively complete model of worker-firm employment relationships.

One motivation for this research is related to recent observations regarding shifts in the Beveridge curve, which is the relationship between job vacancies and job searchers. While the unemployment rate in the U.S. has been markedly higher from 2008 and beyond,[1] reported vacancies remain high. This mismatch phenomenon has been investigated through a variety of modeling frameworks (see, e.g., Cairo (2013), Lindenlaub (2013)), typically by allowing some shift in the demand for workers' skills. In our modeling framework, such a shift could be viewed as a shift downward in the distribution of initial match productivities. Given the absence of individuals with the desired skill sets, the obvious question is why workers and firms do not engage in on-the-job investment so as to mitigate the mismatch in endowments. Using our model, we can theoretically and empirically investigate the degree to which a decentralized labor market with search frictions is able to offset deterioration in the initial match productivity distribution. This will lead us to consider policies that could promote increased investment activities, some of which are described below.

Another motivation for our research is to provide a richer model of the path of wages

---

[1]While the unemployment rate has declined recently, the employment rate in the population is at a historic low. Many of those counted as out of the labor force are in fact willing to take a reasonable job offer, and hence should be considered to be "untemployed" in the true sense of the term.

on the job and a more complete view of the relationship between workers and firms. In this model, firms offer workers the opportunity to make mutually advantageous investments in the worker's skills, both of the general and specific (to the job) type. While investing, the worker devotes less time to productive activities, which is the only cost of investment that we include in the model.[2] Wage changes over the course of the employment spell are produced by changes in general skill levels, changes in specific skill levels, and changes in investment time. In model specifications that include on-the-job search possibilities, which is the case for ours, wages during an employment spell may also increase due to the presence of another firm bidding for the employee's services, as in Postel-Vinay and Robin (2002), Dey and Flinn (2005), and Cahuc et al. (2006). When there exists the potential for other firms to "poach" the worker from her current firm, investments in match-specific skills will be particularly attractive to the current employer, since high levels of match-specific skills will make it less likely that the worker will exit the firm for one in which her initial match-specific skill level is higher. Other things equal, these differences in the retention value component of specific-skill investment implies that firms will reduce the employee's wage less for a given level of specific-skill investment than for general skill investment.

Why would firms ever partially finance improvements in general ability? In our modeling framework the flow productivity of a match is given by $y(a, \theta) = a\theta - \zeta$, where a is the general ability of the worker, θ is match productivity, and ζ is a flow cost of employment, which may be thought of as the rental rate on capital equipment required for the job. The gain in flow productivity from a small change in a is simply θ and the gain in flow productivity from a small change in θ is simply a. Jobs for which a is relatively high in comparison to θ will experience bigger productivity gains from investment in θ and conversely for job matches in which a is relatively low in comparison with θ. Thus, strictly from the productivity standpoint, there will an incentive to more evenly balance a and θ in the investment process. This, coupled with the fact that there exist search frictions, will lead firms to be willing to finance part of the investment in general human capital, even if this does not change the expected duration of the match.

We believe that our modeling framework may be useful in understanding the link between initial labor market endowments and earnings inequality over the labor market career. Flinn and Mullins (2015) estimated a model with an identical specification for flow productivity as the one employed here and examined the pre-market entry schooling decision.[3] In their model, initial ability endowments were altered by schooling

---

[2]That is, the only costs of investing in either of the skills is the lost output associated with the investment time. Moreover, these costs are the same for either type of investment. There are no direct costs of investment as in Wasmer (2006), for example. In his case, all investment in skills occurs instantaneously at the beginning of a job spell. Lentz and Roys (2015), instead, assume that general and match-specific skills are binary, and that for a low type worker on either dimension the cost of training is a flow cost that is increasing in the rate at which the transition to the high skill type occurs. They do not allow for depreciation in skills, which we find to be an empirically important phenomenon.

[3]They did not consider flow costs of employment, so in that paper $\zeta = 0$.

decisions, and these decisions were a function of all of the primitive labor market parameters. In their setting, a was fixed over the labor market career and a match draw at a firm was also fixed over the duration of the job spell. In the case of our model, both a and θ are subject to (endogenous) change, although it may well be that a is more difficult and costly to change after labor market entry. This is due to the fact that employers are not equipped to offer general learning experiences as efficiently as are schools that specialize in increasing the cognitive abilities of their students. To the extent that a is essentially fixed over the labor market careers, individuals with large a endowments will be more attractive candidates for investment in match-specific skills than will individuals with low values of a. Even if initial values of θ are drawn from the same distribution for all a types, which is the assumption made in our model, higher values of a at the time of labor market entry would be expected to lead to more investments and a more rapidly increasing wage profile over the course of an employment spell. This offers a mechanism to amplify the differences in earnings generated by the initial variability in a.

Using estimates from the model, we can determine the impact of various types of labor market policies on investment in the two types of human capital. For example, Flinn and Mullins (2015) investigated the impact of minimum wage laws on pre-market investment. They found that for relatively low (yet binding for some low general ability workers) minimum wage levels, the minimum wage could be a disincentive for pre-market investment, since similar wage rates could be obtained without costly education. At high levels of the minimum wage, however, most individuals invested in a to increase their chance of finding a job. For firms to earn nonnegative profit flows in that model, productivity has to be at least $a\theta \geq m$, where m is the minimum wage. For high values of m, workers will invest in general ability to increase their level of a so as to increase the likelihood of generating a flow productivity level that satisfies the firm's nonnegative flow profit condition. In our framework of post-entry investment, the impact of minimum wages is also ambiguous. As in the standard Becker story, a high minimum wage will discourage investment activity if the firm is to achieve nonnegative flow profits. On the other hand, through investment activity that raises the individual's productivity (through a and/or θ), the firm and worker can act to remove the constraint and push the individual's productivity into a region for which $w > m$. These possibilities may mitigate the need to increase pre-market entry investment in a.

Another policy that we will investigate is government subsidization of OTJ training by firms.[4] In our model, there exists only an opportunity cost to training, which is the output lost from the worker investing in skill acquisition during a portion of the work day. In this case, a government subsidy amounts to a transfer to the worker-firm pair that compensates them for a portion of the lost productivity associated with training. We will have these subsidies financed through taxes on labor earnings and firm profits, and will look for constant marginal rate taxes on both that will pay for an efficient

---

[4]Investment ubsidies are not considered in the current draft..

subsidy program.

In terms of related research, the closest paper to ours is probably Wasmer (2006). He presents a formal analysis of the problem in a framework with search frictions and firing costs. His model is stylized, as is the one we develop below, and is not taken to data. He assumes that human capital investments, be it of the general or specific kind, are made as soon as the employment relationship between a worker and a firm begins. Investment does not explicitly involve time or learning by doing, which we believe to be an important part of learning on the job. However, due to the simplicity of the investment technology, Wasmer is able to characterize worker and firm behavior in a general equilibrium setting, and he provides elegant characterizations of the states of the economy in which workers and firms will choose only general, only specific, or both kinds of human capital investment. The goal of our paper is to estimate a partial equilibrium version of this type of model with what we think may be a slightly more realistic form of the human capital production technology, one in which time plays the central role.

Another related paper is Bagger et al. (2014). This paper examines wage and employment dynamics in a discrete-time model with deterministic growth in general human capital in the number of years of labor market experience. There is no match-specific heterogeneity in productivity, but the authors do allow for the existence of firm and worker time invariant heterogeneity. There is complementarity between the worker's skill level and the productivity level of the firm, so that it would be optimal to reallocate more experienced workers to better firms. The authors allow for renegotiation of wage contracts between workers and firms when an employed worker meets an alternative employer, and due to the generality of human capital, the more productive firm always wins this competition. The model is estimated using Danish employer-employee matched data. Key distinctions between our approach and the one taken in that paper are the lack of firm heterogeneity but the presence of worker-firm match heterogeneity, the value of which can be changed by the investment decisions of the worker-firm pair. This paper also allows for worker heterogeneity that is an endogenous stochastic process partially determined through the investment decisions of workers and firms.

Lentz and Roys (2015) also examine general and specific human capital accumulation in a model that features worker-firm renegotation and the ability of firms to make lifetime welfare promises to workers in the bargaining stage. There is firm heterogeneity in productivity, and the authors find that better firms provide more training. The nature of the contracts offered to workers is more sophisticated than the ones considered here, and the authors explicitly address the issue of inefficiencies in the training and mobility process. They assume that there are only four training states in the economy (an individual can be high or low skill in general and specific productivity), which greatly aids in the theoretical analysis at the cost of not being able to generate wage and employment sample paths that can fit patterns observed at the individual level. They also assume that there is no skill depreciation, which aids in the theoretical analysis of the model.

6

The plan of the paper is as follows. In Section 2, we analyze a partial equilibrium search model with general and specific human capital and subsequently extend it to a general equilibrium framework. Section 3 discusses the data used in the estimation of the model, and presents descriptive statistics. Section 4 discusses econometric issues such as the model specification used in our estimation, the estimator we use, and identification. In Section 5, we present the estimation results and discuss the details of the estimated model, such as parameter values, within-sample fit and policy rules. Section 5 also presents a discussion regarding the implications of our estimated model for sources of wage growth and provides a unique perspective on the interpretation of the standard Mincer wage regression. In Section 6, we conduct a minimum wage experiment to determine the impact of minimum wages on general and specific human capital investment decisions, in a partial as well as general equilibrium framework. Section 7 concludes.

2   Modeling Framework

Individuals are characterized in terms of a (general) ability level a, with which they enter the labor market.[5] There are M values of ability, given by

$$0 < a_1 < ... < a_M < \infty.$$

When an individual of type a_i encounters a firm, she draws a value of θ from the discrete distribution G over the K values of match productivity θ, which are given by

$$0 < \theta_1 < ... < \theta_K < \infty.$$

The flow productivity value of the match is given by

$$y(i,j) = a_i\theta_j - \zeta,$$

where ζ is a flow cost of the job, which we think of as the rental rate on capital equipment that must be used in the production process in addition to the labor input. In the general equilibrium version of the model, to be discussed below, firms are assumed to pay flow posting fees while holding a vacancy open. One rationale for such a cost could be that firms rent a piece of capital equipment on which an individual's skills at the job can be assessed when they apply. In this case, it also seems reasonable to assume that a piece of capital equipment is required in the case in which the individual is hired. Under our assumptions on the distributions of a and θ, we obtain an estimate of ζ that

---

[5]Flinn and Mullins (2015) examine pre-market entry education decisions in a search environment in which a hold-up problem exists. We will not explicitly model the pre-market entry schooling decision, but will merely assume that the distribution of an individual's initial value of a at the time of market entry is a stochastic function of their completed schooling level. In estimation, we will distinguish three schooling levels.

is large and positive, and that significantly improves the fit of the model. It also serves to produce what we consider to be more reasonable human capital investment policy functions than when $\zeta = 0$.

We consider the case in which both general ability and match productivity can be changed through investment on the job. The investment level, along with the wage, are determined cooperatively in a model with a surplus division rule. At every moment of time, the individual and firm can devote a proportion τ_a of time to training in general ability, in the hope of increasing a. Similarly, they can invest a proportion of time τ_θ in job-specific training, in the hope of increasing θ.

We will assume that the two stochastic production technologies for a and θ are independent, in the sense that the likelihood of an improvement in a depends only on τ_a and not on τ_θ, and that the likelihood of an improvement in θ depends only on τ_θ and not τ_a. Given that the level of a is currently a_i, the rate of improvement in a is given by

$$\varphi_a(i, \tau_a),$$

with $\varphi_a(i, \tau_a) \geq 0$, and $\varphi_a(i, 0) = 0$ for all i. We restrict the improvement process to increase the value of i to $i + 1$ in the case of a successful investment. We will also allow for reductions in the value of a. This depreciation rate is assumed to be constant and equal to δ_a for all $i > 1$. In the case of one of these Poisson shocks, the level of a will decrease from i to $i - 1$, except when $i = 1$, when the individual is already at the lowest ability level. Since the rate of decreases in a are independent of investment time, the implication is that at the highest level of a, a_M, no investment in a will occur.

For purposes of estimation, we further restrict the function φ_a to have the form

$$\varphi_a(i, \tau_a) = \varphi_a^0(i)\varphi_a^1(\tau_a),$$

where φ_a^1 is strictly concave in τ_a, with $\varphi_a^1(0) = 0$. The term $\varphi_a^0(i)$ can be thought of as total factor productivity (TFP) in a way, and we place no restriction on whether $\varphi_a^0(i)$ is increasing or decreasing in i, although the functional form we utilize in estimation will restrict this function to be monotone.[6]

There is an exactly analogous production technology for increasing match-specific productivity, with the rate of increase from match value j to match value $j + 1$ given by

$$\varphi_\theta(j, \tau_\theta) = \varphi_\theta^0(j)\varphi_\theta^1(\tau_\theta),$$

with φ_θ^1 strictly concave in τ_θ, and $\varphi_\theta^1(0) = 0$. There is no necessary restriction on the TFP terms, as above. As is true for the a process, there is an exogenous depreciation

---

[6]By this we mean that either

$$\varphi_a^0(1) \leq \varphi_a^0(2) \leq ... \leq \varphi_a^0(M)$$

or

$$\varphi_a^0(1) \geq \varphi_a^0(2) \geq ... \geq \varphi_a^0(M).$$

rate associated with all θ_j, $j > 1$, which is equal to δ_θ. If one of these shocks arrive, then match productivity is reduced from θ_j to θ_{j-1}. As was true in the case of a, if match productivity is at its highest level, θ_K, then $\tau_\theta = 0$.

The only costs of either type of training are foregone productivity, since total productivity is given by $(1 - \tau_a - \tau_\theta)y(i,j)$. The gain from an improvement in either accrues to both the worker and firm, although obviously, gains in general human capital increase the future value of labor market participation (outside of the current job spell) to the individual only. As noted by Wasmer (2006), this means that the individual's bargaining position in the current match is impacted by a change in a to a greater extent than it is due to a change in θ. Motives for investment in the two different types of human capital depend importantly on the worker's surplus share α, but also on all other primitive parameters characterizing the labor market environment.

2.1 No On-the-Job Search

We first consider the case of no on-the-job search in order to fix ideas. In defining surplus, we use as the outside option of the worker the value of continued search in the unemployment state, given by $V_U(a)$, and for the firm, we will assume that the value of an unfilled vacancy is 0, produced through the standard free entry condition (FEC). We can write the problem as

$$\max_{w,\tau} \left(\tilde{V}_E(i,j;w,\tau_a,\tau_\theta) - V_U(i) \right)^\alpha \tilde{V}_F(i,j;w,\tau_a,\tau_\theta)^{1-\alpha},$$

where $\tilde{V}_E$ and $\tilde{V}_F$ functions are the value of employment to the worker and to the firm, respectively, given the wage and investment times.

We first consider the unemployment state. We will assume that the flow value of unemployment to an individual of type a_i is proportional to a_i, or ba_i, $i = 1, ..., M$. Then we can write

$$V_U(i) = \frac{ba_i + \lambda_U \sum_{j=r^*(i)+1} p_j V_E(i,j)}{\rho + \lambda_U \tilde{G}(\theta_{r^*(i)})} \tag{1}$$

where the critical (index) value $r^*(i)$ is defined by

$$
\begin{aligned}
V_U(i) &\geq V_E(i, \theta_{r^*(i)}) \\
V_U(i) &< V_E(i, \theta_{r^*(i)+1}).
\end{aligned}
$$

An agent of general ability a_i will reject any match values of $\theta_{r^*(i)}$ or less, and accept any match values greater than this.[7]

Given a wage of w and a training level of τ_a and τ_θ, the value of employment of

---

[7]Note that we assume that there are no shocks to the individuals' ability level during unemployment.

type a_i at a match of θ_j is

$$\tilde{V}_E(i,j;w,\tau_a,\tau_\theta) = (\rho + \varphi_a(i,\tau_a) + \varphi_\theta(j,\tau_\theta) + \tilde{\delta}_a(i) + \tilde{\delta}_\theta(j) + \eta)^{-1} \times$$
$$[w + \varphi_a(i,\tau_a)Q(i+1,j) + \varphi_\theta(j,\tau_\theta)V_E(i,j+1) + \tilde{\delta}_a(i)Q(i-1,j)$$
$$+\tilde{\delta}_\theta(j)Q(i,j-1) + \eta V_U(a_i)],$$

where $\tilde{\delta}_k(i) = 0$ if $i = 1$ and $\tilde{\delta}_k(i) = \delta_k$ if $i > 1$, for $k = a, \theta$. The term

$$Q(i,j) \equiv \max[V_E(i,j), V_U(i)],$$

allows for the possibility that a reduction in the value of a or θ could lead to an endogenous termination of the employment contract, with the employee returning to the unemployment state. It also allows for the possibility that an increase in a from a_i to a_{i+1} could lead to an endogenous separation. This could occur if the reservation θ, $r^*(i)$, is increasing in i. In this case, an individual employed at the minimally acceptable match $r^*(i) + 1$, may quit if a improves and $r^*(i+1) \geq r^*(i) + 1$.

The corresponding value to the firm is

$$\tilde{V}_F(i,j;w,\tau_a,\tau_\theta) = (\rho + \varphi_a(i,\tau_a) + \varphi_\theta(j,\tau_\theta) + \tilde{\delta}_a(i) + \tilde{\delta}_\theta(j) + \eta)^{-1} \times$$
$$[(1 - \tau_a - \tau_\theta)y(i,j) - w + \varphi_a(i,\tau_a)Q_F(i+1,j) + \varphi_\theta(j,\tau_\theta)V_F(i,j+1)$$
$$+\tilde{\delta}_a Q_F(i-1,j) + \tilde{\delta}_\theta Q_F(i,j-1)]$$

where $Q_F(i,j) = 0$ if $Q(i,j) = V_U(i)$ and $Q_F(i,j) = V_F(i,j)$ if $Q(i,j) = V_E(i,j)$. Then the solution to the surplus division problem is given by

$$\{w^*(i,j), \tau_a^*(i,j), \tau_\theta^*(i,j)\} = \arg\max_{w,\tau_a,\tau_\theta} \left(\tilde{V}_E(i,j;w,\tau_a,\tau_\theta) - V_U(i)\right)^\alpha$$
$$\times \tilde{V}_F(i,j;w,\tau_a,\tau_\theta)^{1-\alpha};$$
$$V_E(i,j) = \tilde{V}_E(i,j;w^*(i,j),\tau_a^*(i,j),\tau_\theta^*(i,j)),$$
$$V_F(i,j) = \tilde{V}_F(i,j;w^*(i,j),\tau^*(i,j),\tau_\theta^*(i,j)).$$

More specifically, the surplus division problem is given by

$$\max_{w,\tau_a,\tau_\theta} (\rho + \varphi_a(i,\tau_a) + \varphi_\theta(j,\tau_\theta) + \tilde{\delta}_a(i) + \tilde{\delta}_\theta(j) + \eta)^{-1}$$

$$\times \left[\begin{array}{c} w + \varphi_a(i,\tau_a)[Q_E(i+1,j) - V_U(i)] + \varphi_\theta(j,\tau_\theta)[V_E(i,j+1) - V_U(i)] \\ +\tilde{\delta}_a(i)[Q(i-1,j) - V_U(i)] + \tilde{\delta}_\theta(j)[Q(i,j-1) - V_U(i)] - \rho V_U(a_i) \end{array} \right]^\alpha$$

$$\times \left[\begin{array}{c} (1-\tau)y(i,j) - w + \varphi_a(i,\tau_\alpha)Q_F(i+1,j) + \varphi_\theta(j,\tau_\theta)V_F(i,j+1) \\ +\tilde{\delta}_a(i)Q_F(i-1,j) + \tilde{\delta}_\theta(j)Q_F(i,j-1) \end{array} \right]^{1-\alpha}.$$

The first order conditions for this problem can be manipulated to get the reasonably

standard wage-setting equation,

$$w^*(i,j) = \alpha\{(1 - \tau_a^* - \tau_\theta^*)y(i,j) + \varphi_a(i,\tau_a^*)Q_F(i+1,j) + \varphi_\theta(j,\tau_\theta^*)V_F(i,j+1)$$
$$+ \tilde{\delta}_a(i)Q_F(i-1,j) + \tilde{\delta}_\theta(j)Q_F(i,j-1)\}$$
$$+ (1-\alpha)\{\rho V_U(i) - \varphi_a(i,\tau_a^*)(V_E(i+1,j) - V_U(i))$$
$$- \varphi_\theta(j,\tau_\theta^*)(V_E(i,j+1) - V_U(i))] - \tilde{\delta}_a(i)Q(i-1,j) - \tilde{\delta}_\theta(j)Q(i,j-1)\}.$$

The first order conditions for the investment times τ_a and τ_θ are also easily derived, but are slightly more complex than the wage condition. The assumptions regarding the investment technologies φ_a and φ_θ have important implications for the investment rules, obviously. The time flow constraint is

$$1 \geq \tau_a + \tau_\theta,$$
$$\tau_a \geq 0$$
$$\tau_\theta \geq 0.$$

Depending on the parameterization of the production technology, it is possible that optimal flow investment of either type is 0, that one type of investment is 0 while the other is strictly positive, and even that all time is spent in investment activity, whether it be in one kind of training or both. In such a case, it is possible to produce the implication of negative flow wages, and we shall not explicitly assume these away by imposing a minimum wage requirement in estimation. In the case of internships, for example, which are supposed to be mainly investment activities, wage payments are low or zero. Including the worker's direct costs of employment, the effective wage rate may be negative. What is true is that no worker-firm pair will be willing to engage in such activity without the future expected payoffs being positive, which means that the worker would generate positive flow profits to the firm at some point during the job match.

2.2 On-the-Job Search

In the case of on-the-job search, individuals who are employed are assumed to receive offers from alternative employers at a rate λ_E, and it is usually the case that $\lambda_E < \lambda_U$. If the employee meets a new employer, the match value at the alternative employer, $\theta_{j'}$, is immediately revealed. Whether or not the employee leaves for the new job and what the new wage of the employee is after the encounter depends on assumptions made regarding how the two employers compete for the individual's labor services. In Flinn and Mabli (2009), two cases were considered. In the first, in which employers are not able to commit to wage offers, the outside option in the wage determination problem always remains the value of unemployed search, since this is the action available to the employee at any moment in time. This model produces an implication of efficient mobility, in that individuals will only leave a current employer if the match produc-

tivity at the new employer is at least as great as current match productivity (general productivity has the same value at all potential employers). An alternative assumption, utilized in Postel-Vinay and Robin (2002), Dey and Flinn (2005), and Cahuc et al. (2006), is to allow competing employers to engage in Bertrand competition for the employee's services (this model assumes the possibility of commitment to the offered contract on the part of the firm). In this case, efficient mobility will also result, but the wage distribution will differ in the two cases, with employees able to capture more of the surplus (at the same value of the primitive parameters) in the case of Bertrand competition. We begin our discussion with the Bertrand competition case in this section, although we estimate the model under both scenarios. For reasons explained at the end of this section, the empirical work will emphasize the no-renegotiation case.

Under either scenario, a_i has no impact on mobility decisions, since it assumes the same value across all employers. In the Bertrand competition case (as in Dey and Flinn (2005), for example), the losing firm in the competition for the services of the worker is willing to offer all of the match surplus in its attempt to retain the worker. For example, let the match value at the current employer be θ_j, and the match value at the potential employer be $\theta_{j'}$. We will denote the maximum value to an employee of type a_i of working at a firm where her match value is θ_j by $\bar{V}(i,j)$, which is the case in which the employee captures all of the match surplus (since the value of holding an unfilled vacancy is assumed to be equal to 0, by transferring all of its surplus to the employee, the firm is no worse of than it would be holding an unfilled vacancy). In the Bertrand competition case then, and assuming that $j' \leq j$, the losing firm offers $\bar{V}(i,j')$ for the individual's labor services. The winning firm then divides the surplus with the employee, where the employee's outside option becomes $\bar{V}(i,j')$. Note that in the case that $j' = j$, the individual would be indifferent between the two firms, the two firms would be indifferent with respect to hiring her or not, and whichever offer it accepted, the employee would capture the entire match value, that is, $V_E(i,j) = \bar{V}(i,j)$. Because of the investment possibilities, it is not generally the case that the wage at the winning firm will be equal to $a_i\theta_j - \zeta$, which would be true when there are no investment possibilities.

In the case of on-the-job search with Bertrand competition between employers, we denote the value of the employment match to the worker and the firm by $V(i,j,j')$ and $V_F(i,j,j')$, respectively. The first argument denotes the individual's general ability type, a_i, and the second denotes the value of the match at the employer. The third argument in the function is the highest match value encountered during the current employment spell (which is a sequence of job spells not interrupted by an unemployment spell) at any other employer. Since mobility decisions are efficient, we know that $j' \leq j$. When the individual has encountered no other match values during the current employment spell that exceeded the value $r^*(i) + 1$, then we will write $V(i,j,j^*(i))$. When an individual encounters a new firm with a new match draw j'', then the individual's

new value of being employed is given by

$$
\begin{array}{lll}
V(i,j'',j) & \text{if} & j'' > j \\
V(i,j,j'') & \text{if} & j \geq j'' > j' \\
V(i,j,j') & \text{if} & j' \geq j''
\end{array}
\tag{2}
$$

In the first row, the individual changes employer, and now the match value at the current employer becomes the next best match value during the current employment spell. In the second row, the employee stays with her current employer, but gains more of the total surplus associated with the match, which implies an increase in her wage at the employer. In the third row, the individual does not report the encounter to her current employer, since it doesn't increase her outside option.

To see these effects more formally, we first consider the case in which a worker with a current match value of θ_j who has previously worked at a job with a match value of θ_k, $k \leq j$, and where there was no intervening unemployment spell. In this case, we write the worker's value given wage w and training time τ as

$$
\tilde{V}_E(i,j,k;w,\tau_a,\tau_\theta) = \frac{N_E(w,\tau_a,\tau_\theta;i,j,k)}{D(\tau_a,\tau_\theta;i,j,k)},
$$

where

$$
N_E(w,\tau_a,\tau_\theta;i,j,k) = w + \lambda_E[\sum_{s=k+1}^{j} p_s V_E(i,j,s) + \sum_{s=j+1} p_s V_E(i,s,j)]
$$
$$
+\varphi_a(i,\tau_a)Q(i+1,j,k) + \varphi_\theta(j,\tau_\theta)V_E(i,j+1,k)
$$
$$
\tilde{\delta}_a(i)Q(i-1,j,k) + \tilde{\delta}_\theta(j)Q(i,j-1,k) + \eta V_U(i);
$$

$$
D(\tau_a,\tau_\theta;i,j,k) = \rho + \lambda_E \tilde{G}(\theta_k) + \varphi_a(i,\tau_a) + \varphi_\theta(j,\tau_\theta)
$$
$$
\tilde{\delta}_a(i) + \tilde{\delta}_\theta(j) + \eta.
$$

The term $Q(i+1,j,k) = \max\{V(i+1,j,k), V_U(i+1)\}$, indicating the possibility that an increase in a could lead to an endogenous separation depending on the value of θ_j. The term $Q(i-1,j,k) = \max\{V(i-1,j,k), V_U(i-1)\}$, indicating that the value of unemployed search has decreased as well. Finally, we have $Q(i,j-1,k) = \max\{V(i,j-1,\min(j-1,k)), V_U(i)\}$. In the case where $j = k$, this implies that the value of the outside option is reduced with the current match value. We impose this convention so as to keep the surplus division problem well-defined. Other assumptions could be made regarding how the negotiations between and employer and employee are impacted when the match value decreases.

The value to the firm is given by

$$
\tilde{V}_F(i,j,k;w,\tau_a,\tau_\theta) = \frac{N_F(w,\tau_a,\tau_\theta;i,j,k)}{D(\tau_a,\tau_\theta;i,j,k)},
$$

13

where

$$N_F(w, \tau_a, \tau_\theta; i, j, k) = y(i, j)(1 - \tau_a - \tau_\theta) - w + \varphi_a(i, \tau_a)Q_F(i + 1, j, k)$$
$$+\varphi_\theta(j, \tau_\theta)V_F(i, j + 1, k) + \tilde{\delta}_a(i)Q_F(i - 1, j, k)$$
$$+\tilde{\delta}_\theta(j)Q_F(i, j - 1, \min(j - 1, k)) + \lambda_E \sum_{s=k+1}^{j} p_s V_F(i, j, s).$$

Now the surplus division problem is

$$\max_{w, \tau_a, \tau_\theta} D(\tau_a, \tau_\theta; i, j, k)^{-1}[N_E(w, \tau_a, \tau_\theta; i, j, k) - \bar{V}(i, k)]^\alpha$$
$$\times N_F(w, \tau_a, \tau_\theta; i, j, k)^{1-\alpha},$$

which is only slightly more involved than the problem without OTJ search, but the generalization yields another fairly complex dependency between the current value of the match and the training time decisions. It is clear that the value of match-specific investment to the employer in the case of OTJ search is even higher than in the no OTJ case, since it also increases (in expected value) the duration of the match, and this value always exceeds the value of an unfilled vacancy, which is 0. The value of either type of training is also enhanced from the point of view of the worker, since in addition to increasing her value at her current employer, higher values of a or θ enhance her future bargaining position during the current employment spell, and, in the case of a, even beyond the current employment spell. Once the employment spell ends, the bargaining advantage from the match history ends, including gains accumulated through investment in match-specific productivity. On the other hand, the value of previous investments in general productivity is carried over, in a stochastic sense, which is what makes this type of human capital particularly valuable from the worker's perspective, and accounts for her disproportionate costs of funding these investments. Finally, the value of unemployed search is given by

$$V_U(i) = \frac{ba_i + \lambda_U \sum_{j=r^*(i)+1} p_j V_E(i, j, j^*(i))}{\rho + \lambda_U \tilde{G}(\theta_{r^*(i)})}.$$

As we have seen, in the case of Bertrand competition, there is some arbitrariness in defining the employment state when the outside option and current match values are equal and there is depreciation in the current match value. In this case, we have simply assumed that the employee continues to receive the entire surplus of the match, although this total surplus has decreased due to the decrease in match-specific productivity from θ_j to θ_{j-1}. The other case of employer-employee interaction we consider is when employers do not respond to outside offers. This would be the case when outside offers cannot be observed and verified. Moreover, even if they were, employers have an incentive to cheat on the employment contract agreed to once the outside offer is no longer available. When the outside offer is removed, the employee's only alternative

is to quit into unemployed search, so that this is the outside option considered when deciding upon wage-setting and the amount of work time devoted to investment.

In this case, decisions are considerably simplified. As in the case of no OTJ search, the employment contract is only a function of the individual's type and the current match value, (i, j). The property of efficient turnover decisions continues to hold, with the employee accepting all jobs with a match value $j' > j$, and refusing all others. The formal structure of the problem is modified as follows.

$$\tilde{V}_E(i, j; w, \tau_a, \tau_\theta) = \frac{N_E(w, \tau_a, \tau_\theta; i, j)}{D(\tau_a, \tau_\theta; i, j)},$$

where

$$N_E(w, \tau_a, \tau_\theta; i, j) = w + \lambda_E \sum_{s=j+1} p_s V_E(i, s)$$

$$+\varphi_a(i, \tau_a) Q(i + 1, j) + \varphi_\theta(j, \tau_\theta) V_E(i, j + 1)$$

$$+\tilde{\delta}_a(i) Q(i - 1, j) + \tilde{\delta}_\theta(j) Q(i, j - 1) + \eta V_U(a_i);$$

$$D(\tau_a, \tau_\theta; i, j) = \rho + \lambda_E \tilde{G}(\theta_j) + \varphi_a(i, \tau_a) + \varphi_\theta(j, \tau_\theta)$$

$$+\tilde{\delta}_a(i) + \tilde{\delta}_\theta(j) + \eta.$$

The term $Q(i + 1, j) = \max\{V(i + 1, j), V_U(i + 1)\}$, indicating the possibility that an increase in a could lead to an endogenous separation depending on the value θ_j. The term $Q(i - 1, j) = \max\{V(i - 1, j), V_U(i - 1)\}$, indicating that the value of unemployed search has decreased as well. Finally, we have $Q(i, j - 1) = \max\{V(i, j - 1), V_U(i)\}$.

The value to the firm conditional on the wage and investment decisions is given by

$$\tilde{V}_F(i, j; w, \tau_a, \tau_\theta) = \frac{N_F(w, \tau_a, \tau_\theta; i, j)}{D(\tau_a, \tau_\theta; i, j)},$$

where

$$N_F(w, \tau_a, \tau_\theta; i, j) = y(i, j)(1 - \tau_a - \tau_\theta) - w + \varphi_a(i, \tau_a) Q_F(i + 1, j)$$

$$+\varphi_\theta(j, \tau_\theta) V_F(i, j + 1) + \tilde{\delta}_a(i) Q_F(i - 1, j)$$

$$+\tilde{\delta}_\theta(j) Q_F(i, j - 1),$$

and where $Q_F(i + 1, j) = V_F(i + 1, j)$ if $V(i + 1, j) > V_U(i + 1)$ and equals 0 otherwise, $Q_F(i - 1, j) = V_F(i - 1, j)$ if $V(i - 1, j) > V_U(i - 1)$ and equals 0 otherwise, and $Q_F(i, j - 1) = V_F(i, j - 1)$ if $j - 1 > r^*(i)$. Now the surplus division problem becomes:

$$\max_{w, \tau_a, \tau_\theta} D(\tau_a, \tau_\theta; i, j)^{-1} [N_E(w, \tau_a, \tau_\theta; i, j) - V_U(i)]^\alpha$$

$$\times N_F(w, \tau_a, \tau_\theta; i, j)^{1-\alpha}.$$

15

The value of unemployed search in this case is simply

$$V_U(i) = \frac{ba_i + \lambda_U \sum_{j=r^*(i)+1} p_j V_E(i,j)}{\rho + \lambda_U \tilde{G}(\theta_{r^*(i)})}.$$

In what follows, we will emphasize the estimates associated with the no renegotiation model. This is due to its relative simplicity, and the fact that in other studies (Flinn and Mabli (2009), Flinn and Mullins (2015)) and in this one, we have found that the no renegotiation model fits the sample characteristics used to define our Method of Simulated Moments (MSM) estimator better than does the Bertrand competition model. Of course, in a model without investment options, the model without renegotiation implies that wages will be constant over a job spell of an individual and a particular firm. The Bertrand competition assumption in a stationary search setting implies that wage gains may be observed over the course of a job spell, but never wage declines. Of course, in the data we see a number of wage decreases over a job spell. No doubt, many of these are due solely to measurement error, or the fact that wages fixed in nominal terms across interview dates will imply real wage declines in the face of inflation. Our model, with endogenous productivity shocks in both general and specific human capital, is capable of generating both types of wage fluctuations without relying on the use of difficult to verify bargaining protocols.

2.3   Equilibrium Model

The model described to this point is one set in partial equilibrium, with contact rates between unemployed and employed searchers and firms viewed as exogenous. The model can be closed most simply by employing the matching function framework of Mortensen and Pissaridies (1994). We let the measure of searchers be given by $S = U + \xi E$, where U is the steady state measure of unemployed and E is the measure of the employed ($= 1 - U$, since we assume that all individuals are participants in the labor market). The parameter ξ reflects the relative efficiency of search in the employed state, and it is expected that $0 < \xi < 1$. We denote the measure of vacancies posted by firms by v. The flow contact rate between workers and firms is given by

$$M = S^\phi v^{1-\phi},$$

with $\phi \in (0, 1)$.[8] Letting $k \equiv v/S$ be a measure of labor market tightness, we can write the rate at which searchers contact firms holding vacancies by

$$\frac{M}{v} = k^\phi.$$

The proportion of searchers who are employed is given by $\xi E/S$, so that the mass

---

[8]We have fixed $TFP = 1$ in the Cobb Douglas matching function due to the impossibility of identifying this parameter given the data available.

of matches that involve an employed worker is simply $\xi E / S \times M$, which means that the flow rate of contacts for the employed is

$$
\begin{aligned}
\lambda_E &= \frac{\xi E}{S} \frac{S^\phi v^{1-\phi}}{E} \\
&= \xi k^{\phi-1}.
\end{aligned}
$$

A similar argument is used to find the contact rate for unemployed searchers,

$$
\lambda_U = k^{\phi-1}.
$$

A fact that will be utilized in the estimation of demand side parameters below is that $\xi = \lambda_E / \lambda_U$.

Let the flow cost of holding a vacancy be given by $\psi > 0$. The steady state distributions of a among the unemployed and (a, θ) among the employed are complex objects that have no closed form solution, due to the (endogenous) dynamics of the a and θ processes in the population. However, these distributions are well-defined objects, the values of which can be obtained through simulation. The way we obtain the steady state distributions through simulation is given in Appendix A.

Let the steady state distribution of a among the unemployed be given by $\{\pi_i^U\}$, $i = 1, ..., M$, and the steady state distribution of (a, θ) among the employed by $\{\pi_{i,j}^E\}$, $i = 1, ..., M, j = 1, ..., K$. Then the expected flow value of a vacancy in the steady state is given by

$$
\begin{aligned}
-\psi + \frac{k^\phi}{S} &\times \{U \sum_i \sum_{j=r^*(i)+1} p_j V_F(i,j)\pi_i^U \\
&+ \xi E \sum_i \sum_{j'} \sum_j p_{j'} V_F(i,j')\pi_{i,j}^E\}.
\end{aligned}
$$

By imposing a free entry condition on firms that equates this value to zero, the equation can be solved for equilibrium values of λ_U and λ_E given knowledge of the parameters ψ, ϕ, and ξ.

3 Data

We utilize data from the National Longitudinal Survey of Youth 1997 (NLSY97) to construct our estimation sample. The NLSY97 consists of a cross-sectional sample of $6,748$ respondents designed to be representative of people living in the United States during the initial survey round and born between January 1, 1980, and December 31, 1984, and a supplemental sample of $2,236$ respondents designed to oversample Hispanic, Latino and African-American individuals. At the time of first interview, respondents' ages

range from 12 to 18, and at the time of the interview from the latest survey round, their ages range from 26 to 32.

For our analysis, we use a subsample of 1,994 respondents from the NLSY97. We obtain this sample through three main selection criteria: (1) The oversample of Hispanic, Latino and African-American respondents are excluded so that the final sample comprises only of the nationally representative cross-sectional sample, (2) The military sample is excluded, (3) All females and high-school dropouts are excluded. A respondent who satisfies these criteria enters our sample after completing all schooling.

The estimation sample is constructed this way since our model is not designed to explain behavior while in school and staying in school or continuing education are not endogenous choices. These sample selection criteria give us an unbalanced sample of 1,994 individuals and 661,452 person-week observations. The proportion of high school graduates is 37 percent and the proportion of those with some college and those with a college degree are 30 and 33 percent, respectively.

NLSY97 provides detailed retrospective data on the labor market histories and wage profiles of each respondent. This retrospective data is included in the employment roster, which gives the start/end dates of each employment spell experienced by the respondent since the last interview, wage profiles and other characteristics of the each employment (or unemployment) episode. We use the employment roster to construct weekly data on individual labor market histories. This information provides us with some of the key moments that identify the parameters of the search environment faced by the agents in our model. Some of these moments are transitions between employment states and transitions between jobs, average wages and wage changes during employment transitions, and wage growth within jobs.

While we make extensive use of the weekly data constructed retrospectively from the NLSY97 employment rosters for obtaining moments related to training, we mostly use information collected from respondents on interview dates for our empirical analysis on wages and wage transitions. This is mainly because of the potential measurement problems inherent in the weekly employment data in NLSY97 due to its retrospective nature. More specifically, in each annual survey round, the NLSY97 respondents answer detailed questions about current and previously held jobs and this data is collected about every employer for whom the respondent worked since her previous interview so that a complete picture of the respondent's employment can be constructed. That there is usually approximately a year between each interview date brings into question the accuracy of respondents' answers regarding the detailed questions they are asked about especially the wages that pertain to each employment episode experienced during the course of the year. For these reasons, while we still take advantage of some of the retrospective weekly data in the employment roster, we use it in conjunction with information collected on interview dates and put more emphasis on the latter especially for moments related to wages and wage transitions.

Some of the implications of restricting our data to interview dates can be seen in Table 1, which displays the percentage of job spells by the number of interview dates they span. The proportions for each number of interview date are shown separately

18

for each schooling level and they closely mirror the actual duration distribution of jobs obtained from the retrospective data from employment rosters. They are not exactly the same since the start date of the job and the calendar date of the interview also matter in addition to the actual duration of the job spell. However, the maximum discrepancy should be one year. We see in Table 1's second row that for high school graduates, about 61 percent of all observed job spells cover no interview dates at all. This is due to the fact that a large proportion of job spells for this group end in less than a year and do not last long enough to coincide with an interview date. For the same education group, 23 percent of all observed job spells span one interview date and 7 percent last long enough to span two interview dates. The proportion of job spells with longer durations increase by education level. Consequently, it can be seen from Table 1 that the proportion of higher number of interview dates covered increases by education level. For example, the proportion of job spells that span two interview dates is 11 percent for individuals with a college degree. In the context of the model, the duration of a job spell is indicative of the value of the match between worker and the firm. This value is a result of initial match-specific quality, initial individual level worker productivity as well as the specific or general training taken on by the firm-worker pair during the course of the job spell.

In addition to key labor market variables, NLSY97 contains a wealth of information that is of central importance to the focus of our analysis, which is human capital growth and factors that govern firm and worker incentives to engage in different types of human capital investment. In our model, human capital investment on the job explicitly involves time and learning by doing. For this aspect of our analysis, we use NLSY97's training roster, where respondents are asked about what types of training they receive over the survey year and the start/end dates of training periods by source of training.[9] Combining the information from the employment and training rosters, we construct a weekly event history of employment and training for each respondent. We do not make assumptions regarding the specificity of human capital acquired during a training episode. Instead, we use the empirical relationship between the patterns of training and previous/future employment and wage transitions in order to make inferences about the degree of specificity in the human capital accumulation process.

Tables 2-3 present some descriptive statistics on the training patterns observed in our sample. More specifically, these tables display the incidence of training by schooling and the timing of training spells by job tenure. The proportion of respondents with at least one training spell is 18 percent for high school graduates. For higher schooling categories, this proportion is 13 percent. Moreover, most respondents experience only one training spell during the time they are observed in the sample: It can be seen in Table 3 that proportion with one training spell (conditional on having at least one training spell throughout the labor market history) is 72 percent.

---

[9]Some examples to sources of training are business colleges, nursing programs, apprenticeships, vocational and technical institutes, barber and beauty schools, correspondence courses, and company training. Training received in formal regular schooling programs is included in the schooling variables.

We next discuss training in relation to employment and wage transitions in our sample. Table 4 provides detailed information about job-to-job transitions between interview dates and the impact of these transitions on wage levels. In Table 4, the first two rows of Panel A focuses on transitions where the worker is employed in both interview dates $t-1$ and t. These rows distinguish between two types of transitions: (1) employment transitions that involve a chance in employer (job switchers), and (2) employment transitions that do not involve a change in employer (i.e. worker simply stays at her current firm). For example, we see in the first row of Panel A that out of all high school graduates who remain employed at interview dates $t-1$ and t, only 19 percent had a different employer at t relative to employer at $t-1$. The first row of Panel A also shows that this proportion is lower for the higher schooling categories.

The last two rows of Panel A distinguish between different types of job switchers according to whether the workers experience an intervening period of nonemployment between their employment episodes. These are the cases where individuals are observed to be not working for a period of at least 13 weeks between the first employment episode that covers their interview date $t-1$ and second employment episode that covers their interview date t. For example, consider an individual who is interviewed in September 2002 and also interviewed in November of the following year. Suppose that this individual is observed to be currently employed on both of these interview dates. The nonemployment spell in between these dates is determined according to this individual's weekly employment status in the periods between these dates. If the individual is observed to leave the first job in May 2003 and not start the second employment episode until August 2003, then this person is considered to be a job switcher with an intervening nonemployment spell between the interview dates 2002 and 2003. It can be seen that a large proportion experience no nonemployment between their consecutive employment spells. For example, for high school graduates, the proportion of job switchers who experience an intervening nonemployment spell is 4 percent and the proportion who do not experience such a spell is 15 percent. Note that the sum is 19 percent, which is the overall percentage of job switchers among the employment-to-employment transitions. Unlike the overall decomposition between job switchers and stayers, there is no clear pattern by education level for job switchers with nonemployment spells.

As discussed previously, wage growth within and across job spells are important indicators of which type of human capital investment behavior workers engage in. Panel B shows the difference between log wages for employment-to-employment transitions between interview dates. These transitions are separated into groups of job switchers and stayers as well as job switchers with or without nonemployment spells in between in the same manner described in Panel A. One of the patterns that stand out in Panel B is the fact that mean log wage difference ($logw_t - logw_{t-1}$) experienced by job switchers increases by education level. More specifically, average log wage difference is 0.11, 0.15 and 0.20 for high school graduates, individuals with some college education and college graduates, respectively.

Table 5 shows mean log wage differences ($logw_t - logw_{t-1}$) by training status be-

20

tween consecutive interview dates for transitions between employment spells, where an employment spell is defined as a sequence of jobs not interrupted by an unemployment spell. As explained above, we do not use retrospective weekly wage information for these moments. An individual is considered to be a part of the group with training if he/she receives general or match-specific training by the interview date $t-1$. We see that for stayers, the average log wage differences by training status are equal. On the other hand, for job switchers, the average lof wage difference between t and $t-1$ for those individuals who obtained some form of training in the first job spell is smaller. In terms of what these differences mean for whether the training acquired is general or match-specific, it is more informative to look at the wage differences for job switchers distinguishing between whether the individual experienced a period of nonemployment in between the two job spells. These are displayed in the last two rows of Table 5. We see that for individuals who moved to their next job with no intervening spell of nonemployment, the average log wage difference is 0.15 if they received no training in the previous job and it is 0.09 otherwise.

In the context of the model, wage growth for job switchers occurs through four channels. The first is human capital accumulation that arises through training that takes place during the course of the first job spell (the job that is still ongoing at interview date $t-1$). The second is due to the fact that the return to the worker's general human capital may be different at the new firm compared to the previous firm. This is because of bargaining, which allows the worker to potentially acquire a different share of the total match surplus at the new firm. The third is due to the possibility that the worker's initial match quality with the new firm may be higher than the one he/she had by the end of his/her previous job spell. The fourth channel is the difference in the training time between the two jobs. The fact that wage growth seems to be lower for job changes that follow an employment spell with training (0.15 vs. 0.09 as summarized above), might mean that the training that took place prior to the move is more general rather than match-specific. It is difficult to draw definitive conclusions based on one moment, but all else equal, the theoretical model shows that there should be a positive correlation between the level of match-specific capital in the previous firm and the wage acquired at the new firm since the wage change required for moving to a new firm will be higher if the match quality at the incumbent firm is high.

4 Econometric Issues

4.1 Empirical Implementation of the Model

We make several assumptions in order to solve the model, which does not produce closed form solutions. We restrict workers and firms to choose training times from a discrete choice set consisting of multiples of five percent of the worker's total time ($\tau_a, \tau_\theta \in \{.00, .05, .10, ...1.00\}$). The production functions are assumed to have the

following functional forms. Recall that there are M values of a, $0 < a_1 < ... < a_M$. There is no ability to increase ability if an individual is already at the highest level, so the hazard rate for improvements for a_M is equal to 0. For $i < M$, we have that the hazard rate to level $i + 1$ is given by

$$\varphi_a(i, \tau_a) = \delta_a^0 \times a_i^{\delta_a^1} \times (\tau_a)^{\delta_a^2},$$

where δ_a^0, δ_a^1, and δ_a^2 are scalar constants. Similarly, there are K values of θ, $0 < \theta_1 < ... < \theta_K$, and no possibility to increase match productivity when $\theta = \theta_K$. For a worker with $i < K$ who spends a fraction τ_θ of his time in firm-specific training, the value of the match increases at rate

$$\varphi_\theta(j, \tau_\theta) = \delta_\theta^0 \times \theta_j^{\delta_\theta^1} \times (\tau_\theta)^{\delta_\theta^2},$$

where again δ_θ^0, δ_θ^1, and δ_θ^2 are scalar constants.

Because it is difficult to separately identify the level of general ability and match quality, we attempted to make the support of the distributions of a and θ as symmetric as possible. Therefore, we choose identical grids for for a_i and θ_j. We chose grid points to cover the range of likely values of θ including the possibility that workers with high values of θ will receive match-specific training that will produce match values above the set of values that they would naturally receive from searching. In the end, we use a grid containing 24 points which are spaced logarithmically from 2.5 standard deviations below the mean of the theta distribution to 3.5 standard deviations above it. At the estimated parameters of our baseline model, moving up by one grid point in either a or θ corresponds to a roughly 9 percent increase in productivity.

Several model parameters are fixed outside the estimation. We choose $\alpha = 0.5$, giving the worker and firm equal bargaining weight. All rate parameters are expressed at a weekly frequency and we set the discount factor $\rho = 0.0016$, corresponding to to a four percent annual discount rate.

Training observed in the data is likely a very rough proxy for the amount of time spent developing workers' human capital. To relate our observed measures of training in the data to the training time chosen in the model simulations, we assume that a worker who spends a fraction of time τ engaged in training is observed to receive training is that period with probability

$$\text{Prob(Training observed} \mid \tau) = \Phi(\beta_0 + \beta_1 \tau)$$

where Φ is the c.d.f. for the normal distribution. In calculating τ from the simulations, we compute the average fraction of time spent training over each six month period, or, for job spells lasting less than six months, over the entire job spell. We estimate the parameters β_0 and β_1 jointly with the other parameters of the model, giving us a total of 21 parameters to estimate.

4.2   Estimator

4.2.1   Estimation of Supply-Side Parameters

We utilize a method of simulated moments estimator (MSM) in order to estimate all of the parameters of the model with the exception of those characterizing firms' vacancy decisions. Under the data generating process (DGP) of the model, there are a number of sharp restrictions on the wage and mobility process that are generally not consistent with the empirical distributions observed. In such a case, measurement error in wage observations is often added to the model. This is not really a feasible alternative here given that we are already trying to estimate a convolution, so that the addition of another random variable to the wage and mobility processes can only exacerbate the difficulty of separately identifying the distributions of a and θ, particularly given their endogeneity with respect to investment decisions. We chose to use a moment-based estimator which employs a large amount of information characterizing wage distributions within and across jobs, often by schooling class, as well as some training information, as was described in the previous section.

The information from the sample that is used to define the estimator is given by M_N, where there are N sample observations. Under the DGP of the model, the analogous characteristics are given by $\tilde{M}(\omega)$, where ω is the vector of all identified parameters (which are all parameters and decision rules except ρ). Then the estimator is given by

$$\hat{\omega}_{N,W_N} = \arg\min_{\omega \in \Omega}(M_N - \tilde{M}(\omega))'W_N(M_N - \tilde{M}(\omega)),$$

where W_N is a symmetric, positive-definite weighting matrix and Ω is the parameter space. The weighting matrix, W_N, is a diagonal matrix with elements equal to the inverse of the variance of the corresponding element of M_N. Under our random sampling assumption, we have that $\text{plim}_{N\to\infty}M_N = M$, the population value of the sample characteristics used in estimation. Since W_N is a positive-definite matrix by construction, our moment-based estimator is consistent since $\text{plim}_{N\to\infty}\hat{\omega}_{N,Q} = \omega$ for any positive-definite matrix Q. We compute bootstrap standard errors using 100 replications.

4.2.2   Demand-Side Parameter Estimator

It is most often the case that the parameters characterizing firms' vacancy creation decisions are not identified. Our estimator of an employment cost parameter may enable identification of a Cobb-Douglas matching function parameter. In this section we explore how this can be accomplished.

Recall that the matching function was defined as

$$M = \nu^\phi S^{1-\phi},$$

with $\nu \in (0,1)$, and the measure of searchers was given by $S = U + \xi E$, where U is

the measure of unemployed and E is its complement. The parameter ξ is a measure of the search efficiency of employed agents relative to that of the unemployed, and it is expected that $\xi \in (0,1]$. M is the flow matching rate, and we note that we have assumed that the TFP parameter is equal to 1 to aid in identification (the number of matches is unobserved, so that this essentially amounts to a normalization). The rate at which employers with vacancies contact applicants is

$$
\begin{aligned}
\lambda_F &= \frac{M}{\nu} \\
&= \nu^{\phi-1} S^{1-\phi} \\
&= k^{\phi-1},
\end{aligned}
$$

where $k \equiv \nu/S$ is a measure of labor market tightness.

The proportion of matches that involve an unemployed worker is given by

$$
\frac{U}{U + \xi E} M,
$$

so that the contact rate per unemployed searcher is

$$
\begin{aligned}
\lambda_U &= \frac{U}{S} \frac{M}{U} \\
&= k^{\phi}.
\end{aligned}
$$

The contact rate for employed searchers is

$$
\begin{aligned}
\lambda_E &= \frac{\xi E}{S} \frac{M}{E} \\
&= \xi k^{\phi}.
\end{aligned}
$$

Proposition 1 *If a consistent estimator of the cost of posting a vacancy, ψ, is available, then the matching function parameter ϕ can be consistently estimated.*

Proof Our first stage MSM estimator produces consistent estimates of λ_U and λ_E. Then a consistent estimator of ξ is given by

$$
\hat{\xi} = \hat{\lambda}_E / \hat{\lambda}_U.
$$

Using the model estimates, we can compute consistent estimates of the steady state values of U and E, which are denoted by $\hat{U}$ and $\hat{E}$. The free entry condition (FEC) implies that

$$
0 = -\psi + \lambda_F p(\mathrm{A}) E(V_F | \mathrm{A}),
$$

where the event A denotes job acceptance. From the first stage estimates, we can consistently estimate the probability of acceptance, $p(\mathrm{A})$, and the expected value of a new filled vacancy, $E(V_F | \mathrm{A})$, where the estimated values are given by $\hat{p}(\mathrm{A})$ and $\widehat{E(V_F | \mathrm{A})}$,

and let $B \equiv p(\text{A})E(V_F|\text{A})$. Then $\hat{B} = \hat{p}(\text{A}) \times \widehat{E(V_F|\text{A})}$ is a consistent estimator of B. We can write

$$\lambda_F = \lambda_U S / \nu.$$

The FEC is rewritten as

$$\psi = \frac{\lambda_U S}{\nu} B,$$

and after substituting consistent estimators, we have

$$\psi = \frac{\hat{\lambda}_U \hat{S}}{\nu} \hat{B}$$

$$\Rightarrow \nu = \frac{\hat{\lambda}_U \hat{S}}{\psi} \hat{B}.$$

If a consistent estimator of ψ is available, $\hat{\psi}$, then a consistent estimator of the steady state vacancy rate is

$$\hat{\nu} = \frac{\hat{\lambda}_U \hat{S}}{\hat{\psi}} \hat{B}.$$

Given this estimate of ν, we have

$$\hat{\lambda}_U = \hat{k}^{\phi},$$

where $\hat{k} = \hat{\nu}/\hat{S}$. Then a consistent estimator of ϕ is given by

$$\hat{\phi} = \frac{\ln \hat{\lambda}_U}{\ln \hat{k}}.$$

∎

In our modeling framework, we assume that costs of employment and vacancies are identical, with costs in both cases consisting of the flow rental rate of capital required to produce output and to evaluate the productivity of applicants arriving at random points in time. Under this assumption, we can recover the Cobb-Douglas matching function parameter ϕ.

Note that the computation of $E(V_F|\text{A})$ requires us to solve for the steady state distribution of general and match-specific levels. In Appendix A we derive this distribution. While we have not yet proven that this distribution is unique, simulations of the steady state distribution and the steady state distribution derived in the appendix are equivalent.

In practice, we found that our estimate of the employment cost parameter ζ was insufficiently large to produce an estimate of the Cobb-Douglas parameter ϕ that lie in the unit interval. In this case, it appears that we must reject the assumption that the cost of posting a vacancy is the same as the cost of capital in an employment match. From previous analyses (e.g., Flinn (2006), Flinn and Mullins (2015)), we know that the implied value of ψ is typically much larger than our estimate of ζ. In this case, the

parameters of the demand side are not identified, and we follow the usual approach for recovering an estimate of ψ. Under the assumption of a given value of the Cobb-Douglas parameter, ϕ, we first find an estimator for unobserved vacancies, ν. We have

$$
\begin{aligned}
\lambda_U &= k^\phi \\
&= (\nu/S)^\phi \\
\Rightarrow \nu &= S(\lambda_U)^{\frac{1}{\phi}}.
\end{aligned}
$$

Using consistent estimates of the relevant parameters, we have that a consistent estimate of ν is

$$
\hat{\nu} = \hat{S}(\hat{\lambda}_U)^{\frac{1}{\phi}}.
$$

Of course, consistency of $\hat{\nu}$ is based on the assumption that we have used the true matching function parameter, ϕ. In practice, we utilize the value of 0.5, which is common in the literature (see Petrongolo and Pissarides (2001)).

Using this estimator of ν, we then find a consistent estimator of λ_F, which is simply

$$
\hat{\lambda}_F = (\hat{v}/\hat{S})^{\phi-1}.
$$

We then find a consistent estimator of ϕ, which is given by

$$
\hat{\psi} = \hat{\lambda}_F \hat{p}(A) \widehat{E(V_F|A)}.
$$

The estimate of ψ is used in our counterfactual experiments involving the minimum wage.

5   Estimation Results

5.1   Parameter Estimates

The estimated parameter values are shown in Table 6 together with bootstrapped standard errors. We start with a discussion of the parameters that control employment transition rates. First, the flow value of unemployment for a worker of ability a is estimated to be $\hat{b}a = 4.93a$, very close to the output of that worker at a firm with the median match quality, $exp(\hat{\mu}_\theta) \cdot a = 4.37a$. For unemployed workers, an offer arrives at a rate of $\hat{\lambda}_u = .145$ or approximately once every seven weeks. Workers with medium levels of general ability accept 32 percent of job offers, implying that the average unemployment spell lasts 21 weeks. Conversely, matches are exogenously dissolved at a rate of $\hat{\eta} = .0033$, or approximately once every six years. Matches may also be dissolved endogenously, if a shock to general ability or match quality makes unemployment preferable to the worker's current match. To assess the relative importance of these two shocks, we observe that the overall unemployment rate in the model is 14.9 percent, close to the data target of 14.0 percent. However, much of this unemployment occurs

along the transition path as the model moves towards steady state and the steady state unemployment rate in the model is just 8.8 percent. Together with the job finding rate and exogenous job separation rate, this equilibrium unemployment rate implies that approximately one quarter of separations are endogenous. For employed workers, new offers arrive at rate $\hat{\lambda}_e = 0.074$, or approximately once every three months, about half as frequently as for unemployed workers.

The parameters $\mu_a(e)$ and σ_a control the distribution of starting values for general ability, where e denotes the education level of the worker. The estimated values imply that workers with some college education begin their labor force careers with 20 percent more human capital than high school graduates, on average, and those with at least a bachelor's degree begin with an additional 25 percent. These parameters are identified largely from wages of new workers entering the labor force. We match the starting wages of workers is the two higher education groups almost exactly. For workers with only a high school degree, starting wages are slightly higher in the model than in the data but subsequently increase at a slow rate. The variance for the initial distribution of ability $\hat{\sigma}_a^2 = 0.036$, which is approximately one quarter the variance in the distribution of match qualities.

The parameters that govern the technologies for the rate of increase in general ability are δ_a^0, δ_a^1 and δ_a^2. As specified in Section 4.1, for an individual with general ability a_i, the hazard rate of improvement to ability level a_{i+1} is given by

$$\varphi_a(i, \tau_a) = \delta_a^0 \times a_i^{\delta_a^1} \times (\tau_a)^{\delta_a^2} \quad i < M$$

with the analogous expression specified for the θ process. In the estimated model, $\hat{\delta}_a^0$ and $\hat{\delta}_\theta^0$ are basically equal to each other, at 0.015. However, the remaining components of the general and match-specific skill processes look considerably different. In Table 6, we see that $\hat{\delta}_a^1$ is -0.050, whereas $\hat{\delta}_\theta^1$ is 0.702. In other words, the parameter estimates show that general training becomes less productive as a increases, whereas match-specific training becomes more productive with increases in θ. This is a reasonable finding since a is likely to be more difficult and costly to change after labor market entry due to the fact that employers are not equipped to offer general learning experiences as efficiently as are schools that specialize in increasing students' cognitive abilities. These parameter estimates also provide a bridge between this model and the Flinn and Mullins (2015) specification, where a is assumed to be fixed over the labor market career. In our model, we allow a to change over the labor market career, but the estimated model shows that it can indeed be thought of as quasi-fixed since it is so difficult to change after labor market entry anyway.

As we wrote earlier, the training observed in the data is likely a very rough proxy for the amount of time spent developing workers' human capital. Despite the predictions of our model that most workers are generally receiving some kind of training, only five percent of workers in the data report training in their current job. The parameters β_0 and β_τ control the relationship between training in the model and the probability that we observe a worker to be receiving training in the data. Although the median

worker in our model spends 25 percent of her time training, we expect that this worker will be observed to be involved in training only 0.5 percent of the time. For a worker engaged in full-time training, we would expect to observe this training in the data only 25 percent of the time.

To help us identify the separate roles of general and firm-specific training in generating the patterns we see in the data, we estimate alternative versions of the model where we allow only one kind of training. First, we consider a model with only firm-specific training. The parameter estimates for this version of the model are shown in the second and third columns of Table 7. When all wage growth within a job spell is attributed to improvements in the match between the worker and firm, the quality of the match must increase relatively quickly and it becomes less likely that an outside offer will dominate her current match. As a result, we estimate a higher arrival rate of offers for on-the-job search (λ_e) in order to match the rate of job-to-job transitions. However, at higher match values, the expected increase in θ from a new outside offer is smaller and the model generates wage changes for job-to-job transitions that are smaller than those observed in the data. Our baseline model that includes general training as well does a better job matching the wage gains over these transitions. We conclude from this exercise that the amount of general training is at least partially identified from the growth in wages across job transitions.

Alternatively, we estimate a version of the model with only general training. This version of the model does a better job matching the average wage growth within and across job spells. However, without the match-specific training, it is unable to match the differences in wage growth and separation rates that we observe between short and long job spells. In all versions of the model, longer job spells are associated with higher match qualities. In the baseline model, the productivity of match-specific training rises with the quality of the match since $\hat{\delta}_2^\theta > 0$. This explains why jobs with better matches experience more wage growth. In addition, the increase in match quality over time due to match-specific training explains the decrease in the job-to-job transition rate with increasing job tenure. The alternative model with only general training is unable to match these features of the data. Therefore, we conclude that the amount of firm-specific training is at least partially identified by the wage growth and job-to-job transition rates for workers with different tenures.

5.2 Policy Rules

We next explore the choices of workers and firms implied by these parameter values. First, we consider the worker's decision to accept a match. A worker of ability a who receives an offer with match value θ will accept the offer if $\theta > \theta^*(a)$, and will otherwise remain unemployed and continue to search. For our estimated parameters, we plot $\theta^*(a)$ in Figure 1. At very low values of a, a high value of θ is required for the match to cover the firm's employment cost and also deliver more value to the worker than the value of unemployment. As a increases, $\theta^*(a)$ decreases as matches of lower value become feasible. The value of $\theta^*(a)$ remains flat and then begins to increase at

higher values of a. To understand the reason for this increase, we need to examine the choice of how much general training the firm provides at each combination of a and θ, which is plotted in Figure 2 (analagous plot for θ in Figure 3). At medium values of a, for values of θ just above $\theta^*(a)$, workers spend 20 percent of their time engaged in general training. This suggests that these marginal matches become feasible only because of the opportunity they provide for the workers to build their general human capital. As a increases, general training becomes less productive (because $\delta_1^a < 0$), these marginal draws no longer deliver positive surplus relative to unemployment, and workers raise their reservation value of θ.

Next, we consider the choices of how much general and match-specific training firms and workers choose to provide for different levels of (a, θ). First, to understand how much workers value each kind of training relative to simply receiving wages, Figure **??** shows combinations of training and wages that solve the bargaining problem between the worker and the firm. Near the actual solution, the wages decrease quickly as the firm chooses to provides more general training, suggesting that workers regard general training as a good substitute for wages. In contrast, increasing match-specific training results in a much smaller decrease in wages, implying that the worker's value from additional match-speciific training is small, and that most of the value from match-specific training goes to the firm. However, the worker does receive some benefit from the match-specific training, which is reflected in his willingness to trade off some wages for more match-specific training.

To show the outcome of the bargaining over training, in Figures 2 and 3, we plot the amount of general and match-specific training that workers receive at different combinations of a and θ. For ease of illustration, both states are shown on a log scale and the lines on the graph show contours along which the amount of training remains constant. The bottom of the figures, corresponding to low values of θ, are combinations for which workers will not accept the job offer.

Looking first at the policy for firm-specific training plotted in Figure 3, we see that the amount of firm-specific training is essentially a function of the current value of θ with very little dependence on the worker's level of general ability. At values of θ just above the minimum $\theta^*(a)$ threshold, the amount of training is small. Firm-specific training increases for higher values of θ, reaching a maximum of 25 percent of the worker's time at roughly the 85th percentile of the distribution of acceptable θ draws. Two different mechanisms contribute to this pattern. First, in the estimated model, $\hat{\delta}_\theta^1 > 0$ so firm-specific training is more productive at higher values of θ. Second, at higher values of θ, the expected duration of the current match increases as it becomes less likely that the worker will leave to take an outside offer. Because firm-specific training increases future output only for as long as the worker remains with her current employer, this increase in expected duration raises the value of match-specific training. Offsetting these effects is the incentive for the firm to provide match-specific training in order to raise the value of θ and thereby increase the length of the current match. This incentive is stronger at lower values of θ because the density of potential job offers is higher so that increase in θ yields a greater reduction in the fraction of outside

offers that would cause the worker to leave. This mechanism should also give firms an incentive to provide more match-specific training at higher values of a where the benefit of the match is higher. However, in practice, firms seem to have little ability to increase the match duration by providing firm-specific training. Also, as described above, workers with higher general ability do not receive noticeably more match-specific training. Finally, we find that in a counter-factual experiment with no on-the-job search, the amount of match-specific training actually increases. All of this evidence suggests that the firm's ability to retain workers by providing match-specific training is quite limited. Instead, the decision to provide firm-specific training seems to depend on the productivity of that training

Next, we look at the amount of general training provided to the workers, which we plot in Figure 2. At values of θ just above the $\theta^*(a)$ cutoff, workers spend about 20 percent of their time engaged in general training. The amount of training decreases at higher values of either a or θ. Also, unlike the match-specific training discussed above, a worker retains her accumulated general human capital even after the current match is dissolved, so general training does not become more valuable as the expected duration of the match increases. Rather, as in a standard Ben-Porath model, the benefits of general training flow largely to the worker and the amount of general training is determined by worker's trade off in allocating time between production and the accumulation of general human capital. In the context of our model, this implies that negotiations over the amount of general training should look similar to the negotiations over wages. In states where the bargaining process yields higher compensation for the worker, she will choose to receive some of this compensation as higher wages and some as general training. Indeed, in Figure 4, we plot the fraction of worker's output that is paid in wages and we observe that it follows the same pattern as the general training shown in Figure 2.

5.3 Within Sample Fit

In this section, we go through the simulated moments from the estimated model and look at how they compare to the data. In parts of the discussion, we also include a comparison to the simulated moments from different estimations of the model. In particular, in addition to the baseline, we estimate two other versions. One is the same model but without the possibility of investment in general human capital and the other is without the possibility of investment in match-specific human capital. We include the predictions from these two other estimated models in order to demonstrate how the training component built into the model is necessary to capture certain aspects of some key employment and wage patterns in the data. These comparisons also help to see which aspects of data on labor market transitions and wage growth patterns help identify the degree of 'specificity' in training and human capital investment.

Figures 5-7 show that the model simulations for wages provide a good fit to observed wages in the data. More specifically, the model accurately captures the variation by schooling level and job tenure. Figure 5 contains the histogram plots for log wages

of workers with job tenure of 0-2 years. There are three panels in this figure and each panel displays a histogram plot for workers of a different education group. For example, in Figure 5, the first panel is for high school graduates, the panel to the right of it is for individuals with some college and the histogram in the lower panel is for the log wage distribution of individuals with a BA or higher degree. All three panels of Figure 5 show that for workers of this job tenure level (0-2 years), the model predictions for log wages of each of the three education groups are consistent with the data. While all are consistent with the data, the model simulations do a slightly better job for the log wages of workers with some college degree and workers with a college in comparison to high school graduates. For example, the average log wage level for high school graduates is 2.36 in the data and 2.47 in the estimated model so that the model understates the log wages for this education group. On the other hand, for workers with some college degree, the average log wage is 2.72 in both the estimated model as well as in the data. Also, for workers with a college degree, the average log wage is 3.04 in the data and 3.03 in the model.

Figure 6 shows the histogram plots for log wages of workers with job tenures of 3-5 years. Again, the three panels correspond to the three education groups. The average log wage levels can be seen in Table 12 for the same groups of workers. The first aspect of note here is that the model predictions do a very good job in matching the distribution of wages for all education groups. This can be seen in Figure 6. For example, the average log wage is 2.67 in data and 2.71 in model for high school graduates, 2.98 in the data and 2.95 in the model for workers with some college degrees, and 3.25 in both the data and model for workers with college degrees. Lastly, Figure 7 shows the histogram plots for log wages of workers with job tenures of 6-8 years. We see that the wage fit is very good for this job tenure category as well.

Table 8 shows the incidence of training for the data and estimated model. We see that the model does a good job in this dimension as well. The first column shows the overall proportion of individuals who are observed to acquire some form of training (general or match-specific) during the total time they are observed in the sample. We see that in the data this proportion is 15 percent and in the model simulations it is 16 percent. The next three columns display the proportion of training for workers of different education groups. In the data, 18 percent of workers with a high school degree acquire training at least once during the sample period and this number is 13 percent for workers with some college education as well as workers with a college or higher degree. In the estimated model, we see that these numbers are 20, 15 and 13 percent, respectively. It can be seen that the model does a good job in matching the decreasing proportion of training with education.

Which features of the estimated model allow us to capture the fact that it is the individuals with lower education levels who are more likely to get training throughout their labor market career? The first thing to note here is that in the model, individuals of different educational attainment levels are allowed to differ only in terms of the mean of their initial a distributions. Parameter estimates show that there are indeed considerable differences between the means: μ_a is 1.07, 1.28 and 1.52, for workers with

31

high school degree, workers with some college education and workers with a college or higher degree, respectively. Consequently, on average the initial values of a drawn at the time of labor market entry are higher for workers with more education.

While it is easy to see which aspect of the estimated model allows us to capture differences in training by education, the explanation behind the exact direction of the relationship between training and education is more involved. To find the factors that give rise to the negative relationship between training and education, we look at some of the policy rule Figures 2 and 3, which were previously discussed in Section 5.2. These figures display decision rules for training against different pairs of (θ, a) values. It can be seen from Figure 3 that the amount of firm-specific training is mostly an increasing function of the current value of θ (above a certain θ threshold) with little dependence on a. One could expect some dependence of firm-specific training on a due to the complementarity between a and θ in the production technology, since higher values of a imply larger marginal returns to an increase in θ. However, we see from Figure 3 that this is not the case, which leads us to conclude that the positive effect of a on match-specific investment that arises due to complementarity in the production technology is clearly being offset by a countervailing force. High a workers are more attractive to other firms relative to lower a workers and therefore might have higher exit rates and lower expected job duration.[10]. This decreases the value of any training on high a individuals. It seems that these two countervailing factors offset each other resulting in a match-specific training policy rule that is seemingly independent of a.

In addition, it can be seen from Figure 2 that the amount of general training provided to the workers decrease with θ as well as a (more so for θ). High a does not lead to more investment in a because general training becomes less productive at higher values of a. Through Figures 2 and 3 together with an analysis of some of the parameter estimates, we have looked at the reasons the estimated model predicts decreasing training (specific or general) with values of a. We also mentioned how the values of the mean of the initial a distribution decrease by education. The combination of these factors together provide the explanation for how the estimated model gives rise to decreasing training with education.

The plots for the policy rules together with the above explanation of some of the mechanisms at work in the estimated model show the complexity of how various aspects of match-specific heterogeneity and training opportunities manifest themselves as various wage and employment patterns in equilibrium. In particular, we see that the observed policy rules and consequent wage growth patterns through transitions are a result of countervailing incentives at play vis-a-vis human capital investment on-the-job. For example, for investment in general ability, important channels at play are the complementarity between a and θ as well as the fact that the productivity of general

---

[10]This is true regardless of whether there is renegotiation or not. When the outside option is always unemployed search, workers with high a's are still more likely to exit into unemployment or to another job. They have a higher value of unemployed search overall and the value of their current job is more likely to be exceeded by an alternative option

training decreases with a. On the other hand, parameter estimates show that productivity of specific training increases with values of θ. Then, for high values of θ, the complementarity in the production technology gives an incentive to engage in general training to balance the value of θ and a, while the estimated values of $\hat{\delta}_1^\theta \geq 0$ and $\hat{\delta}_1^a \leq 0$ mean that there is a large opportunity cost to spending the total available time in general training (since $\hat{\delta}_1^\theta \geq 0$, it is more productive to spend that time in specific investment). Similar to the description above of the patterns by education, this leads to two offsetting effects as well: The increased incentive to invest in more a at high values of θ due to production complementarities is offset by the high opportunity cost of general training. This highlights the importance of taking into account different parts of the estimated model together in understanding the policy rules generated in equilibrium.

Table 9 compares some moments that pertain to employment transitions between the data and the estimated model. It consists of three panels and each are defined in the same way as in the corresponding table in the data section, where the first panel displays the proportion of job-to-job transitions between interview dates, and the second and third panels display log wage growth for these transitions. More specifically, these are transitions that involve individuals who were employed at the time of the interview in both survey rounds $t-1$ and t. We see that overall, the estimated model provides a good fit to the data. For example, the proportion of stayers for high school graduates in the data is 81 percent and it is 78 percent in the model. While the model simulations get the decreasing proportion of stayers with education, it does not match the proportion of stayers for the higher education groups as well as it does for high school graduates. This can be seen in the second and third columns of the first row in Table 9. In the data, the proportion of stayers is 85 and 88 percent for these two education categories, whereas they are 75 and 73 percent in the model.

The last two moments in Panel A of Table 9 distinguish between two different kinds of job switches: (1) Employment spells interrupted by an unemployment spell, and (2) employment spells uninterrupted by an unemployment spell.[11] In the data, there are stark differences between these two transitions both in terms of their overall rates as well as in terms of the wage growth that occurs during these transitions. Moreover, these differences contain a wealth of information about the relative contributions of general and specific human capital to productivity, output and wages. For example, for high school graduates, the proportion of employment spells with and without an intervening unemployment spell is 8 and 14 percent, respectively, and the corresponding data moments are close, at 5 and 15 percent.

For the higher education groups, transition rates and the corresponding wage transitions look similar. Therefore, in what follows, we mainly base our discussion on the moments for high school graduates in the first row of Table 9. Panel B shows wage growth during these employment spells. In the data, for job switchers with an intervening unemployment spell, the average log wage difference is 0.06, and for job switchers

---

[11]The precise definition of these transitions can be found in the data section.

with no intervening unemployment spell, it is 0.12. In the model, when individuals receive an alternative job offer while working, they only leave their current employer if the match productivity θ at the new employer is at least as great as their current match productivity. On the other hand, when an unemployment spell is experienced between the job at $t-1$ and the job at t, the total density of acceptable jobs at t is larger than the case when there is no unemployment spell in between. This is how the model is able to match the fact that the average wage growth for job switchers with an intervening unemployment spell is lower than those without an intervening unemployment spell. While the estimated model captures the direction of the implications of an intervening unemployment spell (i.e. that the wage change is larger for the case with no intervening unemployment spell), it does not do so well in terms of the exact levels. In the estimated model, the corresponding simulated moments are -0.17 and 0.21. Below, we discuss aspects of the estimated model that may be behind the failure of the model to capture some of these wage growth moments as well as aspects of the estimated model that enable the model simulations to do well for some of the wage growth moments.

In the model, values of θ and any investments made in match-specific productivity do not carry over into future employers, whereas general human capital does. In other words, any advantage that a worker has at a particular firm due to high values of θ (that result from high initial θ and/or investment in θ) ends once an employment spell ends and the worker becomes unemployed. The wage difference during a job transitions with an intervening unemployment spell therefore includes the value of the loss in the match-specific productivity. It can be seen in Panel C that the proportion of job switches that entailed a negative wage change between $t-1$ and t is 54 percent with an intervening unemployment and it is 18 percent without. This explains the negative wage change observed in Panel B.

As explained above, for employment spells with an intervening unemployment spell, the worker loses the value from his previous match-specific productivity and for the next job in the sequence, there is a large density of match values that are acceptable to her. On the other hand, for a worker who directly switches to a new job, the only acceptable jobs are those that have θ values higher than her current value. Therefore, it makes sense that the proportion of negative wage transitions is higher for the latter case (the case without an intervening unemployment spell). The model in fact does capture this pattern so that the model matches the fact that the proportion of negative wage transitions is lower for the cases with no intervening unemployment spell. Due to the fact that it overstates this proportion for the case with an intervening unemployment spell, we see in Panel B that the wage change in the model predictions is not a good fit with the data.

We see more evidence of general and specific types of human capital investment contributing differentially to employment and wage transitions in Table 10, which reports the same transition moments as in Table 9 for alternative estimations of the model. We estimate the model under two alternate scenarios. The first is one where we only allow for investment in match-specific human capital and the second is when

we only alow for investment in general human capital. Panel C shows that the proportion of negative wage transitions decrease from 54 in the baseline estimation to 49 percent in the estimation with no θ investment. With no possibility of investment in match-specific productivity, any loss incurred during a job transition with an unemployment spell occurs due to the initial value of θ that the firm-worker pair draw when they meet. Hence, the additional loss that occurs due to past investment in θ does not happen in this scenario and this is the reason why the proportion of negative wage transitions is lower. However, the fact that it is not considerably lower shows that a large part of the variation in θ in the original estimated model (i.e. baseline) comes from the variation in the initial distribution of match values. On the other hand, in the estimation with no a investment, the proportion of negative wage transitions actually increase from 54 to 60 percent. With the only possibility of investment being θ, there is no increase in the general skill of the worker during a job spell. Hence, with a decrease in the component of wages that the worker can carry over (decrease in a through lack of investment activity in a), the proportion of negative wage transition increases.

The average log wage differences for job switchers displayed in Panel C of Table 10 support these explanations as well. This table shows that the average log wage difference for the estimations with no θ is lower in absolute value compared to the baseline (-0.08 in comparison to -0.17). This confirms that the main reason the negative wage difference for these transitions is overstated is the investments made in θ.

The comparison between the estimations with no investment in a and no investment in θ help us see the importance of transition moments in the identification of θ and a distributions and the degree to which skills and investment activities on-the-job are firm-specific. Similar to the annual transition dates that only look at changes that occur between interview dates, Table 11 shows the comparison of weekly transition rates between baseline and the two estimations with only θ and only a investments. We see that job finding probability is 4.04 percent in the baseline, 3.95 percent for the estimation with no a investment and 5.09 percent for the estimation with no θ investment. Hence, the first aspect to note is the similarity between the job finding probability estimates between the baseline estimation and the estimation with no investment in a. On the other hand, the job finding probability estimate for the estimation with no investment in θ is much higher relative to the baseline estimation. Without any possibility of investing in θ, workers will, on average, have higher endowments of a and therefore be more attractive to firms. Therefore, it will take a shorter time to find a job for any worker.

Lastly, Figures 8-10 show average log wages for the three different estimations of the model (baseline, no a and no θ). Each figure corresponds to a different education level. We see the same pattern for all three education groups: The initial accepted wages (wages that correspond to Year 0-2 of labor market tenure) are lower in the estimated model with no investment in θ and higher in the estimated model with no investment in a. The baseline is in between. On the other hand, this order gets reversed over the course of the worker's labor market tenure. In other words, by Year 6-8, the average log wages that correspond to the estimated model with no investment in θ overtakes

35

the other two and the average log wages for the case with no investment in a ends up as the smallest average among all three estimated models. In the estimation with no θ investment, there is only the possibility of investing in a worker's general human capital, a. Hence, throughout the worker's labor market career, she carries over the benefits of any training incurred in her jobs and the resulting wage growth throughout the labor market career of a workers is consequently larger (due to higher accumulation of a). On the other hand, for the case with no a investment, the opposite is true. The only training that a worker is able to engage in is training in θ, which is something that she cannot carry over to future jobs. Hence, the amount of benefits that she can accrue and transfer to future periods is smaller in the model with no a investment and consequently overall wage growth throughout the labor market career is much lower. This is why the average log wages for the case with no a ends up being the lowest at the end of the Year 6-8 tenure profile.

The comparison between these wage plots shows us that the difference between wage growth rates during the course of a worker's career within a firm and over the course of a worker's overall labor market career is an important indication of the occurrence of either form of training. We exploit these differences by including in the estimation moments such as the average wages by number of years on a job as well as number of years in the labor market and wage growth at different levels of firm-specific or overall labor market tenure.

5.4   Training and Wage Growth in the Model Simulations

Given these decision rules, how much training do workers actually receive? In Figure 11, we plot the fraction of time that workers spend training as they move through the first years of their careers. When workers first enter the labor force, initial match qualities are relatively low and therefore most of the training takes the form of general training. In the model simulations, workers in their first year in the labor force spend about 15 percent of their time in general training and 10 percent in firm-specific training. Over time, match quality increases as workers sort into jobs of higher match quality. Because the expected duration of these jobs is higher and also because firm-specific training becomes more productive at higher values of θ, more of the training becomes firm-specific and the total amount of training increases, peaking one to two years after labor market entry. As match quality increases further over time, the amount of time spent on both general and firm-specific training begins to decline as workers spend less time training and more time engaged in production.

We next address the question of what drives wage growth in our model. The model contains four possible sources of wage growth. First, workers can increase their productivity by building general human capital through on-the job training. Second, workers can increase productivity by improving the quality of the match with their employers. Third, as workers spend less time training and more time engaged in

production, some of this increased output will flow to them in the form of higher wages. (Alternatively, a shift towards less training and higher wages could be interpreted as a shift in the workers' compensation towards higher current wages and away from expected higher future wages.) Finally, the bargaining between workers and firms can result in different shares of worker output being paid as wages depend on the value of the workers' outside options. The current section aims to quantify the importance of each of these channels.

In the absence of employment costs, output y would equal $a \cdot \theta \cdot (1 - \tau_a - \tau_\theta)$ and we could formally write the wage w as the product of the four pieces described above:

$$w = a \cdot \theta \cdot (1 - \tau_a - \tau_\theta) \cdot (w/y),$$

or in logs,

$$\log(w) = \log(a) + log(\theta) + \log(1 - \tau_a - \tau_\theta) + \log(w/y)$$

In Figure 12, we plot the evolution of each of these four components, together with the total wage, as workers move through the early years of their careers. The figure shows that in the worker's first few years in the labor force, most of the growth in wages can be attributed to the development of match-specific human capital. General ability grows more slowly, contributing less to wage growth at the beginning of a worker's career but accounting for a larger fraction of total wage growth as the rise in match-specific human capital slows over time. As described above, the time spent training increases at the very start of a worker's career, but thereafter, the reduction in training time begins to contribute noticeably to the worker's overall output and therefore to her wage. The last component, the fraction of output represented by the worker's wage, is quite flat and has almost no effect on the evolution of wages over time.

5.5   Interpreting Mincer Regressions

Our structural model allows us to interpret the coefficients of a standard wage regression in terms of the different source of wage growth discussed above. As an example, we consider a simple Mincer wage regression of the form

$$log(wage_{it}) = \sum_j \beta_j^w X_{it}^j + \varepsilon_{it}^w$$

where $log(wage_{it})$ is the log of the wage for person i at time t. Specifically, we estimate

$$log(wage_{it}) = \beta_0^w + \sum_e \beta_e^w ed_{ie} + \beta_y^w years_{it} + \beta_t^w tenure_{it} + \varepsilon_{it}^w. \tag{3}$$

where ed_{ie} is a dummy variable indicating that person i has education level e (for each level of education except HS graduate), $years$ denotes the number of years in the labor force and $tenure$ the length of time with current employer. We first estimate this

model on the actual NLSY data and on the simulated data from the model. Results are shown in the first two lines of Table 13. As expected, more education, more years in the labor force and greater job tenure are all associated with higher wages. Comparing the regressions on the real and simulated data, we find that additional education is associated with less of an increase in wages in the simulated data cokmpared less of in increase in wages in than in the data. Also, relative to the data, more of the wage growth in the model is attributed to tenure with particular employers and less to overall labor-market experience.

Focusing on the regression using the simulated data, we next aim to understand how the increases in wages associated with education, labor market experience and job tenure reflect the different determinants of wages present in the model. Similar to the decomposition described in the previous section, we can decompose log wages in the model as a sum of the logs of i) general ability, ii) match quality, iii) time spent not training and iv) wage as a fraction of output. Additionally, it is useful in this context to decompose θ into two separate components, as $\theta = \theta_0 \cdot \theta_\tau$ where θ_0 the match quality at the start of the match, and θ_τ is the additional match quality accumulated trough match-specific training. Finally, because we are interested in the level of wages rather than just the growth rate, we express the worker's general ability as a combination of his initial endowment (a_0) and the additional human capital he accumulates through training (a_τ). This defines six components of wages, which we denote $Y_{it}^k, k = 1, ..., 6$, allowing us to write

$$
\begin{aligned}
\log(w_{it}) \quad &= \sum_{k=1}^{6} \log(Y_{it}^k) \\
&= \log(a_{0,it}) + \log(a_{\tau,it}) + \log(\theta_{0,it}) + \log(\theta_{\tau,it}) + \log(1 - \tau_{a,it} - \tau_{\theta,it}) + \log(\tfrac{w_{it}}{y_{it}})
\end{aligned} \tag{4}
$$

In order to measure how education, labor-market experience and job tenure affect each of these components, we repeat the regression from Equation 3 on each of these six pieces separately, i.e. we estimate

$$
\log(Y_{it}^k) = \sum_j \beta_j^k X_{it}^j + \varepsilon_{it}^k, \quad k = 1, ..., 6.
$$

It is straight forward to show that for each covariate (indexed by j), the sum of the regression coefficients from these six regressions must equal the coefficient for regression using the total log wage, i.e.

$$
\beta_j^w = \sum_{k=1}^{6} \beta_j^k.
$$

This allows us to interpret each of the coefficients β_j^w from Equation 3 as reflecting different combinations of the components of wages defined in Equation 4. The results of this exercise, which we present in Table 13, are all quite reasonable. The increase in wages with more education is largely picking up differences in initial ability and

to a lesser extent the fact that more educated workers are paid a larger share of the output, possibly because the constant employment costs consume a smaller fraction of their output. The positive coefficient on labor market experience is mostly capturing general ability learned through training and to some extent workers' ability to find better matches over time and their receiving a larger fraction of output. Finally, the increase in wages for workers with more tenure mostly reflects improved match quality from training. To a lesser extent, wages also appear to increase with tenure because workers engage in less training as they remain with a firm longer, and there is an additional selection effect whereby longer-tenured workers received better initial matches with their firms. Overall, we feel our model provides a novel and interesting way to interpret the results of standard wage regressions.

6   Policy Analysis: The Minimum Wage and Investment

6.1   Minimum Wage in Partial Equilibrium

To study the effect of a minimum wage in the context of our model, we resolve the model imposing a minimum wage of $15 per hour (in 2014 dollars, corresponding to $10.17 in the our model and data, where we express wage rates in 1994 dollars).

As is common in the literature (e.g. Flinn and Mullins (2015)), we find that imposing a minimum wage increases the unemployment rate by rendering low quality matches unprofitable for firms. Figure 13 shows the minimal acceptable match quality draw in the baseline model and under the minimum wage. Because the minimum wage is more likely to be binding for low ability workers, the policy raises the value of the In our simulations, imposing the minimum wage raises the equilibrium unemployment rate from 8.8 to 9.7 percent with most of the increase concentrated among less educated workers.

In addition to the effect on employment, our model shows how the minimum wage can also affect the amount of training provided to employed workers. Because employers must pay workers a higher wage, they decrease the amount of compensation that is provided in the form of general training. Figure 14 compares the amount of training that workers receive in the baseline model and with the minimum wage. With a minimum wage in place, employed workers spend 5-10 percent less time on general training than they do in the baseline model. The long-run effect of this decrease in training is to reduce the average amount of general ability in the population by about half a percent. At the bottom of the distribution, where the impact of a minimum wage is largest, general ability is about three percent lower than in the baseline.[12]

---

[12]As shown in figure 14, the amount of match-spefic training is slightly higher in the presence of a miniumum wage. The reason for this effect is that the minimum wage raises the average match quality of accepted offers and match-specific training is more productive for workers with higher match qualities.

Conditional on employment, workers receive higher wages with a minimum wage in place than they do in the baseline model. Average wages are six percent higher for workers entering the labor force and remain 1.5 percent higher after several years. Part of this increase is due to selection on both general ability and match quality as the presence of a minimum wage raises the distribution of acceptable match values and also disproportionately keeps low-ability workers unemployed. In addition to the selection effects, the lower amount of training discussed above means that workers are spending more of their time engaged in production and some of this additional output naturally flows to the worker in the form of a higher wage. Finally, when the minimum wage is binding, employers must pay workers a higher fraction of their output then they otherwise would in order to raise their wages to the required level. Quantitatively, this total increase in wages can be attributed to the increase in match quality, the decrease in training time and increase in the workers' wages as a fraction of total output, each of which contribute an equal amount.

Overall, the welfare effects from the minimum wage are negative. However, the losses are small as the loss to workers from a higher unemployment rate and lower amounts of general training is largely offset by the higher wages they receive. At the lowest value of general ability, the welfare losses to an unemployed worker from facing a minimum wage is one percent. For high school graduates on average, the welfare loss is just 0.02 percent and it is even smaller for those with more education. Finally, because the minimum wage reduces the value of being unemployed, it has additional effect of decreasing the value of the worker's outside option when she bargains with a potential employer. As a result, low-ability workers receive slightly lower wages even when the minimum wage is not binding, though the difference in wages is small.

6.2   Minimum Wage in General Equilibrium

The results presented above consider a partial equilibrium response to a minimum wage, ignoring the effect of the policy change on vacancy creation and thereby on the job offer rate for workers. Flinn (2006) and Flinn and Mullins (2015), however, find that minimum wage impacts on labor market outcomes and welfare tend to be greatly changed after relaxing the assumption of constant transition rates. To investigate minimum wage impacts more robustly, we study its effects within a simple Mortensen-Pissarides framework as described in Section 2.3. In order to move to a general equilibrium framework, we need to define the steady-state distribution of workers. We do this by assuming that workers are subject to a retirement shock that causes them to leave the labor force at a rate calibrated so that the average career last 45 years.[13] This addition makes the workers expected lifetime finite and allows us to simulate the model until a point where the remaining fraction of non-retired workers approaches zero. Under assumptions of stationarity and zero-population growth, these simulations represent an approximation to the steady-state distribution of workers in an overlapping generations

---

[13]In the model solution, this retirement shock increases the worker's discount factor by $\upsilon = 1/(45 \text{ years})$.

framework.

Having identified the baseline steady-state distribution of workers, the next challenge is to is to identify the firm's cost for posting a vacancy (ψ) and the value of the Cobb-Douglass parameter in the matching function for workers and firms (θ). Following the discussion in Sections 2.3 and 4.2.2, we attempt to estimate these parameters in several different ways. First, as described in Section 4.2.2, we assume that the vacancy posting cost is equal to the estimated employment cost $\hat{\zeta}$ and back out the value of Cobb-Douglas parameter. Unfortunately, this approach fails under our estimated parameters, yielding an negative estimate of θ. Alternatively, as described in Section 4.2.2, we fix the Cobb-Douglas parameter at $\theta = 0.5$ and back out the vacancy cost. This procedure produces an estimate of the vacancy cost $\psi = 594$.

In general equilibrium, the minimum wage constraint reduces the firm's incentive to post a vacancy and therefore decreases the number of vacancies and the rate at which workers receive job offers. With a minimum wage of \$15 (in 2014 dollars), the minimum wage binds on a very small fraction of matches in the steady state distribution so the impact on transition rates is small: the change in job-finding rates is smaller than our uncertainty in our estimates of these rates. As a result, the general equilibrium response of firms to the imposed minimum wage does not noticeably change the partial equilibrium results presented above. Alternatively, we consider higher minimum wages equal to \$20 and \$25.[14] With the minimum wage set at \$20, the steady state unemployment rate increases from 11.3% to 13.0%. In addition, because low-ability workers are employed less and receive less training, the very bottom of the general ability distribution shifts downward with the first percentile of the distribution moving nine percent lower. As we increase the minimum wage to \$25, we get even larger effects with the steady state unemployment rate climbing to 17.5% and the first percentile of the ability distribution falling to the lowest grid point, a total decline of 27% from the baseline. In addition, the higher minimum wage now has effects further up the ability distribution so that the tenth percentile of general human capital declines by nine percent relative to the baseline.

7   Conclusion

We have developed an estimable model of investment in both (completely) general and (completely) match-specific human capital while individuals are active members of the labor force, which we assume follows the completion of formal full-time schooling. While other researchers have examined investment decisions in a search, matching, and borrowing framework, ours is perhaps the first to attempt to estimate such a model in a reasonably general framework. Perhaps the greatest challenge we face in estimation is to attain credible identification of such a model when human capital stocks and investments essentially are unobservable. In this we are aided by having access

---

[14]These values are again expressed in 2014 dollars. Expressed in the our units of 1994 dollars, the wage rates we actually impose in the model are \$13.56 and \$16.95 repsectively.

to (self-reported) data on whether a worker engaged in formal training during a job spell. We heavily exploit this information in our moment-based estimation procedure. Furthermore, our assumptions regarding the specificity of human capital imply that changes in the stock of general human capital have no impact on the future mobility decisions of an individual during the employment spell. This stands is stark contrast to changes in the stock of match-specific human capital, which strictly reduce the likelihood of accepting a job with another firm during the employment spell. Thus job-to-job mobility along with wage changes during the current job spell can be utilized to infer whether the wage change was the result of general or match-specific investment.

Our estimates of the human capital production technology exhibit decreasing returns to investment in both types of human capital, of approximately the same degree. Our production technology also includes a TFP term that captures how the current level of both types of human capital affect the returns to investment. Here we find that the payoffs from time investment in match-specific human capital are increasing in its current level, while there is no impact of the current level of general human capital on the return to investment in it. These results imply fairly complex dynamic patterns in investment and wage growth. They also serve to produce a job acceptance probability from the unemployment state that is non-monotone in the individual's level of general human capital.

We use our estimates to examine the impact of minimum wage policy on investment in human capital and equilibrium outcomes. Our experiments are conducted in both partial and general equilibrium frameworks, where the general equilibrium specification relies on the topical matching function approach. Unlike previous estimates of these minimum wage effects, as in Flinn (2006) and Flinn and Mullins (2015), the deleterious effects on high minimum wages on job finding rates can partially be alleviated by increases in the productivity of workers, which is obtained by investment in general and match-specific human capital on the job. In the general equilibrium setting, we find that a minimum wage of $15 does little to affect the labor market equilibrium, since the steady state distributions of worker and match quality used to determine the vacancy creation decisions of firms already produce a vast majority of matches that have productive quality levels greater than this amount. A minimum wage of $20 dollars an hour also has fairly minor impacts on unemployment and other features of the labor market equilibrium. On the other hand, the impact of a minimum wage of $25 dollars an hour does have notable impacts on unemployment and the steady state distribution of human capital and wages. To some extent, these results parallel those found in Flinn and Mullins (2015), which examined minimum wage impacts within a model of formal schooling decisions that did not allow post-schooling investment in human capital of either type.

In our future research, we intend to endogenize the formal schooling decision, as in Flinn and Mullins. This will provide us with a relatively complete model of human capital investment over the entire life-cycle, and will allow us to examine the relationship between the human capital acquired during the formal schooling phase and that acquired while in the labor market. Our belief is that formal school training is to

some extent similar to what we are calling general human capital in this paper, but that the two are not perfect substitutes in production. We believe that much of what one acquires during formal schooling is a technology for learning, which impacts the production of both general and match-specific human capital during the individual's labor market career. In this view, skills acquired or not acquired during early periods of development and formal schooling will have long-lasting effects on labor market outcomes and lifetime welfare.

Appendix

A   Deriving the Steady State Distribution in the Labor Market

In order to simplify notation, we first expand the space of match values to include 0, which signifies that the agent is unmatched, that is, unemployed. All employed individuals at an arbitrary point in time are characterized by the labor market state (j, k), which signifies a_j and θ_k. Let the probability of (j, k) be denoted by $\pi(j, k)$. The steady state marginal distributions of a and θ are given by $\pi_a(j)$ and $\pi_\theta(k)$, respectively. There are J possible values of a, and K possible values of θ (for employed agents), with

$$
\begin{aligned}
0 &< a_1 < ... < a_J \\
0 &< \theta_1 < ... < \theta_K.
\end{aligned}
$$

From our estimates, we have determined the minimal value of θ that is acceptable when an agent characterized by a_j is in the unemployment state, which we denote by $\underline{k}(j)$ (that is, $\theta_{\underline{k}(j)}$ is the minimal acceptable match value to an individual with ability level a_j). We define the indicator variable

$$
d(j, k) = \left\{ \begin{array}{ll} 1 & \text{if } k \geq \underline{k}(j) \\ 0 & \text{if } k < \underline{k}(j) \end{array} \right. , \ j = 1, ..., J; k = 1, ..., K.
$$

Tautologically, $\pi(j, k) = 0$ for all (j, k) such that $d(j, k) = 0$, $j = 1, ..., J$, $k = 1, ..., K$.

Unemployed agents of type a_j occupy the state $(j, 0)$, with the probability of a type j individual being unemployed given by $\pi(j, 0)$. Then we have

$$
\pi_a(j) = \pi(j, 0) + \sum_{k>0} \pi(j, k)
$$

is the marginal distribution of a in the population. The conditional probability that a type j individual is unemployed is $U(i) = \pi(j, 0)/\pi_a(j)$.

We begin by considering movements in the probability of being unemployed for an agent of ability type j. We have

$$
\begin{aligned}
\dot{\pi}(j, 0) \ = \ & \eta \sum_{k>0} \pi(j, k) \\
& + \tilde{\delta}_a(j+1) \Big[\sum_{k>0} \pi(j+1, k)(1 - d(j, k)) \Big] \\
& + \sum_{k>0} \tilde{\delta}_\theta(k) \pi(j, k)(1 - d(j, k-1)) \\
& - \lambda_U \pi(j, 0) \sum_{k>0} p_\theta(k) d(j, k).
\end{aligned}
$$

The right hand side terms correspond to the following events. On the first line is the rate at which jobs of any acceptable type (all k for which $d(j, k) = 1$) are destroyed times the probability that type j individuals are employed. The second line represents inflows in to the unemployment state that result from depreciation in general skills that are associated with "endogenous" quits into unemployment. In this case, an individual employed with skills $(j + 1, k)$ will quit into unemployment if they would not accept employment at (j, k). The third line represents inflows into unemployment from individuals with skills (j, k) when their match skill level depreciates to $k - 1$ and $(j, k-1)$ is not an acceptable match. The last line is the outflow from the unemployment state, which is the the product of the rate of receiving job offers in the unemployed state, the probability of being in the state $(j, 0)$, and the probability of receiving an acceptable job offer. Then, in the steady state,

$$
\begin{aligned}
\pi^*(j, 0) \quad = \quad & [\lambda_U \sum_{k>0} p_\theta(k)]^{-1} \\
& \times \{ \eta \sum_{k>0} \pi^*(j, k) \\
& + \tilde{\delta}_a(j+1)[\sum_{k>0} \pi^*(j+1, k)(1 - d(j, k))] \\
& + \sum_{k>0} \tilde{\delta}_\theta(k)\pi(j, k)(1 - d(j, k-1))\},
\end{aligned}
\tag{5}
$$

for $j = 1, ..., J$.

Now consider the determination of the probabities associated with employment, those for which $k > 0$. The generic expression for the time derivative of $\pi(j, k)$ is

$$
\begin{aligned}
\dot{\pi}(j, k) \quad = \quad & \pi(j-1, k)\varphi_a(j-1, k) + \pi(j, k-1)\varphi_\theta(j, k-1) \\
& + \tilde{\delta}_a(j+1)\pi(j+1, k) + \tilde{\delta}_\theta(k+1)\pi(j, k+1) \\
& + \lambda_U \pi(j, 0)p_\theta(j) + \lambda_E \sum_{l<k} \pi(j, l) \\
& - [\eta + \tilde{\delta}_a(j) + \tilde{\delta}_\theta(k) + \lambda_E \sum_{l>k} \pi(j, l)]\pi(j, k).
\end{aligned}
$$

In terms of the expressions on the right hand side of this equation, the first line represents improvements resulting in attaining state (j, k) from the states $(j - 1, k)$ and $(j, k - 1)$. The second line represents inflows from human capital depreciations from the states $(j + 1, k)$ and $(j, k + 1)$. The third lines represent inflows from the unemployment state and from contacts with other employed individuals of ability type j who are currently working at jobs in which there match value is less than k. The final line represents all of the ways in which individuals from (j, k) exit the state. These are the exogenous dismissals, decreases in j or k, or finding another job for which match

productivity is greater than k. Then in the steady state we have

$$
\begin{aligned}
\pi^*(j,k) \;=\; & [\eta + \tilde{\delta}_a(j) + \tilde{\delta}_\theta(k) + \lambda_E \sum_{l>k} \pi^*(j,l)]^{-1} \\
& \times \{\varphi_a(j-1,k)\pi^*(j-1,k) + \varphi_\theta(j,k-1)\pi^*(j,k-1) \\
& + \tilde{\delta}_a(j+1)\pi^*(j+1,k) + \tilde{\delta}_\theta(k+1)\pi^*(j,k+1) \\
& + \lambda_U \pi^*(j,0)p_\theta(k) + \lambda_E \sum_{l<k} \pi^*(j,l)\},
\end{aligned}
$$

for $j = 1, ..., J,\ k = 1, ..., K$.

We can vectorize the π matrix, and define the column vector

$$
\Pi = \begin{bmatrix} \pi(1,\cdot) \\ \pi(2,\cdot) \\ \vdots \\ \pi(J,\cdot) \end{bmatrix},
$$

where $\pi(j,\cdot) = (\pi(j,0)\ \pi(j,1)\ ...\ \pi(j,K))'$. Some elements of this vector are identically equally to 0, those for which $d(j,k) = 0$. Let the number of nonzero elements of Π be denoted $N(\Pi)$, where $N(\Pi) \leq J \times (K+1)$. Denote the entire system of equations by $D(\Pi)$. Then we seek

$$
\Pi^* = D(\Pi^*).
$$

With no on-the-job search, this mapping is monotone on a compact space, and hence the solution is unique. With on-the-job search, it is clear that an equilibrium always exists, although we have not yet proven uniqueness. Simulations of the model and computation of the fixed point have consistently agreed, however, so that we are confident in the uniqueness property.

References

Acemoglu, D. and J. S. Pischke (1999) "The Structure of Wages and Investment in General Training," *Journal of Political Economy* 107(3): 539-572.

Acemoglu, D. (1996) "A Microfoundation for Social Increasing Returns in Human Capital," *Quarterly Journal of Economics* 111: 779804.

Acemoglu, D. (1997) "Training and Innovation in an Imperfect Labor Market," *Review of Economic Studies* 64: 44564.

Acemoglu, D. and R. Shimer (1999) "Holdups and Inefficiencies with Search Frictions," *International Economic Review* 40: 827-849.

Bagger, J., F. Fontaine, F. Postel-Vinay and J-M Robin (2014), "Tenure, Experience, Human Capital and Wages: A Tractable Equilibrium Search Model of Wage Dynamics," *American Economic Review* 104(6): 1551-96.

Bartolucci, C. (2013) "Gender Wage Gaps Reconsidered: A Structural Approach Using Match Employer-Employee Data," *Journal of Human Resources* 48: 998-1034.

Becker G. (1964) *Human Capital*, 2nd edition, Columbia University Press, New York.

Bound, J., C. Brown, G. Duncan, and W. Rodgers (1994) "Evidence on the Validity of Cross-Sectional and Longitudinal Labor Market Data," *Journal of Labor Economics* 12: 345-368.

Brown, J. N. (1989) "Why do Wages Increase with Tenure? On-the-Job Training and Life-Cycle Wage Growth Observed within Firms," *American Economic Review*, 79(5): 971-991.

Cahuc, P., F. Postel-Vinay, and J.-M. Robin (2006) "Wage Bargaining with On-the Job Search: Theory and Evidence." *Econometrica* 74: 323-64.

Card, D. (1992) "Do Minimum Wages Reduce Employment? A Case Study of California, 1987-1989," *Industrial and Labor Relations Review* 46: 38-54.

Dey, M. and C. Flinn (2005) "An Equilibrium Model of Health Insurance Provision and Wage Determination," *Econometrica* 73: 571-627.

Eckstein, Z., and K. Wolpin (1995) "Duration to First Job and the Return to Schooling: Estimates from a Search-Matching Model," *Review of Economic Studies* 62: 263- 286.

Flinn, C. (2006) "Minimum Wage Effects on Labor Market Outcomes under Search, Bargaining, and Endogenous Contact Rates," *Econometrica* 74: 1013-1062.

Flinn, C., and J. Heckman (1982) "New Methods for Analyzing Structural Models of Labor Force Dynamics." *Journal of Econometrics* 18: 115-68.

Flinn, C., and J. Mabli (2009) "On-the-Job Search, Minimum Wages, and Labor Market Outcomes in an Equilibrium Bargaining Framework, Department of Economics, NYU. Downloadable at: http://www.nyu.edu/econ/user/flinnc/papers/flinnmabliberlin.pdf

Flinn, C., and J. Mullins (2015) "Labor Market Search and Schooling Investment," *International Economic Review* 56: 359-398.

Gavrel, F., J-P. Guironnet and I. Lebon. (2012) "Mismatch, On-the-job Training, and Unemployment," Working Paper, CREM-CNRS.

Heckman, J., L. Lochner, and C. Taber. (1998) "Explaining Rising Wage Inequality: Explanations with a Dynamic General Equilibrium Model of Earnings with Heterogeneous Agents," *Review of Economic Dynamics* 1: 1-58.

Hosios, A. (1990) "On the Efficiency of Matching and Related Models of Search and Unemployment," *Review of Economic Studies* 57: 279-298.

Keane, M., and K. Wolpin (1997) "The Career Decisions of Young Men," *Journal of Political Economy* 105: 473-522.

Krueger, A. (1999) "Measuring Labors Share," *American Economic Review* 89(2): 45-51.

Lee, D. (2005) "An Estimable Dynamic General Equilibrium Model of Work, Schooling, and Occupational Choice," *International Economic Review* 46: 1-34.

Lee, D. and K. Wolpin (2006) "Intersectoral Labor Mobility and the Growth of the Service Sector," *Econometrica* 47: 1-46.

Lentz, R. and N. Roys (2015), "Training and Search On the Job," Manuscript, U. of Wisconsin-Madison.

Lindenlaub, I. and J. Eeeckhout (2013) "Unemployment Cycles," Manuscript, Institute for Fiscal Studies.

Loewenstein, M. A., and J. R. Spletzer (1998) "Dividing the Costs and Returns to General Training," *Journal of Labor Economics* 16(1): 142-171.

Malcomson, J. (1997) "Contracts, Hold-Up, and Labor Markets," *Journal of Economic Literature* 35: 1916-1957.

Mortensen, D. and C. Pissarides (1994) "Job Creation and Job Destruction in the Theory of Unemployment," *Review of Economic Studies* 61: 397-415.

Neumark, D. and O. Nizalova (2007) "Minimum Wage Effects in the Longer Run," *Journal of Human Resources* 42: 435-452.

Neumark, D. and W. Wascher (1995) "Minimum Wage Effects on Employment and School Enrollment," *Journal of Business and Economics Statistics* 13: 199-206.

Neumark, D. and W. Wascher (2003) "Minimum Wages and Skill Acquisition: Another Look at Schooling Effects," *Economics of Education Review* 22: 1-10.

Petrongolo, B. and C. Pissarides (2001) "Looking into the Black Box: A Survey of the Matching Function," *Journal of Economic Literature* Vol. XXXIX: pp. 390431.

Pissarides, C. (2000) *Equilibrium Unemployment Theory*, 2nd edition. Cambridge, MA: MIT Press.

Postel-Vinay, F., and J-M Robin (2002) "Equilibrium Wage Dispersion with Heterogeneous Workers and Firms." *Econometrica* 70: 2295-1350.

Shimer, R., R. Rogerson and R. Wright (2005) "Search Theoretic Models of the Labor Market: A Survey," *Journal of Economic Literature*, 43: 959-988.

Table 1: **Proportion of Job Spells by the number of interview dates they span**

	High School Graduates	Some College	College Graduates
0	61%	55%	53%
1	23%	24%	21%
2	7%	9%	11%
3	3%	4%	6%
4	2%	3%	3%
5	1%	2%	3%
6	3%	3%	3%

Table 2: **Incidence of Training**

	All	Some HS	College College	or More
% who got training at least once	15%	18 %	13 %	13%
% who got training at the start of job spell	6%	10 %	5 %	3%

Table 3: **Proportion by Number of Training Spells (Conditional on Having Participated At Least Once)**

	Percentage over All Workers with at Least One Training Spell
1	72 %
2	18 %
3	6 %
4	3 %
5	1 %

Table 4: **Annual Labor Turnover Rates and Wage Growth**

	HS	Some College	College or More
Panel A: job-to-job transitions btw interview dates $t-1$ and t			
% of stayers	81 %	85 %	88%
% of job switchers	19 %	15 %	12%
..... % of job switchers with nonemployment spell btw $t-1$ and t	4 %	3%	2%
..... % of job switchers with no nonemployment spell btw $t-1$ and t	15%	12%	10%
Panel B: Wage Growth btw Interview Dates $t-1$ and t			
stayers	0.08	0.08	0.09
job switchers	0.11	0.15	0.20
..... job switchers with nonemployment spell btw $t-1$ and t	0.06	0.06	0.23
..... job switchers with no nonemployment spell btw $t-1$ and t	0.12	0.17	0.20
Panel C: % of Negative Wage Growth btw Interview Dates $t-1$ and t			
stayers	17 %	20 %	23%
job switchers	21 %	22 %	15%
..... job switchers with nonemployment spell btw $t-1$ and t	30 %	36%	18%
..... job switchers with no nonemployment spell btw $t-1$ and t	18%	18%	14%

Table 5: **Log Wage Difference btw Interview Dates** $t-1$ **and** t **by Training**

	$logw_t - logw_{t-1}$	
	No Training	Got Training
stayers	0.08	0.08
job switchers	0.14	0.10
job switchers with nonemployment spell btw $t-1$ and t	0.08	0.15
job switchers with no nonemployment spell btw $t-1$ and t	0.15	0.09

Table 6: Parameter Estimates

PARAMETERS FOR EMPLOYMENT TRANSITIONS			
flow value of unemployment	b	4.93	(0.064)
job offer rate - unemployed	λ_u	0.145	(.004)
job offer rate - employed	λ_e	.074	(.002)
exogenous job separation rate	η	.0033	(.0001)
PARAMETERS OF INVESTMENT FUNCTIONS			
General ability investment TFP	δ_a^0	.0150	(.0003)
Firm-specific investment TFP	δ_θ^0	.0151	(.0003)
State-dependence of general ability investment	δ_a^1	-.050	(.010)
State-dependence of firm-specific investment	δ_θ^1	.702	(.006)
Curvature of general ability investment	δ_a^2	.354	(.008)
Curvature of firm-specific investment	δ_θ^2	.493	(.007)
Rate of decrease in general ability	$\tilde{\varphi}_a^-$	.0011	(.00004)
Rate of decrease in match quality	$\tilde{\varphi}_\theta^-$	.0144	(.0004)
PARAMETERS OF INITIAL ABILITY DISTRIBUTIONS			
Mean of initial general ability - High School	$\mu_a(e_i = 1)$	1.07	(.024)
Mean of initial general ability - Some College	$\mu_a(e_i = 2)$	1.28	(.020)
Mean of initial general ability - BA or higher	$\mu_a(e_i = 3)$	1.53	(.051)
Variance of initial general ability	σ_a	.190	(.005)
PARAMETERS OF JOB OFFERS			
Mean of match quality distribution	μ_θ	1.48	(.016)
Variance of match quality distribution	σ_θ	.362	(.006)
Employment cost	ζ	4.51	(.120)
PARAMETERS GOVERNING TRAINING OBSERVATION			
Intercept for training observation	β_0	-3.19	(.025)
Coefficient on τ for training observation	β_a	2.35	(.019)

Table 7: **Parameter Estimates: Baseline vs. Estimations with No a/No θ**

		Baseline	No a	No θ
PARAMETERS FOR EMPLOYMENT TRANSITIONS				
flow value of unemployment	b	4.93	2.92	5.20
job offer rate - unemployed	λ_u	0.145	0.148	0.161
job offer rate - employed	λ_e	.074	0.104	0.044
exogenous job separation rate	η	.0033	0.0037	.0043
PARAMETERS OF INVESTMENT FUNCTIONS				
General ability investment TFP	δ_a^0	.0150	0	0.0164
Firm-specific investment TFP	δ_θ^0	.0151	0.0204	0
State-dependence of general ability investment	δ_a^1	-.050	-	0.083
State-dependence of firm-specific investment	δ_θ^1	.702	0.696	-
Curvature of general ability investment	δ_a^2	.354	-	0.261
Curvature of firm-specific investment	δ_θ^2	.493	0.610	-
Rate of decrease in general ability	$\tilde{\varphi}_a^-$	.0011	0	0.0012
Rate of decrease in match quality	$\tilde{\varphi}_\theta^-$	.0144	0.0118	0
PARAMETERS OF INITIAL ABILITY DISTRIBUTIONS				
Mean of initial general ability - High School	$\mu_a(e_i=1)$	1.07	1.05	1.24
Mean of initial general ability - Some College	$\mu_a(e_i=2)$	1.28	1.24	1.41
Mean of initial general ability - BA or higher	$\mu_a(e_i=3)$	1.53	1.50	1.62
Variance of initial general ability	σ_a	0.19	0.29	0.17
PARAMETERS OF JOB OFFERS				
Mean of match quality distribution	μ_θ	1.48	1.36	1.44
Variance of match quality distribution	σ_θ	0.36	0.36	0.36
Employment cost	ζ	4.51	3.79	6.15
PARAMETERS GOVERNING TRAINING OBSERVATION				
Intercept for training observation	β_0	-3.19	-3.04	-2.99
Coefficient on τ for training observation	β_a	2.35	2.26	2.63

Figure 1: Minimum Acceptible Match-Quality

This figure shows the $\theta^*(a)$, the lowest match quality that workers with each level of general ability will accept.
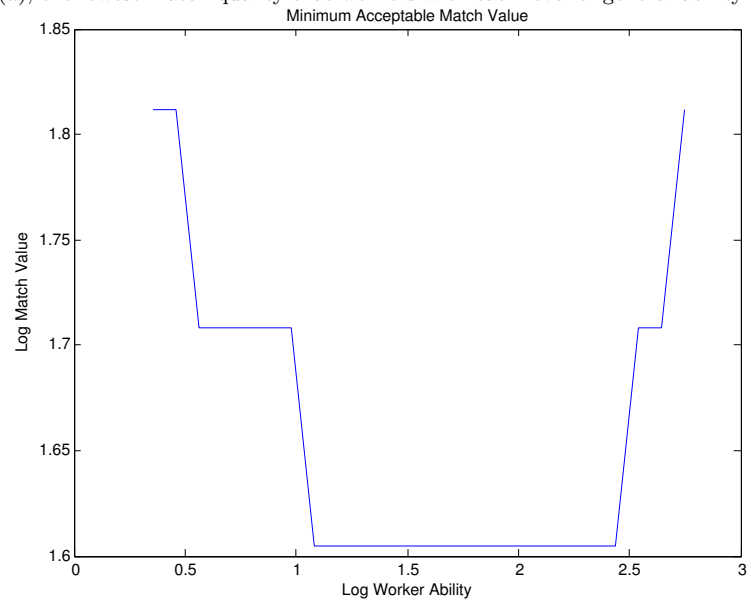

Minimum Acceptable Match Value

Figure 2: General Training

This figure shows the amount of general training that wokers receive at different combinations on a and θ. Lines on the graph show contours along which the amount of training remains constant. The blank area below the black line, shows states for which workers will not accept the job offer.

Figure 3: Match-Specific Training

This figure shows the amount of match-specific training that wokers receive at different combinations on a and θ. Lines on the graph show contours along which the amount of training remains constant. The blank area below the black line, shows states for which workers will not accept the job offer.
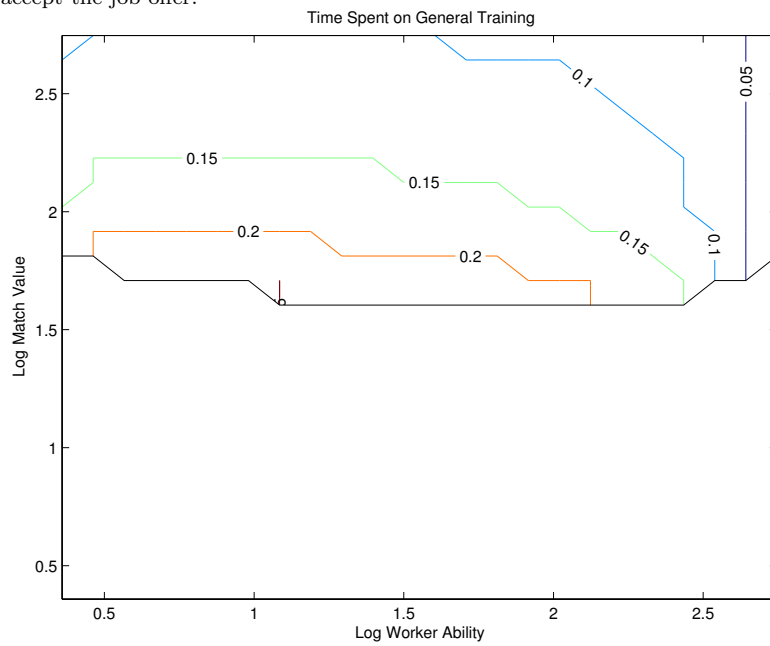

Time Spend on Match–Specific Training

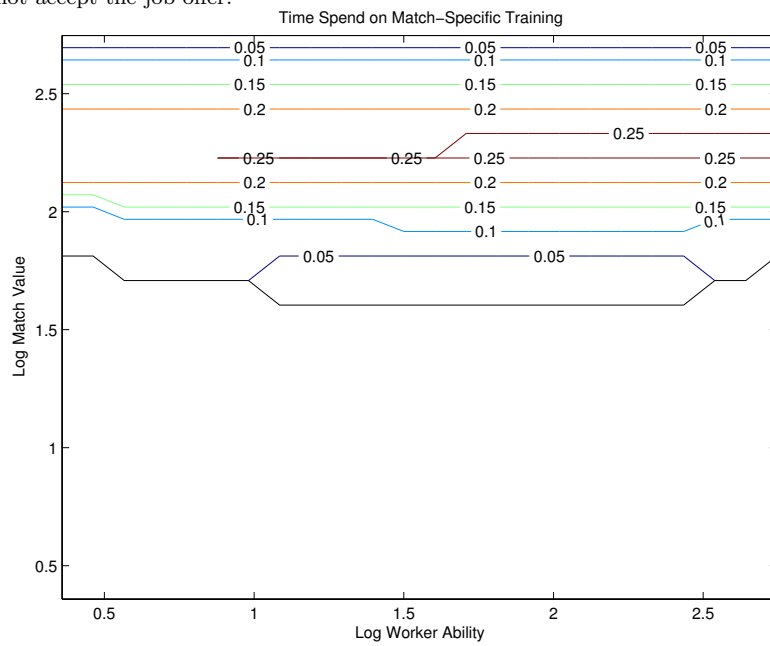## Figure 4: Wage as Fraction of Output

This figure shows the worker's wage as a fraction of her total output at different combinations on a and θ. Lines on the graph show contours along which the fraction remains constant. The blank area below the black line, shows states for which workers will not accept the job offer.
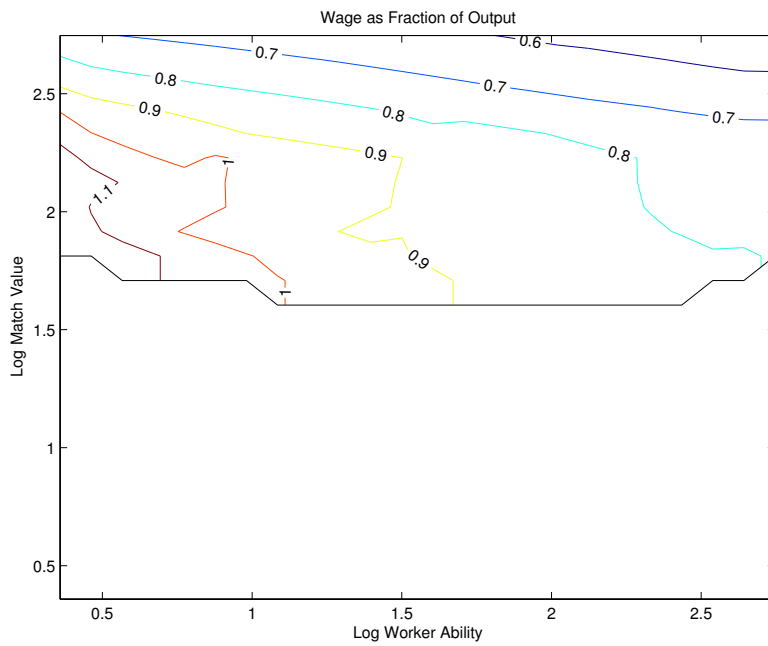
Figure 5: **Model Fit: Log Wage Distribution - Year 0-2**
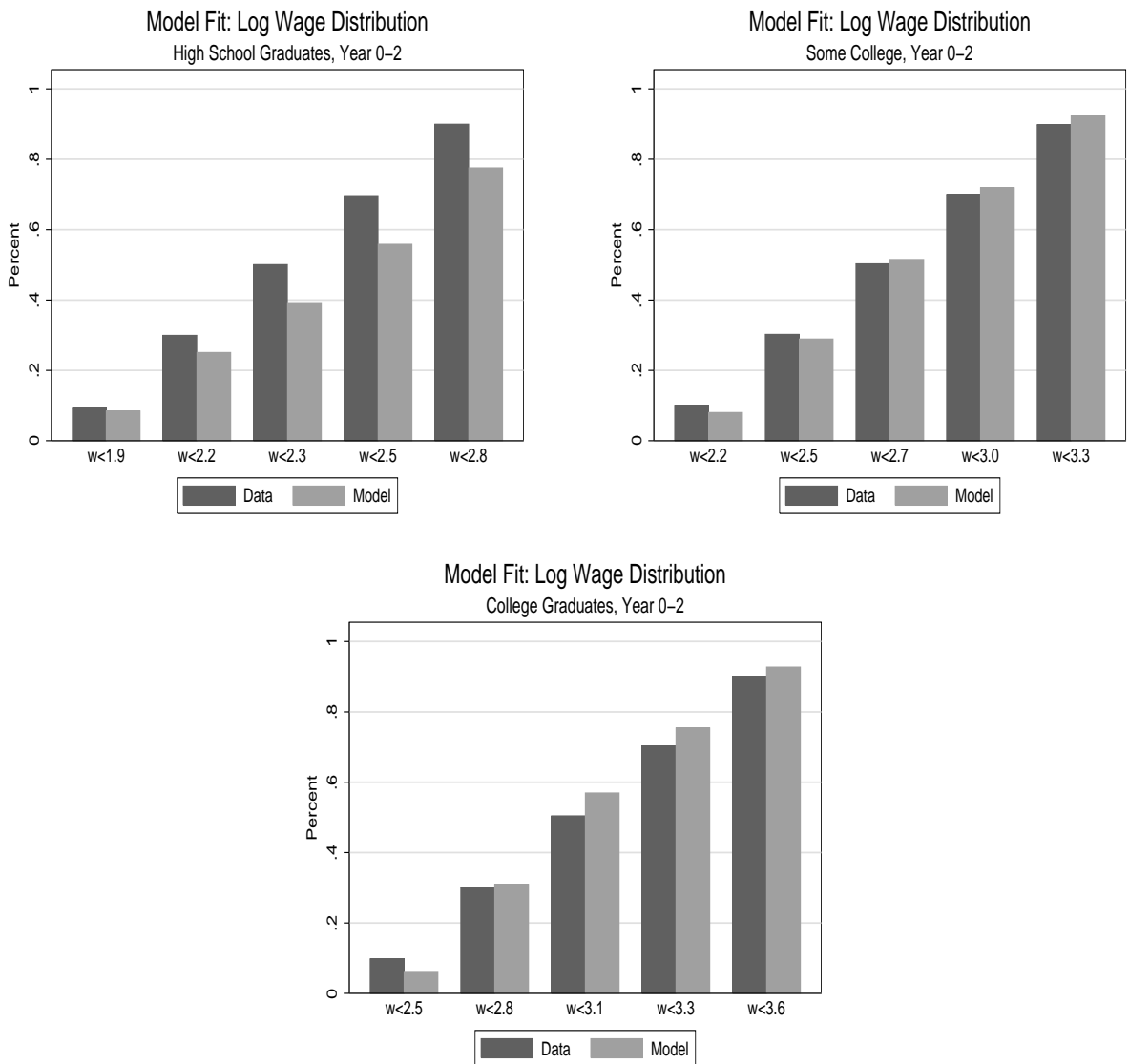
Figure 6: **Model Fit: Log Wage Distribution - Year 3-5**

Model Fit: Log Wage Distribution
High School Graduates, Year 3–5

Model Fit: Log Wage Distribution
Some College, Year 3–5

Model Fit: Log Wage Distribution
College Graduates, Year 3–5
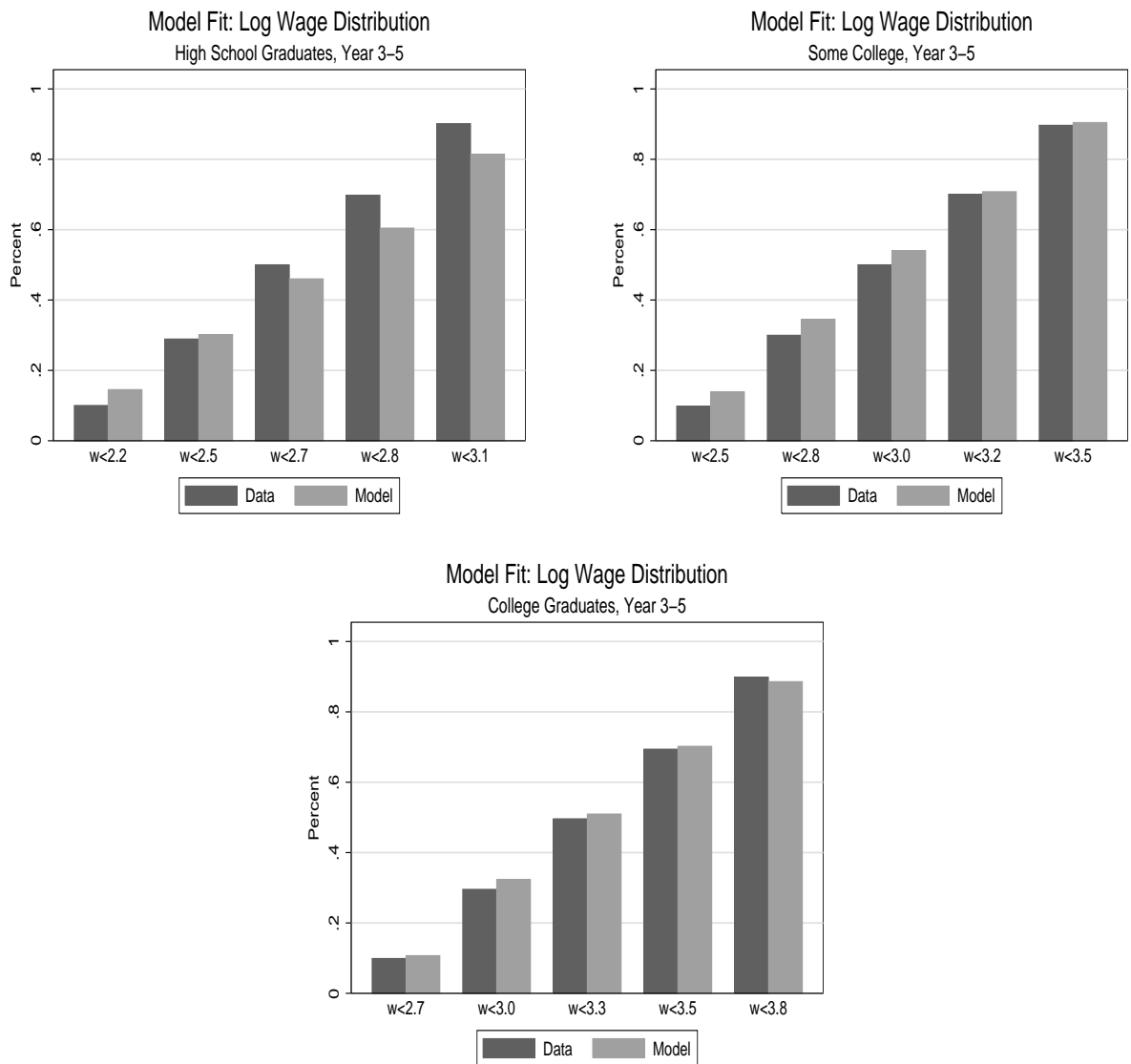
Figure 7: **Model Fit: Log Wage Distribution - Year 6-8**

Model Fit: Log Wage Distribution
High School Graduates, Year 6–8

Model Fit: Log Wage Distribution
Some College, Year 6–8

Model Fit: Log Wage Distribution
College Graduates, Year 6–8
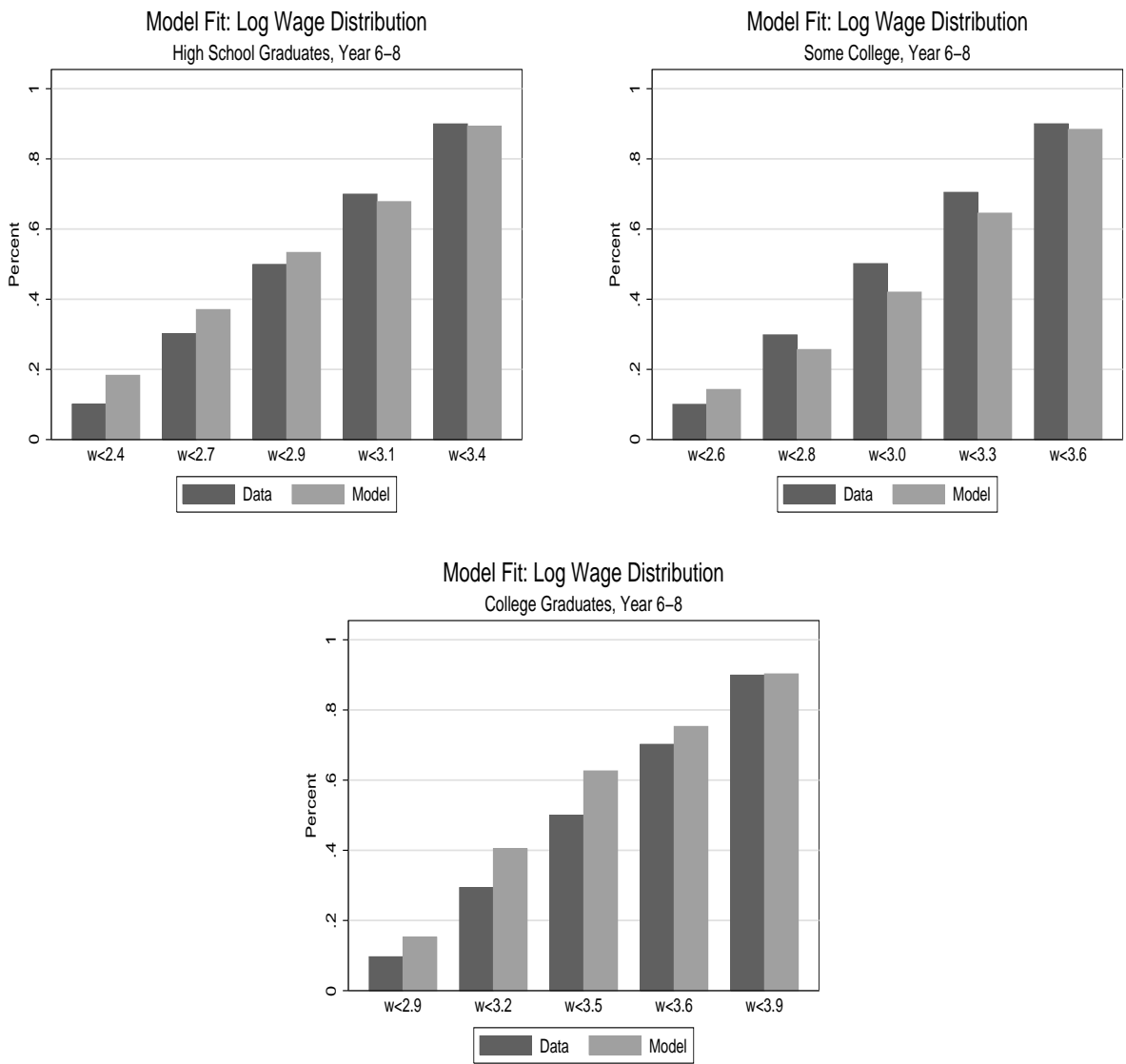
Table 8: **Model Fit: Incidence of Training**

	All	HS	Some College	College or More
% who got training at least once				
.... Data	15%	18 %	13 %	13%
.... Model	16%	20 %	15 %	13%
% who got training at the start of job spell				
.... Data	6%	10 %	5 %	3%
.... Model	4%	5 %	4 %	4%

Table 9: **Model Fit: Annual Labor Turnover Rates and Wage Growth**

	HS	Some College	College or More
Panel A: job-to-job transitions btw interview dates $t-1$ and t			
% of stayers			
.... Data	81 %	85 %	88%
.... Model	78 %	75 %	73%
% of job switchers			
.... Data	19 %	15 %	12%
.... Model	22 %	25 %	27%
% of job switchers with nonemployment spell btw $t-1$ and t			
.... Data	4 %	3%	2%
...... Model	8 %	8%	9%
% of job switchers with no nonemployment spell btw $t-1$ and t			
...... Data	15%	12%	10%
...... Model	14%	16%	18%
Panel B: Wage Growth btw Interview Dates $t-1$ and t			
stayers			
.... Data	0.08	0.08	0.09
.... Model	0.08	0.08	0.07
job switchers			
.... Data	0.11	0.15	0.20
.... Model	0.07	0.11	0.12
job switchers with nonemployment spell btw $t-1$ and t			
.... Data	0.06	0.06	0.23
.... Model	-0.17	-0.12	-0.08
job switchers with no nonemployment spell btw $t-1$ and t			
.... Data	0.12	0.17	0.20
.... Model	0.21	0.22	0.22
Panel C: % of Negative Wage Growth btw Interview Dates $t-1$ and t			
stayers			
.... Data	17 %	20 %	23%
.... Model	39 %	39 %	39%
job switchers			
.... Data	21%	22%	15%
.... Model	31%	27%	26%
job switchers with nonemployment spell btw $t-1$ and t			
.... Data	30 %	36%	18%
.... Model	54 %	50%	47%
job switchers with no nonemployment spell btw $t-1$ and t			
.... Data	18%	18%	14%
.... Model	18%	16%	16%

Table 10: **Simulations: Baseline vs. Estimations with No a/No θ**

		Some	College
	HS	College	or More

Annual Labor Turnover Rates and Wage Growth

Panel A: job-to-job transitions btw interview dates $t-1$ and t

	HS	Some College	College or More
% of stayers			
.... Baseline	78 %	75 %	73%
.... No a	78 %	76 %	72%
.... No θ	72 %	72 %	71%
% of job switchers			
.... Baseline	22 %	25 %	27%
.... No a	22 %	24 %	28%
.... No θ	28 %	28 %	29%
% of job switchers with nonemployment spell btw $t-1$ and t			
.... Baseline	8 %	8%	9%
.... No a	8 %	8 %	9%
.... No θ	10%	10 %	10%
% of job switchers with no nonemployment spell btw $t-1$ and t			
.... Baseline	14%	16%	18%
.... No a	14%	16%	19%
.... No θ	18%	19%	19%

Panel B: Wage Growth btw Interview Dates $t-1$ and t

	HS	Some College	College or More
stayers			
.... Baseline	0.08	0.08	0.07
.... No a	0.07	0.08	0.08
.... No θ	0.06	0.06	0.06
job switchers			
.... Baseline	0.07	0.11	0.12
.... No a	-0.008	0.03	0.05
.... No θ	0.17	0.19	0.17
job switchers with nonemployment spell btw $t-1$ and t			
.... Baseline	-0.17	-0.12	-0.08
.... No a	-0.22	-0.17	-0.14
.... No θ	-0.08	-0.05	-0.009
job switchers with no nonemployment spell btw $t-1$ and t			
.... Baseline	0.21	0.22	0.22
.... No a	0.12	0.12	0.13
.... No θ	0.31	0.31	0.27

Panel C: % of Negative Wage Growth btw Interview Dates $t-1$ and t

	HS	Some College	College or More
stayers			
.... Baseline	39 %	39 %	39%
.... No a	40%	39%	38%
.... No θ	40%	40%	40%
job switchers			
.... Baseline	31%	27%	26%
.... No a	38%	34%	31%
.... No θ	27%	24%	24%
job switchers with nonemployment spell btw $t-1$ and t			
.... Baseline	54 %	50%	47%
.... No a	60%	55%	52%
.... No θ	49%	46%	43%
job switchers with no nonemployment spell btw $t-1$ and t			
.... Baseline	18%	16%	16%
.... No a	26%	23%	21%
.... No θ	14%	13%	13%

Table 11: **Simulations: Baseline vs. No a/No θ**

| | Weekly Transition Rates Between $t-1$ and t | | |
	Baseline	No a	No θ
EU probability	0.42%	0.42%	0.44%
% not employed - At Interview Date	14.88%	15.25%	12.23%
% not employed - All Periods	13.59%	13.91%	11.27%
tenure x EU probability	41.45%	44.02%	41.35%
EU probability by Tenure			
... ¡6 mths	0.52%	0.51%	0.45%
... 6-12 mths	0.46%	0.46%	0.43%
... 1-2 yrs	0.41%	0.41%	0.44%
... 2-4 yrs	0.36%	0.37%	0.43%
... ¿4 yrs	0.33%	0.37%	0.42%
EE transition probability	0.38%	0.42%	0.39%
tenure x EE probability	21.22%	18.16%	30.19%
job finding probability	4.04%	3.95%	5.09%
À			
time not employed x job finding probability	91.29%	90.07%	90.21%

Simulations: Baseline vs. No a/No θ

Figure 8: **Avg. Log Wages by Tenure**

Model Simulations: Baseline vs. No Training in A or Theta
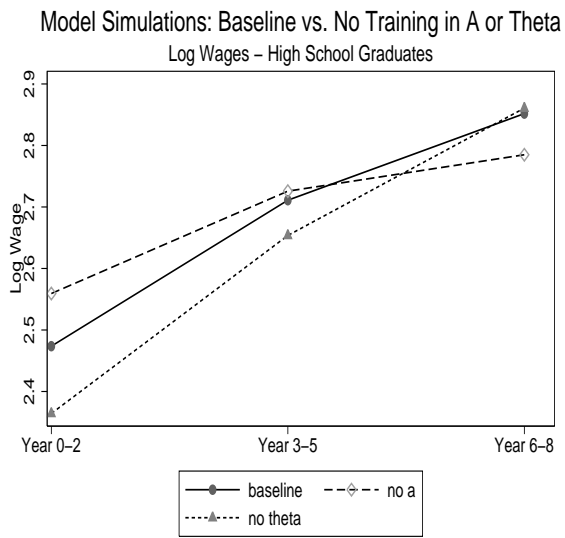Log Wages – High School Graduates

Figure 9: **Avg. Log Wages by Tenure**

Model Simulations: Baseline vs. No Training in A or Theta
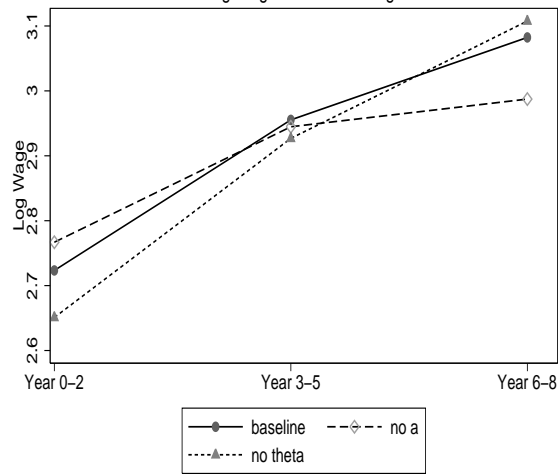
Log Wages – Some College

Figure 10: **Avg. Log Wages by Tenure**

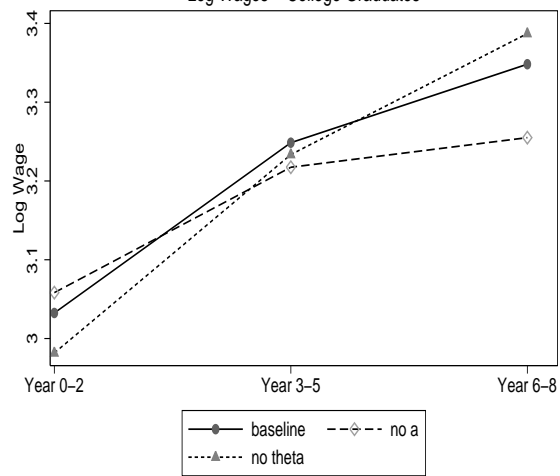Model Simulations: Baseline vs. No Training in A or Theta

Log Wages – College Graduates

Table 12: **Simulations: Baseline vs. No a/No θ**

| Years in Labor Mkt. | Education | Log Wages by Labor Market Tenure and Education | | | |
		Data	Baseline	No a	No θ
0-2	High School	2.36	2.47	2.56	2.36
3-5	High School	2.67	2.71	2.73	2.65
6-8	High School	2.90	2.85	2.78	2.86
0-2	Some College	2.72	2.72	2.77	2.65
3-5	Some College	2.99	2.96	2.94	2.93
6-8	Some College	3.04	3.08	2.99	3.11
0-2	College or Higher	3.03	3.03	3.06	2.98
3-5	College or Higher	3.25	3.25	3.22	3.23
6-8	College or Higher	3.43	3.35	3.26	3.39

Figure 11: Training by Years in Labor Market

This figure shows the average fraction of their time that workers spend on general and match specfic training in the model simulation as a fucntion of the number of years in the labor market.
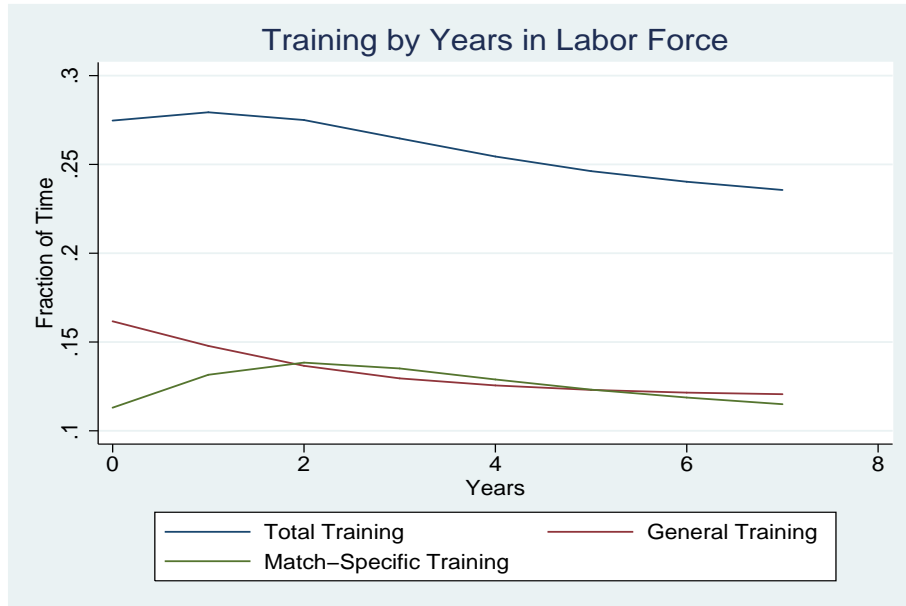
Figure 12: Sources of Wage Growth by Years in Labor Market

This figure shows the average of the log of general human capital, the average amount of matc-specific capital, the log of the fraction of time they spend not training, and the log of the average wage as a fraction of worker output for simulated workers as a fucntion of the number of years in the labor market. In the absence of employment costs, these four componets would add up to the total log wage, which is also shown.
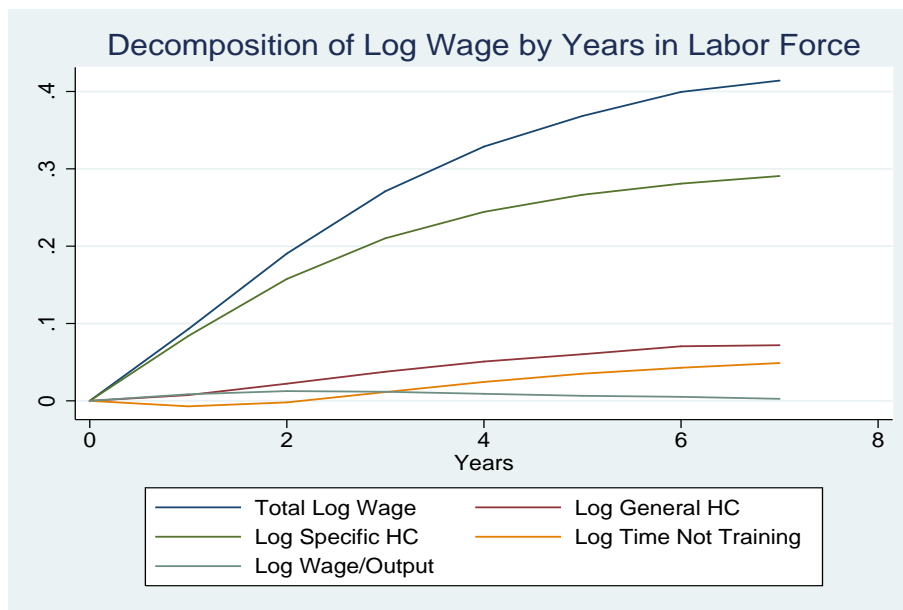
Table 13: **Mincer Regressions**

	Const	Some College	BA	Years in LF	Tenure
DATA					
log wage	2.141	0.316	0.796	0.051	0.045
	(0.098)	(0.098)	(0.098)	(0.002)	(0.003)
MODEL					
log wage	2.421	0.257	0.563	0.031	0.075
starting a	1.067	0.208	0.459	0.001	-0.002
a from training	0.016	-0.001	-0.005	0.020	0.002
θ from seach	2.011	-0.006	-0.009	0.005	0.017
θ from training	0.028	0.001	0.002	-0.001	0.046
hours worked	-0.345	0.005	0.012	0.001	0.019
wage/output	-0.356	0.049	0.104	0.005	-0.008

Figure 13: Mininimum Acceptable Wage with Minimum Wage

This figure shows the $\theta^*(a)$, the lowest match quality that workers with each level of general ability will accept. The solid line shows $\theta^*(a)$ for the baseline model, the dashed line when we impose a minimum wage.
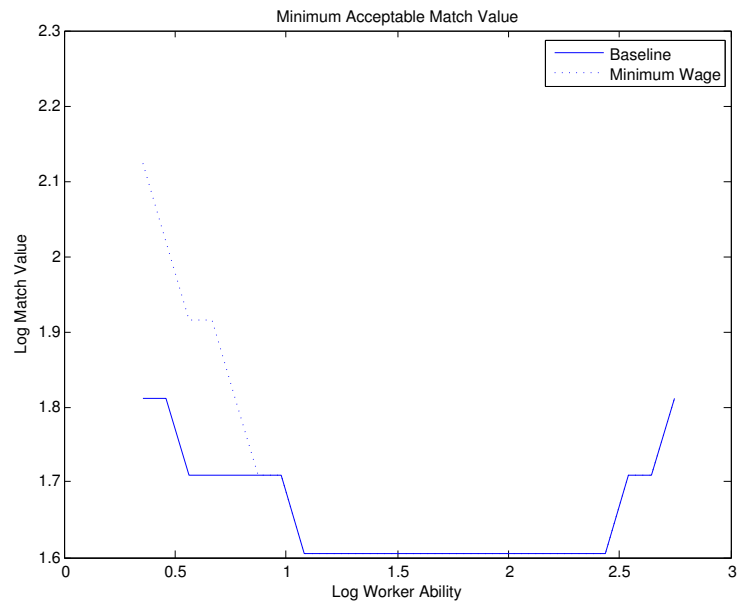
Figure 14: Training by Years in Labor Market with Minimum Wage

This figure shows the average fraction of their time that workers spend on general and match specfic training in the model simulation as a fucntion of the number of years in the labor market. The solid lines show the amount of training in the baseline model. The corresponding dashed lines show the amount of training when we impose a minimum wage.

74