Finance and Economics Discussion Series
Divisions of Research & Statistics and Monetary Affairs
Federal Reserve Board, Washington, D.C.

Search, Matching and Training

Christopher Flinn, Ahu Gemici and Steven Laufer

2016-075

Please cite this paper as:
Flinn, Christopher, Ahu Gemici and Steven Laufer (2016). "Search, Matching and Training," Finance and Economics Discussion Series 2016-075. Washington: Board of Governors of the Federal Reserve System, http://dx.doi.org/10.17016/FEDS.2016.075.

Search, Matching and Training

Christopher Flinn
New York University
Collegio Carlo Alberto

Ahu Gemici
Royal Holloway, U. of London

Steven Laufer*
Federal Reserve Board

July 13, 2016

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.

*We thank participants at the HCEO conference on Human Capital and Inequality, the IZA/Crest/Bonn Conference on "Employment Policies and Heterogeneity in the Labor Market", the 2014 SED in Toronto, and seminar participants at Duke, London School of Economics and George Washington for useful feedback. We have also received helpful suggestions from Jesper Bagger, Jim Heckman, Thibaut Lamadon, Rasmus Lentz, Jean-Marc Robin, Chris Taber and Tony Yezer. Christopher Flinn thanks the National Science Foundation for partial support of this research. All remaining errors are our own. The views expressed herein are those of the authors and not necessarily those of the Board of Governors or others within the Federal Reserve System.

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 introduce training decisions into what is otherwise is a reasonably standard search model with general and specific human capital. The training data we use to estimate the model, 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) and 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 downward movement 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.

Another motivation for our research is to provide a richer model of the path of wages on the job and a more complete view of the relationship between workers and firms.

[1] Although 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 "unemployed" in the true sense of the term.

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 models 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.

In our modelling framework, the complementarity between general and specific human capital gives firms an incentive to partially finance improvements in general ability. Specifically, 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 be an incentive to "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 school-

[2]That is, the only costs of investing in either of the skills is the lost output associated with the investment time. Moreover, we will assume that 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. Depreciation in skills is not considered in either paper.

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

ing decisions, and these decisions were a function of all of the primitive parameters characterizing the labor market. 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 and exogenous) 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 and noncognitive abilities of their students. To the extent that a is essentially fixed over the labor market career, 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 that we make below, higher values of a at the time of labor market entry could 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, it is possible to examine 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 a 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 higher levels of the minimum wage, however, most individuals invested in a to increase their probability 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 for competitive labor markets, 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 make the constraint nonbinding by pushing 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.

In terms of related research, the closest paper to ours is probably Wasmer (2006). He presents a formal analysis of the human capital investment problem after market entry 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. One of the goals of our paper is to estimate both partial and general equilibrium versions 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 serves to simplify the theoretical analysis of the model.

The model that we develop and estimate below has several notable features. As noted above, it endogenizes the general-productivity level of the individual and the match-productivity level of the worker-firm pair using a cooperative model of worker-firm interactions. In most of the literature on worker-firm sorting (e.g., Abowd et al. (1999), Postel-Vinay and Robin (2002), Cahuc et al. (2006)), match productivity is ignored and worker and firm types are assumed to be time-invariant. The focus of much of this literature is on worker-firm sorting patterns. Within our framework, it is clear that the total productivity of an employee at a particular firm is a fluid object, with "mismatches" being potentially rectified through cooperative investment choices by the worker and firm. Viewing the productivity of a worker at a given firm as an endogenous stochastic process is an important point of differentiation of our model from most of the literature.

The fact that productivity can be altered forces us to reconsider the usual constraints on firm hiring that are implied by models without endogenous productivity. When productivity is fixed, then for firms to at least break even on the employment contract, the flow profit of a firm should be non-negative. However, when productivity can be increased through investment, it is possible (and occurs given our model estimates) that firms could earn negative flow profits for some part of the employment contract. This is an especially important consideration when evaluating the impact of policies such as a mandated minimum wage. In Flinn (2006), for instance, establishing a binding minimum wage of m in a market (in which there was not one previously) immediately caused the loss of all jobs for which the worker's flow productivity was less than m. With endogenous productivity, this may no longer be the case, since through investment, the worker's productivity could be raised sufficiently so that the firm earns positive flow profits. The extent to which this phenomenon occurs is an empirical matter, which we can investigate using our model estimates.

Another literature to which this paper contributes is that concerned with the decomposition of the sources of wage growth over the life cycle, a literature the genesis of which traces back at least to Mincer, whose work on wage determination is summarized in Mincer (1974). While much of the work in this literature focuses on the estimation of the return to schooling,[4] our contribution is to the interpretation of the part of the earnings process associated with the worker's total labor market experience and tenure on their current job. The model generates positive dependencies between the duration of time in the market and at the current job with wages since both duration measures are positively related to the values of general and match-specific human capital. As individuals age, there is a tendency for general human capital, a, to increase. Also, because the turnover decision is determined through the comparison between the match productivity value θ at the current firm, with the initial match productivity draw θ' at the competing firm, longer employment durations at the firm indicate higher current values of match productivity. Even though the Mincer earnings function is largely an atheoretical construct, it is of interest to determine the extent to which the data generating process associated with our model produces relationships between wages, schooling, general experience, and job tenure broadly consistent with what would be found when estimating a Mincerian wage function using data from our sample. We find that, by and large, the wage functions estimated using the sample and those using simulated data from the model are roughly in agreement, even though the sample regression was not used in estimating the model parameters. This gives us some confidence that the model estimated does not generate empirical implications at variance with the results in literature focused on the estimation and interpretation of Mincerian earnings functions.

[4]Heckman et al. (2006) present an exhaustive consideration of the estimation of Mincer-type earnings functions, although their main focus is on the consistent estimation of rates of return to schooling.

The model estimates we present are obtained using data from the 1997 cohort of the National Longitudinal Survey of Youth (NLSY97), which consists of individuals between the ages of 12-16 at the end of 1997. The advantage of using these data are that we can observed individuals from the beginning of their labor market careers, which minimizes initial conditions problems. Moreover, it is well-known that job changing behavior and wage growth is more pronounced at the onset of the labor market career. One negative aspect of estimating the model using only relatively recent labor market entrants is that we are likely to get predictions for later career events that are at odds with the data. This also implies that the steady state distributions implied by the model should be interpreted with this caveat in mind. Our future research plan is to extend the model to include endogenous pre-market entry schooling decisions and to estimate the model using samples with a larger range of ages and participation histories.

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 estimates, model 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 novel 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 the paper.

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

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

[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.

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

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

We denote the c.d.f. of θ by G, and we define $p_j = \Pr(\theta = \theta_j)$, $j = 1, ..., K$.

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 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 the model using 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 the arrival one of these Poisson-distributed 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 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, with total productivity 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 $V_U(i)$, 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

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

[6] By this we mean that either

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

or

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

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, $i = 1, ..., M$, .where b is a scalar parameter. The value of unemployed search can be written as

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

where λ_U is the rate of arrival of potential employment opportunities to the individual. The discount rate ρ is the sum of the subjective discount rate of the individual, ρ_0, and a constant death rate of ℓ, so that $\rho = \rho_0 + \ell$.[7] It is assumed that the value associated with the state of death is 0. 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.[8]

Given a wage of w and a training level of τ_a and τ_θ, the value of employment of type a_i at a match of θ_j is

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

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 $r^*(i + 1) \geq r^*(i) + 1$.

[7] The death rate is introduced so as to produce more reasonable steady state distributions than the ones we generate using model estimates and an assumption that $\ell = 0$. Less dramatically, we can think of this state as corresponding to retirement, but it should be borne in mind that we assign the value of this absorbing state to be 0.

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

The corresponding value to the firm is

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

where $Q_F(i,j) = 0$ if $Q(i,j) = V_U(i)$ and $Q_F(i,j) = V_F(i,j)$ if $Q(i,j) = V_E(i,j)$.[9] Then the solution to the surplus division problem is given by

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

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 will obviously have important implications for

[9]Note that the discount rate ρ is the same for both workers and firms, even if we do not think of firms as being subject to a death shock. In this framework, there is essentially one worker per firm, and the worker's "death" terminates the match just as does a shock dissolving that particular job, η, assumed to arise due to changes in demand conditions. In both cases, the firm is left without an employee, and we use the FEC associated with vacancies to apply the value of 0 in either case.

the investment rules. The time flow constraint is

$$\begin{aligned} 1 &\geq \tau_a + \tau_\theta, \\ \tau_a &\geq 0 \\ \tau_\theta &\geq 0. \end{aligned}$$

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 productivity 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, altough for reasons explained at the end of this section, the empirical work will emphasize the no-renegotiation case.

2.2.1 On-the-Job Search with Bertrand Competition

Whether or not we assume that firms engage in Bertrand competition, 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 possibilies.

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

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

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

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

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

where

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

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

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

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}.

2.2.2 On-the-Job Search with No Renegotiation

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

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

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

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

$$\begin{aligned} N_F(w,\tau_a,\tau_\theta;i,j) = y(i,j)(1-\tau_a-\tau_\theta) - w + \varphi_a(i,\tau_a)Q_F(i+1,j) \\ +\varphi_\theta(j,\tau_\theta)V_F(i,j+1) + \tilde{\delta}_a(i)Q_F(i-1,j) \\ +\tilde{\delta}_\theta(j)Q_F(i,j-1), \end{aligned}$$

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:

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

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. 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 ($E = 1 - U$, 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 $\phi \in (0, 1)$.[10] 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

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

[10] We have fixed $TFP = 1$ in the Cobb Douglas matching function due to the impossibility of identifying this parameter given the data available. The number of matches is unobserved, so that this essentially amounts to a normalization.

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

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

By a similar argument, the mass of matches involving an unemployed worker is $U/S \times M$, and the contact rate for unemployed searchers is

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

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

Turning to the firm's problem, let the flow cost of holding a vacancy be given by $\psi > 0$. The distribution of potential hires is determined by the steady state distributions of a among the unemployed and (a, θ) among the employed, which 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

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

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 is excluded so that the final sample comprises only the nationally representative cross-sectional sample, (2) the military sample is excluded, and (3) all females and high-school dropouts are excluded. A respondent who satisfies these criteria enters our sample after having completed 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.

The NLSY97 provides detailed retrospective data on the labor market histories and the wage profiles of each respondent. This retrospective data is included in the employment roster, which gives the start and end dates of each employment spell experienced by the respondent since the last interview, wage profiles and other characteristics of 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, including transitions between jobs.

While we make extensive use of the weekly data constructed retrospectively from the NLSY97 employment rosters for obtaining moments related to employment transition, our empirical analysis of wages uses information collected from respondents about current wages as of each interview date. This information is likely to have fewer measurement problems than wage information in the employment roster, which is collected as part of a set of retrospective questions about all current and previously held jobs since her previous interview.

In the context of the model, the duration of a job spell is indicative of the value of the match between worker and firm. Therefore, looking at job spells of different lengths provides information about how wages and training differ at different match values. Our decision to only use wage observations from interview dates suggests defining the length of a job spell as the number of annual wage observations rather than using length from the employment roster. Table 1 shows the percentage of job spells by the number of interview dates they span for each schooling level. This distribution closely mirrors the actual duration distribution of jobs obtained from the employment rosters, suggesting that the two approaches should yield similar conclusions. For high school graduates, the table shows that about 61 percent of all observed job spells cover no interview dates at all; while 23 percent spells span one interview date, 7 percent last long enough to span two interview dates and 9 percent span three or more. The proportion of job spells with longer durations, and therefore spanning more interview dates, increases by education level. For example, for individuals with a college degree, 15 percent of job spells span more than two interview dates.

Given the importance of schooling on the labor market environment faced by the agents in our model, we distinguish between three groups of individuals in our empirical analysis: (1) individuals who have a high-school degree, (2) individuals who have attended college but who do not have a college degree, and (3) Individuals with a college degree or more. In what follows, we refer to these three levels as low, medium and high education groups, respectively.

In addition to key labor market variables, NLSY97 contains a wealth of information about training, which in our model is the way workers and firms invest to build human capital. 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 and end dates of training periods by source of training.[11] 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, 13 and 13 percent for workers with low, medium and high education, respectively. This suggests a negative relationship between schooling and training.

We next discuss training in relation to employment and wage transitions in our sample. Table 4 provides detailed information about employment and wage transitions between interview dates. We distinguish between three types of employment-to-employment transitions that may occur between interview dates $t-1$ and t: (1) transitions that do not involve a change in employer, (2) transitions that involve a change in employer, with no intervening spell of non-employment between the two jobs, and (3) transitions that involve a change in employer, with an intervening spell of non-employment.[12] The transitions that involve a change in employer are usually referred to as job-to-job transitions in the literature and we follow the same definitions

[11] 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.

[12] Using the employment rosters, we determine that there was an intervening non-employment spell between two consecutive jobs, if 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.

in our discussion. Among workers who are employed in two successive interviews, the fraction of workers who change jobs decreases with education. For example, among high school graduates who remain employed at consecutive interview dates $t-1$ and t, 19 percent had a different employer, compared to only 12 percent of college graduates. Among workers who do change jobs, those with more education are more likely to do so without an intervening spell of non-employment.

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 of Table 4 shows the difference between log wages for employment-to-employment transitions between interview dates. Again, we distinguish between the three types of transitions described above. We observe that log wage difference ($logw_t - logw_{t-1}$) for job-to-job transitions increases by education level: the average log wage difference is 0.11, 0.15 and 0.20 for low, medium and high education groups, respectively. The differences by education are particularly large for job-to-job transitions with an intervening non-employment spell.

Finally, Table 5 shows average log wage difference between consecutive interview dates $t-1$ and t, broken down by whether the worker receives training at the job he held at $t-1$. An individual is considered to have received training if the training roster reports him to have been enrolled in a training program during a week that was (1) before interview date $t-1$ (2) while he was also employed at the job he held at time $t-1$. We see that for employment transitions that do not involve a change in employer, the average log wage difference does not change by whether the worker obtained any training in the past. On the other hand, for employment transitions that do entail an employer change (i.e. job-to-job transitions), the average log wage difference between t and $t-1$ for those individuals who obtained some form of training in the first job spell is smaller. It is also instructive to examine the wage differences for job-to-job transitions that involve an intervening non-employment spell. 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 non-employment, the average log wage difference is 0.09 if they they received training in the previous job and 0.15 if they did not.

These moments will serve as a basis for our estimation of the parameters in the model described above.

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 $j < K$ who spends a fraction τ_θ of her 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. Finally, we set the death shock to produce an average career length of 45 years, $\ell = 1/(45 \cdot 52) = 0.00043$.

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, with the variance of this measurement error estimated together with the other model parameters. 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.[13] 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 proportional 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 50 replications.

4.2.2 Demand-Side Parameter Estimator: Method 1

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

[13] We do introduce measurement error into the wage observations, but the variance of this error is fixed rather than being estimated with the other model parameters. Including some realistic measurement error allows us to consider higher moments of the wage distribution, while imposing a fixed variance prevents the estimator from having to separately identify yet another source of variation in observed wages. Following Bound *et. al*) (1994), we set the standard deviation of the measurement error on observed log wages at $\sigma_e = 0.15$

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 $\xi \in (0,1)$. 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 our 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

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

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

$$\begin{aligned} \psi &= \frac{\hat{\lambda}_U \hat{S}}{\nu} \hat{B} \\ \Rightarrow \nu &= \frac{\hat{\lambda}_U \hat{S}}{\psi} \hat{B}. \end{aligned}$$

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.[14]

[14]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.

4.2.3 Demand-Side Parameter Estimator: Method 2

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 ψ, 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 $\hat{b}a = 4.94a$, very close to the output of that worker at a firm with the median match quality, $(\exp(\hat{\mu}_\theta) \cdot a = 4.59a)$. For unemployed workers, an offer arrives at a rate of $\hat{\lambda}_u = .219$ or approximately once every five weeks. Workers with medium levels of general ability accept 23 percent of job offers, implying that the average unemployment spell lasts 20 weeks. Conversely, matches are exogenously dissolved at a rate of $\hat{\eta} = .0036$, or approximately once every five 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 13.1 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 9.2 percent. Together with the job finding rate and exogenous job separation rate, this equilibrium unemployment rate implies that approximately 15 percent of separations are endogenous. For employed workers, new offers arrive at rate $\hat{\lambda}_e = 0.094$, or approximately once every 11 weeks, 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 24 percent more human capital than high school graduates, on average, and those with at least a bachelor's degree begin with an additional 24 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 slower rate. The variance for the initial distribution of ability $\hat{\sigma}_a^2 = 0.043$, a bit less than half 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

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

with the analogous expression specified for the θ process. In the estimated model, $\hat{\delta}_a^0$ and $\hat{\delta}_\theta^0$ are very similar: $\hat{\delta}_a^0 0.021$, and $\hat{\delta}_\theta^0 = .016$. 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 -0.132, whereas $\hat{\delta}_\theta^1$ is 0.673. 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 difficult to change after labor market entry.

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 20 percent of her time training, we expect that this worker will be observed to be involved in training only one percent of the time. For a worker engaged in full-time training, we would expect to observe this training in the data only 35 percent of the time.

5.2 Model Fit

In this section, we compare the fit of the simulations from our estimated model to the corresponding moments in the data. We begin by comparing the model-predicted and observed wage distributions. Overall, the model-predicted wage distributions are close to the empirical ones. Conditioning on tenure and education, we construct histogram plots from the simulated and observed wages, shown in Figures 1, 2 and 3 for workers with 0-2, 3-5 and 6-8 years of tenure, respectively.

Next, we evaluate the ability of the model to replicate the training patterns in the data. In Table 7, we see that the observed and model-predicted proportion of individuals who have participated in at least one training spell during the time they are observed is 15 and 17 percent, respectively. Columns (2) and (3) of Table 7 further reveal that the model accurately captures the decreasing pattern of training with education, albeit with a small tendency to overstate the incidence of training for low-education workers: the proportion of individuals who get training at least once is 18, 13 and 13 percent for the low, medium and high education workers, respectively; whereas in the model simulations these moments are 20, 16 and 13 percent.

One of the distinguishing features of our model is our focus on within-job spell investment in human capital. This investment behavior, whether general or match-specific, impacts transition rates and wage processes within and across job spells. In the estimation, we match moments related to the joint distribution of wages between consecutive interview dates *within* a job spell as well as the joint distribution of wages between consecutive interview dates *between* different job spells. Table 8 displays how the model performs in generating some of these transition moments. Here, we limit our discussion to events that span only two consecutive interview dates, $t-1$ and t, and to workers who are employed at both dates. As described in the data section, we consider three possible events that may occur between $t-1$ and t: 1) no job change, 2) job-to-job transition with an intervening spell of non-employment, and 3) job-to-job transition with no intervening spell of non-employment.[15]

First, we consider the the proportion of job-to-job transitions, shown in Panel A of Table 8. In the data, the proportions of job-to-job transitions are 15, 12 and 10 percent for the low, medium and high-education groups, respectively. The corresponding model-predictions for these moments are 14, 16 and 17 percent, respectively. Hence, the model-predicted transition rates are reasonably close to the actual ones in the data for the lower education workers. However, the discrepancy between the model predictions and the data seems to increase with education.

Next, we examine the distribution of wage changes between consecutive interview dates, shown in Panels B and C of Table 8. We focus our discussion on the moments for high-school graduates, shown in Column (1), though the patterns for the other two education groups, shown in Columns (2) and (3), are similar. Panel B shows the average wage growth during these employment spells. In the data, for job-to-job transitions with an intervening non-employment spell the average log wage difference is 0.06. For job-to-job transitions with no intervening non-employment spell, it is 0.12. The corresponding moments in the model simulations are -0.14 and 0.19, respectively. These numbers show that the estimated model captures correctly the direction of the implications of an intervening non-employment spell, but that it performs poorly in matching the levels of these changes. Similarly, Panel C shows that the in the data, the proportion of job-to-job transitions that are associated with a wage decrease is 39 percent with an intervening unemployment, and only 30 percent without. The model matches the proportion for transitions with no intervening non-employment but overstates the fraction of negative changes for transitions that do include a spell of non-employment. Nevertheless, the model correctly captures the fact that transitions spanning a non-employment spell are more likely to have negative wage growth than those that do not.

To understand the intuition behind these moments, we recall that in the model, values of θ and any investment made in match-specific productivity do not carry over into future employment, whereas general human capital does. 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, the average wage gains from a job-to-job transition are high and the proportion of negative wage transitions is low (and would be zero without any measurement error). In contrast, for transitions that do involve an intervening non-employment spell, the worker loses the value from her accumulated match-specific human capital. Once she enters unemployment, she is willing to accept offers with a wider range of θ values, including values that are lower than the match quality at her recent job. Her willingness to accept such offers decreases the average wage change across such transitions and results in a higher fraction of wage changes that

[15] As described in the data section, we define a job-to-job transition to not involve a intervening non-employment spell if the time between the end of first job and beginning date of second job is a non-employment spell of 4 weeks or less. This allows us to distinguish between instantaneous turnovers from those that involve a period of search between consecutive jobs.

are negative. The wage losses predicted by this mechanism in the model are evidently larger than those observed in the data.

5.3 Identification

To better understand the source of identification in the model, we reestimate the model under two restrictions on human capital investment. First, we eliminate the possibility of investment in general human capital, and second, the possibility of match-specific investment. We then compare the baseline parameter estimates and model predictions with the ones obtained from the constrained estimations. The purpose of this comparison is to demonstrate what aspects of the data and the model help distinguish between general and match-specific human capital in the estimation.

First, we compare the model predictions and parameter estimates we obtain from the constrained estimation with no possibility of investment in general human capital to the ones we obtain from the baseline estimation. In the absence of general human capital accumulation, all wage growth within a job spell is attributed to growth in the worker's match-specific human capital, which must increase more quickly than in the baseline model to match the overall rate of wage growth. Because workers have more match-specific human capital, fewer potential offers would cause them to leave their current employers and we would therefore expect, holding all else constant, a decrease in the job-to-job transition rates. However, as shown in Table 9, our estimate of the job offer rate (λ_e) increases to 0.108 from a baseline value of 0.094, so the model still matches the empirically observed transition rate. A second consequence of the increase in match-specific human capital is that the the wage gains from job-to-job transitions tend to be smaller.[16] As a result, the restricted model is less able to match the wage growth across job-to-job transitions. This reveals an important aspect of identification between general and math-specific human capital in the model estimation: transition data alone is not sufficient to distinguish between the different types of human capital and we need transition date in conjunction with moments that pertain to the joint distribution of wages within and across job spells in order to isolate one form of investment from the other.

Next, we reestimate the model with only general training. This version of the model does a better job of matching the average wage growth within and across job spells. However, without match-specific training, the model is unable to capture 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

[16]Formally, θ is log-normally distributed so $E(\theta' - \theta | \theta' > \theta)$ is decreasing in θ.

features of the data.

The results from the constrained estimations discussed above demonstrate the empirical content of general and match-specific investments. As the results illustrate, general and match-specific investment have distinct implications for subsequent employment and wage transitions. These distinct implications can be summarized through plotting the life-cycle wage profiles obtained from each estimation. These graphs are displayed in Figures 4-6. 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, with the baseline in between. This order gets reversed as the worker accumulates more labor market experience. In other words, by Year 6-8, the average log wages that correspond to the estimated model with no investment in θ overtakes 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 estimated model with only general human capital investment, the benefits of any training undertaken by the worker can be carried over to other jobs. The resulting wage growth throughout the labor market career of a worker is consequently larger relative to the estimated model with only θ investment.

The opposite is true for the estimated model with no a investment: 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 estimated 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 type of training as well.

5.4 Policy Rules

5.4.1 Acceptable Job Offers

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 7. 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 $\theta^*(a)$ begins to increase again 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 10. At lower values of a, for values of θ just above $\theta^*(a)$, workers spend a full 15 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). As background to this discussion, we want to understand how much workers value each kind of training relative to simply receiving wages. To this end, Figure 8 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 provide 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-specific training is small, and that most of the value from match-specific training goes to the firm. However, the worker does seem to receive some benefit from the match-specific training, which is reflected in her willingness to trade off some wages for more match-specific training.

5.4.2 Training Policies

Having identified the trade-offs between wages and training, we next examine the three outcomes of the bargaining process: the two types of training and the wage. In Figures 9 to 11, we plot the amount of general and match-specific training and the wage 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 wages or 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 9, 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 15 percent of the worker's time at roughly the 75th 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.

Next, we look at the amount of general training provided to the workers, which we plot in Figure 10. At low values of a and θ just above the $\theta^*(a)$ cutoff, workers spend about 15 percent of their time engaged in general training. The amount of training decreases at higher values of either a or θ. General training decreases with a because general graining becomes less productive at higher values of a ($\hat{\delta}_a^1 < 0$). Meanwhile, the decrease in general training at higher θ seems to reflect the decrease in overall benefits flowing to the worker at higher values of θ. To understand this, we recall our earlier discussion where we showed that the benefits of general training flow largely to the worker. 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 lower compensation for the worker, she will choose to lose some of this compensation by receiving lower wages and some as a decrease in general training. Indeed, in Figure 11, we plot the fraction of worker's output that is paid in wages and we observe that it also decreases at higher values of θ. The decrease in both wages and general training is consistent with a decrease in the worker's overall bargaining position as θ increases.[17]

Given these policy rules, how much training do workers actually receive? Based on the simulations from the estimated model, Figure 12 plots 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 12 percent of their time in general training and 9 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.

5.5 Sources of Wage Growth

In this section, we examine the factors that drive wage growth in the model and provide a link between our structural approach and the well-known results from the

[17] The corresponding increase in the firm's bargaining position as θ increases provides an additional incentive for it to provide match-specific training to the worker.

literature on Mincer-type earnings functions. The model contains five 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 searching for a new job with a higher match quality. Third, workers can improve the quality of the match with their current employers by engaging in firm-specific training. Fourth, 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.[18] 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 $a \cdot \theta \cdot (1 - \tau_a - \tau_\theta)$. Additionally, it is useful 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 (net of depreciation). This allows us to formally write the wage w as the product of the five pieces described above:

$$w = a \cdot \theta_0 \cdot \theta_\tau \cdot (1 - \tau_a - \tau_\theta) \cdot (w/y)$$

or in logs,

$$\log(w) = \log(a) + \log(\theta_0) + \log(\theta_\tau) + \log(1 - \tau_a - \tau_\theta) + \log(w/y)$$

In Figure 13, we plot the evolution of each of these five 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, the two most important sources of wage growth are the development of match-specific human capital through search and through training. 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.

More formally, our structural model also allows us to interpret the coefficients of a standard wage regression in terms of the different source of wage growth discussed above. This analysis helps relate our results to the well-known results from the literature on earnings dynamics. As an example, we consider a simple Mincer wage regression of the form

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

[18] 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.

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

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

where ed_{ie} is a dummy variable indicating that person i has education level e (for each level of education except HS graduate), $years_{it}$ denotes the number of years in the labor force and $tenure_{it}$ the length of time with current employer. We first estimate this model on the actual NLSY data and then on the simulated data from the model. Results are shown in the first two lines of Table 10. 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 than the actual 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 at the start of the match iii) additional match quality accumulated through match-specific training, iv) time spent not training and v) wage as a fraction of output. Additionally, 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 her initial endowment (a_0) and the additional human capital she accumulates through training (a_τ). This defines six components of wages, which we denote $Y_{it}^k, k = 1, ..., 6$, allowing us to write

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

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 10, 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.

In addition to studying Mincer wage regressions on the entire sample, we can run these regressions separately on workers with different amounts of education to help understand differences in wage growth between these groups. Results from this exercise are shown in Table 11. The first panel shows the difference in the constant term. The model captures the higher wages of more educated workers and, as one would expect, most of this difference is captured simply by the higher larger amounts of general human capital that more educated workers have upon entering the labor force. In addition, a portion of the differences is explained by the ability of more highly educated workers to capture a larger share of the output, perhaps because they tend to produce more surplus above the fixed employment costs. These sources of positive association between wages and education are offset slightly by the fact that more educated workers accept offers at lower match values, which tends to lower their output and therefore their wages.

In the middle panel of Table 11, we show the effect of increasing labor market experience on wages for the three different education groups. In the data, high school graduates receive a larger increase in wages for each year in the labor market than do more educated workers. The model is unable to explain these differences; wages in the simulations increase about four percent for each additional year of labor market experience, regardless of education.

The final panel of Table 11 describes the returns to increasing tenure for each of the three education groups. Here, the model matches the data well. In both the data and the model, workers with more education have higher returns to tenure. Each year of additional employment is associated with a two percent larger increases in wages for college educated workers compared to those with only a high school diploma in the data, though just once percent more in the model. In the model, most of this difference is accounted for by more educated workers building more match-specific human capital from training during their job spell. Offsetting this effect is the fact that, with increasing tenure, workers receive a smaller fraction of their output as wages and this decrease happens more quickly for more educated workers.

6 Policy Analysis: The Minimum Wage and Investment

In competitive markets, minimum wages are expected to impact human capital formation in relatively immediate ways (Leighton and Mincer (1981), Hashimoto (1982)). In the case of general human capital, the returns to which acrue to the worker, the theory implies that the sum of the cost of investment and the wage paid to the worker should be equal to the worker's (marginal) productivity. Restrictions on the wage that must be paid to the worker act as a constraint on the amount of general human capital investment that the individual can undertake, which typically results in lower levels of investment early in the labor market career than would be efficient. Within our modeling framework, in which there exist two types of human capital and search frictions, these types of considerations still apply, but in a much more subtle manner.

One of the more interesting implications of our model regarding the impact of minimum wage laws is that, given the existence of search frictions and the possibility of improving the productivity of the match, the introduction of a minimum wage that exceeds the current flow productivity of the match need not lead to the termination of the employment contract. Models in which the productivity of the match is fixed produce the implication that all matches with flow productivities less than the value of the minimum wage will be terminated (Flinn (2006), Flinn and Mullins (2015)). The possibility of increasing the flow productivity means that the decision to keep the match alive involves a comparison of the expected profits of the match (to the firm) with its outside option of zero, while the worker must receive an expected value of continuing on the job at the binding minimum wage that exceeds the value of unemployed search under the new minimum wage. In general, not all matches with current flow net productivity $a\theta - \zeta < m$, where m is the newly imposed minimum wage, will be terminated. The likelihood of termination in such a case will be a function not only of net productivity, but also the mix of general and match-specific capital possessed by the worker. For a given shortfall in net productivity with respect to the minimum wage, firms will be more likely to continue the match when match-specific capital is greater. To eventually earn positive flow profits from the match, the employee will have to be likely to remain with the firm, and this is an increasing function of the current level of θ. Moreover, this consideration will make it more likely for the worker-firm pair to invest in match-specific human capital than general human capital. The extent of these effects will depend on the parameters characterizing the model.

As has been found in Flinn (2006) and Flinn and Mullins (2015), the impact of the minimum wage is likely to vary significantly depending on whether we use the partial or general equilibrium version of the model. In the partial equilibrium version, contact rates (λ_U and λ_E) are assumed to be fixed. In the general equilibrum version of the model, these contact rates are a function of the job vacancy creation decisions of firms. As minimum wages increase, the share of the match surplus accruing to firms shrinks, which decreases firms' incentives to create vacancies. This corresponds, roughly, to a shift downward in the demand function, and exacerbates the minimum wage's negative impact on the employment rate. We begin our analysis with the partial equilibrium version of the model.

6.1 Minimum Wage in Partial Equilibrium

In order to study the effect of a minimum wage in the context of our model, we use the estimates of the model parameters and then solve the model imposing a minimum wage of $15 per hour (in 2014 dollars, corresponding to $10.17 in the 1994 dollars we use in our analysis). As expected, we find that imposing a minimum wage increases the unemployment rate by rendering low quality matches unprofitable for firms.[19] Figure 14 shows the minimal acceptable match quality draw in the baseline model and under the minimum wage. In our simulations, imposing the minimum wage raises the equilibrium unemployment rate from 9.2 to 9.6 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 15 compares the amount of training that workers receive in the baseline model and under the minimum wage. With a minimum wage in place, employed workers spend 3-8 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 with most of the impact concentrated at the bottom of the distribution.

While the decreased training time ultimately reduces workers' accumulation of skill, the flexibility of employers to adjust the amount of worker training reduces the impact of the minimum wage on unemployment. In particular, some matches that were feasible in the baseline model become infeasible if employers are forced to maintain the same level of training while also paying the higher minimum wage. However, because employers in our model are able to reduce the amount of training they provide, they continue to be able to form matches with some of these workers by giving them less training when the minimum wage is raised. As described above, imposing a $15 minimum wage raises the unemployment rate in our model from 9.2 to 9.6 percent while simultaneously reducing the amount of employer-provided training. Our model predicts that if employers were not able to adjust the amount of training, the unemployment rate would rise an additional 0.3 percentage points to 9.9 percent. Our model therefore shows how endogenously determined investment in human capital can mitigate the effect of minimum wage laws on unemployment.

Conditional on employment, workers receive higher wages with a minimum wage in place than they do in the baseline model. Average wages are eight percent higher

[19] Recall that our model does not include a labor force participation decision. An increase in unemployment translates into a decrease in employment, since these rates sum to one.

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 higher wages. 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, the selection of higher ability workers into employment and the increase in the workers' wages as a fraction of total output contribute the most to the the total increase in wages.

Looking in more detail at the effect of the minimum wage on the distribution of wages, Figure 16 shows how various percentiles of the wage distribution evolve with labor market experience, with and without the minimum wage. As expected, most of the of effect of the minimum wage occurs at the bottom of the distribution. For workers entering the labor market, the minimum wage raises the first percentile of the wage distribution by 50 percent and this difference persists even as workers gain more experience. At the tenth percentile, wages start out 20 percent higher for new workers but the effect fades after workers have been working for about four years. The effect on workers further up the distribution is insignificant.

Overall, the welfare effects from the minimum wage 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. Workers at the very lowest value of general ability experience a welfare loss of half a percent as they are most impacted by the lower job-finding rates. For high school graduates on average, the welfare loss is just 0.1 percent and it is even smaller for those with more education. These relatively small effects are due to the minimum wage being set at a relatively low level taking into account the baseline marginal wage distribution.

6.2 Minimum Wage in General Equilibrium

We now consider the impact of minimum wage increases in a simple general equilibrium version of the search and matching model in which firms' vacancy creation decisions and the measure of unemployed and employed searchers determine the contact rates λ_U and λ_E. In order to move to the general equilibrium framework, we need to define the steady-state distribution of workers, which is utilized in computing the expected return to a filled vacancy when solving for firms' vacancy creation decisions. We describe the computation of this steady-state distribution in Appendix A.

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-Douglas parameter in the aggregate matching function, ϕ. Following the discussion in Sections 4.2.2 and 4.2.3, 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 a negative estimate of ϕ. Alternatively, as described in Section 4.2.3, we fix the Cobb-Douglas parameter at $\phi = 0.5$ and back out the vacancy cost. This procedure produces an estimate of the vacancy cost $\hat{\psi} = 182$.

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. However, 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 that the impact on transition rates is small: the decrease in job-finding rates is just 0.4 percent and the effects on unemployment, human capital and welfare are little changed from the partial equilibrium model.

The general equilibrium effects become more significant at higher values of the minimum wages. As an illustration, we consider minimum wages equal to $20 and $25.[20] With the minimum wage set at $20, the job offer rate falls by one and a half percent and the steady state unemployment rate increases an additional 1.4 percentage points to 11 percent. In addition, because low-ability workers are employed less and receive less training, the amount of general human capital in the economy falls 3.8 percent for high school graduates and an average of 1.7 percent across all workers compared to the baseline. Together, the increased unemployment rate and loss of human capital produce net average welfare loss of 1 percent for high school graduates and 0.3 percent for those with some college. As we increase the minimum wage further to $25, we get even larger effects with the job offer rate declining to 4.7 percent below its baseline value and the steady state unemployment rate climbing to 17.2 percent. General human capital declines an average of 13 percent relative to the baseline for high school graduates and 6.4 percent for all workers. Finally, the welfare losses continue to increase for workers at the bottom of the distribution. For high school graduates, the average welfare loss is 6 percent, while for workers with some college, average welfare decreases by 2.1 percent.

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

[20]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 respectively.

and investments essentially are unobservable. In this we are aided by having access 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 in 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 (i, j), which signifies a_i and θ_j. Let the probability of (i, j) be denoted by $\pi(i, j)$. The steady state marginal distributions of a and θ are given by $\pi_a(i)$ and $\pi_\theta(j)$, respectively. There are M possible values of a, and K possible values of θ (for employed agents), with

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

From our estimates, we have determined the minimal value of θ that is acceptable when an agent characterized by a_i is in the unemployment state, which we denote by $r^*(i) + 1$ (that is, $\theta_{r^*(i)}$ is the maximal unacceptable match value to an individual with ability level a_i). We define the indicator variable

$$d(i,j) = \begin{cases} 1 & \text{if} \quad k > r^*(i) \\ 0 & \text{if} \quad k \leq r^*(i) \end{cases}, \; i = 1, ..., M; j = 1, ..., K.$$

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

Unemployed agents of type a_i occupy the state $(i, 0)$, with the probability of an unemployed individual if type i being given by $\pi(i, 0)$. Then we have

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

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

We assume a constant death rate $\ell > 0$ in the population. Since we have assumed that population size is constant, the death rate is equal to the birth rate. When a new individual is born, she is assigned a value of schooling according to the distribution of schooling types observed in the data, and an initial value of a is chosen from the schooling-specific initial distribution of a. The individual then enters the labor market as an unemployed individual with this general ability type.

We begin by considering movements in the probability of being unemployed for an agent of ability type i. We have that the change in the proportion of unemployed of ability type i is given by

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

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 j for which $d(i,j)=1$) are destroyed times the probability that type i individuals are employed. The second and third lines represent inflows in to the unemployment state that result from depreciation and appreciation of general skills that are associated with "endogenous" quits into unemployment. In these cases, an individual employed with skills $(i+1,j)$ or $(i-1,j)$ will quit into unemployment if they would not accept employment at (i,j). The fourth line represents inflows into unemployment from individuals with skills (i,j) when their match skill level depreciates to $j-1$ and $(i,j-1)$ is not an acceptable match. The penultimate 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 $(i,0)$, and the probability of receiving an acceptable job offer. The last line consists of two terms, the first of with is the rate of births into this population multiplied by the likelihood that they begin life with general ability level i, with $\hat{\pi}(i)$ denoting the probability of beginning ones labor market career with general ability level i. The second term is the death rate times the probability of being in state $(i,0)$.

Then, in the steady state,

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

for $i = 1, ..., M$.

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

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

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

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

for $i = 1, ..., M,\ j = 1, ..., K.$

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

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

where $\pi(i, \cdot) = (\pi(i, 0)\ \pi(i, 1)\ ...\ \pi(i, K))'$. Some elements of this vector are identically equally to 0, those for which $d(i, j) = 0$. Let the number of nonzero elements of Π be denoted $N(\Pi)$, where $N(\Pi) \leq M \times (K + 1)$. 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.

Hashimotoe, M. (1982) Minimum Wage Effects on Training on the Job,” *American Economic Review* 72: 1070-87.

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.

Heckman, J., L. Lochner, and P. Todd. (2006) “Earnings Functions, Rates of Return and Treatment Effects: The Mincer Equation and Beyond,” *Handbook of Economics of Education*, Volume 1.

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.

Leighton, L., and J. Mincer. ”Effect of Minimum Wages on Human Capital Formation,” in *The Economy of Legal Minimum Wages*, edited by Simon Rottenberg, Washington, D.C.: American Enterprise Institute, 1981.

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	HS	Some 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: Employment-to-Employment (EE) transitions btw $t-1$ and t			
% of EE transitions *with no* job change	81 %	85 %	88%
% of EE transitions *with* job change (job-to-job transitions)	19 %	15 %	12%
..... % of job-to-job transitions *with* non-employment btw $t-1$ and t	4 %	3%	2%
..... % of job-to-job transitions *with no* non-employment btw $t-1$ and t	15%	12%	10%
Panel B: Wage Growth btw $t-1$ and t			
EE transitions *with no* job change	0.08	0.08	0.09
EE transitions *with* job change (job-to-job transitions)	0.11	0.15	0.20
..... job-to-job transitions *with* non-employment btw $t-1$ and t	0.06	0.06	0.23
..... job-to-job transitions *with no* non-employment btw $t-1$ and t	0.12	0.17	0.20
Panel C: % of Negative Wage Growth btw $t-1$ and t			
EE transitions *with no* job change	17 %	19 %	23%
EE transitions *with* job change	32 %	32 %	28%
..... job-to-job transitions *with* non-employment btw $t-1$ and t	40 %	47%	27%
..... job-to-job transitions *with no* non-employment btw $t-1$ and t	30%	28%	28%

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

	$logw_t - logw_{t-1}$	
	No Training	Got Training
transitions with no job change	0.08	0.08
transitions with job change (job-to-job transitions)	0.14	0.10
job-to-job transitions with non-employment spell btw $t-1$ and t	0.08	0.15
job-to-job transitions with no non-employment spell btw $t-1$ and t	0.15	0.09

Table 6: Parameter Estimates

PARAMETERS FOR EMPLOYMENT TRANSITIONS			
flow value of unemployment	b	4.94	(0.067)
job offer rate - unemployed	λ_u	0.219	(.007)
job offer rate - employed	λ_e	.094	(.004)
exogenous job separation rate	η	.0036	(.0001)
PARAMETERS OF INVESTMENT FUNCTIONS			
General ability investment TFP	δ_a^0	.0209	(.0002)
Firm-specific investment TFP	δ_θ^0	.0163	(.0003)
State-dependence of general ability investment	δ_a^1	-.132	(.009)
State-dependence of firm-specific investment	δ_θ^1	.673	(.008)
Curvature of general ability investment	δ_a^2	.260	(.007)
Curvature of firm-specific investment	δ_θ^2	.432	(.013)
Rate of decrease in general ability	$\tilde{\varphi}_a^-$	.0008	(.00004)
Rate of decrease in match quality	$\tilde{\varphi}_\theta^-$	.0123	(.001)
PARAMETERS OF INITIAL ABILITY DISTRIBUTIONS			
Mean of initial general ability - High School	$\mu_a(e_i = 1)$	0.92	(.01)
Mean of initial general ability - Some College	$\mu_a(e_i = 2)$	1.16	(.01)
Mean of initial general ability - BA or higher	$\mu_a(e_i = 3)$	1.40	(.03)
Variance of initial general ability	σ_a	.208	(.008)
PARAMETERS OF JOB OFFERS			
Mean of match quality distribution	μ_θ	1.52	(.01)
Variance of match quality distribution	σ_θ	.315	(.002)
Employment cost	ζ	4.51	(.136)
PARAMETERS GOVERNING TRAINING OBSERVATION			
Intercept for training observation	β_0	-3.02	(.02)
Coefficient on τ for training observation	β_a	2.64	(.03)

Figure 1: **Model Fit: Log Wage Distribution - Year 0-2**

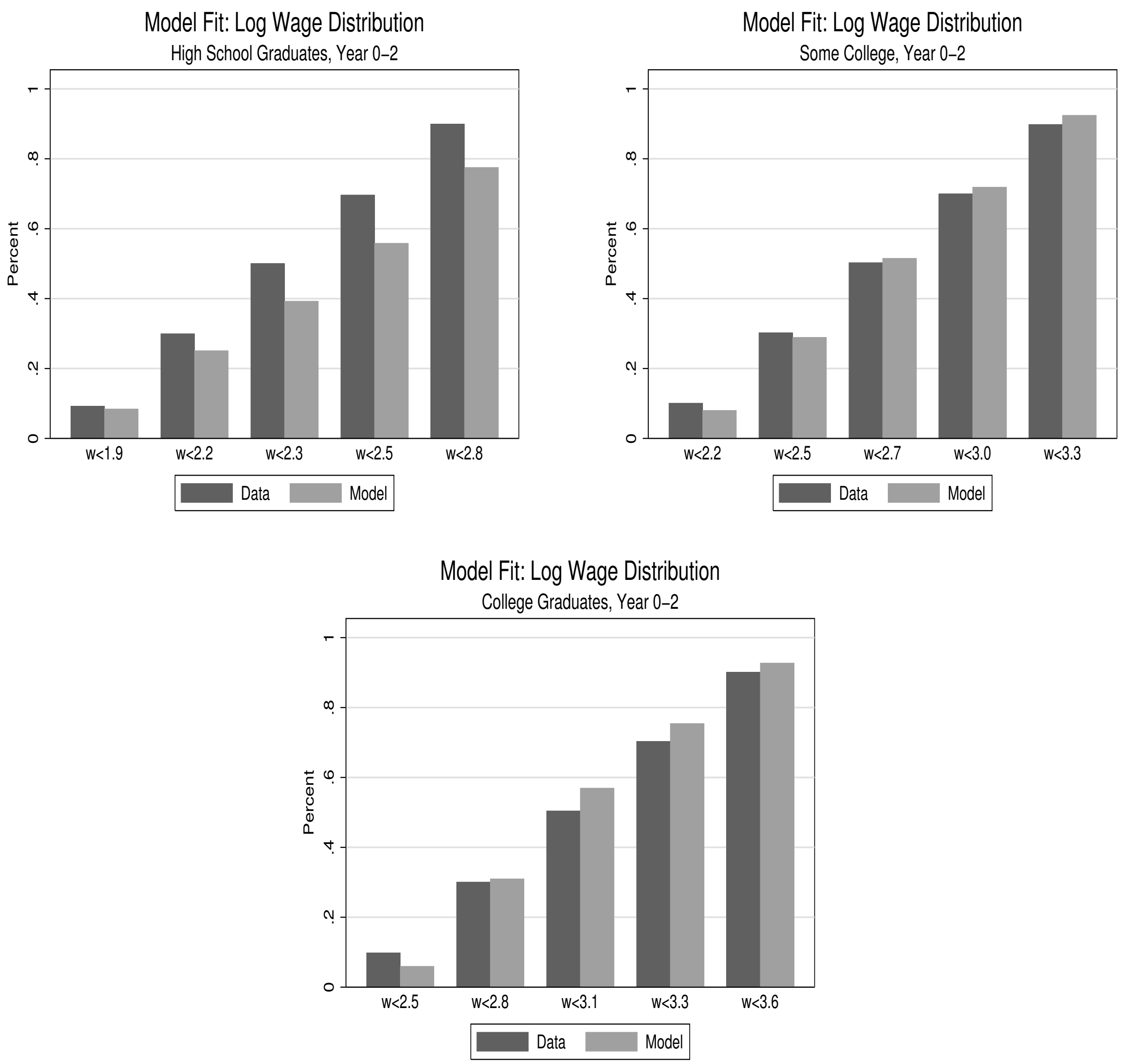

Figure 2: **Model Fit: Log Wage Distribution - Year 3-5**

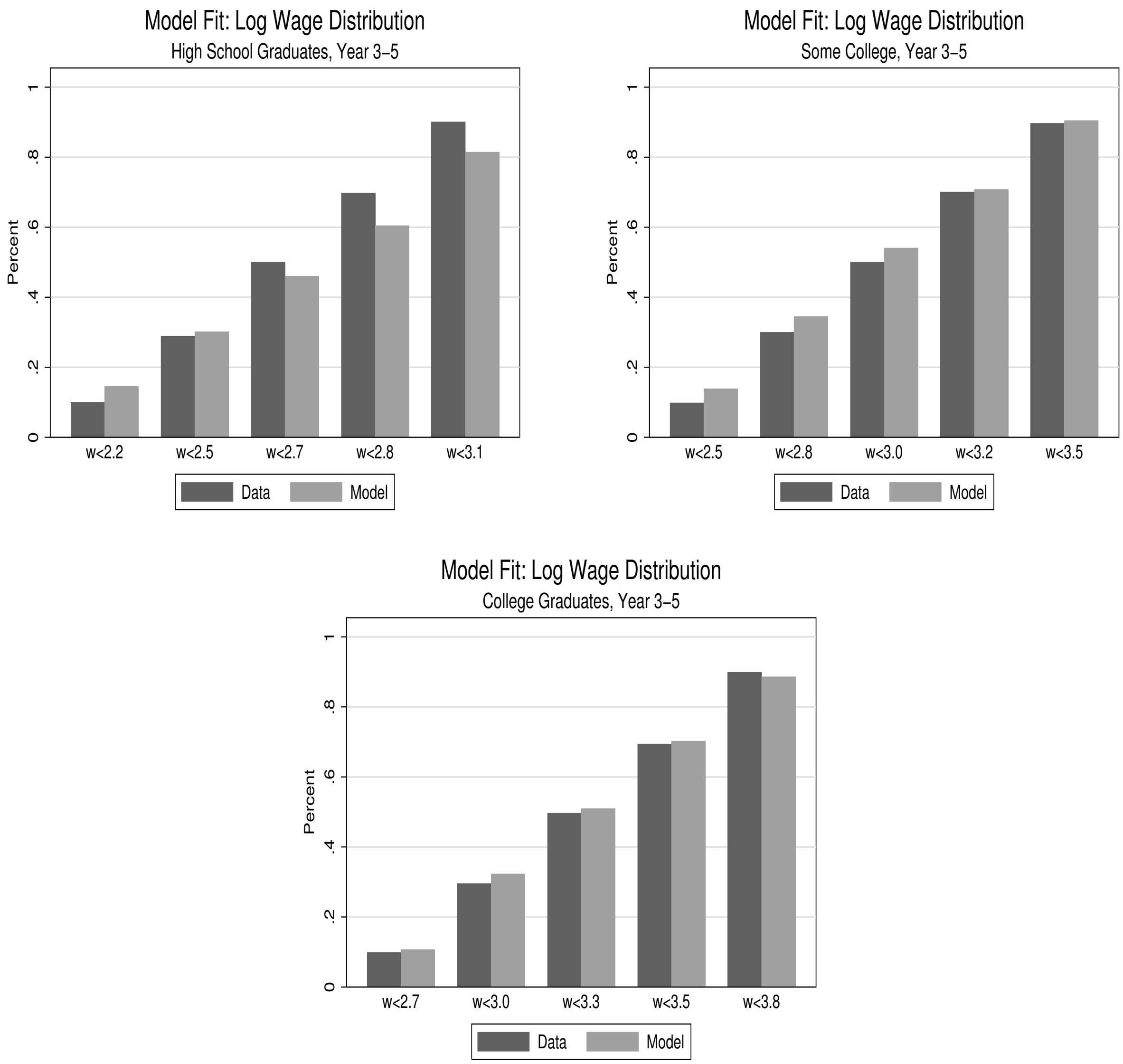

Figure 3: **Model Fit: Log Wage Distribution - Year 6-8**

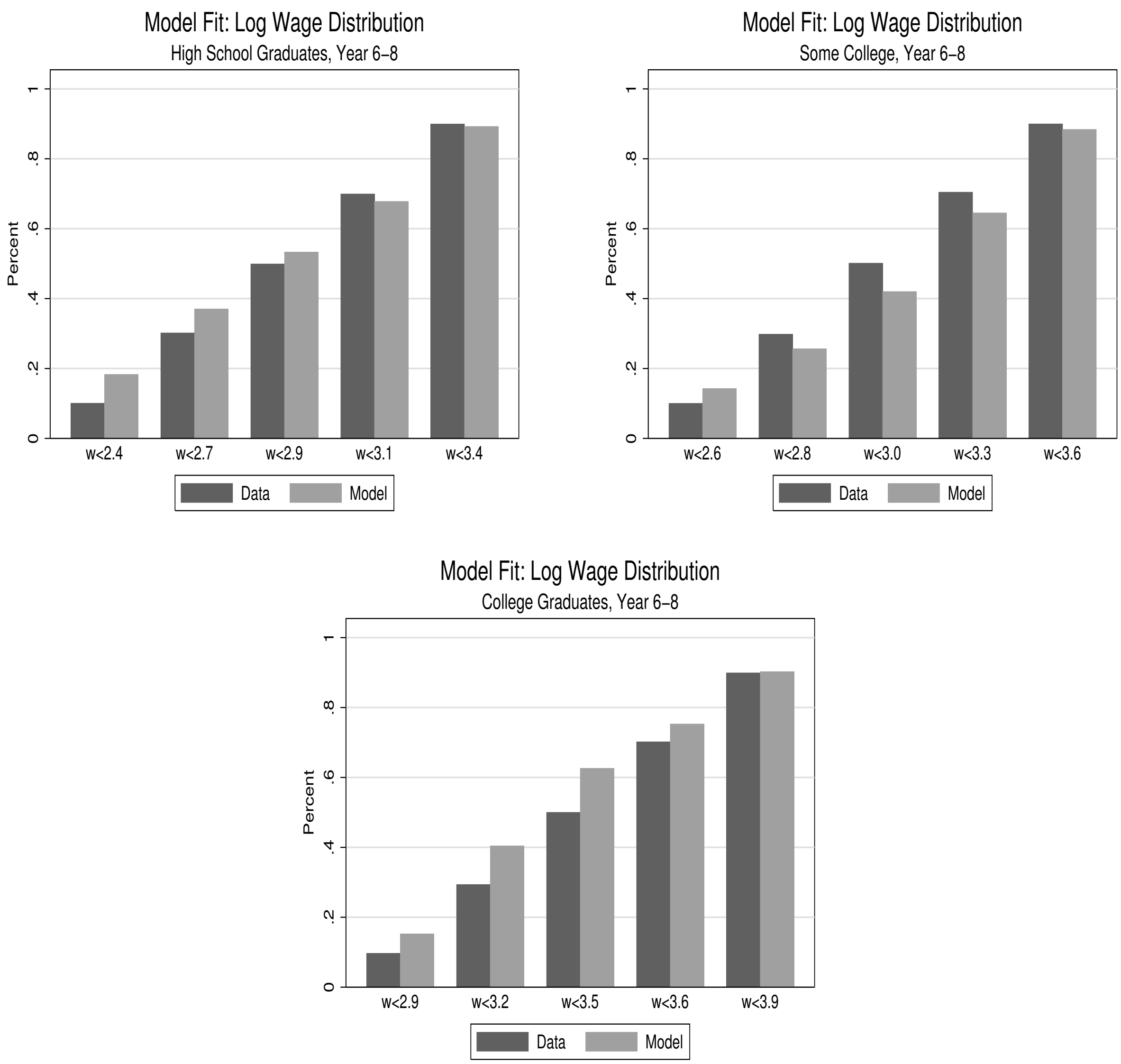

Table 7: **Model Fit: Incidence of Training**

	All	HS	Some College	College or More
% who got training at least once				
.... Data	15%	18 %	13 %	13%
.... Model	17%	20 %	16 %	13%
% who got training at the start of job spell				
.... Data	6%	10 %	5 %	3%
.... Model	5%	6 %	4 %	4%

Table 8: **Model Fit: Annual Labor Turnover Rates and Wage Growth**

	HS	Some College	College or More
Panel A: Employment-to-Employment (EE) transitions btw $t-1$ and t			
% of EE transitions *with no* job change			
.... Data	81 %	85 %	88%
.... Model	77%	75 %	74%
% of EE transitions *with* job change (job-to-job transitions)			
.... Data	19 %	15 %	12%
.... Model	23 %	25 %	26%
% of job-to-job transitions *with* non-employment btw $t-1$ and t			
.... Data	4 %	3%	2%
.... Model	8 %	8%	9%
% of job-to-job transitions *with no* non-employment btw $t-1$ and t			
.... Data	15%	12%	10%
.... Model	14%	16%	17%
Panel B: Wage Growth btw $t-1$ and t			
EE transitions *with no* job change			
.... Data	0.08	0.08	0.09
.... Model	0.08	0.08	0.08
EE transitions *with* job change			
.... Data	0.11	0.15	0.20
.... Model	0.07	0.09	0.10
job-to-job transitions *with* non-employment btw $t-1$ and t			
.... Data	0.06	0.06	0.23
.... Model	-0.14	-0.08	-0.05
job-to-job transitions *with no* non-employment btw $t-1$ and t			
.... Data	0.12	0.17	0.20
.... Model	0.19	0.18	0.18
Panel C: % of Negative Wage Growth btw $t-1$ and t			
EE transitions *with no* job change			
.... Data	17 %	19 %	23%
.... Model	38 %	38 %	38%
EE transitions *with* job change			
.... Data	32%	32%	28%
.... Model	41%	37%	36%
job-to-job transitions *with* non-employment btw $t-1$ and t			
.... Data	40 %	47%	27%
.... Model	63 %	58%	55%
job-to-job transitions *with no* non-employment btw $t-1$ and t			
.... Data	30%	28%	28%
.... Model	27%	26%	26%

Table 9: **Parameter Estimates: Baseline vs. Constrained Estimations with No a/No θ**

		Baseline	No a	No θ
PARAMETERS FOR EMPLOYMENT TRANSITIONS				
flow value of unemployment	b	4.94	4.06	5.57
job offer rate - unemployed	λ_u	0.219	0.127	0.180
job offer rate - employed	λ_e	.094	0.108	0.045
exogenous job separation rate	η	.0036	0.0031	.0038
PARAMETERS OF INVESTMENT FUNCTIONS				
General ability investment TFP	δ_a^0	.021	0	0.024
Firm-specific investment TFP	δ_θ^0	.016	0.022	0
State-dependence of general ability investment	δ_a^1	-.132	-	0.024
State-dependence of firm-specific investment	δ_θ^1	.673	0.678	-
Curvature of general ability investment	δ_a^2	.260	-	0.065
Curvature of firm-specific investment	δ_θ^2	.432	0.548	-
Rate of decrease in general ability	$\tilde{\varphi}_a^-$	.001	0	0.0002
Rate of decrease in match quality	$\tilde{\varphi}_\theta^-$	.012	0.019	0
PARAMETERS OF INITIAL ABILITY DISTRIBUTIONS				
Mean of initial general ability - High School	$\mu_a(e_i = 1)$	0.92	0.56	1.14
Mean of initial general ability - Some College	$\mu_a(e_i = 2)$	1.16	0.94	1.35
Mean of initial general ability - BA or higher	$\mu_a(e_i = 3)$	1.40	1.25	1.56
Variance of initial general ability	σ_a	0.21	0.29	0.16
PARAMETERS OF JOB OFFERS				
Mean of match quality distribution	μ_θ	1.52	1.60	1.50
Variance of match quality distribution	σ_θ	0.31	0.35	0.24
Employment cost	ζ	4.50	3.79	8.07
PARAMETERS GOVERNING TRAINING OBSERVATION				
Intercept for training observation	β_0	-3.02	-3.07	-2.59
Coefficient on τ for training observation	β_a	2.64	2.36	3.21

Simulations: Baseline vs. No a/No θ

Figure 4: **Avg. Log Wages by Tenure**

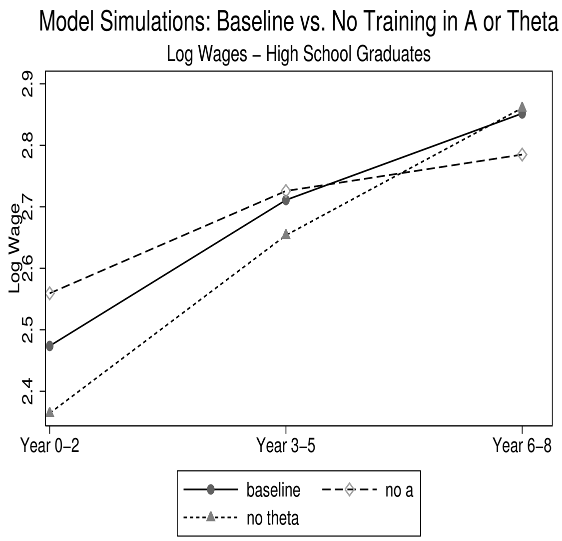

Figure 5: **Avg. Log Wages by Tenure**

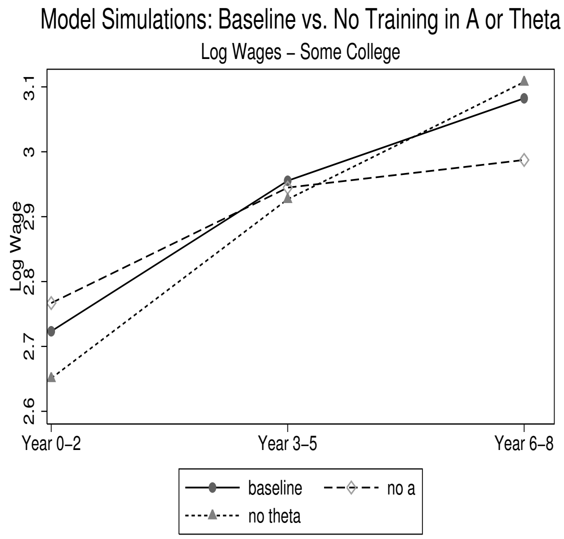

Figure 6: **Avg. Log Wages by Tenure**

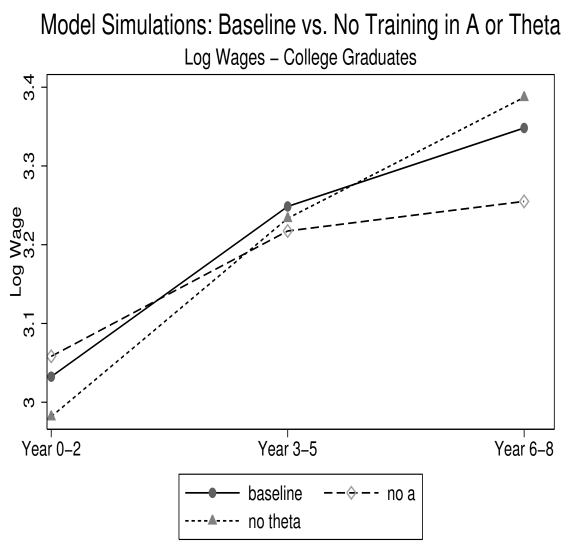

Figure 7: Minimum Acceptable Match-Quality

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

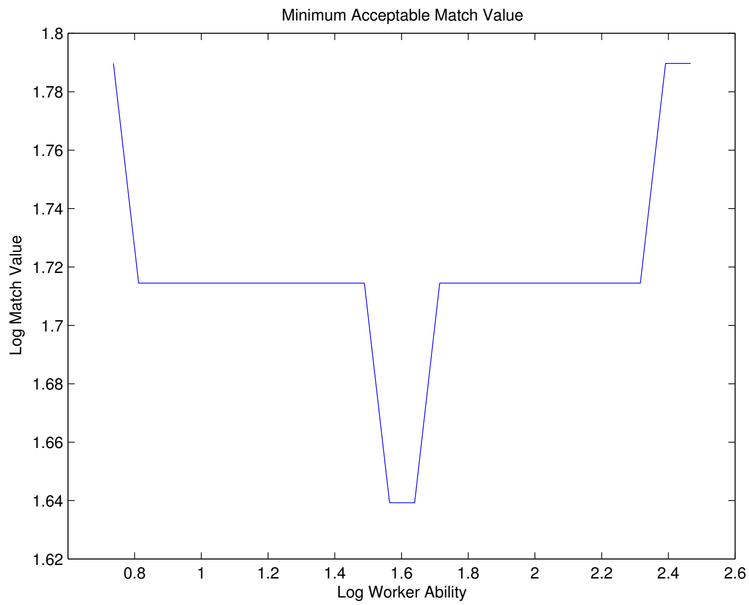

Figure 8: Trade-Off between Wages and Training in Bargaining Problem

This figure shows combinations of general training (τ_a), match specific training (τ_θ) and wages that solve the bargaining problem between the worker and the firm at the median values of a and θ. Lines on the graph show combinations of τ_a and τ_θ along which the negotiated wage remains constant. The blue dot shows the surplus maximizing combination, which is the model solution for τ_a and τ_θ.

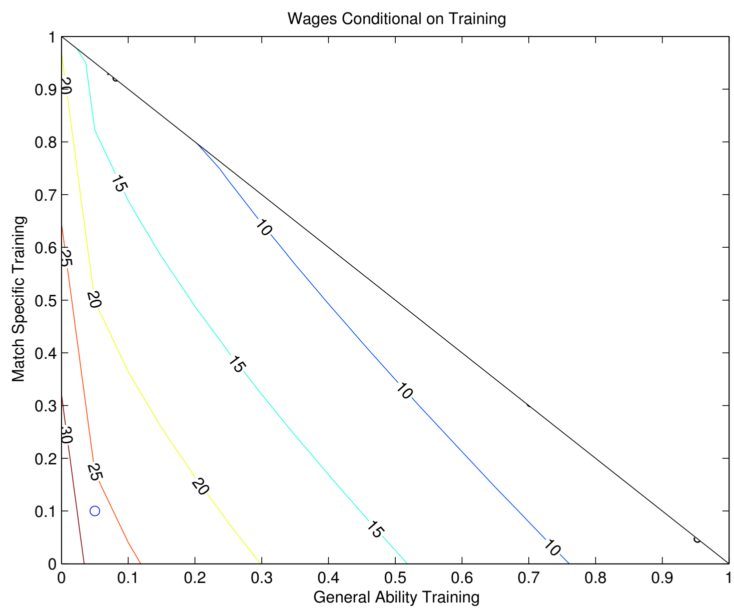

Figure 9: Match-Specific Training

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.

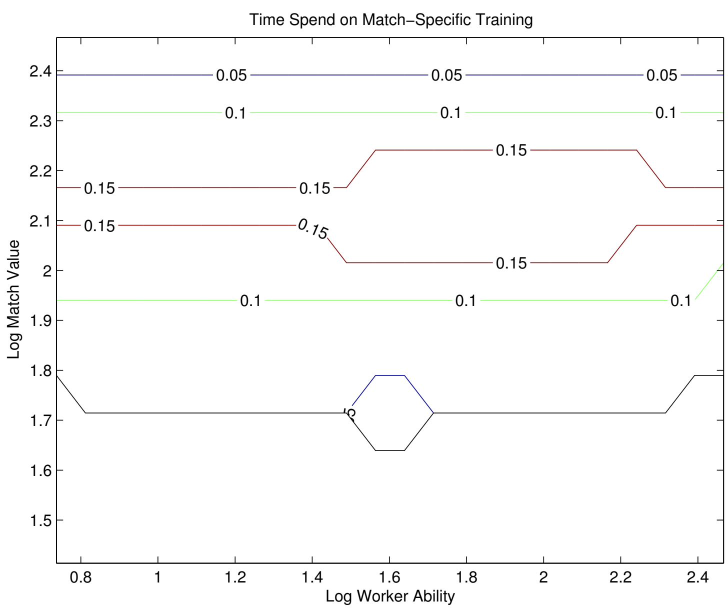

Figure 10: General Training

This figure shows the amount of general training that workers 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.

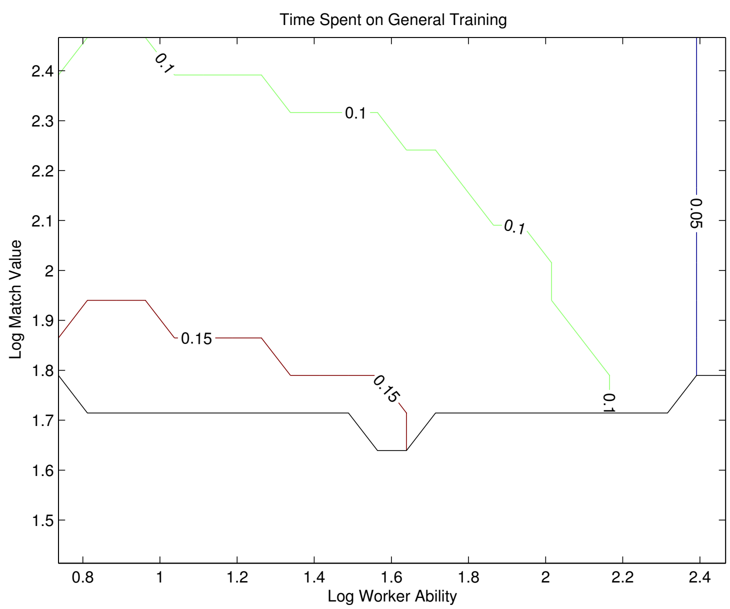

Figure 11: Wage as Fraction of Output

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.

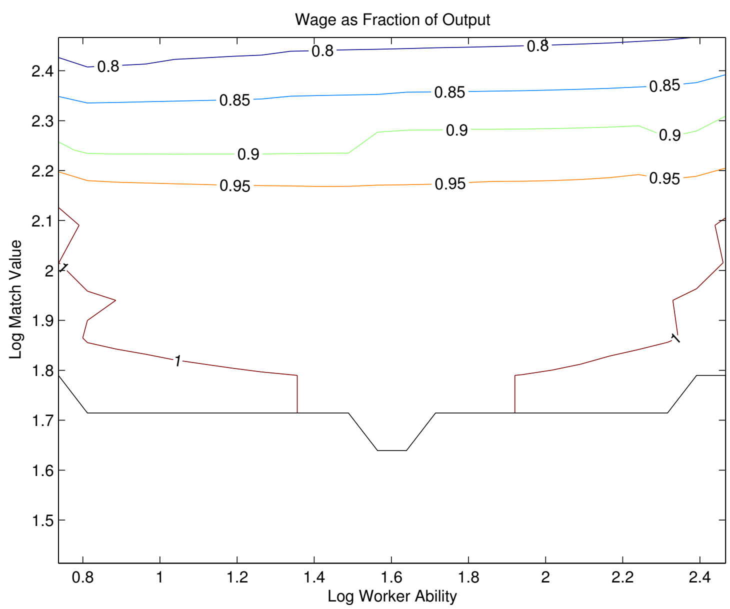

Figure 12: Training by Years in Labor Market

This figure shows the average fraction of their time that workers spend on general and match-specific training in the model simulation as a function of the number of years in the labor market.

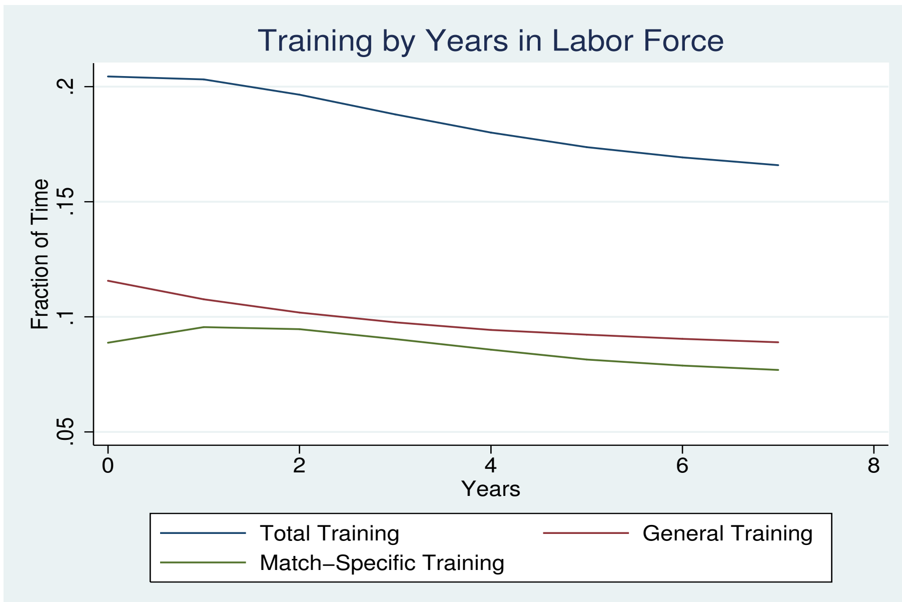

Figure 13: Sources of Wage Growth by Years in Labor Market

This figure shows the average of the log of general human capital, the amount of match-specific capital accumulated through search and through training, 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 function of the number of years in the labor market. In the absence of employment costs, these five components would add up to the total log wage, which is also shown.

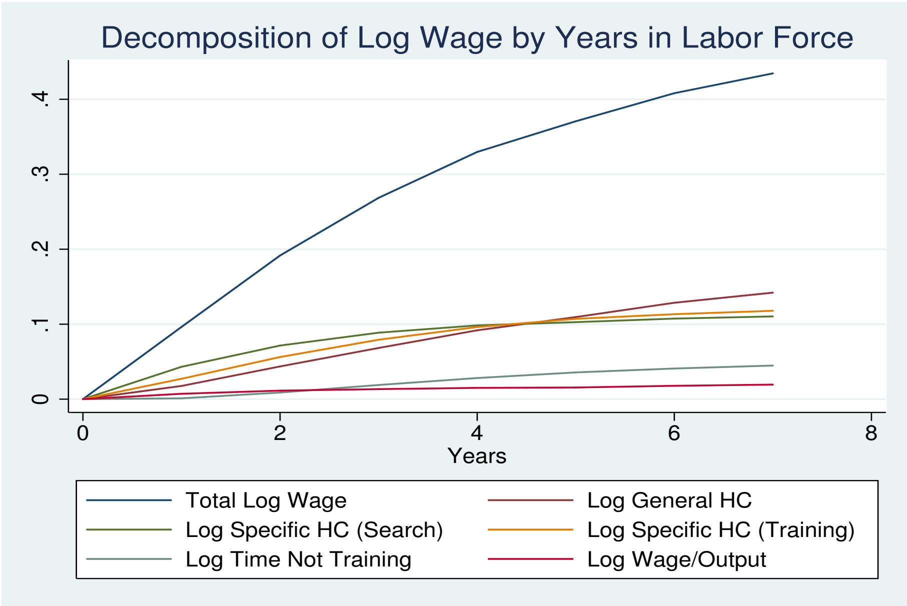

Table 10: **Mincer Regressions**

	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	2.401	0.260	0.558	0.042	0.058
starting a	0.943	0.207	0.454	0.001	-0.001
a from training	0.022	-0.001	-0.010	0.028	0.003
θ from search	2.023	-0.004	-0.007	0.005	0.010
θ from training	0.026	0.004	0.005	-0.001	0.038
hours worked	-0.241	0.006	0.013	0.002	0.014
wage/output	-0.371	0.052	0.104	0.007	-0.006

Table 11: **Mincer Regressions by Education**

	Constant			Years in LF			Tenure		
	HS	College	BA	HS	College	BA	HS	College	BA
DATA									
log wage	2.305	2.667	2.951	0.069	0.039	0.044	0.033	0.052	0.053
	(0.011)	(0.015)	(0.015)	(0.002)	(0.004)	(0.006)	(0.003)	(0.006)	(0.007)
MODEL									
log wage	2.408	2.658	2.947	0.042	0.042	0.044	0.056	0.061	0.064
starting a	0.944	1.150	1.400	0.001	0.001	0.002	-0.005	-0.001	-0.003
a from training	0.018	0.019	0.017	0.029	0.028	0.027	0.003	0.003	0.002
θ from search	2.032	2.015	2.000	0.003	0.006	0.009	0.010	0.010	0.012
θ from training	0.034	0.024	0.017	-0.001	-0.001	-0.001	0.035	0.042	0.048
hours worked	-0.239	-0.236	-0.232	0.001	0.002	0.002	0.013	0.014	0.016
wage/output	-0.380	-0.314	-0.252	0.008	0.006	0.004	-0.005	-0.007	-0.011

Figure 14: 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.

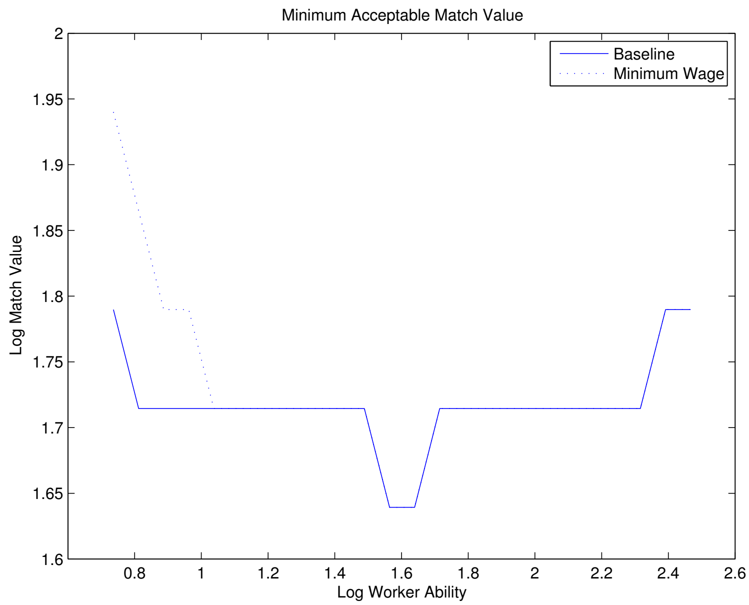

Figure 15: Training by Years in Labor Market with Minimum Wage

This figure shows the average fraction of their time that workers spend on general and match specific training in the model simulation as a function 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.

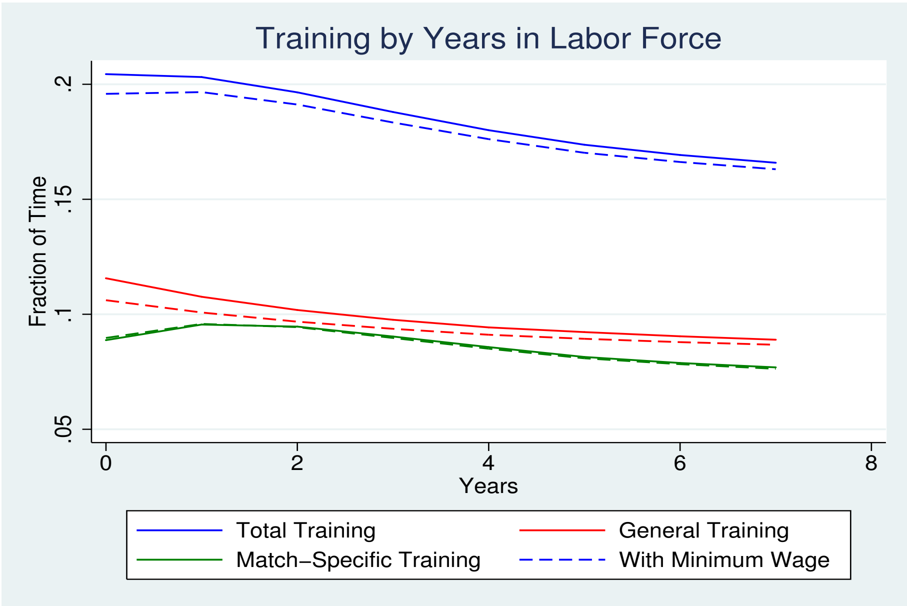

Figure 16: Log Wage Distribution by Years in Labor Market with Minimum Wage

This figure shows the distribution of log wages in the model simulation as a function of the number of years in the labor market. The solid lines show percentiles of the distribution in the baseline model. The corresponding dashed lines show the same percentiles when we impose a minimum wage.