Job Search, Human Capital and Wage Inequality

Carlos Carrillo-Tudela ^{*†}

February 2010

Abstract

The objective of this paper is to construct and quantitatively assess an equilibrium search model with on-the-job search and human capital accumulation. In the model workers enter the labour market with different abilities and firms differ in their labour productivities. Wages are dispersed because of search frictions (firms pay workers of the same productivity different wages) and workers' productivity differentials (workers of different productivities earn different wages). Further, there is positive sorting between workers and firms and this increases wage dispersion. The model generates a simple (log) wage variance decomposition that is used to measure the importance of productivity differentials, search frictions and sorting dynamics between workers and firms. I calibrate the model to match spell durations and wage variation of a (relative) homogeneous sample of workers using UK household level data. I show that wage variation due to productivity difference explains around 60 percent, search frictions around 25 percent and sorting dynamics the remainder 15 percent. The model is then used to analyse the average wage-experience profile of workers and shows the importance of human capital accumulation in shaping such a profile.

Keywords: Search, wage dispersion, contracts, turnover

JEL: J63, J64, J41, J42.

^{*}Correspondence: Department of Economics, University of Leicester, University Road, Leicester, LE1 7RH; email: cct9@le.ac.uk

[†]I would like to thank Ken Burdett, Guido Menzio, Iourri Manvoski and Gregory Jolivet for their comments and insights. I would also like to thank Annette Jäckle and Mark Bryan for their comments and help with the data. Most of this research was conducted while I visited the Department of Economics of the University of Pennsylvania and the University of Essex. I thank these institutions for their hospitality. The usual disclaimer applies.

1 Introduction

Differentials in worker productivities have long been recognized as an important source of wage inequality. Since the pioneering work of Becker (1964) and Mincer (1962), human capital theory has been used as the norm to analyse the wage growth of workers over the life cycle and the cross sectional wage distribution. Under this theory, wage differentials accrue through difference in innate ability, schooling and the accumulation of general and firm specific skills. Recent evidence suggests, however, that along side productivity differentials a significant proportion of the observed wage dispersion is also due to the frictional nature of labour markets. In particular, evidence finds that equally productive workers are paid differently and that these pay differentials are large and persistent (see Abowd, Kramarz and Margolis, 1999, Mortensen, 2003, and Hornstein, Krusell and Violante, 2009). Search theory has provided an elegant and powerful explanation for this phenomena. Firms differentiate their wage policies as an optimal response to workers' job shopping behavior. In spite of the prominence of these two explanations, it is surprising to find how little work has been done in constructing a unified framework to jointly analyse and quantitatively assess the effects of human capital differentials and search frictions on wage inequality. Indeed, without such a framework it seems difficult to gauge the relative importance of these theories and the impact of labour market policies that aim to reduce wage inequality.

The present paper considers an equilibrium search model somewhat similar to Burdett and Mortensen (1998) (henceforth B/M) in which workers with different abilities search and learn on the job and employers differ in their labour productivities. The main objective is to analyse the relative contribution of search frictions and human capital accumulation on the cross sectional distribution of wages and the average wage-experience profile. In the model, workers enter the labour market with different abilities and accumulate general human capital through learning-by-doing while employed (see Rosen, 1972). The labour market is frictional in that there is imperfect information about the location of job opportunities and workers are constantly searching for jobs. Firms, on the other hand, differ in their labour productivities and operate under a constant returns to scale technology. Each firm pays their employees on a piece rate basis and offers the same piece rate to any worker it meets. Frictional wage dispersion arises in equilibrium due to the same reason as in B/M. Firms differentiate their pay policies as an optimal reaction to workers' on-the-job search behavior. Hence, at any point in time, wages are dispersed because of differentials in worker and firm productivities and earned piece rates. Further, since in equilibrium more productive firms offer higher paying jobs, wages increase over time because (i) workers becomes more productive while employed and (ii) they move from less to more productive jobs when the opportunity arises.

The model highlights a new source of wage dispersion that arises due to workers' sorting dynamics. In particular, the interaction of on-the-job search and learning-by-doing implies that more productive workers end up employed in more productive firms; and more productive firms end up employing a more productive workforce. Wages become more disperse

because of the positive correlation between workers' experience, human capital accumulation and the productivity of their jobs. Furthermore, these sorting dynamics also affect frictional wage dispersion. Since employment provides two source of wage growth (accumulating human capital and climbing the firm productivity distribution), workers are willing to reduce their reservation wages and leave unemployment as quickly as possible. In equilibrium, however, low productivity firms react to this behavior by offering even lower paying jobs and thus increasing the wage dispersion across equally productive workers.

The model provides a remarkable simple decomposition of the variation of (log) wages into its constituent parts. This decomposition is one of the key element for the quantitative analysis and it is used to assess the impact of human capital accumulation, search frictions and sorting dynamics on the cross sectional wage distribution in the UK. To do so I calibrate the model using the labour market histories of a sample of full-time, white, male, private sector workers drawn from the British Household Panel Survey (BHPS) for the period 1991-2004. I further subdivide this sample into low, medium and high skill groups based on the individuals' educational qualifications. For each skill group the calibration targets (i) the observed average job, employment and unemployment spell duration and potential work experience; (ii) Hornstein et al. (2009) measure of frictional wage dispersion, the Mean-min ratio; and (iii) the kernel estimate of the cross sectional (hourly) wage density distribution. Applying the variance decomposition to the latter, I show that across skill groups around 60 percent of variation in (log) wages is due to productivity differentials between workers (initial ability and human capital accumulation), 25 percent of the variation is due to search frictions and the remainder 15 percent is due to workers' sorting dynamics. I also show that the latter increases frictional wage dispersion by around 55 percent from a model in which workers only search on the job.¹

To explore the model's implications for the average wage-experience profile, I simulate workers employment histories using the model as the data generating process given the parameters obtained from the variance decomposition. This exercise is useful as it gives a sense of how much of the wage-experience profile is due to employed workers' job shopping behaviour. Here the model produces mixed results. The model reproduces quite well the average returns to experience for the first 15 years of the worker's career, but overestimates the the rest. The main reason for this is that, although on-the-job search generates a concave wage-experience profile (see Burdett, 1978), the degree of concavity generated by job to job transitions is to low.

Although the model does not consider wage-tenure contracts (see Stevens, 2004, and Burdett and Coles, 2003) or accumulation of firm specific skills (see Hashimoto, 1982), the model generates positive tenure effects when estimating a Mincer wage equation on simulated data. These tenure effects are purely driven by the fact that (in the model) workers employed in high paying jobs will be less likely to quit and hence have longer tenures. There is a

¹Since the Diamond (1971) outcome applies to this model when we shut down workers on-the-job search; i.e. there is no frictional wage dispersion. The mean-min ratio is one in that case.

prominent literature that for the past 20 years has been trying to establish how much of the observed tenure effects in the cross-sectional or longitudinal data is due to the process identified here (see Abraham and Farber, 1987, Altonji and Shakotko, 1987, Topel, 1991, Altonji and Williams, 2007, Dustmann and Meghir, 2005, among others). The model's simulations indicate that on-the-job search can generate a substantial bias, explaining at least 50 percent of the observed tenure effects in the BHPS sample considered here. In this sense, the simulations give support to those studies that obtain small "true" returns to tenure.

A closely related paper is that of Burdett, Carrillo-Tudela and Coles (2009), who also incorporate human capital accumulation into the B/M framework. I extend that framework by introducing heterogeneity in firm productivities and letting the arrival rate of job offers depend on workers' employment status. Further, I provide a quantitative assessment of the model's implications. Adding these extensions is important for my purposes as there is now a great deal of evidence both showing that firm productivity dispersion is large and persistent (see Bartelsman and Doms, 2000, and Lentz and Mortensen, 2008) and that employed workers receive job offers at a different (lower) rate than unemployed workers (see Jolivet, Postel-Vinay and Robin, 2006). From a theoretical standpoint I provide a full analytical characterisation of the (more general) model and show how the introduction of firm heterogeneity generate the possibility of multiple equilibria.

There are also a few other papers that have investigated learning-by-doing effects within a search environment. Bunzel, Christensen, Kiefer and Korsholm (2000) analyzed a model of human capital accumulation using the B/M framework. However, they assume workers are initially homogeneous, all firms have the same productivity and workers lose all their human capital when laid off. This leads to very different results. Rubinstein and Weiss (2007), Barlevy (2008) and Liu (2009) estimate the wage process identified here but do not consider equilibrium. Fu (2009) considers a similar model, but studies whether homogeneous firms can provide general human capital for its employees or not.

Bagger, Fontaine, Postel-Vinay and Robin (2006) and Yamaguchi (2009) instead extend the offer matching framework developed by Postel-Vinay and Robin (2002a,b) to incorporate learning on-the-job with individual productivity shocks. Their focus is on estimating the resulting wage process over a worker's life cycle. Here I do not allow for productivity shocks and assume that a worker simply quits if he/she receives a preferred outside offer. Although offer matching arises in certain markets, such as the academic market for economists, it may not be a good description of behavior in other labour markets (see for example Mortensen, 2003). Further, the focus of this paper is to study wage dynamics and wage dispersion. I provide a tractable analysis with close form solutions of the endogenous steady state distributions that describe the interaction between human capital accumulation and on-the-job search.

The paper is outlined as follows. Section 2 describes the model. Section 3 defines and

characterises the equilibrium. Sections 4 and 5 analyse equilibrium sorting and the equilibrium distribution of wages. Section 6 and 7 describe the data, the calibration procedure and show the results. Section 7 concludes. All proofs are relegated to a technical Appendix.

2 The Model

We assume time is continuous with an infinite horizon and the economy is in a steady state. There is a continuum of both firms and workers, each of measure one. 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 ϕdt . ϕ also describes the inflow of new labour market entrants.

Assume that workers differ on their initial ability ε . In particular, let A denote the cumulative distribution function of these abilities with positive support $[\underline{\varepsilon}, \bar{\varepsilon}]$. 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 ε worker's productivity is $y = \varepsilon e^{\rho x}$. An unemployed worker's productivity y remains constant through time. Firms, on the other hand, operate using a constant returns to scale technology and differ in their productivity, p . Let Γ denote the cumulative distribution function of firms' productivities with finite and positive support $[\underline{p}, \bar{p}]$. A worker with productivity y employed at a firm with productivity p generates flow output yp while employed. I normalise the price of the production good to one, so yp also describes flow revenue.²

A firm pays each of its employees the same piece rate θ . Thus given an employee with productivity y in a firm with productivity p , the worker is paid flow wage $w = yp\theta$. A type p firm's profit flow is given by $yp(1 - \theta)$. Letting $z = p\theta$ denote the firm's specific component to a worker's wage, it will be useful to describe job offers in terms of z rather than on θ . Although formally a type p firm offers θ , a worker decides whether to accept a job offer based on z . Given this re-normalisation, the associated wage is then $w = yz$, while the firm's profit flow is given by $y(p - z)$. Let $F(z | p)$ denote the proportion of firms of type p offering a z' no greater than z . Integrating across firm types then yields $F(z)$, the proportion of all firms offering a z' no greater than z . Further, let $\underline{z}, \bar{z}$ denote the infimum and supremum of the support of F .

Assume that workers and firms meet randomly. In particular, each unemployed and employed worker receive job offers according to a Poisson process with parameter $\lambda_u > 0$ and $\lambda_e > 0$, respectively. Random matching then implies $F(z)$ describes the probability of receiving an offer z' no greater than z . There are also job destruction shocks in that

²In principle, the rate of on-the-job learning can depend on the ability of the worker and the productivity of the firm he is employed. In the theory section of the paper, I avoid this complication to keep the model tractable. However, in the quantitative section I consider three distinct labour markets based on the workers' educational level and obtain the value of ρ for each of them. To the extent that difference in the rate of human accumulation across workers are explained by differences their school attainment, then the model is able to capture such differences.

each employed worker is displaced into unemployment according to a Poisson process with parameter $\delta > 0$. I assume that unemployment is equivalent to working in home production, where the latter can be thought of as a firm with productivity $p_h \in (0, \bar{p})$ offering a piece rate of $b \leq 1$ and $z_b = bp_h$.³ A worker with productivity y then enjoys a flow payoff yz_b while unemployed.

I 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. If a job offer is rejected, the worker remains in his/her current state and there is no recall. Let $\phi > \rho$ to ensure total expected lifetime payoffs are finite.

All agents are risk neutral. For simplicity assume workers do not discount the future - they maximize expected lifetime income. Each firm maximize steady state flow profit, taking into account the search strategies of workers by choosing a z . The corresponding piece rate is then given by $\theta = z/p$.

2.1 Optimal Search Strategies

For a given offer distribution F , consider optimal worker behavior. Let $W^U(y)$ denote the expected lifetime payoff of an unemployed worker with productivity y using an optimal search strategy. Let $W^E(y, z)$ denote the expected lifetime payoff of a worker with productivity y , currently employed at a firm offering z , when using an optimal search strategy. As $W^E(y, z)$ is strictly increasing in z an employed worker will quit to an outside offer z' if and only if $z' > z$. Thus, the flow Bellman equation for employed workers implies:

$$(\phi + \delta)W^E(y, z) = zy + \rho y \frac{\partial W^E}{\partial y} + \lambda_e \int_z^{\bar{z}} [W^E(y, z') - W^E(y, z)] dF(z') + \delta W^U(y). \quad (1)$$

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

$$\phi W^U(y) = z_b y + \lambda_u \int_z^{\bar{z}} \max[W^E(y, z') - W^U(y), 0] dF(z'). \quad (2)$$

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

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

Moreover, the Bellman equation (2) implies the unemployed worker's optimal strategy is to

³For a similar assumption see Postel-Vinay and Robin (2002b).

accept any offer z' satisfying $W^E(y, z') \geq W^U(y)$. As W^E is increasing in z' , the worker, will accept any offer $z' \geq z_R$ where z_R is given by $\alpha^E(z_R) = \alpha^U$. The crucial property here is that all unemployed workers have the same reservation value z_R . The next result characterises, $\alpha^U, \alpha^E(\cdot)$ and z_R . It is convenient first to define

$$q(z) = \phi + \delta + \lambda_e(1 - F(z)),$$

which describes the separation rate of a worker employed at a firm offering a payoff of z .

Proposition 1: *Optimal job search implies*

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

$$\frac{d\alpha^E}{dz} = \frac{1}{q(z) - \rho}, \quad (3)$$

with boundary condition $\alpha^E(z) = (\bar{z} + \delta\alpha^U)/(\delta + \phi - \rho)$ at $z = \bar{z}$,

(ii) (α^U, z_R) satisfy the pair

$$\rho\alpha^U = z_b - z_R + (\lambda_u - \lambda_e) \int_{z_R}^{\bar{z}} \frac{1 - F(z)}{q(z) - \rho} dz, \quad (4)$$

$$\phi\alpha^U = z_b + \lambda_u \int_{z_R}^{\bar{z}} \frac{1 - F(z)}{q(z) - \rho} dz. \quad (5)$$

Further, $\bar{z} > z_b(\phi - \rho)/\phi$ is sufficient to guarantee that a solution exists, is unique and implies $\alpha^U > 0$ and $z_R < \bar{z}$.

2.2 Firms Profits

Now consider firm behavior given an F and $z_R < \bar{z}$ solving the conditions in Proposition 1. Two important remarks are in order: (i) Given z_R and Γ , there may exist a set of firms with productivity $p \in [\underline{p}, z_R)$ that would obtain negative profits if they offered jobs. Since not offering jobs yields zero profit, these firms will not be active in the labour market. Let Γ_0 denote the productivity distribution of active firms such that

$$\Gamma_0(p) = \frac{\Gamma(p) - \Gamma(p_0)}{1 - \Gamma(p_0)} \quad (6)$$

describes the probability that an active firm has productivity $p' \leq p$, where $p_0 = \max\{z_R, \underline{p}\}$ denotes the lowest productivity of an active firm and $1 - \Gamma(p_0)$ denotes the measure of active firms. (ii) $F(\cdot | p)$ and F are defined with respect to the productivity distribution of active firms.

Define three steady-state objects - (a) U_ε : the fraction of type ε workers who are unemployed, (b) $N_\varepsilon(\cdot)$: the distribution function describing productivities across type ε workers who are unemployed, and (c) $H_\varepsilon(y, z)$: the joint distribution function describing produc-

tivities and earned z across type ε workers who are employed. Next consider a firm with productivity p offering $z \geq z_R$. As there is no discounting, steady state flow profit equals the hiring rate of the firm, multiplied by the expected profit of each hire. This firm's steady state flow profit is given by

$$\Omega(z; p) = \frac{p - z}{q(z) - \rho} \int_{\underline{\varepsilon}}^{\bar{\varepsilon}} \left[\lambda_u U_\varepsilon \int_{y'=\varepsilon}^{\infty} y' dN_\varepsilon(y') + \lambda_e (1 - U_\varepsilon) \int_{z'=\underline{z}}^z \int_{y'=\varepsilon}^{\infty} y' dH_\varepsilon(y', z') \right] dA(\varepsilon). \quad (7)$$

Such a firm then chooses z to maximise $\Omega(z; p)$. Since $\theta = p/z$, the choice of z then gives the corresponding offered piece rate θ . Let $\bar{\Omega}(p) = \max \Omega(z; p)$ for each $p \geq p_0$. Of course, in equilibrium the optimal choices of z for each firm with $p \geq p_0$ must imply a distribution $F(z)$ consistent with the optimal strategy of workers as described by z_R . I now formally define such an equilibrium.

3 Market Equilibrium

A Market Equilibrium is a set $\{z_R, U_\varepsilon, N_\varepsilon(\cdot), H_\varepsilon(\cdot, \cdot), F(\cdot \mid p)\}$ for all $\varepsilon \in [\underline{\varepsilon}, \bar{\varepsilon}]$ and $p \in [p_0, \bar{p}]$ such that

Given optimal worker search strategies as described by z_R :

- (i) The productivity distribution of active firms is given by (6) where $p_0 = \max\{z_R, \underline{p}\}$.
- (ii) Given Γ_0 , the constant profit condition is satisfied for all active firms; i.e.,

$$\begin{aligned} \Omega(z; p) &= \bar{\Omega}(p) & \text{for all } z \text{ where } dF(z \mid p) > 0; \\ \Omega(z; p) &\leq \bar{\Omega}(p) & \text{for all } z \text{ where } dF(z \mid p) = 0, \end{aligned}$$

for all $p \in [p_0, \bar{p}]$.

- (iii) Given Γ_0 and $F(\cdot \mid p)$, the distribution $F(\cdot)$ satisfies

$$F(z) = \int_{p_0}^{\bar{p}} F(z \mid p) d\Gamma_0(p).$$

- (iv) $U_\varepsilon, N_\varepsilon(\cdot)$ and $H_\varepsilon(\cdot, \cdot)$ are consistent with steady state turnover given $F(\cdot)$.

Given such an F :

- (v) z_R solves the conditions in Proposition 1.

Before solving for a Market Equilibrium consider some preliminary remarks. Note that since active firms have productivities $p \geq z_R$, any active firm of type p will not choose a $z > p$ (and hence $\theta > 1$) as doing so will yield negative profits. It then follows that equilibrium steady state profit flows $\bar{\Omega}(p) > 0$ for all $p > p_0$ and $\bar{\Omega}(p) \geq 0$ at $p = p_0$. Further, when the least generous firms have productivity $p > z_R$, equilibrium implies $\underline{z} = z_R$. It is not optimal to offer a $z > z_R$, as these firms can strictly increase flow profit by reducing z without affecting their hiring or retention rates. Similarly, it is not optimal to offer a $z < z_R$ as this

z does not attract any unemployed worker and this implies zero profits. When the least generous firms have productivity $p = z_R$ offering $z > z_R$ implies negative profits. However, these firms are indifferent between offering a $z \leq z_R$, as these offers yield zero profits. In this case, consider equilibria in which $\underline{z} = z_R$.

Finally, the arguments described in B/M to establish that any equilibrium $F(\cdot | p)$ and F must be continuous (no mass points) and have connected support also apply here. To obtain this result, however, one must rule out productivity distributions in which there is no positive mass of firms with productivity $p = z_R$. For simplicity, the present analysis abstracts from these type of productivity distributions.

3.1 Steady state measures

Given Γ_0 and the implied equilibrium F , I now construct the three steady state objects defined above, U_ε , $N_\varepsilon(\cdot)$ and $H_\varepsilon(\cdot, \cdot)$. First consider the steady state pool of type ε unemployed workers. It is straightforward to verify that steady state turnover implies that the unemployment rate for type ε worker is

$$U_\varepsilon = \frac{\phi + \delta}{\phi + \delta + \lambda_u}.$$

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

The next result solves for the measure of type ε unemployed workers with productivity no greater than y (and $y \geq \varepsilon$) and the pool of type ε employed workers who have productivity no greater than y and receive a payoff no greater than z .

Lemma 1: *A Market Equilibrium implies distribution functions*

$$N_\varepsilon(y) = 1 - \frac{\lambda_u \delta}{(\phi + \lambda_u)(\phi + \delta)} \left(\frac{y}{\varepsilon}\right)^{-\left(\frac{\phi(\phi+\delta+\lambda_u)}{\rho(\phi+\lambda_u)}\right)} \text{ for all } y \geq \varepsilon, \quad (8)$$

$$H_\varepsilon(y, z) = \frac{(\phi + \delta)F(z)}{q(z)} \left[1 - \left(\frac{y}{\varepsilon}\right)^{-\frac{q(z)}{\rho}}\right] - \frac{\delta F(z)}{q(z) - \phi F(z)} \left[\left(\frac{y}{\varepsilon}\right)^{-\left(\frac{\phi(\phi+\delta+\lambda_u)}{\rho(\phi+\lambda_u)}\right)} - \left(\frac{y}{\varepsilon}\right)^{-\frac{q(z)}{\rho}}\right] \quad (9)$$

for all $z \in [\underline{z}, \bar{z}]$ and $y \geq \varepsilon$.

3.2 Characterisation

A Market Equilibrium requires that a firm of productivity $p \in [p_0, \bar{p}]$ chooses a $z \leq p$ to maximise $\Omega(z; p)$. Using the solutions for $N_\varepsilon(\cdot)$ and $H_\varepsilon(\cdot, \cdot)$ identified in Lemma 1 to solve for the integrals in (7), steady state profits can be expressed as

$$\Omega(z; p) = \tilde{\varepsilon} l(z)(p - z) = \tilde{\varepsilon} \frac{k_0}{(q(z) - \rho)^2} \left[k_1 - k_2 \frac{(\lambda_u - \lambda_e)F(z)(1 - F(z))}{q(z) - \phi F(z)} \right] (p - z), \quad (10)$$

where $\tilde{\varepsilon}$ describes the workers' average (initial) ability, $l(z)$ describes the firm's steady state workforce when offering z and

$$k_0 = \frac{\lambda_u \phi}{\phi(\phi + \delta + \lambda_u) - \rho(\phi + \lambda_u)}, \quad k_1 = (\phi + \delta - \rho)(\phi + \delta + \lambda_e - \rho) \quad \text{and} \quad k_2 = \frac{\lambda_e \delta \rho}{\phi + \delta + \lambda_u},$$

are positive constants.

Since $\phi > \rho$, some algebra establishes that $l(z) > 0$ for all $z \geq \underline{z}$. Further, $\delta \geq \phi$ and $\lambda_u \geq \lambda_e$ are sufficient restrictions to guarantee that $l(\cdot)$ is increasing in z . As it is well known, these restrictions are in line with much of the evidence available from empirical analysis on job search models (see Jolivet, Postel-Vinay and Robin, 2006). Given $l(\cdot)$ increases with z , it is then straightforward to show that in any equilibrium more productive firms offer higher z and hence have larger steady state workforces (see B/M for a proof).

To solve the firms' maximisation problem, consider the case in which Γ is continuous and (twice) differentiable. As established by B/M and Bontemp, Robin and Van den Berg (2000) a continuous productivity distribution and that, in equilibrium, more productive firms offer higher z , imply that firms of the same productivity offer the same z . That is, there exists a unique function $\zeta : [p_0, \bar{p}] \rightarrow [\underline{z}, \bar{z}]$ such that $\zeta(p)$ is increasing for all $p \geq p_0$ and the offered piece rate by a firm with productivity $p \geq p_0$ is $\theta(p) = \zeta(p)/p \leq 1$. Given such a $\zeta(\cdot)$, the offer distribution is then given by $F(z) = \Gamma_0(\zeta(p))$ for all $p \geq p_0$. The next result solves for ζ .

Proposition 2: *Given $\delta \geq \phi$ and $\lambda_u \geq \lambda_e$, a firm with productivity $p \in [p_0, \bar{p}]$ maximises steady state profit flow by offering a $z = \zeta(p)$, where*

$$\zeta(p) = p - [q(p) - \rho]^2 \alpha(p) \int_{\underline{z}=z_R}^p \frac{dx}{[q(x) - \rho]^2 \alpha(x)}, \quad (11)$$

$q(p) = \phi + \delta + \lambda_e(1 - \Gamma_0(p))$ and

$$\alpha(p) = \frac{[q(p) - \phi \Gamma_0(p)]}{k_1 q(p) - \Gamma_0(p)[k_1 \phi + k_2(\lambda_u - \lambda_e)(1 - \Gamma_0(p))]} \quad (12)$$

3.3 Existence

Note that the offered $\zeta(p)$ described in (11) is derived for a given z_R and p_0 . Showing an equilibrium exists requires showing that z_R solves the conditions in Proposition 1 given ζ satisfies Proposition 2 and that $p_0 = \max\{\underline{p}, z_R\}$. Given p_0 and noting that in equilibrium $F(\zeta(p)) = \Gamma_0(p)$, equations (4) and (5) imply that z_R is given by

$$\phi z_R = z_b(\phi - \rho) + [\lambda_u(\phi - \rho) - \phi \lambda_e] \int_{p_0}^{\bar{p}} \frac{1 - \Gamma_0(x)}{q(x) - \rho} \zeta'(x) dx.$$

Differentiation of (11) and integration by parts then implies that z_R solves $T(z_R; p_0) = 0$, where

$$T(z_R; p_0) \equiv \phi z_R - z_b(\phi - \rho) - [\lambda_u(\phi - \rho) - \phi \lambda_e] \int_{p_0}^{\bar{p}} \left[\frac{(\phi + \delta - \rho)(p_0 - z_R)}{(\phi + \delta + \lambda_e - \rho)} + \int_{p_0}^x \frac{ds}{(q(s) - \rho)^2 \alpha(s)} \right] \beta(x) dx, \quad (13)$$

and the parametric restrictions $\delta \geq \phi$ and $\lambda_u \geq \lambda_e$ are sufficient to guarantee that

$$\beta(x) = [1 - \Gamma_0(x)][2\lambda_e \Gamma'_0(x)\alpha(x) - (q(x) - \rho)\alpha'(x)] > 0. \quad (14)$$

Denote $z_R(p_0)$ the solution to $T(z_R; p_0) = 0$. Since $p_0 = \max\{\underline{p}, z_R\}$, there are two possible cases. If $\underline{p} > z_R(\underline{p})$, then $p_0 = \underline{p}$ and the lowest productivity of the set of active firms coincides with the infimum of the support of Γ . Using (13) this condition is equivalent to require that

$$\underline{p} > \frac{z_b(\phi - \rho) + [\lambda_u(\phi - \rho) - \phi \lambda_e] \int_{\underline{p}}^{\bar{p}} \left[\frac{(\phi + \delta - \rho)\underline{p}}{(\phi + \delta + \lambda_e - \rho)} + \int_{\underline{p}}^x \frac{ds}{(q(s) - \rho)^2 \alpha(s)} \right] \beta(x) dx}{\phi + \frac{[\lambda_u(\phi - \rho) - \phi \lambda_e](\phi + \delta - \rho)}{(\phi + \delta + \lambda_e - \rho)} \int_{\underline{p}}^{\bar{p}} \beta(x) dx}. \quad (15)$$

On the other hand, if $\underline{p} \leq z_R(\underline{p})$, then $z_R = p_0 \geq \underline{p}$ and some firms will not be active in the labour market. In this case (13) implies p_0 satisfies $p_0 = \tilde{T}(p_0)$, where

$$\tilde{T}(p_0) = \frac{z_b(\phi - \rho)}{\phi} + \frac{[\lambda_u(\phi - \rho) - \phi \lambda_e]}{\phi} \int_{p_0}^{\bar{p}} \left[\int_{p_0}^x \frac{ds}{(q(s) - \rho)^2 \alpha(s)} \right] \beta(x) dx.$$

Since $\tilde{T}$ is continuous, $\tilde{T}(\underline{p}) > \underline{p}$ and $\bar{p} \in (z_b, \infty)$, there exists a $p_0^* \in (\underline{p}, \bar{p}]$ such that $p_0^* = \tilde{T}(p_0^*)$. Uniqueness of p_0^* is guaranteed when $\tilde{T}$ is monotonic. Since the latter depends on the properties of Γ , in general $\tilde{T}$ may have more than one fixed point.

Given a $p_0 \leq \bar{p}$ always exists, z_R is then the unique solution to (13), $\zeta(\cdot)$ is uniquely characterised by (11) in Proposition 2 and $F(\zeta(p)) = \Gamma_0(p)$ for all $p \in [p_0, \bar{p}]$, where Γ_0 is given by (6). Furthermore, since Proposition 2 implies no firm with productivity $p \in [p_0, \bar{p}]$ will offer a different z , as doing so yields lower steady state profits, this establishes the following result.

Theorem 1: *Given $\delta \geq \phi$ and $\lambda_u \geq \lambda_e$, there exists a Market Equilibrium.*

Although uniqueness of equilibrium is not guaranteed when $p_0 = z_R$, uniqueness is guaranteed when $p_0 = \underline{p}$ and all firms are active. In this case z_R is described by the RHS of expression (15).

4 Employment sorting dynamics

The above equilibrium implies positive sorting between workers and firms. In particular, the interaction of on-the-job search and human capital accumulation implies that more

productive workers end up employed in more productive firms. To show this consider $\tilde{H}_\varepsilon(p | y)$, the probability that a worker of productivity y is employed at a firm with productivity no greater than p given that this worker entered the labour market with ability ε . Using (9) to construct $H(z | y, \varepsilon)$ and noting that $F(\zeta(p)) = \Gamma_0(p)$ yields

$$\tilde{H}_\varepsilon(p | y) = \Gamma_0(p) \left[\frac{\delta(\phi + \delta + \lambda_u) + (\phi + \delta + \lambda_e)(\phi + \lambda_u)(1 - \Gamma_0(p)) \left(\frac{y}{\varepsilon}\right)^{-\left(\frac{\delta\lambda_u + \lambda_e(\phi + \lambda_u)(1 - \Gamma_0(p))}{\rho(\phi + \lambda_u)}\right)}}{(\phi + \delta + \lambda_u)[(\phi + \lambda_e)(1 - \Gamma_0(p)) + \delta]} \right]. \quad (16)$$

Inspection establishes that $\tilde{H}_\varepsilon(\cdot | y)$ is first order stochastically increasing in y ; i.e. $\partial \tilde{H}_\varepsilon(\cdot | y) / \partial y < 0$. That is, more productive workers are indeed more likely to be employed in more productive firms. The intuition for this result is quite simple. In an economy with on-the-job search and on-the-job learning, workers who spend more time employed in the labour market will have had more time to accumulate human capital and have had more chances of finding better paying jobs. Given that more productive firms offer better paying jobs (ζ is increasing in p), it follows that more experience and hence productive workers end up employed in more productive firms. The interaction between human capital accumulation and on-the-job search in the presence of firm productivity heterogeneity generates positive sorting among workers and firms for any initial ability type.

Now consider the distribution of workers' productivities employed in firms with productivity p conditional on these workers having an initial ability of ε . Using (9) to construct $H_\varepsilon(y | z)$ and noting that $F(\zeta(p)) = \Gamma_0(p)$ then implies that the probability that a firm with productivity p employs workers of initial ability ε with a productivity no greater than y is given by

$$\begin{aligned} \tilde{H}_\varepsilon(y | p) = 1 - \left(\frac{y}{\varepsilon}\right)^{-\frac{q(p)}{\rho}} - \frac{\delta q(p)^2}{(\phi + \delta)[q(p) - \phi\Gamma_0(p)]^2} \left[\left(\frac{y}{\varepsilon}\right)^{-\left(\frac{\phi(\phi + \delta + \lambda_u)}{\rho(\phi + \lambda_u)}\right)} - \left(\frac{y}{\varepsilon}\right)^{-\frac{q(p)}{\rho}} \right] \\ - \frac{\lambda_e \phi \Gamma_0(p)(1 - \Gamma_0(p))q(p)}{\rho(\phi + \delta)[q(p) - \phi\Gamma_0(p)]} \left(\frac{y}{\varepsilon}\right)^{-\frac{q(p)}{\rho}} \ln \left(\frac{y}{\varepsilon}\right). \end{aligned}$$

Differentiation of $\tilde{H}_\varepsilon(\cdot | p)$, the restriction $\lambda_u \geq \lambda_e$ and noting that

$$\left(\frac{y}{\varepsilon}\right)^{\left[\frac{q(p)}{\rho} - \left(\frac{\phi(\phi + \delta + \lambda_u)}{\rho(\phi + \lambda_u)}\right)\right]} - 1 \geq \left[\frac{q(p)}{\rho} - \left(\frac{\phi(\phi + \delta + \lambda_u)}{\rho(\phi + \lambda_u)}\right) \right] \ln \left(\frac{y}{\varepsilon}\right)$$

implies first order stochastic dominance in p ; i.e. $\partial \tilde{H}_\varepsilon(\cdot | p) / \partial p < 0$. The interaction of on-the-job search and human capital accumulation also implies a steady state allocation in which more productive firms end up employing a more productive workforce.

5 Equilibrium wage dispersion

The above arguments imply that in this economy wages are dispersed due to (i) the distribution of initial abilities A ; (ii) worker productivity grows through on-the-job learning; (iii) firms' differential pay policies or frictional wage dispersion; and (iv) sorting dynamics. Letting $h_\varepsilon(y, z)$ denote the joint density of (y, z) across type ε employed workers, the equilibrium wage distribution is given by:

$$G(w) = \int_{\underline{\varepsilon}}^{\bar{\varepsilon}} \int_{z=\underline{z}}^{\bar{z}} \left[\int_{y'=\varepsilon}^{w/z'} h_\varepsilon(y', z') dy' \right] dz' d\varepsilon, \quad (17)$$

where

$$h_\varepsilon(y, z) = \begin{cases} \partial^2 H_\varepsilon(\cdot, \cdot) / \partial y \partial z & \text{for } y \geq \varepsilon \text{ and } z \in [\underline{z}, \bar{z}], \\ 0 & \text{otherwise.} \end{cases}$$

In the Appendix I provide a detailed derivation of this wage distribution. It is straightforward to verify that this wage distribution retains the same properties as the one derived in Burdett et al. (2009). In particular, it exhibits a density with an interior mode, it is right skewed and has a long and “fat” right tail.

5.1 Decomposition of (log) wages

The key insight here is that the structure of the model provides a simple way to decompose the variation of (log) wages implied by (17). Since the wage of any worker i at time t , employed at firm j with experience x can be described by the wage equation

$$\log w_{ijt} = \log \varepsilon_i + \rho x_{ijt} + \log z_j, \quad (18)$$

the contribution of each of these factors can be analysed using the following (log) wage variance decomposition

$$\text{var}(\log w) = \text{var}(\log \varepsilon) + \rho^2 \text{var}(x) + \text{var}(\log z) + 2\rho \text{cov}(\log z, x). \quad (19)$$

The first two terms in (19) describe the contribution of workers' productivities (initial abilities and human capital accumulation) in explaining wage dispersion. The third term measures the extent to which firms' differential pay policies explain wage dispersion; i.e. frictional wage dispersion. Since (16) establishes that x and z are positively correlated but orthogonal to ε , the last term then measures the contribution of sorting to wage dispersion. It is useful to stress that in the event of no on-the-job learning ($\rho = 0$) or no on-the-job search ($\lambda_e = 0$) or both, the sorting effect disappears. For example, in an otherwise standard B/M model, wages dispersion will only be explained by variation in z . Although in that model labour market experience and z are also positively correlated (see Burdett, 1978, for an early

example), experience does not have any real effect on wages dispersion. On the other hand, with no on-the-job search, any firm offers the same z to workers (see Diamond, 1971) and wage dispersion is purely driven by the fact that some workers spend more time employed than others. It is clear that in this case there is no sorting between workers and firms.

In the quantitative section of the paper I use these insights and in particular the variance decomposition (19) to analyse the relative contribution of productivity differentials, on-the-job search and sorting in explaining the observed cross-sectional distribution of wages. It is important to note, however, that the last term of in (19) only measures the direct contribution of sorting to total wage dispersion. As it will now be established, by reducing unemployed workers' reservation wages, sorting also has an indirect effect on wage dispersion by increasing frictional wage dispersion.

5.2 Frictional wage dispersion

An important contribution of the B/M model is that it gives a theory of frictional wage dispersion.⁴ Hornstein, et al. (2009) used the ratio between the average wage to the minimum observed wage (or reservation wage) conditional on workers characteristics as a way to measure this type of wage dispersion. In short, they argued 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. Burdett et al. (2009) show that when learning-by-doing is introduced to an otherwise standard B/M model one obtains a lower reservation wage and a higher Mm ratio. Workers are willing to accept wages below their opportunity cost of employment as an investment for future productivity growth through learning-by-doing. The same logic applies here. However, when firms differ in their productivities this investment effect is strengthen as employment becomes even more valuable to the worker due to the possibility of moving to more productive employers through on-the-job search.

This is most easily seen by considering the special case in which the arrival rate of job offers is independent of employment status; i.e $\lambda_u = \lambda_e = \lambda$. Under this restriction, $\alpha(p) = 1/k_1$ and $\zeta(\cdot)$ reduces to

$$\zeta(p) = p - [q(p) - \rho]^2 \int_{z=z_R}^p \frac{dx}{[q(x) - \rho]^2},$$

for all $p \geq p_0$. Further, (13) then implies that

$$z_R = z_b - \frac{\rho}{\phi(\phi + \delta + \lambda - \rho)^2 - \rho\lambda^2} \left[[(\phi + \delta + \lambda - \rho)^2 z_b + \lambda^2(p_0 - z_b)] + \int_{p_0}^{\bar{p}} \left[\frac{\lambda(1 - \Gamma_0(x))}{q(x) - \rho} \right]^2 dx \right]. \quad (20)$$

The first term in the expression in squared brackets is the same as in Burdett et al. (2009) and

⁴For evidence of this empirical regularity see Abowd, Kramarz and Margolis (1999), Mortensen (2003), Postel-Vinay and Robin (2006) and Horstein, Krusell and Violante (2007).

captures the pure investment effect of on-the-job learning on workers' reservation wages. The second term arises due to sorting. In particular, it shows the value unemployed workers are willing to forgo for the possibility of increasing their own productivity (on-the-job learning) and climbing the firm productivity distribution (on-the-job search).⁵

In the model frictional wage dispersion refers to wage variation induced only by variation in $z = \zeta(p)$. For this reason the Mm ratio should be defined by the ratio between the average z , z^M , and the minimum z , z_R , earned by employed workers. Using (4) and (5) in Proposition 2, (9) and following Hornstein et al. (2009) approximation it is shown in the Appendix that the Mm ratio in case of equal arrival rates is given by

$$Mm \cong \left[1 - \frac{\lambda\rho}{\phi(\phi + \delta + \lambda - \rho)} \right] / \left[\frac{\phi - \rho}{\phi} \chi - \frac{\lambda\rho}{\phi(\phi + \delta + \lambda - \rho)} \right], \quad (21)$$

where $\chi = z_b/z^M$ is a constant with

$$z^M \cong \frac{(\phi + \delta - \rho)z_R + \lambda p_0}{(\phi + \delta + \lambda - \rho)} + \frac{(\phi + \delta + \lambda - \rho)}{\lambda} \int_{p_0}^{\bar{p}} \left[\frac{\lambda(1 - \Gamma_0(x))}{q(x) - \rho} \right]^2 dx,$$

and z_R given by (20). Letting $\rho \rightarrow 0$ implies that $z_R \rightarrow z_b$ and the Mm ratio decreases and converges to $1/\chi$ (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. Below it is shown that for plausible parameter values this result extends to the case in which $\lambda_u > \lambda_e$.

6 Quantitative Analysis

I now provide a quantitative analyses of the model and analyse the implications of human capital accumulation, firms' differential pay policies and sorting on wage dispersion. To do so I calibrate the model to match salient features of the UK labour market using the British Household Panel Survey (BHPS).

6.1 Data

The BHPS is an annual survey of individuals, age 16 years or more, members of a nationally representative sample of about 5,500 households. This makes approximately 10,000 individuals interviewed each year. It started in 1991 and collects socio-economic information, including information about household organisation, the labour market, income and wealth, housing, health and socio-economic values. Using this information one is able to reconstruct

⁵When unemployed workers search more frequently than employed workers (i.e. $\lambda_u > \lambda_e$), there is a drop in the expected capital gains from on-the-job search that counters the expected gains from sorting. Indeed, Burdett and Mortensen (1998) show that when $\lambda_u > \lambda_e$ (and there is no on-the-job learning), the reservation wage of unemployed workers is above the flow opportunity cost of employment.

the labour market histories of individuals since leaving full-time education. Maré (2006) provides a comprehensive guide on how to derive consistent histories that summarise individual's transitions between employment, unemployment and non-participation; transitions between jobs; occupational and industry changes; actual and potential work experience; wages and hours worked; and several socio economic characteristics that are standard in household survey data.

I construct individual labour market histories following Maré's (2006) procedure, considering only those individuals that were originally sampled in 1991 and were between 16 and 60 years of age at that time. Further, I restrict attention to white male workers that were only ever in (i) paid (dependent) full-time employment or (ii) non-employment.⁶ I construct their entire employment history since leaving full-time education using retrospective work history information and follow these individuals over time until 2004 (or earlier if they left the sample before). To keep the sample as homogeneous as possible I drop those individuals who became self-employed or entered part-time employment at any point during the period of observation. This leaves me with a sample of 1,373 individuals. I then stratify the sample into low skilled ($N = 456$), medium skilled ($N = 772$) and high skilled ($N = 145$) workers based on the individual's highest educational qualification in January 1992. I consider as my earnings variable the real hourly (gross) wage.⁷

I assume that an individual changed jobs if he/she changed employer. In principle, this could underestimate the number of jobs an individual holds during his/her working life as he/she can change jobs within the same employer. However, to be consistent with the theory, I consider job-to-job transitions as employer-to-employer transitions. Following Jolivet, et al. (2006), I also count as a job-to-job transition a transition in which the individual changed employer but experienced an intervening spell of non-employment of less than a month. If the individual experiences a non-employment spell longer than a month, then he/she is considered non-employed. Finally, I assume that if a individual exits the labour force at any point during the period of observation, he/she does not re-enter and is recorded as retired (see Bagger, et al., 2006, for a similar assumption).

⁶The non-employment state includes those states in which an individual is not employed but potentially still participating in the labour force. The retired and long term sick are excluded. A worker is considered non-employed when he/she is unemployed, in maternity leave or caring for family, and in government training schemes or back to full-time education. Missing values and the "other" category are also treated as non-employment spells if the workers regains employment shortly afterwards.

⁷Following Dustmann and Pereira (2008), we consider low skilled workers those with no qualification, other qualifications, apprenticeship, CSE, commercial qualifications, no O-levels. Medium skilled workers are those that achieved O-level or equivalent, nursing qualifications, teaching qualifications, A levels, other higher qualifications. High skilled workers are those that achieved at least a higher degree. We also follow these authors and construct real hourly wages by dividing monthly (gross) earnings by 4.33 weeks and then by the average number of hours worked in a week in full-time jobs. We also take into account overtime hours. We use the CPI to deflate nominal wages.

TABLE 1: Spell durations (since leaving full-time education)

	low skilled		medium skilled		high skilled	
	mean	std. error	mean	std. error	mean	std. error
Non-employment duration	24.12	1.81	12.64	0.69	19.12	3.13
Employment duration	173.52	6.07	186.84	4.40	167.23	9.17
Job duration	63.74	1.84	68.57	1.43	67.01	3.26
Actual experience	324.59	7.84	313.90	4.98	272.18	9.64
Potential experience	354.92	7.42	325.67	4.95	288.93	9.72

Table 1 shows the average duration (in months) of the main spell variables for each skill group. These are computed using the individuals' entire labour market history; i.e. since leaving full-time education till 2004. As expected, low skilled workers have longer non-employment spells durations of non-employment and shorter job spell durations than medium and higher skilled workers. Further, since low skilled workers left full time education earlier, they have a higher average actual and potential experience than medium and high skilled workers.

TABLE 2: Transition statistics of the BHPS sample

	low skilled		medium skilled		high skilled	
	mean	std. error	mean	std. error	mean	std. error
# of jobs spells	5.09	0.16	4.58	0.10	4.06	0.17
# of unemp. spells	1.19	0.068	0.90	0.04	0.88	0.09
Emp. to Emp transitions	3.22	0.15	2.90	0.09	2.43	0.15
Emp. to Non-Emp. transitions	1.05	0.06	0.75	0.04	0.67	0.08
Emp. to Non-part. transitions	0.27	0.02	0.19	0.01	0.15	0.03
Unemp. to Emp. transitions	1.02	0.07	0.83	0.04	0.83	0.09

Table 2 shows the mean of the main transitions statistics by skill group also using the individuals' entire labour market history. The data shows that British male workers hold on average 4.5 jobs during their first 30 years in the labour market and experience about three job to job transitions. Further, they experience on average one unemployment spell (other than the one after leaving full-time education) during their working life. Interestingly, low skilled workers have a (slightly) higher job turnover rate than medium and high skilled workers. Overall, these numbers seem consistent with the ones derived by Dustmann and Pereira (2008) and Postel-Vinay and Turon (2008) using the same data and a similar period of study.

6.2 Calibration strategy

Let $\Theta = \{\Theta_T, \Theta_P, \Theta_A\}$ denote the vector of parameters that need to be recovered. In particular, the model implies a set of transition parameters $\Theta_T = \{\phi, \delta, \lambda_u, \lambda_e\}$; a set of productivity parameters $\Theta_P = \{\rho, \underline{p}, \bar{p}, a, p_h, b\}$, where a denotes the set of parameters that describe the firm productivity distribution; and a set of parameters that describe the distribution of workers initial abilities $\Theta_A = \{\underline{\varepsilon}, \kappa\}$, where κ denotes the set of parameters of the ability distribution. To recover these parameters I follow a recursive procedure involving three steps.

Step 1: Transition parameters

To obtain the parameters in Θ_T I target the unemployment, employment and job spell durations implied by the data. In particular, I set the value of ϕ to match the average potential experience duration reported in Table 1. Similarly, I set λ_u to match the average unemployment duration reported in Table 1.⁸ Noting that (9) implies

$$H(\infty, z) = \frac{(\phi + \delta)F(z)}{\phi + \delta + \lambda_e(1 - F(z))}$$

denotes the probability any worker is employed earning a $z' \leq z$, the average unconditional job spell duration is given by

$$JD = \int_{\underline{z}}^{\bar{z}} \frac{dH(\infty, z)}{\phi + \delta + \lambda_e(1 - F(z))} = \frac{\phi + \delta + \lambda_e/2}{(\phi + \delta)(\phi + \delta + \lambda_e)}. \quad (22)$$

The average probability of a job to non-employment transition conditional on job termination is then given by

$$\Pr EN = \int_{\underline{z}}^{\bar{z}} \frac{\delta dH(\infty, z)}{\phi + \delta + \lambda_e(1 - F(z))} = \delta JD, \quad (23)$$

while the average probability of a job to job transition conditional on job termination is given by

$$\Pr EE = \int_{\underline{z}}^{\bar{z}} \frac{\lambda_e(1 - F(z))dH(\infty, z)}{\phi + \delta + \lambda_e(1 - F(z))} = (\phi + \delta) \left[\left(1 + \frac{\phi + \delta}{\lambda_e}\right) \ln \left(1 + \frac{\lambda_e}{\phi + \delta}\right) - 1 \right] \quad (24)$$

Computing the empirical counterpart of these probabilities, using the average job duration shown in Table 1, we obtain δ and λ_e by using (23) and (24). Table 3 shows the set of values for the transition parameters by skill group and reports the empirical probabilities of a job-to-job change and a job-to-non-employment transition we matched. Table 1 also shows the implied average non-employment rate and the average actual experience. Note that these moments were not targeted and the model provides reasonable estimates for low and medium skilled workers. For the case of high skilled workers the fit is less satisfactory, in

⁸The model implies that any unemployed worker accepts the first job offer he/she encounters. Hence the average duration of unemployment is then simply given by $1/\lambda_u$.

particular, the average non-employment rate. This could be the result of the small number of observations in this category.

TABLE 3: Transition parameters

	low skilled	medium skilled	high skilled
ϕ	0.0028	0.0031	0.0035
δ	0.0051	0.0048	0.0066
λ_u	0.0414	0.0791	0.0523
λ_e	0.0142	0.0137	0.0127
U	0.1589	0.0916	0.1608
Exp	318.77	307.30	258.49
Pr EN	0.32	0.33	0.44
Pr EE	0.60	0.58	0.46

Step 2: Firm and worker productivities

Given the transition parameters, I approximate the firms productivity distribution, Γ , by a Pareto distribution. Since I have assumed that $\bar{p}$ is finite we truncate the Pareto distribution from above. In particular,

$$\Gamma(p) = \frac{1 - (\underline{p}/p)^a}{1 - (\underline{p}/\bar{p})^a},$$

where a denotes the shape parameter of the distribution. I choose $a = 3.5$, which is similar to the values chosen in Bontemps, Robin and Van den Berg (1999), Mortensen (2003) and much of the international trade literature using the Melitz (2003) model. I normalise $\underline{p} = 100$ and set $\bar{p} = 1000$.⁹ I assume that household productivity is a proportion m of the average firm productivity, $\tilde{p}$, such that

$$p_h = m\tilde{p} = \frac{m\underline{p}^a}{1 - (\underline{p}/\bar{p})^a} \left(\frac{a}{a-1} \right) \left[\frac{1}{\underline{p}^{a-1}} - \frac{1}{\bar{p}^{a-1}} \right].$$

Let $m = 0.667$ such that $p_h = \underline{p}$.¹⁰ Further, I normalise $b = 1$ so that unemployed workers obtain the entire returns to household production.

To obtain an estimate of the rate at which workers accumulate general human capital, ρ , I use the measure of frictional wage dispersion proposed by Hornstein, et al. (2009), the mean-min (Mm) ratio. This measure considers the ratio between the mean and the minimum wage of workers after controlling for ex-ante observable and unobservable heterogeneity. Here,

⁹The choice of $\bar{p}$ does not really matter for the results as long as $\bar{p} \geq 1000$. The reason is that productivity values for this range imply that $\zeta(p)$ is essentially flat.

¹⁰This is similar to Hall and Milgrom (2008) analysis for the US economy where they obtain an $m = 0.71$. The exercise presented by these authors assumes that the distribution of firm productivity is degenerate at $p = 1$. Given that in this case $\tilde{p} = p = 1$, they choose $b = 0.71$. Our analysis applies this same logic to the case in which firms' productivities are dispersed.

the corresponding Mm ratio is given by (21), extended to different arrival rates.¹¹ It is important to note that for a given set of transition parameters $\{\lambda_u, \lambda_e, \delta, \phi\}$ and productivity parameters, $\{a, \underline{p}, \bar{p}, p_h, b\}$, ρ gives a unique Mm ratio value.

Following Hornstein, et al. (2009) I first estimate the wage equation

$$\log w_{ijt} = \alpha + \beta_2 x_{it} + \beta_3 x_{it}^2 + \eta_{ijt} \quad (25)$$

for each year between 1992-2004 and skill group using OLS, where η denotes white noise and is assumed to be normally distributed. The second step is to eliminate unobserved worker heterogeneity from wages by using the individual residuals $\hat{\eta}_{it}$ and their individual specific mean $\bar{\eta}_i = \sum_{t=1}^{N_i} \hat{\eta}_{it}/N_i$. The vector $\{\bar{\eta}_i\}_{i=1}^N$ then captures the wage variation due to fixed unobserved individual factors (i.e. innate ability). Finally, one uses the estimated distribution of transformed wages, $\tilde{w}_{it} = \exp(\hat{\eta}_{it} - \bar{\eta}_i)$, across individuals and time to calculate the Mm ratio for each skill group. The rate of human capital accumulation is then chosen to match such a measure.

TABLE 4: Mean-min (Mm) ratios

	low skilled	medium skilled	high skilled
Mm	2.22	4.86	2.82
Mm_1	1.66	1.68	1.74
Mm_5	1.37	1.34	1.34
Mm_{10}	1.27	1.24	1.21
ρ	0.0016	0.0022	0.0023
z_R	63.785	63.207	54.475

Table 4 show the set of Mm ratios obtained from the BHPS sample using the minimum observed wage $\tilde{w}$, the wage at the first percentile (Mm_1), fifth percentile (Mm_5) and tenth percentile (Mm_{10}). It is interesting to note that these values are very close to the ones obtained by Hornstein et al. (2009) for the US economy. Given that the wage data has already been trimmed (5 percent on each side) when performing the OLS regressions and that $\tilde{w}$ is still very noisy (at least for the medium skilled category), I choose to match the mean-min ratio using the wage at the first percentile, Mm_1 . The bottom section of Table 4 shows the corresponding rate of human capital accumulation and unemployed workers' reservation wages. First note that ρ is increasing in skill level and implies that low skilled workers accumulate human capital at a rate of 1.93 percent a year; while medium skilled and high skilled workers do so at a rate of 2.67 percent and 2.80 percent a year, respectively. On a 30 year average these numbers are comparable to that obtained by Williams (2009) when controlling for observed and unobserved heterogeneity using the method developed by Altonji and Shakokto (1987) and Topel (1991).

¹¹See the appendix for a derivation of this Mm ratio.

Given the above parametrisation, the last row of Table 4 shows the reservation z for each skill implied by the model. Note that in all three cases, $p_0 = \underline{p}$ and all firms are active in those labour markets. Theorem 1 then implies there is a unique labour market equilibrium for all skill categories.

Step 3: Ability distribution

Given Θ_T and Θ_P , the last step is to recover the parameters of the ability distribution by matching the mean and variance of cross sectional wage distribution implied by the data. I parametrize the ability distribution to follow a three-parameter Weibull distribution. In particular, let

$$A(\varepsilon) = 1 - e^{-\left(\frac{\varepsilon - \underline{\varepsilon}}{\kappa_2}\right)^{\kappa_1}},$$

where κ_1 describes the shape parameter, κ_2 describes the scale parameter and $\underline{\varepsilon}$ describes the location parameter. The choice of this functional form for A is somewhat arbitrary but convenient as it encompass approximations to the Pareto, log-normal and normal distributions. I choose the parameters of this distribution such that the wage density implied by the model approximates the kernel estimate of the cross sectional density of real hourly for skill level. In particular, I calibrate $\underline{\varepsilon}$ by matching the lowest observable wage in the data for each skill group. To reduce the measurement error in wages at the left tail of the distribution, the data is trimmed from below by eliminating the lowest five percent of wage observations. Further, the shape and scale parameters are calibrated to match the mean and variance of the kernel estimate of the real hourly (log) wage distribution. Table 5 shows the calibrated parameters for this case.

TABLE 5: Ability parameters

	low skilled	medium skilled	high skilled
$\underline{\varepsilon}$	3.2	3.5	4.0
κ_1	2.6	3.12	4.1
κ_2	1.2	2.54	7.0

7 Results

Figure 1 depicts the densities of the firm productivity. Figure 2.a depicts the offered z distribution, $F(\cdot)$, and Figure 2.b the earned z distribution, $H(\infty, \cdot)$ for each skill group. Note that both F and $H(\infty, \cdot)$ have densities that exhibit decreasing (short) right tails as one would expect from an otherwise standard Burdett and Mortensen (1998) model where Γ follows a Pareto distribution (see Mortensen, 2003). Figure 3.a and 3.b show the functions $z = \zeta(p)$ and $\theta = \theta(p)$ implied by this calibration for each skill group. As discussed earlier, more productive firms offer better paying jobs, $\zeta' > 0$; while piece rate offers are not necessarily monotonic in p . Figure 4 shows the distribution of workers' abilities for each skill group. Note that we set the ability distribution approximates the shape of a normal

distribution. This result is consistent with other studies that estimate ability distributions within an equilibrium search framework (see Postel-Vinay and Robin, 2002, and Bagger et al., 2006).

Figure 1: Firm Productivity Density.

I use simulated data to match the empirical wage density. In particular, I simulate the employment histories of 30,000 workers using the model as the data generating process and then taking a random sample of the resulting simulated data set to replicate the structure of the 1992-2004 BHPS sample. As mentioned before κ_1 and κ_2 are chosen to match the mean and variance of the distribution of log wages. Figure 5.a and 5.b depicts the kernel estimates of the wage density distribution and the density of log wages from the simulated data from the model and the BHPS. As shown by these figures the model replicates the empirical wage distribution quite well and provides a good fit to the wage data in the sense that one cannot reject the hypothesis that both samples are drawn from the same distribution when using a Kolmogorov-Smirnov test on the samples implied by the data and the simulated data.

7.0.1 Wage Dispersion

Now consider the variance decomposition of log wages described in (19) to quantify the importance of human capital, firms' differential pay policies and sorting in explaining the cross sectional wage dispersion. Table 6 shows the overall picture is that productivity differentials account between 58 and 65 percent of total (log) wage variation. Human capital accumulation plays a prominent role in explaining the lion share of such variation. The distribution of initial abilities explains a small proportion in all three cases. In their well know study, Abowd et al. (1999) show that worker unobserved heterogeneity explained nearly half of the variation in (log) wages. Postel-Vinay and Robin (2002), however, show that when introducing search frictions and model explicitly labour turnover, they find that the variation

due to workers' unobserved heterogeneity diminishes and becomes negligible for low skilled groups. The findings presented here are consistent with that study.

Table 6: Variance Decomposition of Log Wages					
	$var(\log \varepsilon)$	$\rho^2 var(x)$	$var(\log z)$	$2\rho cov(\log z, x)$	$corr(\log z, x)$
Low skill	3.81	55.61	28.37	12.21	0.154
Medium skill	5.79	60.64	20.91	12.65	0.177
High skill	9.64	55.77	24.05	10.54	0.143

Table 6 show that frictional wage dispersion is still an important component of wage variation, explaining between 20 and 30 percent across the three cases. Furthermore, the sorting of more productive workers into more productive firms due to on-the-job search is also important in explaining the variation in log wages. Sorting dynamics explain around between 10 and 13 percent of total wage variation, being more important for the medium skilled group. Note that in all cases the correlation between experience and earned z is positive showing the and is around 0.15.

7.0.2 Wage-experience profile

An alternative way to measure the impact of human capital accumulation and search frictions on wage inequality is by analysis the average wage-experience profile generated by the model in a cross-section of workers. In particular, consider the following modification of (18),

$$\log w_{ijt} = \log \varepsilon_i + \rho_1 x_{ijt} + \rho_2 x_{ijt}^2 + \eta_{ij}, \quad (26)$$

where $\eta \sim N(0, 1)$ reflects added measurement error. Under this specification the coefficients ρ_1 and ρ_2 capture the joint effect of human capital accumulation and sorting on the average wage-experience profile. Although human capital accumulation affects log wages linearly, on-the-job search implies a degree of curvature because of the decreasing rate at which workers climb the productivity ladder.

Given that $corr(\log z, x) > 0$, not controlling for z when estimating the above equation using Ordinary Least Squares (OLS) produces biased estimates of the true returns to human capital accumulation.¹² The strength of the bias can then be used to measure the importance of the interaction between on-the-job search and human capital accumulation and hence of sorting in determining the average wage-experience profile. Table 7 shows the result of this exercise. In particular, using the simulated data sets I estimate (26) using OLS assuming a quadratic in experience and then adding a quadratic in tenure. Note that the bias in the experience effect is around 20 percent of its true value for all skill levels. The reason for this

¹²See Abraham and Farber (1987), Altonji and Shakoto (1987), Topel (1991), Altonji and Williams (2005) for seminal contributions on the problems faced when estimating returns to tenure and experience.

being that the arrival rate of job offers for employed workers is similar for the three skill groups. Also note that this effect is not diminished substantially when introducing tenure effects. When controlling for firm differential pay policies, OLS capture very well the true values the coefficients of the calibration.

An interesting result is that for the low and medium skill groups, the coefficient estimates for experience and tenure are very similar to that obtain using the BHPS. The main difference refers to the concavity of the profile. That is, on-the-job search on its own is not able to generate enough curvature in the wage-experience profile to match the curvature in the BHPS data. This suggest that the human capital accumulation function must be concave to be able to match the empirical profile. A second interesting observation is that for the three skill groups the tenure effects generated using the simulated sample are comparable to that of the BHPS. This suggest that an important proportion of the observed tenure effects in the data is due to selection bias. Workers earning higher wages will have a lower probability of quitting their jobs and hence generating a positive correlation between tenure and wages. This seems in line with the conclusions of Altonji and Williams (2007), Dustmann and Meghir (2005), Dustmann and Pereira (2008), Williams (2009) and others that “true” tenure effects are small.

Table 7: OLS regressions on simulated data and the BHPS							
	Dep. Var.	$\log \varepsilon$	x	x^2	$\log z$	t	t^2
Low Skills							
	Sim						
	$\log w$	1.43	0.0021*	-3.31e-07*	-	-	-
	$\log w$	1.42	0.0020*	-2.63e-07*	-	0.0005*	-6.97e-08*
Sim	$\log z$	4.59	0.0005*	-3.00e-07	-	-	-
	$\log w$	1.52	0.0016*	-2.96e-07*	1.00*	-	-
BHPS	$\log w$	1.24	0.0023*	-3.53e-06*	-	-	-
	$\log w$	1.23	0.0020*	-3.24e-06*	-	0.0008*	-9.11e-07*
Medium Skills							
	Sim						
	$\log w$	1.43	0.0027*	-3.98e-07*	-	-	-
	$\log w$	1.43	0.0026*	-3.10e-07*	-	0.0005*	-1.41e-07*
Sim	$\log z$	4.59	0.0005*	-3.66e-07*	-	-	-
	$\log w$	1.74	0.0022*	-3.31e-08	0.99*	-	-
BHPS	$\log w$	1.35	0.0029*	-4.25e-06*	-	-	-
	$\log w$	1.34	0.0027*	-3.76e-06*	-	0.0010*	-2.66e-06*
High Skills							
	Sim						
	$\log w$	1.34	0.0028*	-3.06e-07*	-	-	-
	$\log w$	1.33	0.0026*	-2.25e-07*	-	0.0006*	-5.90e-07*
Sim	$\log z$	4.49	0.0005*	-3.50e-07*	-	-	-
	$\log w$	2.32	0.0023*	-3.93e-08	0.99*	-	-
BHPS	$\log w$	1.62	0.0042*	-7.80e-06*	-	-	-
	$\log w$	1.61	0.0038*	-7.15e-06*	-	0.0011*	-2.91e-06*

7.0.3 Reservation wages

Finally consider the effects of on-the-job learning and on-the-job search on the reservation wage of unemployed workers. Recall that with $\lambda_e > 0$ and $\rho > 0$, the calibration implies that the reservation z are given by 63.78, 63.21 and 54.48 for the low, medium and high skilled level groups. Hence for all three cases $z_R < z_b$ and we have a positive investment effect. However, note that these numbers conceal the effects of on-the-job search (relative to searching while unemployed) and of on-the-job human capital accumulation. To consider their effects separately consider first the case of no on-the-job search, $\lambda_e = 0$, while keeping constant the rest of the parameter values. The reservation z in this case is given by $z_R = z_b(1 - \rho/\phi)$. This implies that $z_R = 42.86$, 29.03 and 34.29 for the low, medium and high skilled groups. That is, when on-the-job learning is the only source of wage growth

while employed, workers are willing to give up 57.1, 71 and 65.7 percent, respectively, of the flow value they receive when unemployed. Compared with the benchmark calibration, the possibility of on-the-job search and also increasing wages through job to job transition reduces the impact of human capital accumulation by at least one half across all cases.

Next consider the case of no on-the-job learning, $\rho = 0$. The simulations show that for all three skilled level groups the possibility of on-the-job human capital accumulation generates a strong “investment” effect. One can obtain this investment effect by calculating the percentage decrease in the reservation wage when there is no human capital accumulation. In the latter case, the reservation wages are 132.25, 162.44 and 133.678, for the low, medium and high skilled level groups. The investment effects are 51.8, 61.1 and 59.3 percent respectively. The possibility of human capital accumulation then implies that there exists a set of low productivity firms that are only active in the labour market because workers can use them to increase their human capital.

8 Conclusion

In this paper we have constructed and analyzed a labour market equilibrium in which there is on-the-job search and workers accumulate general human capital through learning-by-doing. The approach yields a new and insightful variance decomposition of wages. Numerical simulations suggest that equilibrium sorting generates an important fraction of wage inequality. New labour market entrants are inexperienced and their first job is likely to yield a low pay rate. Over time those workers not only gain experience and thus become more productive, they also find and quit to better paid employment. Eventually (with luck) they become relatively well-paid employees. Topel and Ward (1992) explicitly document such wage formation behavior for young male workers in the U.S. manufacturing sector. Our numerical example suggests such sorting may more than double the impact of learning-by-doing on measured wage inequality.

This approach seems a highly promising benchmark for future empirical work. It yields (i) a Mincer wage equation, (ii) the ‘right’ time profile of worker wage outcomes - the initial 10 years of a worker’s career are characterized by several job changes and rapid wage growth - and (iii) the implied cross section distribution of wages paid has the ‘right’ structure (the density of wages paid is single peaked with a ‘fat’ Pareto right tail). It also provides a coherent link between the Mincer literature on the determination of wages and income inequality in (dynamic) labour markets.

An important extension, however, is to generalise the human capital accumulation function. Standard Mincer wage equations typically assume decreasing returns to experience; i.e. the return to experience, $\rho = \rho(x)$, is quadratic rather than linear. For tractability the model here has assumed productivity grows with experience at a constant rate. Without this assumption unemployed workers with different experiences would have different returns

to further experience. Intuition suggests more experienced workers, having a lower return to further experience, would then have higher reservation piece rates. Thus when laid-off, more experienced workers would have longer (expected) spells of unemployment (which appears a reasonable prediction). Such an extension would be valuable as it would yield a more flexible Mincer wage equation

$$\log w_{ijt} = \log y_i + \log \theta_j + \rho(x_{it}).$$

Unfortunately disperse reservation piece rates severely complicates the aggregation problem and solving for the equilibrium distribution functions F, H becomes problematic. Nevertheless these latter distribution functions will still be useful for the quantitative analyses.

References

- [1] Abraham, Katherine and Henry Farber. 1987. "Job Duration, Seniority, and Earnings". *American Economic Review*. 77(3): 278-297.
- [2] Abowd, John M., Francis Kramarz and David N. Margolis. 1999. "High Wage Workers and high Wage Firms." *Econometrica*, 67(2): 251-334.
- [3] Altonji, Joseph G. and Robert A. Shakotko. 1987. "Do Wages Rise with Job Seniority?" *Review of Economic Studies*, 54(3): 437-459.
- [4] Altonji, Joseph G. and Nicolas Williams. 2005. "Do Wages Rise with Job Seniority? A Reassessment." *Industrial and Labor Relations Review*, 58(3): 370-397.
- [5] Bagger, Jesper, Francois Fontaine, Fabien Postel-Vinay and Jean-Marc Robin. 2006. "A Feasible Equilibrium Search Model of Individual Wage Dynamics with Human Capital Accumulation." <http://www.efm.bris.ac.uk/ecfybpv/wp/humancapital.pdf>
- [6] Bartelsman, Eric and Mark Doms. 2000. "Understanding Productivity: Lessons from Longitudinal Microdata". *Journal of Economic Literature*. 38:569-94.
- [7] Barlevy, Gadi. 2008, "Identification of Search Models using Record Statistics." *Review of Economic Studies*, 75(1): 29-64.
- [8] Becker, Gary S. 1964. *Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education*. Chicago, IL: University of Chicago Press.
- [9] Bontemps, Christian, Jean-Marc Robin and Gerard J. van den Berg. 1999. "An Empirical Equilibrium Job Search Model with Search on the Job and Heterogeneous Workers and Firms." *International Economic Review*, 40(4): 1039-1074.
- [10] Bontemps, Christian, Jean-Marc Robin and Gerard J. van den Berg. 2000. "Equilibrium Search with Continuous Productivity Dispersion: Theory and Nonparametric Estimation". *International Economic Review*, 41(2): 305-358.
- [11] Bunzel, Henning, Bent J. Christensen, Nicholas M. Kiefer and Lars Korsholm. 2000. "Equilibrium Search and Human Capital Accumulation." in H. Bunzel, B.J. Christensen, N.M. Kiefer and D.T. Mortensen, eds. *Panel Data and Structural Labor Market Models*, North-Holland Publishing.
- [12] Burdett, Kenneth. 1978. "A Theory of Employee Job Search and Quit Rates". *American Economic Review*. 68: 212-220.
- [13] Burdett, Kenneth and Dale T. Mortensen. 1998. "Wage Differentials, Employer Size, and Unemployment." *International Economic Review*, 39(2): 257-273.

- [14] Burdett, Kenneth and Melvyn G. Coles. 2003. "Equilibrium Wage-Tenure Contracts." *Econometrica*, 71(5): 1377-1404.
- [15] Burdett, Kenneth, Carlos Carrillo-Tudela and Melvyn G. Coles. 2009. "Human Capital Accumulation and Labour Market Equilibrium", IZA working paper 4215. <http://ftp.iza.org/dp4215.pdf>.
- [16] Diamond, Peter. 1971. "A Model of Price Adjustment", *Journal of Economic Theory*, 3(2): 156-168.
- [17] Dustmann, Christian and Costas Meghir. 2005. "Wages, Experience and Seniority", *Review of Economic Studies*, 72 (1): 77-108.
- [18] Dustmann, Christian and Sonia Periera. 2008. "Wage Growth and Job Mobility in the United Kingdom and Germany", *Industrial and Labor Relations Review*, 61(3): 374-393.
- [19] Fu, Chao. 2009. "Training, Search and Wage Dynamics." <http://www.econ.upenn.edu/~chaofu/paper2.pdf>.
- [20] Hall, Robert and Paul Milgrom. 2008. "The Limited influence of Unemployment on the Wage Bargain." *American Economic Review*, 98(2): 1653-1674.
- [21] Hashimoto, Masanori. 1981. "Firm-Specific Human Capital as a Shared Investment". *American Economic Review*. 71(3): 475-482.
- [22] Hornstein, Andreas, Per Krusell and Giovanni L. Violante. 2009. "Frictional Wage Dispersion in Search Models: A Quantitative Assessment." http://www.econ.nyu.edu/user/violante/Workingpapers/frictional_main_september4.pdf.
- [23] Jolivet, Gregory, Fabien Postel-Vinay and Jean-Marc Robin. 2006. "The Empirical Content of the Job Search Model: Labor Mobility and Wage Distributions in Europe and the US." *European Economic Review*, 50: 877-907.
- [24] Liu, Huju. 2009. "Life Cycle Human Capital Formation, Search Intensity, and Wage Dynamics". Mimeo, University of Western Ontario.
- [25] Lentz, Rasmus and Dale T. Mortensen. 2005. "Productivity Growth and Worker Reallocation", *International Economic Review*, 46(3): 731-749.
- [26] Melitz, Marc. 2003. "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity." *Econometrica*, 71: 1695-1725.
- [27] Maré, David. 2006. "Constructing Consistent Work-life Histories: A guide for users of the British Household Panel Survey", ISER Working Paper 2006-39, University of Essex.

- [28] Mincer, Jacob. 1974. *Schooling, Experience and Earnings*. New York: National Bureau of Economic Research.
- [29] Mortensen, Dale T. 2003. *Wage Dispersion: Why are Similar Workers Paid Differently?*, Cambridge, MA: MIT Press.
- [30] Postel-Vinay, Fabien and Jean-Marc Robin. 2002a. "The Distribution of Earnings in a Equilibrium Search Model with State-Dependent Offers and Counteroffers." *International Economic Review*, 43(4): 989-1016.
- [31] Postel-Vinay, Fabien and Jean-Marc Robin. 2002b. "Equilibrium Wage Dispersion with Worker and Employer Heterogeneity." *Econometrica*, 70(6): 2295-2350.
- [32] Postel-Vinay, Fabien and Jean-Marc Robin. 2006. "Microeconomic Search-Matching Models and Matched Employer-Employee Data." in Richard Blundell, Whitney Newey and Torsten Persson. eds., *Advances in Economics and Econometrics, Theory and Applications, Ninth World Congress*, Vol. 2, Cambridge: Cambridge University Press.
- [33] Postel-Vinay, Fabien and Helene Turon. 2008. "On-the-job Search, Productivity Shocks, and the Individual Earnings Process", *International Economic Review*, forthcoming.
- [34] Rosen, Sherwin. 1972. "Learning and Experience in the Labor Market." *Journal of Human Resources*, 7(3): 326-342.
- [35] Rubinstein, Yona and Yoram Weiss. 2007. "Post Schooling Wage Growth: Investment, Search and Learning." in Eric A. Hanushek and Finis Welch. eds., *Handbook of Economics of Education*, Vol. 1, Amsterdam: Elsevier Science B.V.
- [36] Topel, Robert. 1991. "Specific Capital, Mobility, and Wages: Wages Rise with Job Seniority." *Journal of Political Economy*, 99(1): 145-176.
- [37] Topel, Robert and Michael P. Ward. 1992. "Job Mobility and the Careers of Young Men." *Quarterly Journal of Economics*, 107(2): 439-479.
- [38] Van den Berg, Gerard and Geert Ridder. 1998. "An Empirical Equilibrium Search Model of the Labor Market." *Econometrica*, 66(5): 1183-1221.
- [39] Yamaguchi, Shitaro. 2009. "Job Search, Bargaining, and Wage Dynamics". <http://socserv.mcmaster.ca/econ/rsrch/papers/archive/2007-03.pdf>.

APPENDIX

Proof of Proposition 1:

Given the functional forms for W^U and W^E , the Bellman equation (2) is equivalent to

$$\phi\alpha^U = z_b + \lambda_u \int_{z_R}^{\bar{z}} [\alpha^E(z') - \alpha^U] dF(z') \quad (27)$$

and the Bellman equation (1) is equivalent to

$$(\phi + \delta)\alpha^E(z) = z + \rho\alpha^E(z) + \lambda_e \int_z^{\bar{z}} [\alpha^E(z') - \alpha^E(z)] dF(z') + \delta\alpha^U, \quad (28)$$

which is a functional equation for $\alpha^E(\cdot)$. Differentiating (28) with respect to z yields (3) and evaluating (28) at $z = \bar{z}$ yields its boundary value $\alpha^E(\bar{z})$.

I now solve the conditions for z_R and α^U . First evaluate (28) at $z = z_R$. As $\alpha^E(z_R) = \alpha^U$ one obtains

$$\phi\alpha^U = z_R + \rho\alpha^U + \lambda_e \int_{z_R}^{\bar{z}} [\alpha^E(z') - \alpha^U] dF(z').$$

Comparing this equation with (27), integrating by parts and using (3) establishes (4) described in Proposition 1. Similarly, $\alpha^E(z_R) = \alpha^U$, integration by parts and using (3) then yields (5) in Proposition 1. Thus (4) and (5) describe a pair of equations for (α^U, z_R) .

I now establish that a solution exists and is unique. First note that the equation described by (4) has slope

$$\left[\frac{d\alpha^U}{dz_R} \right]_{eqn(4)} = -\frac{1}{\rho} \left[\frac{\phi + \delta + \lambda_u(1 - F(z_R)) - \rho}{q(z_R) - \rho} \right] < 0$$

and implies $\alpha^U = (z_b - \bar{z})/\rho$ at $z_R = \bar{z}$. On the other hand, the equation described by (5) has slope

$$\left[\frac{d\alpha^U}{dz_R} \right]_{eqn(5)} = -\frac{1}{\phi} \left[\frac{\lambda_u(1 - F(z_R))}{q(z_R) - \rho} \right] < 0,$$

for $z_R < \bar{z}$ and zero otherwise and implies that $\alpha^U = z_b/\phi$ at $z_R = \bar{z}$. Note that

$$\left[\frac{d\alpha^U}{dz_R} \right]_{eqn(5)} > \left[\frac{d\alpha^U}{dz_R} \right]_{eqn(4)}$$

for all z_R and hence (5) is always flatter than (4). Continuity of (4) and (5) and the restriction $\bar{z} > z_b(\phi - \rho)/\phi$ then guarantee there exists a single crossing between these two functions such that $\alpha^U > 0$ and $z_R < \bar{z}$.||

Proof of Lemma 1:

Consider the pool of type ε unemployed workers with productivity no greater than y . It

is straightforward to verify that steady-state turnover implies

$$N_\varepsilon(y) = \frac{\phi(\phi + \delta + \lambda_u) + \delta\lambda_u H_\varepsilon(y, \bar{z})}{(\phi + \lambda_u)(\phi + \delta)} \quad (29)$$

for all $y \geq \varepsilon$. Next consider the pool of type ε employed workers who have productivity no greater than y and receive a payoff no greater than z . The arguments in Burdett et al. (2009) imply that $H_\varepsilon(., .)$ satisfies the following partial differential equation,

$$\frac{\partial H_\varepsilon(y, z)}{\partial y} + \frac{q(z)}{\rho y} H_\varepsilon(y, z) = \frac{(\phi + \delta)F(z)N_\varepsilon(y)}{\rho y}, \quad (30)$$

for $z \in [\underline{z}, \bar{z}]$ and $y \geq \varepsilon$. For a given z , integrating over y using the integrating factor $y^{\frac{q(z)}{\rho}}$ and noting that $H_\varepsilon(\varepsilon, z) = 0$ yields

$$H_\varepsilon(y, z) = \frac{(\phi + \delta)F(z)}{\rho} y^{-\frac{q(z)}{\rho}} \int_\varepsilon^y y'^{\frac{q(z)}{\rho}-1} N_\varepsilon(y') dy' \quad (31)$$

for all $y \geq \varepsilon, z \in [\underline{z}, \bar{z}]$.

Using these formulae we now solve for steady state $N_\varepsilon(.)$ and $H_\varepsilon(., .)$. In particular, using (29) to substitute out $N_\varepsilon(.)$ and simplifying yields

$$\frac{\partial H_\varepsilon(y, \bar{z})}{\partial y} = \frac{\phi(\phi + \delta + \lambda_u)}{\rho(\phi + \lambda_u)} \frac{1 - H_\varepsilon(y, \bar{z})}{y}$$

for all $y \geq \varepsilon$. As this differential equation is separable and we have the boundary condition $H_\varepsilon(\varepsilon, \bar{z}) = 0$, integration implies

$$H_\varepsilon(y, \bar{z}) = 1 - \left(\frac{y}{\varepsilon}\right)^{-\left(\frac{\phi(\phi + \delta + \lambda_u)}{\rho(\phi + \lambda_u)}\right)}.$$

Using this in (29) and simplifying yields (8). Using (8) to substitute out $N_\varepsilon(.)$ in (31) and some algebra then establishes (9).||

Derivation of firms' steady state profit flows:

Consider a firm with productivity p offering $z \geq z_R$. This firm's steady state profit is given by

$$\Omega(z; p) = \frac{p - z}{q(z) - \rho} \int_{\underline{\varepsilon}}^{\bar{\varepsilon}} \left[\lambda_u U \int_{y'=\varepsilon}^{\infty} y' dN_\varepsilon(y') + \lambda_e (1 - U) \int_{y'=\varepsilon}^{\infty} \int_{z'=\underline{z}}^z y' \frac{\partial^2 H_\varepsilon(y', z')}{\partial y' \partial z'} dz' dy' \right] dA(\varepsilon).$$

Next use the results in Lemma 1 to solve for the integrals in $\Omega(z; p)$. Consider the first integral in the expression in brackets. Using (8) we obtain

$$\int_{y'=\varepsilon}^{\infty} y' dN_\varepsilon(y') = \varepsilon N_\varepsilon(\varepsilon) + \int_\varepsilon^{\infty} \left(\frac{\phi(\phi + \delta + \lambda_u)}{\rho(\phi + \lambda_u)} \right) \frac{\lambda_u \delta}{(\phi + \lambda_u)(\phi + \delta)} \left(\frac{y'}{\varepsilon} \right)^{-\left(\frac{\phi(\phi + \delta + \lambda_u)}{\rho(\phi + \lambda_u)}\right)} dy'.$$

Integration and some algebra then establish that

$$\int_{y'=\varepsilon}^{\infty} y' dN_{\varepsilon}(y') = \frac{\phi(\phi + \delta + \lambda_u)\varepsilon}{(\phi + \delta)} \left[\frac{(\phi + \delta - \rho)}{\phi(\phi + \delta + \lambda_u) - \rho(\phi + \lambda_u)} \right].$$

Next consider the second integral in $\Omega(z; p)$. Integrating over z' implies

$$\int_{y'=\varepsilon}^{\infty} \int_{z'=\underline{z}}^z y' \frac{\partial^2 H_{\varepsilon}(y', z')}{\partial y' \partial z'} dz' dy' = \int_{y'=\varepsilon}^{\infty} y' \left[\frac{\partial H_{\varepsilon}(y', z')}{\partial y'} \right]_{\underline{z}}^z dy'.$$

As $F(\underline{z}) = 0$, (9) implies $H_{\varepsilon}(y', \underline{z}) = 0$. (30) then implies $\frac{\partial H_{\varepsilon}(y', \underline{z})}{\partial y'} = 0$ and the previous expression reduces to

$$\int_{y'=\varepsilon}^{\infty} \int_{z'=\underline{z}}^z y' \frac{\partial^2 H_{\varepsilon}(y', z')}{\partial y' \partial z'} dz' dy' = \int_{\varepsilon}^{\infty} y' \frac{\partial H_{\varepsilon}(y', z)}{\partial y'} dy'.$$

Now (9) implies

$$\frac{\partial H_{\varepsilon}(\cdot, \cdot)}{\partial y} = \frac{\phi F(z)}{y\rho} \left[1 - \frac{\delta F(z)}{q(z) - \phi F(z)} \right] \left(\frac{y}{\varepsilon} \right)^{-\frac{q(z)}{\rho}} + \frac{\delta F(z)}{q(z) - \phi F(z)} \frac{\phi(\phi + \delta + \lambda_u)}{y\rho(\phi + \lambda_u)} \left(\frac{y}{\varepsilon} \right)^{-\left(\frac{\phi(\phi + \delta + \lambda_u)}{\rho(\phi + \lambda_u)} \right)}.$$

Using this expression, integrating and noting that $\phi > \rho$, yields

$$\begin{aligned} \int_{y'=\varepsilon}^{\infty} \int_{z'=\underline{z}}^z y' \frac{\partial^2 H_{\varepsilon}(y', z')}{\partial y' \partial z'} dz' dy' &= \frac{\phi F(z)\varepsilon}{q(z) - \rho} \left[1 - \frac{\delta F(z)}{q(z) - \phi F(z)} \right] \\ &\quad + \frac{\delta F(z)\varepsilon}{q(z) - \phi F(z)} \left[\frac{\phi(\phi + \delta + \lambda_u)}{\phi(\phi + \delta + \lambda_u) - \rho(\phi + \lambda_u)} \right]. \end{aligned}$$

Collecting terms and simplifying then establishes that

$$\begin{aligned} \int_{\varepsilon}^{\infty} \int_{z'=\underline{z}}^z y \frac{\partial^2 H_{\varepsilon}(y, z')}{\partial y \partial z'} dz' dy &= \frac{\phi F(z)\varepsilon}{q(z) - \rho} \left[\frac{(\phi + \delta + \lambda_u)(\phi + \delta - \rho)}{\phi(\phi + \delta + \lambda_u) - \rho(\phi + \lambda_u)} \right] \\ &\quad - \frac{\phi F(z)\varepsilon}{q(z) - \rho} \left[\frac{\delta \rho(\lambda_u - \lambda_e)(1 - F(z))}{[q(z) - \phi F(z)][\phi(\phi + \delta + \lambda_u) - \rho(\phi + \lambda_u)]} \right]. \end{aligned}$$

Substituting out the expression for the integrals in $\Omega(z; p)$ and re-arranging yields (10) in the text.||

Proof of Proposition 2:

Consider a firm with productivity p . This firm chooses a $z \geq z_R$ to maximise

$$\Omega(z; p) = \tilde{\varepsilon} l(z)(p - z), \tag{32}$$

where $l(z)$ is given by (??). Let $z^* = \zeta(p) = \arg \max \Omega(z; p)$. Assuming the second order condition for a maximum holds, the envelope theorem implies that $\Omega'(\zeta(p)) = l(\zeta(p))$, which describes a first order differential equation for $\Omega(\cdot)$ in terms of p subject to the boundary

condition $\Omega(\zeta(p_0)) = \tilde{\varepsilon}l(\zeta(p_0))(p_0 - z_R)$. Noting that $F(\zeta(p)) = \Gamma_0(p)$, its solution is given by

$$\Omega(\zeta(p)) = \tilde{\varepsilon}l(\zeta(p))(p - \zeta(p)) = \tilde{\varepsilon} \int_{z_R}^p \frac{k_0}{[q(x) - \rho]^2} \left[k_1 - k_2 \frac{(\lambda_u - \lambda_e)\Gamma_0(x)(1 - \Gamma_0(x))}{q(x) - \phi\Gamma_0(x)} \right] dx,$$

where $q(x) = \phi + \delta + \lambda_e(1 - \Gamma_0(x))$. Since $\zeta(p) = p - \Omega(\zeta(p))/\tilde{\varepsilon}l(\zeta(p))$, substituting out for $\Omega(\zeta(p))$ and $l(\zeta(p))$ and some algebra yields (11), the expression for $\zeta(p)$ in the text.

I now show that the second order condition for a maximum is indeed met at $z = z^*$. First note that differentiation of (11) wrt p implies ζ satisfies the differential equation

$$\zeta'(p) = (p - \zeta(p))\Gamma_0'(p) \left[\frac{2\lambda_e}{q(p) - \rho} - \frac{k_2(\lambda_u - \lambda_e)[(q(p) - \phi\Gamma_0(p))(1 - \Gamma_0(p)) - \delta\Gamma_0(p)]}{[q(p) - \phi\Gamma_0(p)][k_1q(p) - \Gamma_0(p)[k_1\phi + k_2(\lambda_u - \lambda_e)(1 - \Gamma_0(p))]]} \right], \quad (33)$$

given the boundary condition $\zeta(p_0) = z_R$. Using $F'(\zeta(p))\zeta'(p) = \Gamma_0'(p)$ and (??) it is straightforward to establish that (33) describes the first order condition of (32). Hence the function ζ implied by (33) satisfies the first order condition for a maximum. Substitution of $p - \zeta(p)$ from (11) and some algebra then establish that $\delta \geq \phi$ and $\lambda_u \geq \lambda_e$ are sufficient condition for $\zeta'(p) > 0$ for all $p \geq p_0$. Recall that these parametric restrictions also imply $l'(\cdot)$ is increasing in z .

Next, let $\Omega(\zeta(\hat{p}); p) = \tilde{\varepsilon}(p - \zeta(\hat{p}))l(\zeta(\hat{p}))$ denote the steady state profit of a firm of productivity p by offering a $z = \zeta(\hat{p})$ and let $\Delta(\hat{p}) = \hat{p} - p$. For $\hat{p} \in (p, \bar{p}]$, the second order condition for a maximum requires that offering such a z should not increase profits or that

$$\begin{aligned} \left[\frac{d\Omega(\zeta(x); p)}{dx} \right]_{x=\hat{p}} &= \tilde{\varepsilon} \left[\frac{dl(\zeta(x))}{dz} \zeta'(x)(p - \zeta(x)) - \zeta'(x)l(\zeta(x)) \right]_{x=\hat{p}} \\ &= \tilde{\varepsilon} \left[\left((x - \zeta(x)) \frac{dl(\zeta(x))}{dz} - l(\zeta(x)) \right) \zeta'(x) - \Delta(x) \frac{dl(\zeta(x))}{dz} \zeta'(x) \right]_{x=\hat{p}} \leq 0. \end{aligned}$$

Since the first order condition implies $(x - \zeta(x)) \frac{dl(\zeta(x))}{dz} - l(\zeta(x)) = 0$ for any $x > p_0$, we obtain that

$$\left[\frac{d\Omega(\zeta(x); p)}{dx} \right]_{x=\hat{p}} = \left[-\varepsilon \Delta(x) \frac{dl(\zeta(x))}{dz} \zeta'(x) \right]_{x=\hat{p}} \leq 0$$

is satisfied when $\delta \geq \phi$ and $\lambda_u \geq \lambda_e$. For $\hat{p} \in [p_0, p)$ a similar argument shows that

$$\left[\frac{d\Omega(\zeta(x); p)}{dx} \right]_{x=\hat{p}} = \left[-\varepsilon \Delta(x) \frac{dl(\zeta(x))}{dz} \zeta'(x) \right]_{x=\hat{p}} \geq 0$$

is satisfied when $\delta \geq \phi$ and $\lambda_u \geq \lambda_e$.

Finally note that a firm with productivity $p = p_0$ will not offer a $z < z_R = \zeta(p_0)$ as doing so will not increase profits. It will strictly decrease profits if $p_0 = \underline{p} > z_R$ and yields the same (zero) profit if $p_0 = z_R$. A firm with $p = \bar{p}$, on the other hand, will not offer a $z > \zeta(\bar{p})$ as doing so does not attract or retain any additional worker, but strictly decreases flow profit $\bar{p} - z$ and hence steady state profits. ||

Derivation of the wage density:

Differentiating (9) one obtains the joint density of workers characteristics, y and z , for each ability type ε ,

$$h_\varepsilon(y, z) = \frac{\phi F'(z)}{\varepsilon \rho} \left(\frac{y}{\varepsilon}\right)^{-\frac{q(z)}{\rho}-1} \left[1 + \frac{\lambda_e}{\rho} F(z) \ln\left(\frac{y}{\varepsilon}\right)\right] \left[\frac{(\phi + \delta + \lambda_e)(1 - F(z))}{q(z) - \phi F(z)}\right] \\ + \frac{\delta \phi (\phi + \delta + \lambda_e) F'(z)}{\varepsilon \rho [q(z) - \phi F(z)]^2} \left[\left(\frac{\phi + \delta + \lambda_u}{\phi + \lambda_u}\right) \left(\frac{y}{\varepsilon}\right)^{-\frac{\phi(\phi + \delta + \lambda_u)}{\rho(\phi + \lambda_u)}-1} - F(z) \left(\frac{y}{\varepsilon}\right)^{-\frac{q(z)}{\rho}-1}\right],$$

where $h_\varepsilon(y, z) \geq 0$ for all $z \in [\underline{z}, \bar{z}]$ and $y \geq \varepsilon$. Otherwise, $h_\varepsilon(y, z) = 0$.

As wages are given by $w = zy$ and restricting $\underline{z} > 0$, we have that $\hat{h}_\varepsilon(w, z) = h_\varepsilon\left(\frac{w}{z}, z\right) \frac{1}{z}$, where $\hat{h}_\varepsilon(w, z)$ describes the joint density of w, z . Using the expression for $h_\varepsilon(\cdot, \cdot)$ then gives

$$\hat{h}_\varepsilon(w, z) = \frac{\phi F'(z)}{z \varepsilon \rho} \left(\frac{w}{z \varepsilon}\right)^{-\frac{q(z)}{\rho}-1} \left[1 + \frac{\lambda_e}{\rho} F(z) \ln\left(\frac{w}{z \varepsilon}\right)\right] \left[\frac{(\phi + \delta + \lambda_e)(1 - F(z))}{q(z) - \phi F(z)}\right] \\ + \frac{\delta \phi (\phi + \delta + \lambda_e) F'(z)}{z \varepsilon \rho (q(z) - \phi F(z))^2} \left[\left(\frac{\phi + \delta + \lambda_u}{\phi + \lambda_u}\right) \left(\frac{w}{z \varepsilon}\right)^{-\frac{\phi(\phi + \delta + \lambda_u)}{\rho(\phi + \lambda_u)}-1} - F(z) \left(\frac{w}{z \varepsilon}\right)^{-\frac{q(z)}{\rho}-1}\right],$$

for all $z \in [\underline{z}, \bar{z}]$ and $w \geq z \varepsilon$. Otherwise, $\hat{h}_\varepsilon(w, z) = 0$. Integrating appropriately over z then yields

$$G'_\varepsilon(w) = \begin{cases} \int_{\underline{z}}^{w/\varepsilon} \hat{h}_\varepsilon(w, z) dz & \text{for all } w \geq z \varepsilon, \\ 0 & \text{otherwise.} \end{cases}$$

Finally integrating over all possible workers' types gives the aggregate wage density of the economy

$$G'_\varepsilon(w) = \int_{\underline{\varepsilon}}^{\bar{\varepsilon}} \int_{\underline{z}}^{w/\varepsilon} \hat{h}_\varepsilon(w, z) dz d\varepsilon.$$

Integrating over w gives (17).

Derivation of the mean-min ratio:

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

$$z^M = z_R + \int_{z_R}^{\bar{z}} [1 - H(\infty, z)] dz.$$

Putting $y = \infty$ in (9) implies

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

Since $\rho < \phi$ by assumption and for reasonable parametrizations of the model ϕ is at least twice as large as ρ (see the quantitative section in the main body of the paper), I follow Hornstein et al. (2009) and approximate $H(\infty, z)$ by

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

Solving for $1 - F(z)$ and using (4) and (5) yields

$$\begin{aligned} z^R &\simeq \frac{z_b(\phi - \rho)}{\phi} - \frac{\lambda_u(\phi - \rho) - \phi\lambda_e}{\phi(\phi + \delta + \lambda_e - \rho)} \int_{z_R}^{\bar{z}} [1 - H(\infty, z)] dz \\ &\simeq \frac{z_b(\phi - \rho)}{\phi} - \frac{\lambda_u(\phi - \rho) - \phi\lambda_e}{\phi(\phi + \delta + \lambda_e - \rho)} (z^M - z_R). \end{aligned}$$

Dividing both sides by z^M we have that

$$Mm \cong \left[1 + \frac{\lambda_u(\phi - \rho) - \phi\lambda_e}{\phi(\phi + \delta + \lambda_e - \rho)} \right] / \left[\frac{\phi - \rho}{\phi} \chi + \frac{\lambda_u(\phi - \rho) - \phi\lambda_e}{\phi(\phi + \delta + \lambda_e - \rho)} \right],$$

where $\chi = z_b/z^M$ with

$$z^M \cong z_R + (\phi + \delta + \lambda_e - \rho) \int_{p_0}^{\bar{p}} \left[\frac{(\phi + \delta - \rho)(p_0 - z_R)}{(\phi + \delta + \lambda_e - \rho)} + \int_{p_0}^x \frac{ds}{(q(s) - \rho)^2 \alpha(s)} \right] \beta(x) dx,$$

$\beta(x)$ and $\alpha(x)$ are given by (14) and (12), and z_R is the solution of (13). Letting $\lambda_u = \lambda_e = \lambda$ implies these expressions reduce to the ones described in the text.

(a) Low Ability

(b) Medium Ability

(c) High Ability

Figure 2: Z-Offer Density

(a) Low Ability

(b) Medium Ability

(c) High Ability

Figure 3: Z-Earned Density

(a) Low Ability

(b) Medium Ability

(c) High Ability

Figure 4: Z-Function

(a) Low Ability

(b) Medium Ability

(c) High Ability

Figure 5: Theta-Function

(a) Low Ability

(b) Medium Ability

(c) High Ability

Figure 6: Ability

(a) Low Ability

(b) Medium Ability

(c) High Ability

Figure 7: Wage Density

(a) Low Ability

(b) Medium Ability

(c) High Ability

Figure 8: Log Wage Density