HUMAN CAPITAL ACCUMULATION AND LABOUR MARKET EQUILIBRIUM

Ken Burdett\(^1\), Carlos Carrillo-Tudela\(^2\), and Melvyn G. Coles\(^3\)

June 2008

\(^1\)Department of Economic, University of Pennsylvania, Philadelphia, PA, US.
\(^2\)Department of Economics, University of Leicester, Leicester, Leicestershire, UK.
\(^3\)Department of Economics, University of Essex, Colchester, Essex, UK.
Abstract

The objective of this paper is to analyse an equilibrium search model with on-the-job search and human capital accumulation. In our model wages are disperse because firms pay workers of the same productivity different wages and workers of different productivities earn different wages. The wage distribution reflects the interaction between sorting dynamics as workers quit for better paid employment, worker productivity grows through on the job learning, and the distribution of initial productivities. We obtain a wage density with similar properties as the one observed in the data and provide a simple decomposition of wage variation into its constituent parts.

Keywords: Search, wage dispersion, human capital accumulation.

JEL: J24, J42, J64.
1 Introduction

We investigate individual wage dynamics in the context of an equilibrium labour market model where workers accumulate human capital while working. It will be shown the analysis leads to new insights into two important areas in labour economics - the nature of equilibrium in search markets, and the study of on-the-job human capital accumulation. Indeed, the work reported here can be perceived as a tentative first step at a more unified approach to these two areas.

Although there is no free lunch, it is accepted by many that individuals accumulate human capital freely by working. Typists become better typists while working as typists, economists become more productive by doing economics, etc. This seems both an important and intuitive idea. A related idea now common among labour economists is that human capital can be dichotomized into general human capital and firm specific human capital. A worker who enjoys an increase in general human capital becomes more productive at all jobs, whereas accumulating firm specific human capital implies a worker is only more productive at that firm. Workers who change job, or those who are laid off, lose their firm specific human capital but keep their general human capital. Putting the above two ideas together, plus assuming a worker’s wage is an increasing function of both his/her general and specific human capital, leads to at least the rudiment of a theory of how the wages of workers evolve through their working lives.

There is now a large literature investigating equilibrium wage formation and labour turnover in a market where firms post wages and workers (both unemployed and employed) search for better job opportunities. A large proportion of such studies have utilized the Burdett and Mortensen (1998) framework (henceforth, B/M). The B/M framework yields a non-degenerate equilibrium distribution of wage offers by firms, where on-the-job search implies workers change employer whenever better job opportunities are discovered. Even though B/M assume no human capital accumulation, they still find that worker wages are positively correlated with both experience and tenure. The former correlation occurs as it takes time to find well paid jobs, and so high wages are correlated with time, and thus experience, in the labour market. The latter correlation occurs as employees of firms that pay high wages find it difficult to find even better paying jobs.

In the present study we consider how on-the-job learning affect the insights ob-
tained in B/M. This, of course, requires that we extend the B/M framework to heterogeneous workers. This extension is important for several reasons. For example, it is now well known that the typical empirical wage density observed, has an interior mode, is skewed with a “fat” right hand tail. Unfortunately, the equilibrium density generated in the simplest B/M model, where both workers are firms are homogeneous, has an increasing density. Adding worker heterogeneity to this simplest of models, however, can imply an equilibrium wage density that has an interior mode and is skewed the right way. Nevertheless, as Mortensen (2003) and Postel-Vinay and Robin (2006) have shown the distribution of firm productivities required to achieve this goal implies an implausible long right tail. Here we assume identical firms and show that for reasonable distributions of worker abilities (including the degenerate case) on-the-job learning and equilibrium not only yields an interior mode, the right tail of wages paid always has the Pareto distribution and so is suitably “fat”.

As we shall show the inclusion of human capital accumulation within the B/M framework yields a more coherent perspective on market wages. To illustrate, suppose different workers have different productivities $y$ which depend on the worker’s initial productivity and labour market experience. Specifically, we assume a worker, who enters the labour market with initial productivity $y_i$ has productivity $y = y_i e^{\rho x}$ after $x$ years of work experience; i.e. learning-by-doing increases productivity at rate $\rho > 0$. Further, firms compete in piece rates so that a firm which offers piece rate $\theta$ pays wage $w = \theta y$ to an employee with productivity $y$. For the same reasons as in B/M, we show that equilibrium implies the distribution of piece rate offers, $F(\theta)$, is non-degenerate. The wage earned by a type $i$ worker employed at firm $j$ at date $t$ with experience $x_{it}$ can then be decomposed as:

$$\log w_{ijt} = \log y_i + \rho x_{it} + \log \theta_j,$$

where $\theta_j$ is the piece rate offered by firm $j$. Wages are thus disperse as workers have different starting productivities (a worker fixed effect), workers have different total labour market experience, and different firms offer different piece rates (a firm fixed effect). Of course the distribution of firm fixed effects are generated endogenously in this framework and depend on the underlying primitives of the market. Nevertheless, the theory yields the following variance decomposition of wages paid in the market:

$$var(\log w) = var(\log y_i) + var(\log \theta) + \rho^2 var(x) + 2\rho cov(x, \log \theta).$$
This very simple decomposition arises as market outcomes such as quit rates, piece rate offers, etc. are orthogonal to a worker’s starting productivity $y_i$. The variance of log wages is then the weighted sum of the variances of the worker fixed effects, firm fixed effects and experience effects, plus the final covariance term. The covariance term arises as workers with long experience have also had some time to find better paid jobs. As described in Topel and Ward (1992), the wages of young workers rapidly increase over time as they accrue experience within firms and quit to better paid work across firms. Simple numerical examples find this latter covariance term can be large, explaining around one-half of the total observed variation in log wages.

Note, given the environment developed, workers value employment not only because of the wage paid. Employment is an investment opportunity as it provides experience which leads to higher wages in the future. A consequence is that unemployed workers may have lower reservation wages than in the standard search model - essentially such workers are willing to pay for work experience. This insight is useful as it can go some way to explain why starting wages can be very small relative to average wage paid in the market. Although not explored in detail here, this consequence can be useful to overcome the lack of “frictional” wage dispersion found in the standard search models by Hornstein, Krusell and Violante (2007).

There are several literatures related to what we present here. There is, for example, a large empirical literature which attempts to decompose wages into experience effects (general human capital) and tenure effects (firm specific capital) (see for example, Altonji and Shakotko, 1987, Topel, 1991, Altonji and Williams, 2005, and Dustmann and Meghir, 2005). The results obtained are still debated. The difficulty faced by this literature is that tenure and experience are perfectly correlated within any employment spell. As it is unreasonable to assume a quit, which resets tenure to zero, is an exogenous outcome that is orthogonal to the wage paid at the previous employer and at the new one, identifying between tenure and experience wage effects requires an equilibrium theory of wage formation and quit turnover. Extending the B/M framework to allow for experience effects, this paper can be considered as the first step in providing that unifying theory.

There are a few theoretical papers which have investigated learning-by-doing effects within an equilibrium turnover framework. Bunzel, Christensen, Kiefer and
Korsholm (2000) analyzed a model of human capital accumulation within the B/M framework. Unlike our approach here, they assume agents are initially homogeneous and workers lose all their human capital when laid off. As we show this difference leads to very different results. In an interesting application of record statistics, Barlevy (2008) estimates the wage process identified here but does not consider equilibrium. Also see Quercioli (2005) who considers equilibrium investment in firm specific human capital within the B/M framework.

Bagger, Fontaine, Postel-Vinay and Robin (2006) instead extend the offer matching framework developed by Postel-Vinay and Robin (2002a,b) to consider human capital accumulation with experience effects. The key feature of that approach is that when an employed worker receives an outside offer, the firms Bertrand compete over wages. That framework not only yields a remarkably tractable structure for econometric work, it would seem consistent with wage formation observed in many job markets, including the academic job market. Here instead we assume firms do not Bertrand compete for employees - a worker simply quits if he/she receives a preferred outside offer. Postel-Vinay and Robin (2004) argue that when on-the-job search effort is endogenous, large wage increases through Bertrand competition motivates employees to go out and get outside offers. A company wage policy which instead refuses to respond to outside job offers, reduces the value of on-the-job search and thus reduces employee quit rates. Of course, it remains an open question whether firms behave one way or another.

2 The Model

We assume time is continuous with an infinite horizon. Keeping things as simple as possible only steady-states are considered. There is a continuum of both firms and workers, each of measure one. All firms are equally productive and have a constant return to scale technology. Any worker’s life in this market can be described by an exponential distribution with parameter $\phi > 0$. Hence, any worker leaves the market for good in any small time period $dt$ with probability $\phi dt$. $\phi$ also describes the inflow of new labour market entrants. Assume there are $I$ types of workers, where each type is defined by his/her initial productivity. In particular, let $y_i$ denote the initial
productivity of a type $i$ worker and assume $y_1 < y_2 < \ldots < y_I$. Let $A$ denote the distribution function of these initial productivities. The steady-state number of type $i$ workers is indicated by $\gamma_i$.

On-the-job learning implies a worker’s productivity increases at rate $\rho > 0$ when working. Thus after $x$ years of work experience, a type $i$ worker’s productivity is $y = y_i e^{\rho x}$. An unemployed worker’s productivity $y$ remains constant through time.

A worker with productivity $y$ generates flow output $y$ while employed. We normalise the price of the production good to one, so $y$ also describes flow revenue. Each firm pays each of its employees the same piece rate $\theta$. Thus given an employee with productivity $y$, the worker is paid flow wage $w = \theta y$. Each firm’s total profit flow is simply total flow output from its employees multiplied by $(1 - \theta)$. As different firms may offer different piece rates, let $F(\theta)$ denote the proportion of firms offering a piece rate no greater than $\theta$. Further, let $\underline{\theta}, \overline{\theta}$ denote the infimum and supremum of the support of $F$. There are job destruction shocks in that each employed worker is displaced into unemployment according to a Poisson process with parameter $\delta > 0$. As discussed above, we assume a worker with productivity $y$, enjoys utility flow by while unemployed, where $0 < b < 1$.

Each employed and unemployed worker receives a job offer according to a Poisson process with parameter $\lambda > 0$. Search is random in that any job offer $\theta$ is considered as a random draw from $F$. If a job offer is rejected, the worker remains in his/her current state and there is no recall. We make the standard tie-breaking assumptions: an unemployed worker accepts a job offer if indifferent to accepting it or remaining unemployed, while an employed worker quits only if the job offer is strictly preferred. We require $\phi > \rho$ to ensure total expected lifetime payoffs are finite.

All agents are risk neutral. For simplicity we assume workers do not discount the future - they maximize expected lifetime income. Each firm chooses piece rate $\theta$ to maximize steady state flow profit, taking into account the search strategies of workers.
3 Optimal Search Strategies

Let $W^U(y)$ denote the expected lifetime payoff of an unemployed worker with productivity $y$ using an optimal search strategy. Further, let $W^E(y, \theta)$ denote the expected lifetime payoff of a worker with productivity $y$, currently employed at piece rate $\theta$, when using an optimal search strategy. As $W^E(y, \theta)$ is strictly increasing in $\theta$ (it is always better to be employed at a firm with a larger piece rate) an employed worker quits to an outside offer $\theta'$ if and only if $\theta' > \theta$. Thus, the flow Bellman equation for employed workers implies:

$$ (\phi + \delta)W^E(y, \theta) = \theta y + \rho y \frac{\partial W^E}{\partial \theta} + \lambda \int_{\theta}^{\theta'} [W^E(y, \theta') - W^E(y, \theta)] dF(\theta') + \delta W^U(y). \quad (1) $$

The first term on the right hand side describes flow earnings, the second describes increased value through on-the-job learning, the third describes the capital gain by receiving a preferred outside offer $\theta' > \theta$, while the last corresponds to the welfare loss through being laid-off.

As there is no on-the-job learning while unemployed (and no depreciation), the flow Bellman equation describing $W^U(y)$ is instead

$$ \phi W^U(y) = by + \lambda \int_{\theta}^{\theta'} \max[W^E(y, \theta') - W^U(y), 0] dF(\theta'), \quad (2) $$

where a job offer is received at rate $\lambda$ and conditional on the realized draw $\theta'$, the worker either accepts it and enjoys welfare gain $W^E(y, \theta') - W^U(y)$, or remains unemployed with productivity $y$.

Note, a worker’s income, whether unemployed or employed, is always proportional to $y$. As on-the-job learning is also proportional to $y$ and workers are risk neutral, the above Bellman equations imply there exists a number $\alpha^U$ and a function $\alpha^E(\cdot)$ such that

$$ W^U(y) = \alpha^U y, \text{ and } W^E(y, \theta) = \alpha^E(\theta)y. $$

We now determine $\alpha^U$ and $\alpha^E(\cdot)$.

The Bellman equation (2) implies the unemployed worker’s optimal strategy is to accept any offer $\theta'$ satisfying $W^E(y, \theta') \geq W^U(y)$. As $W^E$ is increasing in $\theta'$, the worker, will accept any offer $\theta' \geq \theta^R$ where the reservation piece rate $\theta^R$ is given by $\alpha^E(\theta^R) = \alpha^U$. Claim 1 now characterises $\alpha^U$ and $\alpha^E(\cdot)$. It is convenient first to define

$$ q(\theta) = \phi + \delta + \lambda(1 - F(\theta)). $$
which is the rate at which any given employee exits employment from a firm offering piece rate $\theta$.

**Claim 1:** $W_E(y, \theta) = \alpha(E) y$ and $W_U(y) = \alpha(U)$ satisfy the above Bellman equations if and only if

(i) $\alpha(E)$ is the solution to the ordinary differential equation

$$\frac{d\alpha(E)}{d\theta} = \frac{1}{q(\theta) - \rho},$$

with boundary condition $\alpha(E)(\theta) = (\overline{\theta} + \delta \alpha(U))/((\delta + \phi - \rho)$ at $\theta = \overline{\theta}$, and

(ii) $\alpha(U)$ satisfies

$$\phi \alpha(U) = b + \int_{\theta^R}^{\overline{\theta}} [\alpha(E)(\theta') - \alpha(U)]dF(\theta'),$$

with $\theta^R$ satisfying $\alpha(E)(\theta^R) = \alpha(U)$.

**Proof:**

Given the above functional forms for $W_U$ and $W_E$, the Bellman equation (2) is equivalent to (4). The Bellman equation (1) is equivalent to

$$(\phi + \delta)\alpha(E)(\theta) = \theta + \rho \alpha(E)(\theta) + \lambda \int_{\theta}^{\overline{\theta}} [\alpha(E)(\theta') - \alpha(E)(\theta)]dF(\theta') + \delta \alpha(U),$$

which is a functional equation for $\alpha(E)(\theta)$. Differentiating with respect to $\theta$ yields (3) and putting $\theta = \overline{\theta}$ yields its boundary value $\alpha(E)(\overline{\theta})$. This completes the proof of Claim 1.

We next simplify the conditions of Claim 1 by substituting out $\alpha(E)(\theta)$. First put $\theta = \theta^R$ in (5). As $\alpha(E)(\theta^R) = \alpha(U)$ we obtain

$$\phi \alpha(U) = \theta^R + \rho \alpha(U) + \lambda \int_{\theta^R}^{\overline{\theta}} [\alpha(E)(\theta') - \alpha(U)]dF(\theta').$$

Comparing this equation with (4) establishes (6) described in Claim 2 below. Next integrate (4) by parts. Using (3) and $\alpha(E)(\theta^R) = \alpha(U)$ then yields (7) in Claim 2. Thus (6) and (7) describe a pair of equations for $(\alpha(U), \theta^R)$. We now establish a solution exists and is unique.

**Claim 2:** For any $F$, $(\alpha(U), \theta^R)$ satisfy the pair

$$\rho \alpha(U) = b - \theta^R,$$

$$\phi \alpha(U) = b + \int_{\theta^R}^{\overline{\theta}} \frac{\lambda(1 - F(\theta))}{q(\theta) - \rho} d\theta.$$
Further, a solution exists, is unique and implies $\alpha^U > 0$ and $\theta^R < b$.

**Proof:** Consider Figure 1 which plots these two equations, where $(\alpha^U, \theta^R)$ occurs at the intersection

![Figure 1: Existence and Uniqueness of $(\alpha^U, \theta^R)$](image)

The locus labelled LL in Figure 1 corresponds to equation (6): this equation is linear, has slope $-1/\rho$ and $\alpha^U = 0$ at $\theta^R = b$. The locus labelled II in Figure 1 corresponds to equation (7). Differentiating (7) with respect to $\theta^R$ implies the II locus is strictly decreasing with slope

$$\frac{d\alpha^U}{d\theta^R}_{eqn(7)} = -\frac{1}{\phi} \left( \frac{\lambda(1 - F(\theta^R))}{q(\theta^R) - \rho} \right).$$

Simple algebra using the parameter restriction $\phi > \rho$ establishes this latter locus is flatter (strictly) than the first locus for all $\theta$, even when $\theta^R < b$. As locus II finds $\alpha^U$ is strictly positive and finite at $\theta^R = b$, continuity now implies these two loci must have a unique intersection at some $\theta^R < b$. This completes the proof of Claim 2.

No experience effects, $\rho = 0$, implies the reservation piece rate $\theta^R = b$. In that case a worker accepts a job offer if and only if the offered piece rate exceeds benefits received while unemployed. But positive experience effects, $\rho > 0$, instead implies $\theta^R < b$. Indeed it is straightforward to show an increase in $\rho$ (with $F$ given) always leads to a fall in the reservation piece rate $\theta^R$. This occurs as a higher $\rho$ increases the
value of being employed. An unemployed worker is then willing to sacrifice current income to invest in increased productivity, and thus higher wages in the future. This comparative static also implies an increase in $\alpha^U$; the value of being unemployed also increases.

4 Profits and Steady-States

Above we have shown that for any distribution function $F$, all workers have the same reservation piece rate $\theta^R_i$. This Section now characterises equilibrium firm behavior and so determines $F$. To do this we need to define and characterise three steady-state objects - (a) $U^i$ : the fraction of type $i$ workers who are unemployed, (b) $N^i(.)$ : the distribution function describing productivities across type $i$ workers who are unemployed, and (c) $H^i(y, \theta)$ : the joint distribution function describing productivities and piece rates earned across type $i$ workers who are employed.

As there is no discounting, the arguments in Burdett and Coles (2003) imply steady state flow profit equals the hiring rate of the firm, multiplied by the expected profit of each hire. If the firm offers piece rate $\theta < \theta^R_i$, it attracts no employees and so makes zero profit. Consider instead offers $\theta \geq \theta^R_i$. In this case steady state flow profit is

$$\pi(\theta) = \sum_i \left[ \lambda \gamma_i U^i \int_{y=y_i}^{\infty} \left[ \int_{x=0}^{\infty} e^{-q(\theta)x}(1-\theta)ye^{\rho x}dx \right] dN^i(y) \right. \right.$$

$$\left. + \lambda \gamma_i (1-U^i) \int_{y=y_i}^{\infty} \left[ \int_{x=0}^{\infty} e^{-q(\theta)x}(1-\theta)ye^{\rho x}dx \right] dH^i(y, \theta') \right] .$$

Recall $\gamma_i$ is the number of workers in the economy who are type $i$, and thus $\gamma_i U^i$ is the number of type $i$ workers who are unemployed. For each $i$, the first term in the above equation is the steady state flow profit due to attracting type $i$ unemployed workers with productivity $y \in [y_i, \infty)$; $\gamma_i U^i dN^i(y)$ is the hiring inflow of such workers and the inside bracketed integral is the expected discounted profit per hired unemployed worker with productivity $y$, taking into account that productivity on-the-job increases at rate $\rho$ and the worker leaves the firm at rate $q(\theta)$. The second term is the flow profit due to attracting type $i$ employed workers on wage rates $\theta' < \theta$ with productivity $y \in [1, \infty)$; $\gamma_i (1-U^i) dH^i(y, \theta')$ describes the hiring inflow of each such worker and the inside bracketed integral is again the expected discounted profit per hire.
Integrating over $x$ yields:

$$
\pi(\theta) = \frac{\lambda(1 - \theta)}{q(\theta) - \rho} \sum_i \left[ \gamma_i U^i \int_{y_i = y_i}^{\infty} y' dN^i(y') + \gamma_i (1 - U^i) \int_{y_i = y_i}^{\theta} \int_{y' = y_i}^{\infty} y' dH^i(y', \theta') \right].
$$

(8)

As $\theta^R < b$ (Claim 2), a firm can always guarantee strictly positive profit by offering piece rate $\theta = b < 1$. Thus, there is no loss in generality in assuming firms make strictly positive profit in a Market Equilibrium. We now formally define such an equilibrium.

**A Market Equilibrium** is a set $\{\theta^R, U^i, N^i(\cdot), H^i(\cdot), F(\cdot)\}$ with $i = 1, \ldots, I$, such that

(i) $\theta^R$ is the optimal reservation piece rate of any unemployed worker;

(ii) $U^i, N^i(\cdot), H^i(\cdot)$ are consistent with steady state turnover given piece rate offers $F(\cdot)$ and optimal worker search strategies;

(iii) the constant profit condition is satisfied; i.e.,

$$
\pi(\theta) = \pi > 0 \text{ for all } \theta \text{ where } dF(\theta) > 0;
$$

$$
\pi(\theta) \leq \pi \text{ for all } \theta \text{ where } dF(\theta) = 0.
$$

We begin with a standard result which much simplifies the analysis.

**Claim 3:** A Market Equilibrium implies (i) $F(\cdot)$ contains no mass points; (ii) $F(\cdot)$ has a connected support and (iii) $\theta = \theta^R$.

**Proof:** As the arguments are already well known (see B/M) we only sketch the proof.

A contradiction argument establishes there cannot be a mass point in $F$. If there were, say at $\theta = \theta^m$, steady state would imply a mass of employees on piece rate $\theta^m$. But offering piece rate $\theta = \theta^m + \varepsilon$, where $\varepsilon > 0$ but very small, would yield a slightly lower profit per hire but a large increase in the hiring rate, and this deviation would strictly increase profit (and so contradict equilibrium). A contradiction argument also establishes the support of $F$ must be connected. Otherwise if there were a hole, say for $\theta \in [\theta_L, \theta_H]$, then the offer $\theta = \theta_L$ would yield strictly greater profit (as both offers attract the same number of workers) which contradicts $\theta_H$ being an optimal offer. Finally a contradiction argument also establishes $\theta = \theta^R$. $\theta < \theta^R$ would imply a firm offering $\theta = \theta$ obtains zero profit, which contradicts strictly positive profit. $\theta > \theta^R$ would instead imply offering $\theta = \theta^R$ makes strictly greater profit than offering $\theta$ which contradicts optimality of $\theta$. This completes the proof.
We now construct the three steady state objects defined above, $U^i$, $N^i$, and $H^i$. First consider the steady state pool of type $i$ unemployed workers. Claim 3(iii), that $\theta = \theta^R$, implies each unemployed worker accepts the first job offer received. Hence the flow of type $i$ workers out of the unemployment pool is $(\lambda + \phi)\gamma_i U^i$. As the inflow is $\phi \gamma_i + \delta \gamma_i (1 - U^i)$, and as steady state turnover implies these flows must be equal, the type $i$ unemployment rate in a steady state is

$$U^i = \frac{\phi + \delta}{\phi + \delta + \lambda}. $$

As $U^i$ is the same for all types, let $U$ denote this common unemployment rate.

Next consider the pool of type $i$ unemployed workers with productivity no greater than $y$ (and $y \geq y_i$). Steady-state turnover requires

$$\gamma_i \phi + \delta \gamma_i (1 - U) H^i(y, \overline{\theta}) = \left[ \phi + \lambda \right] \gamma_i U N(y),$$

where the left hand side describes the inflow (new labour market entrants with productivity $y = y_i$ and employed workers with productivity less than $y$ who lose their jobs) and the right hand side describes the outflow. Solving for $N^i(y)$ (and using the above solution for $U$) yields

$$N^i(y) = \frac{\phi (\phi + \delta + \lambda) + \lambda \delta H^i(y, \overline{\theta})}{(\phi + \lambda)(\phi + \delta)} \text{ for all } y \geq y_i. \quad (9)$$

Finally consider the pool of type $i$ employed workers who have productivity no greater than $y$ and earn piece rate no greater than $\theta$. For $\theta \geq \theta$ and $y \geq y_i$, the total outflow of workers from this pool, over any instant of time $dt > 0$, is

$$\gamma_i (1 - U) H^i(y, \theta) q(\theta) dt + \gamma_i (1 - U) \left[ H^i(y, \theta) - H^i\left(\frac{y}{1 + \rho dt}, \theta\right)\right] + O(dt^2),$$

where $q(\theta) = \phi + \delta + \lambda (1 - F(\theta))$. The first term specifies the number workers in this pool who either leave employment, or quit to a job with $\theta' > \theta$. The second describes those who exit this pool as their productivity increases above $y$ through on-the-job learning. The last term corrects for the fact that some do both but this term has the property $O(dt^2)/dt \rightarrow 0$ as $dt \rightarrow 0$. The inflow into this pool over small time period $dt$, is simply $\gamma_i U N(y) \lambda F(\theta) dt$ : it is those type $i$ unemployed workers with productivity no greater than $y$ who find a job no better than $\theta$. Setting inflow equal
to outflow, rearranging appropriately and then letting $dt \to 0$ yields the following partial differential equation for $H^i$

$$\frac{\partial H^i}{\partial y} + \frac{q(\theta)}{\rho y} H^i = \frac{(\phi + \delta)F(\theta)N^i(y)}{\rho y},$$

(10)

for $\theta \in [\underline{\theta}, \overline{\theta}]$. For given $\theta$, integrating over $y$ using the integrating factor $\frac{\psi(\theta)}{\rho}$ implies

$$\left[H^i y^{\frac{\psi(\theta)}{\rho}}\right]_{y'=y_i}^y = \int_{y'=y_i}^y y^{\frac{\psi(\theta)}{\rho}} \frac{(\phi + \delta)F(\theta)N^i(y')}{\rho y'} dy'.$$

As $H^i(y_i, \theta) = 0$, we obtain

$$H^i(y, \theta) = \frac{(\phi + \delta)F(\theta)}{\rho} y^{\frac{\psi(\theta)}{\rho}} \int_{y_i}^y y^{\frac{\psi(\theta)}{\rho}}^{-1} N^i(y') dy' \text{ for all } y \geq y_i, \ \theta \in [\underline{\theta}, \overline{\theta}].$$

(11)

Using these formulae we now solve for steady state $N^i(.)$ and $H^i(., .)$.

**Claim 4:** A Market Equilibrium implies distribution functions

$$N^i(y) = 1 - \frac{\lambda \delta}{(\phi + \lambda)(\phi + \delta)} \left(\frac{y}{y_i}\right)^{-\left(\frac{\phi(\phi + \delta + \lambda)}{\rho(\phi + \lambda)}\right)} \text{ for all } y \geq y_i,$$

(12)

$$H^i(y, \theta) = \frac{(\phi + \delta)F(\theta)}{q(\theta)} \left[1 - \left(\frac{y}{y_i}\right)^{-\frac{\psi(\theta)}{\rho}}\right] - \frac{\delta F(\theta)}{q(\theta) - \phi F(\theta)} \left[\left(\frac{y}{y_i}\right)^{-\left(\frac{\phi(\phi + \delta + \lambda)}{\rho(\phi + \lambda)}\right)} - \left(\frac{y}{y_i}\right)^{-\frac{\psi(\theta)}{\rho}}\right]$$

(13)

for all $\theta \in [\underline{\theta}, \overline{\theta}]$ and $y \geq y_i$.

**Proof:**

Put $\theta = \overline{\theta}$ in (10). Using (9) to substitute out $N^i(.)$ and simplifying yields

$$\frac{\partial H^i(y, \overline{\theta})}{\partial y} = \frac{\phi (\phi + \delta + \lambda)}{\rho (\phi + \lambda)} \frac{1 - H^i(y, \overline{\theta})}{y} \text{ for all } y \geq y_i,$$

As this differential equation is separable and we have the boundary condition $H^i(y_i, \overline{\theta}) = 0$, integration implies

$$H^i(y, \overline{\theta}) = 1 - \left(\frac{y}{y_i}\right)^{-\left(\frac{\phi(\phi + \delta + \lambda)}{\rho(\phi + \lambda)}\right)}.$$  

(14)

Using this in (9) and simplifying yields the desired result (12). Using (12) to substitute out $N^i(.)$ in (11) and some algebra then establishes (13). This completes the proof of Claim 4.

Note, $y/y_i = \exp(\rho x)$ and so reflects worker experience. Indeed, the distribution functions described in Claim 4 imply the distribution of experience $x$ across employed
and unemployed workers is the same for all worker types. Note that \( H_i(y, \theta) \) describes the distribution of productivities across all type \( i \) employed workers. Interestingly, it follows a Pareto distribution with scale parameter \( y_i \) and shape parameter \( \phi/\rho \)[(\( \phi + \delta + \lambda)/(\phi + \lambda) \)]. \( N^i(y) \) describes type \( i \) unemployed worker productivities. \( N^i(y_i) > 0 \) implies there is a mass of unemployed workers with productivity \( y_i \): For \( y > y_i \), both (14) and (12) imply these productivity densities decline at the same rate, which depends on the turnover parameters \( \lambda, \delta, \phi \) and the on-the-job learning parameter \( \rho \).

## 5 Market Equilibrium

In a Market Equilibrium we require \( \pi(\theta) = \pi \) for all \( \theta \in [\theta, \bar{\theta}] \), where (8) implies

\[
\pi(\theta) = \frac{\lambda(1 - \theta)}{q(\theta) - \rho} \sum_i \left[ \gamma_i U \int_{y' = y_i}^{\infty} y'dN^i(y') + \gamma_i (1 - U) \int_{y' = y_i}^{\infty} \int_{\theta' = \theta}^{\theta} y' \frac{\partial^2 H_i(y', \theta')}{\partial y' \partial \theta'} d\theta' dy' \right].
\]

Using the solutions for \( N^i \) and \( H^i \) identified in Claim 4, the following Claim solves for the integrals in (15).

**Claim 5:**

\[
\int_{y' = y_i}^{\infty} y'dN^i(y') = \frac{\phi(\phi + \delta + \lambda) y_i}{(\phi + \delta)} \left[ \frac{(\phi + \delta - \rho)}{\phi(\phi + \delta + \lambda) - \rho(\phi + \lambda)} \right]
\]

\[
\int_{y' = y_i}^{\infty} \int_{\theta' = \theta}^{\theta} y' \frac{\partial^2 H_i(y', \theta')}{\partial y' \partial \theta'} d\theta' dy' = \frac{\phi F(\theta) y_i}{q(\theta) - \rho} \left[ \frac{(\phi + \delta + \lambda)(\phi + \delta - \rho)}{\phi(\phi + \delta + \lambda) - \rho(\phi + \lambda)} \right].
\]

**Proof:**

See Appendix.

As the distribution of experiences is the same across all types and the productivity of a type \( i \) workers with experience \( x \) is \( y_i e^{\rho x} \) (and therefore proportional to \( y_i \)), the average productivities computed in Claim 5 are also proportional to \( y_i \). Claim 5 and (15) now yield the constant profit condition:

\[
\frac{\lambda(1 - \theta) \bar{y}}{q(\theta) - \rho} \left[ \frac{\phi(\phi + \delta - \rho)}{\phi(\phi + \delta + \lambda) - \rho(\phi + \lambda)} + \frac{\lambda F(\theta) y_i}{q(\theta) - \rho} \left( \frac{\phi(\phi + \delta - \rho)}{\phi(\phi + \delta + \lambda) - \rho(\phi + \lambda)} \right) \right] = \pi,
\]

for all \( \theta \in [\theta, \bar{\theta}] \), where \( \bar{y} = \sum_i \gamma_i y_i \) is the mean of \( A \). This equation determines equilibrium \( F \). Further simplification finds this condition reduces

\[
\lambda(1 - \theta) \bar{y} \left[ \frac{\phi(\phi + \delta - \rho)(\phi + \delta - \rho + \lambda)}{\phi(\phi + \delta + \lambda) - \rho(\phi + \lambda)} \right] = \pi (q(\theta) - \rho)^2.
\]

\(13\)
As this is a quadratic equation for $F$, and noting $F$ must be increasing over $\theta$, (16) now yields the following solution for $F$.

**Claim 6:** A Market Equilibrium implies

$$F(\theta) = \frac{\phi + \delta + \lambda - \rho}{\lambda} - \left[ \frac{\bar{\theta} \phi(\phi + \delta - \rho)[\phi + \delta - \rho + \lambda]}{\lambda \pi \left(\phi + \delta + \lambda - \rho(\phi + \lambda)\right)} \right]^{1/2} (1 - \theta)^{1/2}. \quad (17)$$

Given this solution for $F$, it is straightforward to obtain a closed form solution for a Market Equilibrium.

Recall a Market Equilibrium implies a steady state profit $\pi > 0$. As Claim 3 (no mass points) implies $F = 0$ at $\theta = \hat{\theta}$, (17) implies $\pi$ and $\hat{\theta}$ are related as

$$\frac{\phi + \delta + \lambda - \rho}{\lambda} = \left[ \frac{\bar{\theta} \phi(\phi + \delta - \rho)[\phi + \delta - \rho + \lambda]}{\lambda \pi \left(\phi + \delta + \lambda - \rho(\phi + \lambda)\right)} \right]^{1/2} (1 - \theta)^{1/2}$$

It is convenient to use this condition to substitute out $\pi$ in (17). Thus the equilibrium offer distribution can be written as

$$F = \hat{F}(\theta \mid \bar{\theta}) = \left( \frac{\phi + \delta - \rho + \lambda}{\lambda} \right) \left[ 1 - \left( \frac{1 - \theta}{1 - \bar{\theta}} \right)^{1/2} \right]. \quad (18)$$

Somewhat surprisingly, the functional form describing equilibrium $F$ is the same as in B/M. The goal now is to solve for equilibrium $\hat{\theta}$ and so determine $F$. This requires solving the following problem. Consider an arbitrary value for $\hat{\theta} \leq b$. If $\hat{\theta}$ is an equilibrium value, then (18) implies $F = \hat{F}(\theta \mid \bar{\theta})$ is the equilibrium distribution of offers. Given this explicit solution for $F$, we then solve the conditions of Claim 2 for the workers’ optimal $\theta^R$. Let $\theta^R = \theta^R(\bar{\theta})$ denote that solution. As a Market Equilibrium requires $\theta^R(\bar{\theta}) = \bar{\theta}$ (by Claim 3(iii)) this condition ties down $\bar{\theta}$. In fact, we provide the analytic solution.

**Theorem 1.** For any $\rho < \phi$ a Market Equilibrium exists, is unique and implies

$$\theta = \theta^R = \frac{b(\phi - \rho)(\phi + \delta - \rho + \lambda)^2 - \rho \lambda^2}{\phi (\phi + \delta - \rho + \lambda)^2 - \rho \lambda^2}. \quad (19)$$

Equilibrium $F$ is given by (18) with $\theta$ given by (19), and the steady state distribution functions $N^i, H^i$ are as described in Claim 4.

**Proof:** In the Appendix we show that solving $\theta^R(\bar{\theta}) = \bar{\theta}$ yields the unique solution (19) for $\theta$. Given the equilibrium distribution $F$ as described in Theorem 1 then, by construction, the unemployed worker’s optimal reservation piece rate $\theta^R = \bar{\theta}$. Further
$U^i, N^i, \text{and } H^i$ are consistent with steady state turnover and $F$ ensures the constant profit condition holds for all $\theta \in [\underline{\theta}, \overline{\theta}]$. All that remains is to show there is no other offer which is profit increasing. Offering $\theta < \underline{\theta}$ yields zero profit as all workers reject such offers (as $\theta^R = \underline{\theta}$). Conversely offering $\theta > \overline{\theta}$ yields strictly less profit than offering $\theta = \overline{\theta}$ as it attracts no additional workers. Thus the above identifies the Market Equilibrium. This completes the proof of Theorem 1.

Theorem 1 not only establishes that the equilibrium reservation piece rate $\theta^R$ is independent of the initial distribution of productivities $A$, so is the distribution of piece rates offered, $F$. Nevertheless the distribution of wages paid depends on underlying productivities as wages paid $w = \theta y$.

Note for a given $\theta$, the offer distribution $F = \hat{F}(\theta \mid \theta)$ described in equation (18) is first order stochastically increasing in $\rho$. A greater on-the-job learning rate, $\rho$, implies employees are more valuable and greater wage competition between firms tends to increase equilibrium piece rate offers. But some algebra using (19) implies $\theta^R$, and thus $\underline{\theta}$, decreases with $\rho$. As described in the previous section, a higher $\rho$ implies each unemployed worker is willing to sacrifice current income to invest in increased productivity, and thus the workers’ reservation piece rate decreases. A curious feature of equilibrium, then, is that higher $\rho$ implies firms tend to be more competitive on piece rates (employees are more valuable) while unemployed workers are willing to work for less and so equilibrium $\theta$ falls.

6 Equilibrium Distributions

In B/M with homogeneous agents, the density of wages paid to workers is increasing with wages which is inconsistent with the data. But here we show that on-the-job learning not only implies a distribution of wages whose density declines at high wages, its right tail has the Pareto distribution which would seem consistent with empirical distributions. In this section we also show that introducing on-the-job learning increases the amount of “frictional” wage dispersion that can be generated within the B/M framework. We start by describing how productivity $y$, piece rates $\theta$ and wages $w = \theta y$ are correlated across employed workers. For ease of exposition, from now on assume $b$ sufficiently large that $\theta > 0$. 

15
Recall that the unemployment rate is the same for all types $i$. Consider then a type $i$ worker who is employed. Claim 4 establishes the joint cdf $H^i$ is given by

$$H^i(y, \theta) = \frac{(\phi + \delta)F}{q} \left[ 1 - \left( \frac{y}{y_i} \right)^{-\frac{2}{\rho}} \right] - \frac{\delta F}{q - \phi F} \left[ \left( \frac{y}{y_i} \right)^{-\frac{(\phi + \delta + \lambda)}{\rho(\phi + \delta)}} - \left( \frac{y}{y_i} \right)^{-\frac{2}{\rho}} \right], \quad (20)$$

where $q = \phi + \delta + \lambda(1 - F)$ and Theorem 1 describes equilibrium $F$. Using this expression, it can be shown that conditional on productivity $y$, the distribution of earned piece rates is $^1$

$$H^i(\theta | y) = F \left[ \frac{\delta + (\phi + \lambda)(1 - F) \left( \frac{y}{y_i} \right)^{-\frac{\lambda(\phi - \phi F)}{\rho(\phi + \delta)}}}{\delta + (\phi + \lambda)(1 - F)} \right].$$

Note that $H^i(\theta | y_i) = F$; the first job of new entrants is a random draw from $F$. More interestingly, $H^i(\theta | y)$ is strictly decreasing in $y$; i.e. the conditional distribution of earned piece rates is first order stochastically increasing in $y$. This occurs as workers from time to time quit for better paid work. Thus more experienced workers not only earn more because they are typically more productive, but also because over time they find employment at firms offering better piece rates. As established in B/M, a worker’s earned piece rate $\theta$ is also positively correlated with tenure; workers on higher piece rates are less likely to quit. Thus $\theta$ is positively correlated with experience and tenure. When $\theta$ is unobserved by the econometrician, identifying between experience and tenure effects on wages thus becomes problematic. We shall return to this issue in the Conclusion.

In B/M firms offering higher piece rates attract and retain more workers. But here an additional effect is that offering a higher piece rate also improves the composition of its workforce; i.e., a higher piece rate attracts and retains more experienced and thus more productive workers. Of course in equilibrium, the constant profit condition implies more firms will then offer higher piece rates. To establish this insight consider a type $i$ firm offering piece rate $\theta$ and for ease of exposition, assume $y_i = 1$. Then the conditional distribution of employee productivities is $^2$

$$H^i(y | \theta) = \left( 1 - y^{-\frac{2}{\rho}} \right) - \frac{\delta q^2}{(\phi + \delta) [q - \phi F]^2} \left[ y^{-\frac{(\phi + \delta + \lambda)}{\rho(\phi + \delta)}} - y^{-\frac{2}{\rho}} \right] \frac{\lambda F(1 - F)q}{\rho(\phi + \delta) [q - \phi F]} y^{-\frac{2}{\rho}} \log y.$$

$^1$The easiest way to obtain this condition is to use $H^i(\theta | y) = \int_{\theta'}=\frac{\partial^2 H^i(y, \theta')}{\partial y \partial \theta'} d\theta' / \int_{\theta'}=\frac{\partial^2 H^i(y, \theta')}{\partial y \partial \theta'} d\theta'$. $^2$Use (20) and $H^i(y | \theta) = \int_{y'}=1 \frac{\partial^2 H^i(y', \theta)}{\partial y' \partial \theta} dy' / \int_{y'=1}^{\infty} \frac{\partial^2 H^i(y', \theta)}{\partial y' \partial \theta} dy'$. 

16
This expression is most easily interpreted by first considering the firm with the lowest piece rate, i.e., \( \theta = \bar{\theta} \). Such firms only attract unemployed workers. The first term describes the distribution of worker productivities through hiring new labour market entrants (with initial productivity \( y_i = 1 \)) taking into account that employee productivity grows at rate \( \rho \) and workers exit this firm’s employment at rate \( q \). Given \( \delta > 0 \), the second term captures the composition effect through job destruction. Finally the third term takes into account that a firm offering \( \theta > 0 \); attracts workers from firms offering \( \theta' < \theta \) and those workers also hold previous experience.

A sufficient (but not necessary) condition for first order stochastic dominance is \( \delta > \phi \); i.e. \( \delta > \phi \) guarantees \( \partial[H^i(y|\theta)]/\partial \theta < 0 \). In this case, higher piece rates yield a more productive workforce. Note the restriction \( \delta > \phi \) would appear to be the most reasonable case: it implies each worker expects to be laid off at least once over a working lifetime.

We now turn to the equilibrium wage distribution, which is denoted by \( G(\cdot) \). Let \( h^i(y, \theta) \) denote the joint density of \((y, \theta)\) across type \( i \) employed workers. Thus given \( H^i \) defined by (20) above

\[
h^i = \begin{cases} 
\frac{\partial^2 H^i(\cdot, \cdot)}{\partial y \partial \theta} & \text{for } y \geq y_i \text{ and } \theta \in [\bar{\theta}, \overline{\theta}], \\
0 & \text{otherwise.}
\end{cases}
\]

As wages \( w = \theta y \), then:

\[
G(w) = \sum_i \gamma_i \int_{\theta' = \bar{\theta}}^{\overline{\theta}} \left[ \int_{y' = 0}^{w/\theta'} h^i(y, \theta') dy' \right] d\theta'.
\]

Differentiating with respect to \( w \) identifies the wage density:

\[
G'(w) = \sum_i \gamma_i \int_{\theta' = \bar{\theta}}^{\overline{\theta}} h^i(w/\theta', \theta') \frac{1}{\theta'} d\theta'.
\]
Although it is straightforward to compute $G'$, its solution is long and unwieldy. Nevertheless, two useful insights follow without much trouble. First, the density is initially increasing, and second it finally declines according to a Pareto density. These claims are established below.

The lowest wage paid in the market is $w = \theta y_1$: the least productive worker with no experience who is employed at the firm offering the smallest piece rate. It follows that $G'(w) \to 0$ as $w \to w^+$. In particular, (21) implies

$$G'(w) \to \int_{\theta' = \hat{\theta}}^{w/y_1} h^i(w, \theta') \frac{1}{\theta'} d\theta' \text{ as } w \to w^+,$$

as (i) $h^i = 0$ for all $i > 1$ in this limit (all other type $i$ employees earn strictly more than $w$, and (ii) $h^1(y, \theta') = 0$ for $y = w/\theta' < y_1$. Thus $G'(w) = 0$ at $w = w$ and so the wage density must initially be increasing with $w$.

Consider now the behavior of $G'$ as wages become large. Using (20) above, it is straightforward to show

$$h^i(y, \theta) \to \frac{F^i(\phi + \delta + \lambda)}{y_i (q - \phi F)^2} \left( \frac{\phi + \lambda}{\rho(\phi + \lambda)} \right) \left( \frac{y}{y_i} \right)^{-\left(\frac{\phi + \delta + \lambda}{\rho(\phi + \lambda)}\right) - 1} \text{ as } y \to \infty;$$

i.e. the density of type $i$ worker productivities declines geometrically at rate $-\left(\frac{\phi + \delta + \lambda}{\rho(\phi + \lambda)}\right)$ and is the same for all types. (21) now implies

$$G'(w) = O(w^{-(\phi + \delta + \lambda)/\rho(\phi + \lambda)} - 1) \text{ as } w \to \infty. \quad (22)$$

Although the explicit solution for $G'(.)$ depends on $A$, the right tail of $G'(.)$ declines at the same rate as the productivity distribution (14), which in turn follows a Pareto distribution.\footnote{This is an interesting property of the model as it is well documented that the right tail of the empirical earnings distribution can be approximated by a Pareto distribution (see, for example, von Weizsäcker 1993, and Neal and Rosen, 2000).} Note, the shape of the right tail of the wage density depends only on the growth parameters ($\rho$ and $\phi$) and turnover parameters ($\delta$ and $\lambda$). An increase in on-the-job learning $\rho$ implies the wage density falls more slowly, thus yielding a “fatter” right tail in the density of wages paid. Conversely, an increase in the exit rate $\phi$ or layoff parameter $\delta$ (which implies workers spend more time unemployed and thus not learning on the job) implies a “thinner” right tail.
For ease of exposition the above has considered a finite number of types. Nevertheless, the analysis extends straightforward manner if we assume instead \( A(\cdot) \) describes a continuum of underlying abilities. The only difference is that we instead integrate over \( dA \) in (8) and in (21) rather than sum over \( \gamma_i \). But as both are linear operators, the equilibrium structure is invariant to this extension. Also note that if initial abilities \( y_i \) are bounded (i.e. the support of \( A \) is finite), the argument above; that the right tail of the wage distribution has the Pareto distribution with shape parameter \((\phi/\rho)[(\phi + \delta + \lambda)/(\phi + \lambda)] \), continues to go through. If \( A \) has an unbounded support, this may not necessarily hold.

We now turn to the model’s implications for “frictional” wage dispersion. An important contribution of the B/M model is that it gives a theory of why similar workers are paid differently.\(^6\) Hornstein, Krusell and Violante (2007) used the ratio between the average wage to the minimum observed wage (or reservation wage) as a way to measure this type of wage dispersion. They argued that a reasonably calibrated version of the B/M model fails to account for the amount of “frictional” wage dispersion observed in the data. In short, they argue that the B/M model generates a reservation wage that is too high to match the observed mean-min \((Mm)\) ratio in the US economy for plausible parameter values.

Our analysis has shown that when human capital accumulation is introduced to an otherwise standard B/M model we obtain a lower reservation wage. Workers are willing to accept wages below their opportunity cost of employment as it is an investment that increases their productivity in the future. We now show that this generates a higher \( Mm \) ratio. First note that in our model “frictional” wage dispersion refers to wage variation induced only by firms piece rate policies. For this reason our \( Mm \) ratio should be defined by the ratio between the average piece rate, \( \bar{\theta} \), and the minimum piece rate, \( \theta^R \), earned by employed workers of a given initial productivity. Using (6) and (7) in Claim 2 and (13), we show in the appendix that the \( Mm \) ratio in our model can be approximated by

\[
Mm \simeq \frac{1 - \frac{\lambda \rho}{\phi(\phi - \rho + \delta + \lambda)}}{\frac{\beta(\phi - \rho)}{\phi} - \frac{\lambda \rho}{\phi(\phi - \rho + \delta + \lambda)}},
\]

\(^6\) For evidence of this empirical regularity see Abowd, Kramarz and Margolis (1999), Mortensen (2003), Postel-Vinay and Robin (2006) and Horstein, Krusell and Violante (2007).
where we have set $b = \beta \tilde{\theta}$ with $\beta > 0$. As we are assuming that the offer arrival rate is independent of employment status, the important component here is $\rho$. Letting $\rho \to 0$ implies the $Mm$ ratio decreases and converges to $1/\beta$ (as in B/M). The higher the rate of human capital accumulation the lower the wage unemployed workers are willing to accept to enter employment and hence the higher the $Mm$ ratio.\footnote{Note, however, that our $Mm$ could also be too high or negative as the $\theta^R$ can potentially be negative. In principle this can be offset by allowing for different offer arrival rates. In particular, if the offer arrival rate for employed workers is lower than that of unemployed workers, the latter would not accept such a low wage to enter employment and $\theta^R$ should increase. Here we have considered the case of equal arrival rates for tractability.}

\section{Simulations}

We now perform some numerical simulations to illustrate the properties of the model and its implications for wage distributions. Using a year as a time unit, we set $\phi = 0.025$ so that workers have 40 years of expected working lifetime. Following Jolivet, Postel-Vinay and Robin (2006) estimates for the UK we set $\delta = 0.08$ and $\lambda = 0.126$.\footnote{The offer arrival rate we use is the one estimated for employed workers reported in Table 4.} Finally we set $\rho = 0.009$ and $b = 0.7$ to ensure $\theta^R > 0$ and a reasonable $Mm$ ratio. These set of parameters yield $\theta^R = 0.376$, $\bar{\theta} = 0.883$ and a $Mm$ ratio of 1.943.\footnote{The value of the $Mm$ ratio implies a $\beta = 0.95$.}
Given the parameter values specified above, Figure 2 presents the wage density when $A$ has a mass of one at $y_1 = 1$, i.e., workers are homogeneous initially in productivity. Note that the wage density exhibits similar properties as the empirical wage density. Namely, both are right skewed, exhibit an interior mode, an increasing left tail and a decreasing and long right tail. In our model, the increasing left tail of the wage distribution follows from the offer distribution and resembles the B/M wage density, while the decreasing right tail follows from the distribution of workers’ productivities. Wage dispersion reflects the interaction between workers’ on-the-job search and human capital accumulation.

We now consider worker heterogeneity. When the distribution of initial productivities is disperse, (21) implies the wage density is a mixture of the wage densities for each worker’s type. Assuming $A(.)$ describes a discrete distribution implies the wage density exhibits a number of “spikes” representing the modes of the type-specific wage densities. Increasing the number of types then generates a “smoother” wage density.\footnote{Although not shown here, this is confirmed by additional simulations.} To capture the latter feature we consider the case in which $A(.)$ is continuous. This case is also helpful to evaluate how changes in the variance of $A(.)$ affect wage dispersion in our model.
In what follows we assume that $A(\cdot)$ is distributed according to a Gamma distribution. We chose a Gamma distribution as it provides a simple way to analyse the effects of mean preserving spreads on the wage distribution. In this case the probability density function can be written as

$$A'(x \mid k_0, k_1) = \frac{\left(\frac{x}{k_0}\right)^{k_1-1} e^{-\left(\frac{x}{k_0}\right)}}{k_0 \Gamma(k_1)},$$

where $k_0 > 0$ is the scale parameter, $k_1 > 0$ is the shape parameter and $\Gamma(\cdot)$ is the gamma function. The mean and variance of $A(\cdot)$ are then given by $\mu = k_0 k_1$ and $\sigma^2 = (k_0)^2 k_1$. We consider values of $k_0$ and $k_1$ such that the shape of $A'$ roughly resembles the shape of the distribution of ex-ante worker heterogeneity used in Bontemps, Robin and Van den Berg (1999) and estimated by Postel-Vinay and Robin (2002).\(^{11}\) We apply three mean preserving spreads to $A(\cdot)$ by choosing $k_0$ and $k_1$ appropriately. Namely, (a) $k_0 = 6$, $k_1 = 3.3$; (b) $k_0 = 4$, $k_1 = 5$ and (c) $k_0 = 2$, $k_1 = 10$. These three cases and the resulting three wage densities are illustrated in Figures 3 and 4, respectively.

![Initial productivity densities](image)

Figure 3: Initial productivity densities

\(^{11}\)Since we consider values for $y_i \geq 1$, the values of $k_0$ and $k_1$ also ensure that the probability of $y_i < 1$ is zero.
It is apparent from these figures that introducing heterogeneity in initial productivities adds a new source of wage dispersion, increasing the amount of dispersion that can be generated solely due to workers’ human capital accumulation and firms’ piece rate offers. These figures also show that the overall shape of the wage density resembles that of the density of initial productivities. As $\sigma^2$ decreases, Figure 3 shows $A'(\cdot)$ becomes more symmetric around its mean. Figure 4 shows that a similar property holds for the wage density. More importantly, Figure 4 suggests that, compared with the wage density described in Figure 2, a disperse distribution of initial productivities can help to better approximate the shape of the empirical wage density. In particular, our choice of $A(\cdot)$ clearly improves the shape of the theoretical wage density over the middle range of the support.

The central interest here is to quantify the extent to which the implied distribution of wages paid depends on the underlying distribution of abilities $A(\cdot)$, learning by doing, piece rate dispersion $F(\cdot)$ and quit turnover as workers search for better paid employment. We do this by noting that for a worker with productivity $y = y_i e^{\beta x}$ on

\[\text{Figure 4: Wage density when } A \text{ is continuous.}\]
piece rate \( \theta \), that
\[
\log w = \log y_i + \rho x + \log \theta,
\]
where \( y_i \) and \( \theta \) can be interpreted as the worker and firm fixed effects, respectively. Note that our model then implies that these two fixed effects are orthogonal to each other.\(^{13}\) Moreover since the distribution of worker initial productivities is orthogonal to work experience, total variation in log-wages is given by
\[
\text{var}(\log w) = \text{var}(\log y_i) + \rho^2 \text{var}(x) + 2\rho \text{cov}(x, \log \theta).
\]
The first term of the right-hand side describes the variation due to worker heterogeneity in initial ability. The second term specifies the variance due to firms offering different piece rates, the third due to workers having different productivities through learning by doing. The last term is due to on-the job search: more experienced workers, having spent more time in the labour market, tend to earn higher piece rates (recall the conditional distribution of earned piece rates is first order stochastically increasing in experience).

Table 1 decomposes the wage variation described in Figure 4 into its various components. To construct this table we computed each of the components of \( \text{var}(\log w) \). Their total sum is reported in the second column, while their relative contribution to this sum is given in the subsequent columns. In particular, (18) implies the variance of firms’ pay rate policies is 0.053 and (13) implies the variance of human capital accumulation and on-the-job search is 0.055 and 0.26, respectively.\(^{14}\) The variance of the log of initial productivities is 0.206 with density (a), 0.152 with density (b), and 0.087 with density (c). The final row illustrates the decomposition when initially all workers are the same.

\(^{13}\)Using matched employer-employee data for France and the state of Washington (US), Abowd, Kramarz and Margolis (1999) and Abowd, Finer and Kramarz (1999), respectively, found that workers and firms fixed effects are not correlated with each other. See Mortensen (2003) for a further discussion on this topic.

\(^{14}\)These two values correspond to \( \rho^2 \text{var}(x) \) and \( 2\rho \text{cov}(x, \log \theta) \), respectively.
An interesting feature of Table 1 is that the covariance of workers’ experience and firms’ piece rate policies represents a large proportion of the total variance in distribution of log wages. An alternative way to measure the impact of the interaction between work experience, on-the-job search, and firms’ piece rate policies is by analyzing the lifetime earnings profiles of workers implied by the model. To do so, we simulate workers’ employment histories considering the case when $A$ has a mass of one at $y_1 = 1$ and use this data to estimate

$$\log w = \beta_0 + \rho x + \epsilon,$$

where $\epsilon$ is the error term. Given that this equation does not incorporate the firm specific component, $\log \theta$, the estimate of $\rho$, $\hat{\rho}$, should be biased upwards as $\epsilon$ is correlated with $x$ because on average more experience workers will be employed in higher paying jobs. The difference $\hat{\rho} - \rho$ then gives an approximation of the effect on-the-job search has on workers’ experience profiles. Moreover, as in our model similar workers are paid differently due to on-the-job search, omitting the firm specific component from the above equation also implies that the coefficient of determination, $R$-squared, should give an additional measure of the importance of firms piece rate policies in explaining the variation of log wages.

Using the parameter specification described above our simulations show that on average $\hat{\rho} = 0.011$. This generates an upward bias of $\hat{\rho} - \rho = 0.002$, implying that

\begin{table}
\begin{center}
\begin{tabular}{|l|c|c|c|c|c|}
\hline
Density of $A$ & Total variation & Relative contribution (%) \\
& $\text{var}(\log w)$ & $\text{var}(\log y_i)$ & $\text{var}(\log \theta)$ & $\rho^2 \text{var}(x)$ & $2\rho \text{cov}(x, \log \theta)$ \\
\hline
(a) & 0.574 & 35.89 & 9.23 & 9.58 & 45.30 \\
(b) & 0.52 & 29.23 & 10.19 & 10.58 & 50.00 \\
(c) & 0.455 & 19.12 & 11.65 & 12.09 & 57.14 \\
Degenerate & 0.368 & 0 & 14.40 & 14.95 & 70.65 \\
\hline
\end{tabular}
\end{center}
\end{table}

\footnote{This is robust to other specifications of $A$ and values for $\rho$ and $\lambda$. The alternative specifications of $A$ we used are: Generalised Pareto, Three parameter Weibull and Uniform.}

\footnote{This result is based on the average of 10 cross-sectional samples each of approximately 7,500 employed workers.}
after 30 years of work experience average wages are biased upwards by 8.1 percent due to the correlation between experience and on-the-job search.\textsuperscript{17} In addition, the simulations show that only controlling for experience explains on average 45 percent of total variation in log wages. Introducing differences in earned piece rate explains an additional 53 percent, while induced measurement error explains the remainder 2 percent.

8 Conclusion

In this paper we have constructed and analysed a labour market equilibrium in which there is on-the-job search and workers accumulate general human capital through learning by doing. The resulting equilibrium wage distribution reflects the interaction between sorting dynamics as workers quit for better paid employment, worker productivity grows through learning-by-doing, and the distribution of initial productivities. Together this behaviour generates a wage density that exhibits an increasing left tail and a decreasing long and fat right tail, very much like the one observed in the data. Thus on-the-job search and learning-by-doing would seem a useful approach in explaining empirical wage distributions. This approach also yields a simple and insightful variance decomposition of wages.

Our next aim is to use the arguments developed here to generalise the results of Burdett and Coles (2003) to learning by doing. That paper supposes all firms and workers are ex-ante identical but, in contrast to B/M, firms post contracts where wages paid depend on tenure. It was shown an equilibrium exists where firms offer different contracts, but each firm offers a contract where wages paid increase smoothly with tenure (also see Stevens, 2004). In such a market environment, workers are promoted by seniority: as more senior employees quit or retire, junior workers are promoted to take their place. Furthermore firms ignore outside offers: a junior employee is never promoted above a more senior and no less productive employee. On-the-job learning and a piece rate tenure contract \( \theta = \theta(\tau) \) then yields a wage

\textsuperscript{17}This bias is of the same order of magnitude Williams (2004) found for UK data when estimating returns to experience using OLS and one of the techniques proposed by Topel (1991) to correct for the correlation between experience and workers’ on-the-job search.
equation of the form:

$$\log w_{ijr} = \log y_i + \rho x_{it} + \log \theta_j(\tau),$$

where $\theta_j(.)$ is now the tenure contract offered by firm $j$. Such an extension makes clear the difficulty when trying to decompose wages into experience and tenure effects: the econometrician needs to disentangle the firm fixed effect from firm specific tenure effects. The objective for theory is to identify how this might be done. The arguments developed here provide the necessary techniques for attempting such an extension.

References


APPENDIX

Proof of Claim 5:

(i) Claim 4 implies
\[ \int_{y' = y_i}^{\infty} y'dN_i(y') = N_i(1) + \int_{y' = y_i}^{\infty} \left( \frac{\phi(\phi + \delta + \lambda)}{\rho(\phi + \lambda)} \right) \lambda \delta \frac{dy'}{(\phi + \lambda)(\phi + \delta)} \]
and integrating yields the solution stated in the Claim.

(ii) Integrating over \( \theta' \) implies
\[ \int_{y' = y_i}^{\infty} \int_{\theta' = \theta}^{\theta} y'' \partial^2 H_i(y', \theta') \frac{dy'}{\partial y' \partial \theta'} d\theta' = \int_{y' = y_i}^{\infty} y' \left[ \frac{\partial H_i(y', \theta')}{\partial y'} \right]^\theta_{\theta} dy' \]

As \( F(\theta) = 0 \), (13) then implies \( H_i(y', \theta) = 0 \), and (10) then implies \( \frac{\partial H_i(y', \theta)}{\partial y'} = 0 \), this reduces to
\[ \int_{y' = y_i}^{\infty} \int_{\theta' = \theta}^{\theta} y'' \frac{\partial^2 H_i(y', \theta')}{\partial y' \partial \theta'} d\theta' dy' = \int_{y' = y_i}^{\infty} y' \frac{\partial H_i(y', \theta)}{\partial y'} dy' \]

Now (13) implies
\[ \frac{\partial H_i}{\partial y} = \left( \frac{\phi + \delta}{\rho} \right) \frac{F(\theta)}{q(\theta)} y^{-\frac{q(\theta)}{\rho}} + \frac{\delta F(\theta)}{q(\theta)} \left[ \left( \frac{\phi(\phi + \delta + \lambda)}{\rho(\phi + \lambda)} \right) y^{-\frac{(\phi(\phi + \delta + \lambda))}{\rho(\phi + \lambda)}} - \frac{q(\theta)}{\rho} y^{-\frac{q(\theta)}{\rho} - 1} \right]. \]

Using this in the previous expression and integrating, noting that \( \phi > \rho \), yields
\[ \int_{y' = y_i}^{\infty} \int_{\theta' = \theta}^{\theta} y'' \frac{\partial^2 H_i(y', \theta')}{\partial y' \partial \theta'} d\theta' dy' = \frac{\phi F(\theta)}{q(\theta) - \phi F(\theta)} \left[ 1 - \frac{\delta F(\theta)}{q(\theta) - \phi F(\theta)} \right] \]
\[ + \frac{\delta F(\theta)}{q(\theta) - \phi F(\theta)} \left[ \frac{\phi(\phi + \delta + \lambda)}{\phi(\phi + \delta + \lambda) - \rho(\phi + \lambda)} \right]. \]

Collecting terms and simplifying then yields the second expression in the Claim. This completes the proof of Claim 5.

Proof of Theorem 1 - Derivation of \( \theta^R \):

First note that (6) and (7) in Claim 2 imply \( \theta^R \) is determined by
\[ \theta^R = \frac{b(\phi - \rho)}{\phi} - \frac{\rho}{\phi} \int_{\theta^R}^{\pi} \frac{\lambda (1 - F(\theta))}{q(\theta) - \rho} d\theta, \tag{23} \]
for any given \( F \). Fix a \( \theta^R \) and let \( \tilde{F}(\theta \mid \theta) \), where \( \tilde{F}(\theta \mid \theta) \) is described by (18).

Substituting for \( F \) in the expression describing \( \theta^R \) yields
\[ \theta^R(\theta) = \frac{b(\phi - \rho)}{\phi} - \frac{\rho(\bar{\theta} - \theta^R(\theta))}{\phi} + \frac{\rho(\phi + \delta - \rho)}{\phi(\phi + \delta - \rho + \lambda)} \int_{\theta^R}^{\pi} \left( \frac{1 - \theta}{1 - \theta} \right)^{1/2} d\theta \]
\[ = \frac{b(\phi - \rho)}{\phi} - \frac{\rho(\bar{\theta} - \theta^R(\theta))}{\phi} + \frac{2\rho(\phi + \delta - \rho)(1 - \theta)^{1/2}}{\phi(\phi + \delta - \rho + \lambda)} \left[ (1 - \bar{\theta})^{1/2} - (1 - \theta^R(\theta))^{1/2} \right]. \]
Noting that (18) implies

\[ \bar{\theta} = 1 - \left( \frac{\phi + \delta - \rho}{\phi + \delta - \rho + \lambda} \right)^2 (1 - \theta), \]

solving for \( \theta^R(\bar{\theta}) = \bar{\theta} \) then gives the expression in (19).

**Derivation of the wage density:**

Differentiating (20), we obtain the joint density of workers characteristics, \( y \) and \( \theta \), for each \( i \)

\[
h_i(y, \theta) = \frac{\phi F'(\theta)}{y_i \rho} \left( \frac{y}{y_i} \right)^{-\frac{q(\theta)}{\rho} - 1} \left[ 1 + \frac{\lambda}{\rho} F(\theta) \ln(y) \right] \left[ (\phi + \delta + \lambda)(1 - F(\theta)) \right] \\
+ \frac{\delta \phi (\phi + \delta + \lambda) F'(\theta)}{y_i \rho (q(\theta) - \phi F'(\theta))^{\frac{2}{\phi + \lambda}}} \left[ \left( \frac{\phi + \delta + \lambda}{\phi + \lambda} \right) \left( \frac{y}{y_i} \right)^{-\frac{q(\theta) + (\phi + \delta + \lambda)}{\rho (\phi + \lambda)} - 1} - F(\theta) \left( \frac{y}{y_i} \right)^{-\frac{q(\theta)}{\rho} - 1} \right],
\]

where \( h_i(y, \theta) \geq 0 \) for all \( \theta \in [\bar{\theta}, \bar{\theta}] \) and \( y \geq y_i \). Otherwise, \( h_i(y, \theta) = 0 \). We use this density function as a primitive to derive the next set of distributions. As wages are given by \( w = \theta y \) and noting we are considering cases in which \( \bar{\theta} > 0 \), we have that \( \hat{h}(w, \theta) = h_i\left( \frac{w}{\theta y_i}, \theta \right) \frac{1}{\theta} \), where \( \hat{h}(w, \theta) \) describes the joint density of \( w, \theta \). Using the expression for \( h_i \) then gives

\[
\hat{h}(w, \theta) = \frac{\phi F'(\theta)}{\theta y_i \rho} \left( \frac{w}{\theta y_i} \right)^{-\frac{q(\theta)}{\rho} - 1} \left[ 1 + \frac{\lambda}{\rho} F(\theta) \ln\left( \frac{w}{\theta y_i} \right) + \ln(y_i) \right] \left[ (\phi + \delta + \lambda)(1 - F(\theta)) \right] \\
+ \frac{\delta \phi (\phi + \delta + \lambda) F'(\theta)}{\theta y_i \rho (q(\theta) - \phi F'(\theta))^{\frac{2}{\phi + \lambda}}} \left[ \left( \frac{\phi + \delta + \lambda}{\phi + \lambda} \right) \left( \frac{w}{\theta y_i} \right)^{-\frac{q(\theta) + (\phi + \delta + \lambda)}{\rho (\phi + \lambda)} - 1} - F(\theta) \left( \frac{w}{\theta y_i} \right)^{-\frac{q(\theta)}{\rho} - 1} \right],
\]

for all \( \theta \in [\bar{\theta}, \bar{\theta}] \) and \( w \geq \theta y_i \). Otherwise, \( \hat{h}(w, \theta) = 0 \). Integrating appropriately over \( \theta \) then yields

\[
G'_i(w) = \int_{\frac{w}{\theta y_i}}^{\min(\bar{\theta}, \frac{w}{y_i})} \hat{h}(w, \theta) d\theta \text{ for all } w \geq \theta y_i,
\]

otherwise.

Finally summing up over all possible workers’ types gives the aggregate wage density of the economy

\[
G'(w) = \sum_{i=1}^{I} \gamma_i \int_{\frac{w}{\theta}}^{\min(\bar{\theta}, \frac{w}{y_i})} \hat{h}(w, \theta) d\theta.
\]

**Derivation of the mean-min ratio:**

Since frictional wage dispersion concerns wage dispersion that is not driven by difference is abilities without loss of generality consider the case in which all workers
enter with initial productivity $y_1 = 1$. Next note that $H(\infty, \theta)$ describes the distribution of piece rates across employed workers given the offer distribution $F$. Using integration by parts and $\bar{\theta} = \theta^R$, it can be easily shown that the average piece rate earned by employed workers, $\tilde{\theta}$, is given by

$$\tilde{\theta} = \theta^R + \int_{\theta^R}^{\infty} [1 - H(\infty, \theta)]d\theta.$$ 

Putting $y = \infty$ in (13) implies

$$H(\infty, \theta) = \frac{(\phi + \delta)F(\theta)}{q(\theta)}.$$ 

Since $\rho < \phi$ by assumption and for reasonable parametrizations of the model $\phi$ is at least twice as large as $\rho$ (see the Simulations section in the main body of the paper), we follow Hornstein, Krusell and Violante (2007) and approximate $H(\infty, \theta)$ by

$$H(\infty, \theta) \simeq \frac{(\phi - \rho + \delta)F(\theta)}{q(\theta) - \rho}.$$ 

Solving for $1 - F(\theta)$ and substituting the resulting expression into (23) yields

$$\theta^R \simeq \frac{b(\phi - \rho)}{\phi} - \frac{\lambda \rho}{\phi(\phi - \rho + \delta + \lambda)} \int_{\theta^R}^{\infty} [1 - H(\infty, \theta)]d\theta$$

$$\simeq \frac{b(\phi - \rho)}{\phi} - \frac{\lambda \rho}{\phi(\phi - \rho + \delta + \lambda)} (\tilde{\theta} - \theta^R).$$ 

Letting $b = \beta \tilde{\theta}$ and then solving for $\tilde{\theta}/\theta^R$ yields the $Mm$ ratio shown in the main body of the paper.