DISCUSSION PAPER SERIES

IZA DP No. 6949

Job Search, Human Capital and Wage Inequality

Carlos Carrillo-Tudela

October 2012

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Job Search, Human Capital and Wage Inequality

Carlos Carrillo-Tudela
University of Essex
and IZA

Discussion Paper No. 6949
October 2012

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

IZA Discussion Paper No. 6949
October 2012

ABSTRACT

Job Search, Human Capital and Wage Inequality*

The objective of this paper is to construct and quantitatively assess an equilibrium search model with on-the-job search and general human capital accumulation. In the model workers enter the labour market with different abilities and firms differ in their productivities. Wages are dispersed because of search frictions and workers' productivity differentials. The model generates a simple (log) wage variance decomposition that is used to measure the importance of firm and worker productivity differentials, frictional wage dispersion and workers' sorting dynamics. I calibrate the model using a sample of young workers for the UK. I show that wage inequality among low skilled workers is mostly due to differences in their productivities. Among medium skilled workers frictional wage dispersion and sorting dynamics are, together, as important as workers' productivity differentials. Differences in firms' productivities are also an important source of wage inequality for both skill groups and account for a large share of frictional wage dispersion. Quantitatively the model is able to reproduce the observed cross-sectional wage distribution, the average wage-experience profile and the amount of frictional wage dispersion observed in the data as measured by the Mean-min ratio.

JEL Classification: J63, J64, J41, J42

Keywords: job search, human capital accumulation, wage dispersion, turnover

Corresponding author:

Carlos Carrillo-Tudela
Department of Economics
University of Essex
Wivenhoe Park
Colchester, CO4 3SQ
United Kingdom
E-mail: cocarr@essex.ac.uk

* I would like to thank Ken Burdett, Melvyn Coles, Guido Menzio, Iourri Manvoski, Gregory Jolivet, Fabien Postel-Vinay, Jean-Marc Robin, Gianluca Violante, Ludo Visschers and seminar participants at SUNY Albany, Essex, Leicester, St. Gallen, Mainz, Konstanz, the SED 2010 and the Royal Economic Society 2011 conferences 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. I thank this institution for its hospitality. Financial support for this project was provided by the UK Economic and Social Research Council (ESRC), award reference ES/I037628/1. The usual disclaimer applies.

1 Introduction

1.1 Motivation and Summary

Differentials in worker productivities have long been recognised as an important source of wage inequality. Since the pioneering work of Becker (1964) and Mincer (1974), 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. Empirical evidence suggests, however, that along side workers' productivity differentials a significant proportion of the observed wage dispersion is also due to the frictional nature of labour markets. This evidence finds that equally productive workers are paid differently and that these pay differentials are large and persistent (see Abowd, Kramarz and Margolis, 1999 and Hornstein, Krusell and Violante, 2011). As shown in Mortensen (2003), search theory provides an elegant and powerful explanation for this phenomena. In this paper I study wage inequality using a unified equilibrium framework in which workers accumulate human capital and search for jobs. The main goal is to quantitatively assess the relative contributions of workers' productivity differentials and search frictions in explaining the observed wage distribution.

The theoretical framework combines the ideas of Burdett and Mortensen (1998) (henceforth B/M) and Burdett, Carrillo-Tudela and Coles (2011). 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 must search for them. Firms, on the other hand, differ in their 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 behaviour. 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 as workers becomes more productive and move from less to more productive jobs when the opportunity arises.

The main differences with Burdett, et al. (2011) are the introduction of firm productivity differentials, job offer arrival rates that depend on workers' employment status and the quantitative assessment of the model's implications. Adding the first two extensions is important for my purposes as there is now a great deal of evidence both showing that firm productivity differentials are 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 generates the possibility of multiple equilibria. By quantifying the model's main implications I aim to contribute to a long standing literature that tries to understand the different roles human capital accumulation and search frictions have in shaping wage inequality. The paper also contributes to a growing literature that quantitatively assesses the capacity of job search models to generate the observed amount of frictional wage dispersion as measured by the Mean-min ratio proposed by Hornstein et al. (2011).

In the model, 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 this allows low productivity firms to survive by offering low paying jobs and thus increasing wage dispersion across equally productive workers.

The model provides a simple variance decomposition that relates (log) wages to differences in worker abilities, firm productivities and differences in pay policies through piece rate offers; while also capturing the effects of general human capital accumulation and sorting dynamics on wage variation. Thus the model is able to encompass a similar variance decomposition as analysed by Abowd, et. al (1999) and Postel-Vinay and Robin (2002). At the same time the model captures the effects of labour market experience on wage variation as is the focus in the traditional applied labour literature (see Rubinstein and Weiss, 2007, for an insightful survey). This decomposition is one of the key element for the quantitative analysis and it is used to assess the impact of workers' productivity differentials and search frictions on the cross sectional wage distribution.

To quantitatively assess the model's main implications I calibrate it using the labour market histories of a sample of young workers drawn from the British Household Panel Survey (BHPS). I choose to evaluate the model on young workers as it is precisely at this stage of a worker's labour market history that job mobility is most common (see Hall, 1982, and Topel and Ward, 1992) and a search model, like the one developed here, is more relevant. Since there are strong differences on returns to education across workers, I divide the sample into different educational or skill groups and analyse them separately. I focus on low and medium skill workers. This is because the model's wage determination mechanism seems closer to the ones found in the labour markets for these workers, than in high skilled ones, where firms are more likely to respond to their employees outside offers (see Postel-Vinay and Robin, 2002, and Bagger, Fontaine, Postel-Vinay and Robin, 2011, for models based on this feature).

The calibration results show that the contribution of labour market experience in accounting for wage dispersion is of the order of 30 percent for medium skill workers and of 20 percent for low skilled workers. This difference arises due to the importance of sorting dynamics among medium skilled workers. In particular, I find that low skilled workers tend to reallocate across jobs through unemployment, relying little on job-to-job transitions. Furthermore, these workers tend to be employed in the low paying jobs offered by low productive firms. Medium skilled workers, in contrast, rely as much in job-to-job transitions as in unemployment to employment transitions to change jobs. These workers tend to be employed in firms that exhibit higher productivity and offer higher paying jobs. Overall, however, human capital accumulation still accounts for more than 60 percent of the contribution of labour market experience on wage dispersion across all these workers.

After controlling for labour market experience, I find important differences between the contributions of workers' abilities, firms' productivities and differences in earned piece rates in accounting for wage dispersion across skill groups. For medium skilled workers, firm productivity differentials explain 40 percent of overall wage dispersion and differences in earned piece rates explain 20 percent. Together they imply that frictional wage dispersion contributes 60 percent to the variation of log wages, while workers' ability differentials explains the remainder 40 percent. For low skilled workers, differences in earned piece rates explain only 6 percent, while worker ability and firm productivity differentials explain 50 and 44 percent, respectively.

Taking them together, the different contributions of experience, workers' abilities, firms' productivities and earned piece rates show that wage differentials among low skilled workers are better explained by workers' productivity differentials. Among medium skilled workers, productivity differentials are equally important as search frictions, measured by the sum of the contributions of frictional wage dispersion and sorting dynamics, in explaining wage dispersion.

An important aspect of the calibration is the way workers' human capital accumulation rate and the distribution of firm productivities are recovered. Theoretically I show that the model generates a mapping between the Mean-min ratio and the rate of human capital accumulation; and a mapping between the average wage-experience profile and the distribution of firm productivities. After recovering the transition parameters from duration data, I use the estimated Mean-min ratio and average wage-experience profile for each skill group to calibrate the corresponding human capital accumulation rates and firm productivity distributions. The worker ability distributions can then be obtained by matching the observed wage density distribution. Thus, the calibrated model reproduces the observed labour market transitions, the amount of frictional wage dispersion, the average wage-experience profile and the cross sectional wage distribution. This is important as job search models have been criticised by their inability to match some or all these features of the data.

The rest of the paper is outlined as follows. In the next section I discuss the contribution of the paper with respect to the existing literature. Section 2 describes the model. Section 3 defines and characterises the equilibrium. Section 4 presents a theoretical analysis of the main implications of the model. Sections 5 and 6 describe the data, the calibration procedure and present the main results. Section 7 concludes discussing briefly the main results. All proofs are relegated to a technical Appendix.

1.2 Literature Review

During the past 20 years a large and influential literature has been trying to disentangle the importance of human capital accumulation and worker' job mobility in the estimated returns to labour market experience. This literature has mostly been based on statistical models that span from Mincer's (1974) original work on so-called wage equations. Prominent examples are Abraham and Farber (1987), Altonji and Shakotko (1987), Topel (1991), Altonji and Williams (2007), Dustmann and Meghir (2005), among others. The evidence from this body work, however, is still inconclusive (Rubinstein and Weiss, 2007). This has prompted an alternative approach that uses a unified theoretical framework to quantitatively assess the relative importance of human capital and job search on the dynamics of wages over their life-cycle. Recent examples are Yamaguchi (2010) and Bagger et al. (2011), who use the random search model posit by Postel-Vinay and Robin (2002) and introduce general human capital accumulation with productivity shocks to explain the workers' wage dynamics. Menzio, Telyukova and Visschers (2012) incorporate general human capital accumulation and match-specific productivity to a steady state version of Menzio and Shi (2011) directed search model with the focus of explaining workers' transitions rates over the life-cycle.[1]

The present paper provides an alternative approach to analyse the interaction between human capital accumulation and search frictions based on the B/M random search model. Here the focus is on decomposing the wage distribution, while at the same time providing 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. Furthermore, the quantitative analysis is based on young workers rather than explaining differences across age groups. As opposed to Yamaguchi (2010) and Bagger et al. (2011), whom are closer to this paper than Menzio et al. (2011), I do not allow for productivity shocks and assume that firms do not observe workers' reservation wages or react to their employees' outside offers. Here a worker simply quits if he/she receives a better outside offer. Although offer-matching arises in certain markets where firms have access to a fair amount of information about job applicants, such as the academic market for economists, this may not be a good description of behaviour in other (less skilled) labour markets. Therefore, it seems worthwhile exploring alternative wage determination mechanisms and their effects on the interaction between human capital and search frictions.

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. Manning (2003), Rubinstein and Weiss (2007), Barlevy (2008) and Bowlus and Liu (2012) estimate a wage process similar to the one identified here but do not consider equilibrium. Fu (2011) considers a similar model, but studies whether homogeneous firms can provide general human capital for its employees or not.

Carrillo-Tudela (2009) and Burdett and Coles (2010) consider human capital accumulation in an equilibrium search environment in which firms post wage-tenure contracts. Carrillo-Tudela (2009) considers promotion contracts that depend on workers' experience; while Burdett and Coles (2010)

[1]See Hagedorn and Manovskii (2010) for a recent attempt to measure the extend of search frictions on wage dispersion using a reduced form approach.

extend Burdett et al. (2011) by allowing firms to post piece-rates that depend on their employees' tenure. Although these papers give useful insights about the interaction between human capital accumulation and job search when firms post optimal wage schedules as a respond to their employees search behaviour, they quickly loose their tractability when the distribution of firm productivities is dispersed. Since the main objective of this paper is to provide a quantitative assessment of the importance of human capital and search frictions, in this paper I follow Burdett et al. (2011) and restrict attention to constant piece-rate contracts.

More recently there have been a few studies that have combined human capital accumulation and search frictions to investigate the extend to which their interaction can help in matching the Mean-min ratio for the US economy. These papers present complementary analysis to the one developed here. For example, Ortego-Marti (2012) introduces human capital depreciation into a search and matching model with match-specific productivities (see Pissarides, 2000, chapter 6). His objective is to evaluate the importance of depreciation in unemployed workers' human capital in explaining the observed Mean-min ratio. Tjaden and Wellschmied (2012) add human capital accumulation among employed workers along reallocation shocks, as introduced by Jolivet et al. (2006), and productivity shocks to a similar model as Ortego-Marti (2012). As was suggested by Burdett et al. (2011) and shown in an earlier version of this paper, all these models confirm that adding human capital greatly improves the ability of the job search model to match the observed Mean-min ratio.

2 The Model

Assume time is continuous with an infinite horizon and that the economy is in a steady state. There is a continuum of firms and workers in the labour market, each of measure one. Any worker's life can be described by an exponential distribution with parameter $\phi > 0$. To keep the population of workers constant ϕ also describes the inflow of new labour market entrants.

Workers differ in their ability ε when entering the labour market. In particular, let A denote the cumulative distribution function of these abilities with positive support $[\underline{\varepsilon}, \overline{\varepsilon}]$. Further, learning-by-doing 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 positive support $[\underline{p}, \overline{p}]$. A worker with productivity y employed at a firm with productivity p generates flow output yp while employed. This output is assumed to be sold in a perfect competitive market in which the price of the production good is normalised to one, so yp also describes flow revenue.[2]

[2]Here the productivity of a firm p captures un-modelled factors specific to the firm that increase the productivity of a worker employed in that firm. Workers' initial abilities ε capture differences in innate ability. An important assumption is that the worker's rate of human capital accumulation ρ is independent of the productivity of his

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 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 $F(z \mid p)$ 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 $\underline{z}, \overline{z}$ denote the infimum and supremum of the support of F.

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 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 $p_h \in (0, \overline{p})$ offering a piece rate of $b \leq 1$ and $z_b = bp_h$. A worker with productivity y then enjoys a flow payoff yz_b while unemployed. 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 current state and there is no recall.

All agents are risk neutral. Let $r \geq 0$ denote workers' rate of time preference. Random death then implies workers discount the future at rate $r + \phi$. Further, let $\phi > \rho$ in order to bound human capital accumulation. Workers maximise expected discounted lifetime income. Firms, on the other hand, do not discount the future and maximise 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 behaviour. 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:

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

employer and his initial ability. Although restrictive, this assumption simplifies the analysis considerably and allows the model to be solved in closed form.

As there is no human capital accumulation while unemployed (and no depreciation), the flow Bellman equation describing $W^U(y)$ is instead

$$(r+\phi)W^U(y) = z_b y + \lambda_u \int_{\underline{z}}^{\overline{z}} \max[W^E(y,z') - W^U(y), 0] dF(z').$$

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

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

Moreover, the Bellman equation describing W^U implies the unemployed worker's optimal strategy is to 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, $\alpha^U, \alpha^E(.)$ 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) $\alpha^E(.)$ *is the solution to the ordinary differential equation*

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

with boundary condition $\alpha^E(z) = (\overline{z} + \delta\alpha^U)/(r + \phi + \delta - \rho)$ *at* $z = \overline{z}$,
(ii) (α^U, z_R) *satisfy the pair*

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

$$(r+\phi)\alpha^U = z_b + \lambda_u \int_{z_R}^{\overline{z}} \frac{1 - F(z)}{q(z) + r - \rho} dz. \tag{2}$$

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

Using (1) and (2) we have that the reservation z satisfies the following equation

$$(r+\phi)z_R = z_b(r+\phi-\rho) + [\lambda_u(r+\phi-\rho) - (r+\phi)\lambda_e]\int_{z_R}^{\overline{z}} \frac{1 - F(x)}{q(x) + r - \rho} dx. \tag{3}$$

This equation implies that (for a given F) a higher human capital accumulation rate yields a lower z_R. That is, the higher is the rate at which wages increase due to human capital accumulation, the more valuable employment becomes and the lower is the wage at which workers are willing to become employed. A similar effect is obtained when employed workers face a higher offer arrival rate. In this case, the reduction in z_R responds to an increase in the expected capital gains from on-the-job search. As it will be shown later, the combined effect of on-the-job search and human capital accumulation on z_R implies the model is able to address Hornstein et al. (2011) criticism about the ability of search models to generate enough frictional wage dispersion.

2.2 Firms Profits

Now consider firm behaviour given an F and $z_R < \overline{z}$ 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

$$\Gamma_0(p) = \frac{\Gamma(p) - \Gamma(p_0)}{1 - \Gamma(p_0)} \tag{4}$$

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) $F(. \mid p)$ 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) $N_\varepsilon(.)$: the distribution function describing productivities across type ε workers who are unemployed, and (c) $H_\varepsilon(y, z)$: the joint distribution function describing productivities and earned z across type ε workers who are employed. Next consider a firm with productivity p offering $z \geq z_R$. As firms have a zero discount rate, 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

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

Such a firm then chooses z to maximise $\Omega(z;p)$. Since $\theta = z/p$, the choice of z then gives the corresponding offered piece rate θ. Let $\overline{\Omega}(p) = \max \Omega(z;p)$ 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 $\{z_R, U_\varepsilon, N_\varepsilon(.), H_\varepsilon(.,.), F(. \mid p)\}$ for all $\varepsilon \in [\underline{\varepsilon}, \overline{\varepsilon}]$ and $p \in [p_0, \overline{p}]$ such that

Given optimal worker search strategies as described by z_R :

(i) The productivity distribution of active firms is given by (4) where $p_0 = \max\{z_R, \underline{p}\}$.

(ii) Given Γ_0, the constant profit condition is satisfied for all active firms; i.e.,

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

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

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

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

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

Given optimal firm behaviour as summarised in $F(.)$:

(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 $\overline{\Omega}(p) > 0$ for all $p > p_0$ and $\overline{\Omega}(p) \geq 0$ 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 $F(. \mid p)$ 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, $U_\varepsilon, N_\varepsilon(.)$ and $H_\varepsilon(.,.)$. 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*

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

$$H_\varepsilon(y,z) = \frac{(\phi+\delta)F(z)}{q(z)} \left[1 - \left(\frac{y}{\varepsilon}\right)^{-\frac{q(z)}{\rho}}\right] - \frac{\delta F(z)}{q(z)-\phi F(z)} \left[\left(\frac{y}{\varepsilon}\right)^{-\left(\frac{\phi(\phi+\delta+\lambda_u)}{\rho(\phi+\lambda_u)}\right)} - \left(\frac{y}{\varepsilon}\right)^{-\frac{q(z)}{\rho}}\right] \tag{5}$$

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

3.2 Characterisation

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

$$\Omega(z;p) = \widetilde{\varepsilon} l(z)(p-z) = \widetilde{\varepsilon} \frac{k_0}{(q(z)-\rho)^2} \left[k_1 - k_2 \frac{(\lambda_u - \lambda_e)F(z)(1-F(z))}{q(z)-\phi F(z)}\right](p-z), \tag{6}$$

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

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

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 $l(.)$ is increasing in z. As it is well known, these restrictions are in line with much of the evidence available from the empirical analysis on job search models (see Jolivet et al. 2006) and will be verified below in the quantitative section of the paper. Given $l(.)$ 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 differentiable. As established by B/M and Bontemps, 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 $\zeta : [p_0, \overline{p}] \to [\underline{z}, \overline{z}]$ 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 $\zeta(.)$, 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* $p \in [p_0, \overline{p}]$ *maximises steady* *state profit flow by offering a* $z = \zeta(p)$*, where*

$$\zeta(p) = p - [q(p) - \rho]^2 \alpha(p) \int_{\underline{z}=z_R}^{p} \frac{dx}{[q(x) - \rho]^2 \alpha(x)}, \tag{7}$$

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

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

3.3 Existence

Note that the offered $\zeta(p)$ described in (7) 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 $F(\zeta(p)) = \Gamma_0(p)$, (3) and (7) imply that z_R solves $T(z_R; p_0) = 0$, where

$$\begin{aligned} T(z_R; p_0) \equiv\ & (r + \phi) z_R - z_b(r + \phi - \rho) \\ & - [\lambda_u(r + \phi - \rho) - (r + \phi)\lambda_e] \int_{p_0}^{\overline{p}} \left[\frac{(\phi + \delta - \rho)(p_0 - z_R)}{(\phi + \delta + \lambda_e - \rho)} + \int_{p_0}^{x} \frac{ds}{(q(s) - \rho)^2 \alpha(s)} \right] \beta(x) dx, \end{aligned} \tag{8}$$

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

$$\beta(x) = \frac{q(x) - \rho}{q(x) + r - \rho} [1 - \Gamma_0(x)][2\lambda_e \Gamma_0'(x) \alpha(x) - (q(x) - \rho)\alpha'(x)] > 0.$$

Denote $z_R(p_0)$ the solution to $T(z_R; p_0) = 0$ for any p_0. Since $p_0 = \max\{\underline{p}, z_R(p_0)\}$, however, there are two possible cases. First $p_0 = \underline{p}$ if and only if $\underline{p} > z_R(\underline{p})$. Otherwise, $p_0 > \underline{p}$ and some firms will not be active in the labour market. The next result uses these insights and establishes existence of equilibrium.

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.

4 Implications

4.1 Employment sorting dynamics

The above equilibrium implies that in an economy with on-the-job search and learning-by-doing, 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. These sorting dynamics coupled with the result that more productive firms offer better paying jobs (ζ is increasing in p) imply that more productive workers end up employed in more productive firms. Indeed, $H_\varepsilon(p \mid y)$, 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, is first order stochastically increasing in y; i.e. $\partial H_\varepsilon(. \mid y)/\partial y < 0$.

Now consider $H_\varepsilon(y \mid p)$, the distribution of workers' productivities employed in firms with productivity p conditional on these workers having an initial ability of ε. The restriction $\lambda_u \geq \lambda_e$ implies first order stochastic dominance in p; i.e. $\partial H_\varepsilon(. \mid p)/\partial p < 0$.[3] Thus, workers' sorting dynamics and firm heterogeneity also imply that more productive firms end up employing a more productive workforce. Since more productive firms have larger workforces ($l'(p) > 0$ for all p), the model is consistent with the empirical evidence that documents that larger firms offer higher wages (to equally productive workers) and employ a more productive workforce (see Idson and Oi, 1999). Here the positive correlation between firm size and wages is due to firm and worker characteristics and sorting dynamics.

It is important to stress that the sorting allocation described above is purely due to experience effects (time employed in the labour market). Controlling for workers' experience yields no correlation between workers' (initial) abilities and firms' productivities. There is an important literature that studies the sorting of agents that differ in their exogenous characteristics.[4] In our context this would imply studying under which conditions more (initially) able workers sort themselves into more productive firms. Here I do not consider this type of sorting, but focus on the implications of the sorting dynamics that results from the interaction of on-the-job search and human capital accumulation. As shown below these sorting dynamics have interesting implications for wage inequality.[5]

4.2 Equilibrium wage dispersion

The above arguments imply that in this economy wages are disperse due to (i) the distribution of initial abilities; (ii) worker productivity differences through learning-by-doing; (iii) differences in z that arise due to search frictions; 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:

$$G(w) = \int_{\underline{\varepsilon}}^{\overline{\varepsilon}} \int_{z'=\underline{z}}^{\overline{z}} \left[\int_{y'=\varepsilon}^{w/z'} h_\varepsilon(y', z') dy' \right] dz' d\varepsilon, \tag{9}$$

where

[3] I relegate the derivations of $\partial H_\varepsilon(. \mid y)/\partial y < 0$ and $\partial H_\varepsilon(. \mid p)/\partial p < 0$ to the Appendix B as they are similar to the ones performed by Burdett et al. (2011).

[4] See Lopez de Melo (2009), Lentz (2010), Kircher and Eeckhout (2011) among others.

[5] One way of incorporating into the present analysis sorting due to, for example, workers' innate characteristics is by allowing workers of different abilities to choose their search effort (see Lentz, 2010). Assuming that more able workers are also better at searching, it would be reasonable to conjecture that Lentz (2010) results would also hold here and that one would obtain positive sorting by initial abilities.

$$h_\varepsilon(y,z) = \begin{cases} \partial^2 H_\varepsilon(.,.)/\partial y \partial z & \text{for } y \geq \varepsilon \text{ and } z \in [\underline{z}, \overline{z}] \\ 0 & \text{otherwise.} \end{cases} \tag{10}$$

In Appendix B I provide a detailed derivation of this wage distribution. It is straightforward to verify that this wage distribution retains the same properties as the ones derived in Burdett et al. (2011). In particular, it exhibits a density with an interior mode, it is right skewed and has a long and "fat" right tail. Generating these properties is of particular importance since it is well known that to approximate the empirical wage density in the B/M framework one needs to assume a distribution of firm productivities that has an implausible long right tail or more wage competition.[6] An innovation of incorporating human capital accumulation to the B/M framework is that one can, in principle, generate wage densities with similar properties as the empirical one using plausible firm productivity distributions.

4.3 Variance decomposition of (log) wages

The structure of the model provides a simple way to decompose the variation of cross-sectional (log) wages implied by (9). Since the wage of any worker i at time t, employed at firm j with experience x can be described by the wage equation

$$\log w_{ijt} = \log \varepsilon_i + \rho x_{ijt} + \log z_j, \tag{11}$$

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

$$var(\log w) = var(\log \varepsilon) + \rho^2 var(x) + var(\log z) + 2\rho cov(\log z, x). \tag{12}$$

The first two terms describe the contribution of workers' productivities (initial abilities and human capital accumulation) in explaining wage dispersion. The third term measures the importance of frictional wage dispersion. Recall that here a firm with productivity p offers the same z to all its employees. Since different firms offer different z, variation in z captures the extend to which equally productive workers are paid differently. Further, as x and z are positively correlated but orthogonal to ε, the last term then measures the contribution of sorting dynamics to wage dispersion.[7]

[6] For example, Mortensen (2003) introduces more competition between firms by allowing them to choose their recruiting effort. Postel-Vinay and Robin (2002) also introduce more competition by allowing firms to perfectly discriminate workers by their reservation wages and counter offer their employees' outside offers.

[7] Note that in the event of no on-the-job learning ($\rho = 0$) or no on-the-job search ($\lambda_e = 0$) or both, this 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, experience does not have any real effect on wage dispersion because it does not affect firms pay policies. 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.

Since $z = p\theta$, one can further decompose the variance attributed to frictional wage dispersion into

$$var(\log z) = var(\log p) + var(\log \theta) + cov(\log p, \log \theta). \quad (13)$$

Here the first term captures the direct contribution of differences in firms productivities in explaining wage dispersion. The last two terms captures the effects of difference in firms' piece rate policies on wages. In particular, the second term captures the variation in pay policies that arises under firm homogeneity (see Burdett et. al, 2011); while the last term captures the impact of firm heterogeneity in pay policies differentials as described by $\theta(p) = \zeta(p)/p$.

Taken together (12) and (13) provide a rich decomposition of the variation in (log) wages. On the one hand, it captures the effects due to differences in workers abilities ($var(\log \varepsilon)$), firms productivities ($var(\log p)) + cov(\log p, \log \theta)$) and difference in pay policies through piece rate offers ($var(\log \theta)$); while also capturing experience effects through general human capital accumulation ($\rho^2 var(x)$) and sorting dynamics ($2\rho cov(\log z, x)$). Thus the model is able to encompass a similar variance decomposition as analysed by Abowd, et. al (1999) and Postel-Vinay and Robin (2002), who consider the effects of workers' abilities, firms' productivities and a residual term that Postel-Vinay and Robin attribute to "market frictions". At the same time the model captures the effects of labour market experience on wage variation as is the focus in the traditional applied labour literature (see Rubinstein and Weiss, 2007, for an insightful survey).

In the quantitative section of the paper I use this decomposition to analyse the importance of difference in workers' productivities and search frictions in explaining wage variation. However, before turning to this analysis, I show that the model also has important implications for understanding (i) workers' average wage-experience profile and (ii) the extend of frictional wage dispersion as measured by the Mean-min ratio (see Hornstein et. al, 2011).

4.4 Average wage-experience profile

An alternative way to analyse the wage variation summarised in (9) is by using the average wage-experience profile. In particular, the wage equation (11) implies that the average (log) wage conditional on x is given by

$$E(\log w \mid x) = E(\log \varepsilon) + \rho x + E(\log z \mid x), \quad (14)$$

where the average earned log z conditional on x is given by

$$E(\log z \mid x) = \log z_R + \int_{\underline{z}=z_R}^{\overline{z}=\zeta(\overline{p})} \frac{1 - H(z \mid x)}{z} dz.$$

Here there are two sources by which average (log) wages increase with labour market experience. The first one is due to general human capital accumulation and encompasses the traditional explanation of why more experience workers are paid higher wages. The second one is due to workers' sorting dynamics through on-the-job search and the positive correlation between x and z. The latter implies that average wages are an increasing and concave function of experience. Indeed, noting that $y = \varepsilon e^{\rho x}$, $F(z) = \Gamma_0(p)$ and using a change of variable in $H(z \mid y)$ yields

$$H(z \mid x) = F(z)\left[\frac{\delta(\phi+\delta+\lambda_u) + (\phi+\delta+\lambda_e)(\phi+\lambda_u)(1-F(z))e^{-\left(\frac{\delta\lambda_u+\lambda_e(\phi+\lambda_u)(1-F(z))}{(\phi+\lambda_u)}\right)x}}{(\phi+\delta+\lambda_u)[(\phi+\lambda_e)(1-F(z))+\delta]}\right].$$

Since $\partial H(z \mid x)/\partial x < 0$ and $\partial^2 H(z \mid x)/\partial x^2 > 0$, differentiation implies $E(\log z \mid x)$ is an increasing and concave function of experience and so is $E(\log w \mid x)$. Further, (14) imply that the parameters that govern worker turnover, human capital accumulation and the firm productivity distribution determine the shape of the wage-experience profile.

The turnover parameters and the parameters of the firm productivity distribution influence the shape of the wage-experience profile through $E(\log z \mid x)$. The rate of human capital accumulation, however, influences the average wage-experience profile through the term ρx and through $E(\log z \mid x)$. It can be shown, for example, that decreasing the amount of search frictions faced by employed workers or increasing the rate of human capital accumulation makes the average wage-experience profile steeper, while increasing the job destruction rate makes the average wage-experience profile flatter.[8]

An important implication of the model is that, given a set of transition parameters and a rate of human capital accumulation, one can use the empirical wage-experience profile to recover (parametrically) the distribution of firm productivities. I exploit this insight in the quantitative section of the paper.

4.5 Frictional wage dispersion

Next consider the effects of human capital accumulation and on-the-job search on frictional wage dispersion. In particular, an important contribution of the B/M model is that it gives a theory of why workers with similar characteristics are paid differently. Hornstein, et. al, (2011) 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 (2011) 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. As shown earlier, the same logic applies here.

In the present model frictional wage dispersion refers to wage variationcinduced only by variation in z. For this reason the Mm ratio should be defined by the ratio between the average z, z^M, and

[8]An increase in the job destruction rate decreases the slope of the wage-experience profile at all experience levels when $\lambda_u = \lambda_e$. When $\lambda_u > \lambda_e$, however, the profile might becomes stepper at low experience levels.

the minimum z, z_R, earned by employed workers. Using (3), (5) and following Hornstein, et. al (2011) approximation it is shown in Appendix B that the Mm ratio with on-the-job search and human capital accumulation is given by

$$Mm \cong \left[1 + \frac{\lambda_u(r+\phi-\rho) - (r+\phi)\lambda_e}{(r+\phi)(r+\phi+\delta+\lambda_e-\rho)}\right] / \left[\frac{r+\phi-\rho}{r+\phi}\chi + \frac{\lambda_u(r+\phi-\rho) - (r+\phi)\lambda_e}{(r+\phi)(r+\phi+\delta+\lambda_e-\rho)}\right], \quad (15)$$

where $\chi = z_b/z^M$ denotes the (relative) value of non-market time. Note that (15) collapses to the expressions analysed by Hornstein, et. al (2011).

To illustrate the quantitative performance of the Mm ratio when both on-the-job search and human capital accumulation are allowed, consider the same parametrisation as the one used by Hornstein, et. al (2011) for the US economy. Their baseline calibration implies $\lambda_u = 0.43$, $\delta = 0.03$, $\chi = 0.4$ and $r = 0.0041$. With no on-the-job search or human capital accumulation, (15) yields an $Mm = 1.046$. This is significantly lower than their estimated Mm ratios.[9] As these authors show, allowing for human capital accumulation improves the performance of the job search model, but still falls short from generating an empirically plausible Mm ratio. Indeed, they set $\phi = 0.0021$ (an average working life of 40 years) and $\rho = 0.0017$, which implies an $Mm = 1.076$. On the other hand, they consider on-the-job search on its own (i.e $\rho = \phi = 0$) and set $\lambda_e = 0.13$, generating a monthly job-to-job transition rate of approximately 3.2 percent (see Moscarini and Vella, 2008). In this case, (15) delivers an $Mm = 1.27$. Clearly on-the-job search has the biggest impact in improving the ability of the job search model to match the data, but again falls short from generating an empirically plausible Mm ratio. Now consider both on-the-job search and human capital accumulation. Using all the above parameter values (15) implies an $Mm = 1.49$, which is within the bounds of their estimated Mm ratios.[10] Thus the present framework seems able to generate a plausible measure of frictional wage dispersion under Hornstein et. al (2011) calibration. Below I show that this is also the case under an alternative calibration based on data for the UK labour market.

5 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 dynamics on the cross-sectional distribution of wages. To do so I calibrate the model to match salient features of the UK labour market using the British Household Panel Survey (BHPS).

[9] Hornstein et. al (2007) estimate Mm ratios from the PSID, OES and Census data. For each of these data sets they obtain a summary Mm ratio of 1.46, 1.67 and 1.98, respectively, using the ratio between the mean wage and the wage at the 5th percentile of the "frictional" wage distribution.

[10] Letting $\phi = 0.0021$ implies that to match a job-to-job transition rate of 3.2 one needs to set λ_e slightly lower; ie. $\lambda_e = 0.125$.

5.1 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 was subsumed by the new and bigger survey "Understanding Society" in 2010. The BHPS contains 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 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 white male individuals that were originally sampled in 1991 and were between 16 and 30 years of age at that time. 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). I choose to study a relative young sample of the population for the following reason. To keep the model as parsimonious as possible, I assumed that workers' human capital accumulation process is not subject to decreasing returns. This allows me to solve the model in close form and derive most of the results analytically. As shown below it also allows me to apply a simple calibration procedure to recover the parameters of the model. Excluding decreasing returns to human capital accumulation, however, implies that the model has to capture the concavity of the empirical wage-experience profile through the decreasing returns to on-the-job search. Since employer to employer transitions are more prominent among young workers and tend to fade as workers enter their late twenties or early thirties, it seems appropriate to evaluate the model on this set of workers.

To keep the model's assumptions as consistent as possible with the sample of the population studied, I only focus on low and medium skilled workers.[11] As pointed out by Postel-Vinay and Turon (2011), it seems reasonable to assume that high skilled workers are in labour markets in which firms tend to react to workers' outside offers, a characteristic that is not present in the contracts posted by firms in the present model.[12] I further restrict attention to those spells that lasted at least one month in which the individual was in (i) paid (dependent) full-time employment in the private sector or (ii) unemployment. To keep the sample as homogeneous as possible I only

[11] Following Dustmann and Pereira (2008), I 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, other higher qualifications (but no university degree or higher) and A levels.

[12] Postel-Vinay and Turon (2011), using the BHPS for a similar period, used a sequential auctions model based on Postel-Vinay and Robin (2002) to study the dynamics of wages among high skilled workers and show that this wage mechanism gives a good description of high skilled workers wage dynamics and their average wage-experience profile.

considered those employment and unemployment spells that occur before an individual reported he became (if at all) self-employed, a civil servant, worked for the central or a local government or the armed forces, long-term sick or entered retirement. I also dropped those individuals that had at least one spell in full-time education or in government training.

These restrictions leave me with a sample of 1,722 individuals, where 486 are considered low skilled and 1,236 medium skilled.[13] The average potential experience of low and medium skilled workers is 12.23 and 12.75 years, while their average actual experience is 9.35 and 11.01 years, respectively.[14] 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.[15] I consider as my earnings variable the real hourly (gross) wage of these individuals.[16] It is worth pointing out that wage data is only available as from 1991, the first wave of the BHPS. Overall there are 11,316 spells in the sample, where 3,434 of those are associated with low skilled workers and 7,882 are associated with medium skilled workers. There are 7,984 spells that ended after the BHPS started, where 2,419 are associated with low skilled workers and 5,565 are associated with medium skilled workers.

5.2 Calibration

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

5.2.1 Time Discounting

I start by calibrating the workers' discount rate $\phi + r$. Consider a time period to be equal to a month. Following Hornstein et. al (2011), I let $r = 0.0041$ and fix $\phi = 0.0021$ such that all workers are assumed to have on average 40 years of potential labour market experience. Note that interpreting r as the monthly real interest rate, its value implies an annual interest of 5 percent consistent with the standard values chosen for the US and other OECD countries. The total monthly discount rate

[13] The disparity in the size of these two groups arises primarily due to the high number of workers with at most O-level qualifications, which represents 53.2 percent of the medium skilled group.

[14] Actual experience is obtained from the sum of employment spells, while potential experience is obtained from the sum of employment and unemployment spells.

[15] Since I do not count spells that shorter than a month a transition in which the individual changed employer but experienced an intervening spell of unemployment of less than a month is considered a direct job-to-job transition. If the individual experiences an unemployment spell longer than a month, then he is considered unemployed. See Jolivet et al. (2006) for a similar assumption.

[16] Following Dustmann and Pereira (2008), I 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. I also take into account overtime hours and use the CPI to deflate nominal wages.

Table 1: Transition Parameters and Average Spell Durations in Months

	UD	λ_u	ED	δ	JD	λ_e	PrEE	Urate (%)
Low skilled	12.35	0.081	41.63	0.022	35.68	0.010	0.004	22.88
Medium skilled	7.08	0.141	70.16	0.012	44.66	0.038	0.011	9.16

is then $\phi + r = 0.0062$, similar to the one chosen by Bagger et. al (2011).

5.2.2 Transition Rates

The parameters in Θ_T are obtained using duration data. The model implies that any unemployed worker accepts the first job offer he encounters. Given that for these workers job offers arrive following a Poisson process with rate λ_u, the average duration of unemployment is given by $UD = 1/\lambda_u$. Information on unemployment duration is then sufficient to recover this parameter. Similarly, the model implies that the average duration of an employment spell is $ED = 1/(\phi+\delta)$. For a given value of ϕ, the value of δ is then obtained from information on employment duration. To obtain the value of λ_e we use information on job duration. To do this, note that the average unconditional job duration implied by the model is

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

For given values of ϕ and δ one can then recover the value of λ_e using the average unconditional job duration computed from the data.[17]

Table 1 shows the set of values for the transition parameters, the corresponding spell durations, the implied unconditional probability of a job-to-job transition

$$\Pr EE = (\phi + \delta) \left[\left(1 + \frac{\phi + \delta}{\lambda_e}\right) \ln\left(1 + \frac{\lambda_e}{\phi + \delta}\right) - 1 \right]$$

and the unemployment rate generated by the model for each skill group.[18] It highlights three

[17]A common alternative to recover the transition parameters in Θ_T is to (structurally) estimate them using the likelihood function implied by the theory; see, for example, Bontemps et. al (2000) and Jolivet et. al (2006). In those models, workers decide to change jobs based on the wage posted by the firm and hence one can use information on duration *and* wages to implement the Maximum Likelihood (ML) estimation. In the present model, however, workers decide to change jobs based on $z = p\theta$, which is unobservable to the econometrician, and hence one cannot use wage information to recover the transition parameters. One way to overcome this issue is to use information on spell duration and apply the ML estimation procedure by integrating out wages from the likelihood function (see Jolivet et. al, 2006). Here, however, I follow Hornstein et. al (2007) and use information on spell duration and the model's implications on average spell durations to recover the transition parameters. This is similar as using information on unemployment to employment transition (and vice versa) and job-to-job transitions to estimate the transition parameters by matching the model's transition probabilities to the empirical ones (see Bagger et. al, 2011).

[18]To construct spell durations I have only considered those spells that ended after December 2011 as information

important features that characterise these workers' labour markets.

First, it shows that the reallocation of low skilled workers across jobs is more likely to occur through unemployment than through direct job-to-job transitions, while for medium skilled workers both channels are equally likely. Indeed, low skilled workers, are twice more likely to experience an employment to unemployment transition than receive an outside offer, and five times more likely to experience unemployment than to change jobs. Medium skilled workers are about four times more likely to receive an outside offer than transit to unemployment or to change jobs. Second, low skilled workers have less time to accumulate general human capital than medium skilled workers. Low skilled workers face a risk of unemployment that is about as twice as high as the risk of unemployment faced by medium skilled workers, while at the same time the latter are more likely to leave unemployment faster. Third, the steady state unemployment rate of low skilled workers is more than twice as high as that of medium skilled workers.[19]

5.2.3 Productivity Parameters

The parameters in Θ_P are obtained using wage data. I do this in three steps. Given the transitions parameters, I use the Mm ratio to recover workers' human capital accumulation rate. Then use the empirical average wage-experience profile to recover the distribution of firm productivities. Finally, to recover the distribution of workers' abilities I use the kernel estimate of the log wage density.

Human Capital Accumulation

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. (2011), the mean-min (Mm) ratio, described in (15). Following Hornstein, et al. (2007) I first estimate the wage equation

$$\log w_{ijt} = \beta X_{it} + \eta_{ijt}, \tag{16}$$

for each year of the sample period and skill group using OLS, where X is a vector of covariates consisting of a quartic in actual experience, a dummy for marital status, 8 regional dummies, 8 (one-digit) occupational dummies and 8 (one-digit) industry dummies, and 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 $\widehat{\eta}_{it}$ and their individual specific mean $\overline{\eta}_i = \sum_{t=1}^{N_i} \widehat{\eta}_{it}/N_i$. The vector $\{\overline{\eta}_i\}_{i=1}^{N}$ then captures the wage variation due to fixed unobserved individual factors (i.e. innate ability). Finally, I use the estimated distribution of transformed wages, $\widetilde{w}_{it} = \exp(\widehat{\eta}_{it} - \overline{\eta}_i)$, across individuals and time to calculate the Mm ratio for each skill group.

on wages, which I am going to use to recover the productivity parameters, is not available before 1991.

[19] These results also hold true when spells of part-time work are considered. The transition parameters for low skilled workers in full and part-time jobs are $\lambda_u = 0.091$, $\delta = 0.024$ and $\lambda_e = 0.018$ which imply that $PrEE = 0.0075$ and $Urate = 22.35\%$. For medium skilled workers the transition parameters are $\lambda_u = 0.139$, $\delta = 0.017$ and $\lambda_e = 0.039$ which imply that $PrEE = 0.013$ and $Urate = 12.07\%$.

Table 2: Mean-min Ratios and Human Capital Accumulation

	Low skilled	Medium skilled
Mm	1.82	4.31
Mm_1	1.59	1.70
Mm_5	1.35	1.38
Mm_a	1.47	1.54
ρ	0.0019	0.0020

Table 2 shows the set of Mm ratios obtained using the minimum observed wage Mm, the wage at the first percentile (Mm_1) and fifth percentile (Mm_5), where I have normalised $b = 1$ so that unemployed workers obtain the entire returns to household production (and hence $z_b = p_h$), and assumed $\chi = z_b/z_M = 0.4$ as in Hornstein, et al (2011). Given that the wage data has already been trimmed by 5 percent on each side when performing the OLS regressions and that the minimum observed wage is still very noisy (at least for the medium skilled category), I use as a target Mm_a, the mean-min ratio obtained from averaging the ones obtained for the first and fifth percentile.[20] The last row shows the rate of human capital accumulation consistent with observed frictional wage dispersion, given the set of transition parameters that are consistent with observed spell durations.

First note that the amount of frictional wage dispersion found in the BHPS for young workers is consistent with the ones obtained by Hornstein et al. (2007) for the US. Further, the amount of frictional wage dispersion as measured by the Mean-min ratio is higher for medium skilled workers. Second, the implied rates of human capital accumulation are similar across skill groups. They imply that these workers accumulate general human capital at a rate of around 2.4 percent per year, which is also consistent with $\rho = 0.0017$, the value chosen by Hornstein et al. (2011) as the average human capital accumulation rate among US workers. These results confirm the conjecture made in section 4.5, that a model that incorporates both on-the-job search and general human capital accumulation is consistent with the amount of frictional wage dispersion observed in the data under reasonable parameter values.[21]

To understand why the values of ρ are similar across skill groups consider equation (15). Note that the Mm ratio is increasing in both λ_e and δ as both reduce unemployed workers' reservation wages, and is decreasing in λ_u as it increases unemployed workers' reservation wages. Now consider the transition parameters in Table 1. They show that low skilled workers have a higher job destruction rate and lower job offer arrival rates than medium skilled workers. Quantitatively these differences seem to balanced each other and any difference in the Mean-min ratios of these workers must arise from differences in their human capital accumulation rates. Since both groups

[20]Trimming the data by 5 percent at the left tail of the wage distribution not only helps to reduce measurement error, but also to consider all jobs that pay above the national minimum wage, introduced in the UK in 1999.

[21]If one does not consider human capital accumulation or on-the-job search, the model generates an Mm ratio of 1.18 and 1.07 for low and medium skilled workers. Adding on-the-job search increases the Mm ratios to 1.27 for both skill groups.

Table 3: Productivity Parameters and Returns to Experience

	Low skilled	Medium skilled
Productivity		
$\underline{p}$ (location)	1.670	7.300
κ_1 (shape)	1.050	0.500
κ_2 (scale)	0.450	0.400
p_h	0.718	3.160
Reservation value		
z_r	0.902	4.744
var log$(\widetilde{w}_{it})$		
Data	0.0330	0.0389
Model	0.0326	0.0390
Returns to Exp.	$\log w = \beta_0 + \beta_1 x + \beta_2 x^2 + \eta$	
Data		
β_1	0.00335	0.00395
β_2	-6.26e-06	-7.38e-06
Model		
$\hat{\beta}_1$	0.00310	0.00375
$\hat{\beta}_2$	-3.70e-06	-4.85e-06

face somewhat similar values of Mm_a, they must also exhibit similar human capital accumulation rates.

Firms Productivity Distribution

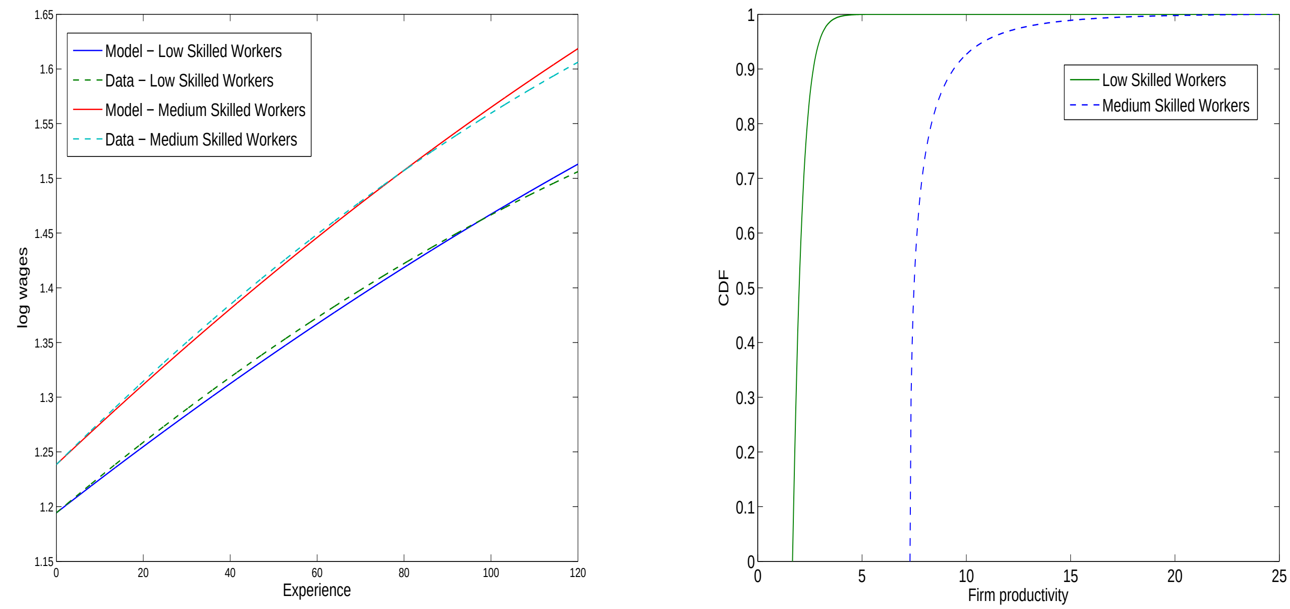
Given the transition parameters and human capital accumulation rates, I calibrate the firm productivity distribution, Γ, for each skill group using their average wage-experience profile. I approximate this distribution using the Weibull distribution

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

where κ_1 describes the shape parameter, κ_2 describes the scale parameter and $\underline{p}$ describes the location parameter. The choice of this functional form follows Bagger et al. (2011), whom also use it to approximate the firm productivity distribution in their sequential action model with human capital accumulation.

I calibrate the shape and scale parameters such that the model is able to replicate the first 10 years of the empirical average wage-experience profile for each skill group. The empirical wage experience profile is obtained by regressing log wages on a quadratic on experience, a quadratic on tenure and the other controls used in (16) to calculate the Mean-min ratio.[22] I choose $\underline{p}$ such

[22] This specification is the same to the one used to compute the Mean-min ratio with the exception that for the



(a) Wage-experience profile

(b) Firm productivity distributions

Figure 1: Workers' Average Wage-experience Profile and Firms' Productivities

that variance of $\log(z)$, used in the variance decomposition exercise as a measure of frictional wage dispersion, matches the variance of the distribution of log wage residuals used in calculating the Mean-min ratio, such that the two measures of frictional wage dispersion are consistent with each other. Details of the simulation procedure can be found in Appendix C.

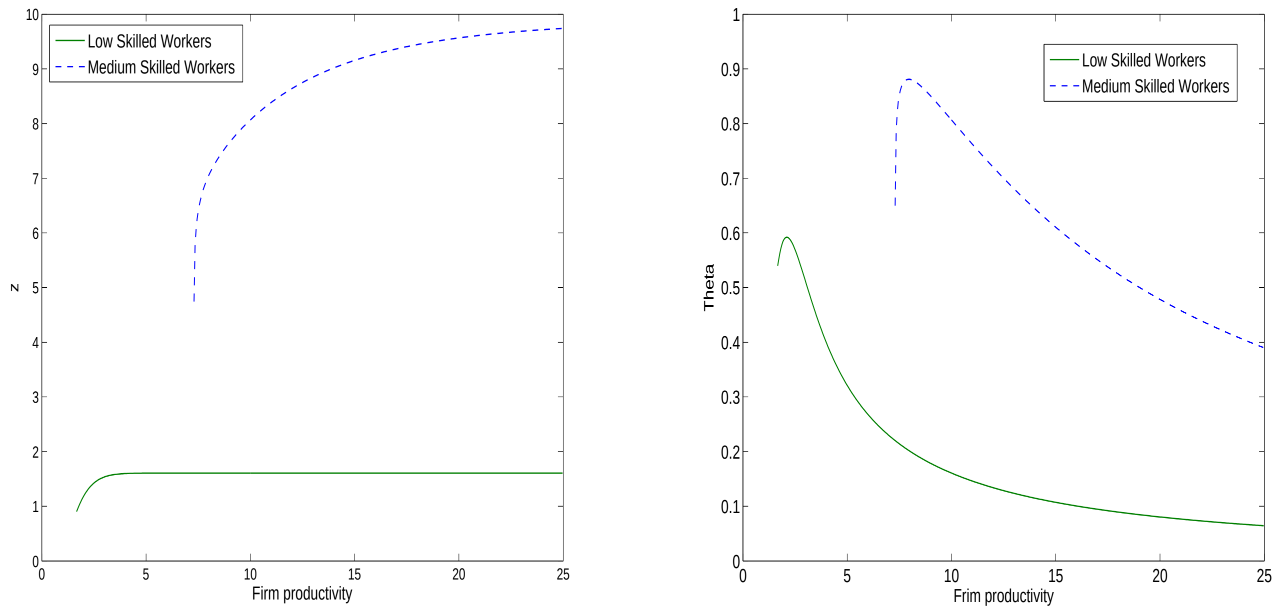
Once these parameters are determined, the household productivity can be obtained by noting that $b = 1$ implies $p_h = \chi z_M$, where z_M can be approximated by

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

and p_h determines the value of z_R as the solution of (8).

Table 3 shows the values of these parameters, the solution for z_R and the estimated coefficients of the average wage-experience profiles, where the coefficients of the wage-experience profiles obtained from the model lie within the 95 percent confident intervals of the ones estimated in the data. Figure 1.a depicts the average wage experience profiles obtained from the model and the data by

latter I did not include a quadratic on tenure and considered a quartic on experience. The reason for the latter is that I wanted to be as close as possible to the specification used by Hornstein et al. (2007). However, if one includes a quadratic on tenure an uses a quadratic on experience to estimate the Mean-min ratio, the results do not change significantly. Indeed, the Mean-min ratios obtained under the latter specification for low skilled workers are $Mm = 1.86$, $Mm_1 = 1.62$, $Mm_5 = 1.36$ and $Mm_a = 1.49$. For medium skilled workers the Mean-min ratios are $Mm = 4.35$, $Mm_1 = 1.66$, $Mm_5 = 1.38$ and $Mm_a = 1.52$.



(a) z-functions

(b) θ-functions

Figure 2: Offered z and θ as Functions of Firm Productivities

skill groups. Note that the values of z_r imply that $p_0 = \underline{p}$ and by virtue of Theorem 1 the above parameter values yield a unique equilibrium for each skill level. That is, all firms at the left tail of the productivity distribution find it profitable to participate in the labour market as unemployed workers are willing to accept a low wage to become employed and accumulate work experience. Further, the values of κ_1 and κ_2 generate productivity distributions with declining densities for all cases, similar to the ones estimated by Bagger et al. (2011). Figure 1.b depicts the firm productivity distributions for low and medium skilled workers. It shows that the firm productivity distribution for medium skilled workers is to the right of the firm productivity distribution for low skilled workers, which implies that medium skilled workers access more productive jobs that low skilled workers.

The differences in the firm productivity distributions across skill groups mainly arise from the difference in the way these workers reallocate across jobs, as both types of workers have similar human capital accumulation rates and their wage-experience profiles are not too dissimilar. Further, Figure 2.a shows that the firm productivity distribution that allows the model to match the observed average wage-experience profile generates very different z functions for low and medium skilled workers. The z function for low skilled workers increases realtively fast and reaches its maximum at relatively low levels of firm productivity. For medium skilled workers, however, the z function increases less abruptly with firm productivity. Their relative position indicate that medium skilled

Table 4: Ability Parameters and the Mean and Variance of Log Wages

	Low skilled	Medium skilled
Ability		
$\underline{\varepsilon}$ (location)	0.474	0.062
α_1 (scale)	1.250	0.070
α_2 (shape)	1.070	2.620
Log Wages		
Variance		
BHPS	0.0835	0.0923
Model	0.0830	0.0918
Mean		
BHPS	$-5.941e-10$	$-3.741e-11$
Model	$-5.375e-03$	$-7.973e-03$

workers not only access more productive jobs than less skilled workers, but that they are paid higher wages than their low skilled counterparts.

Figure 2.b shows the offered piece rates, θ, associated with the z functions. In both cases these functions are hump-shaped but mostly decreasing with firm productivity. This property implies that, except for firms with very low productivities, there is a positive relation between firm productivity and their level of monopsony power. This is similar to the result obtained by Bontemps et al. (2000) when estimating the B/M model with firm heterogeneity. They find that, except for very low productive firms in some sectors, more productive firms have higher values of the monopsony power index $1 - w(p)/p$, where $w(.)$ is an increasing function that associates offered wages to firm of productivity. In the context of the present model this is equivalent to the monopsony power index $1 - \theta(p)$, which increases with p for high enough productivities. Taken together, the above results imply that medium skilled workers end up employed in more productive firms, earning higher wages and facing less monopsony power from their employers than low skilled workers.[23]

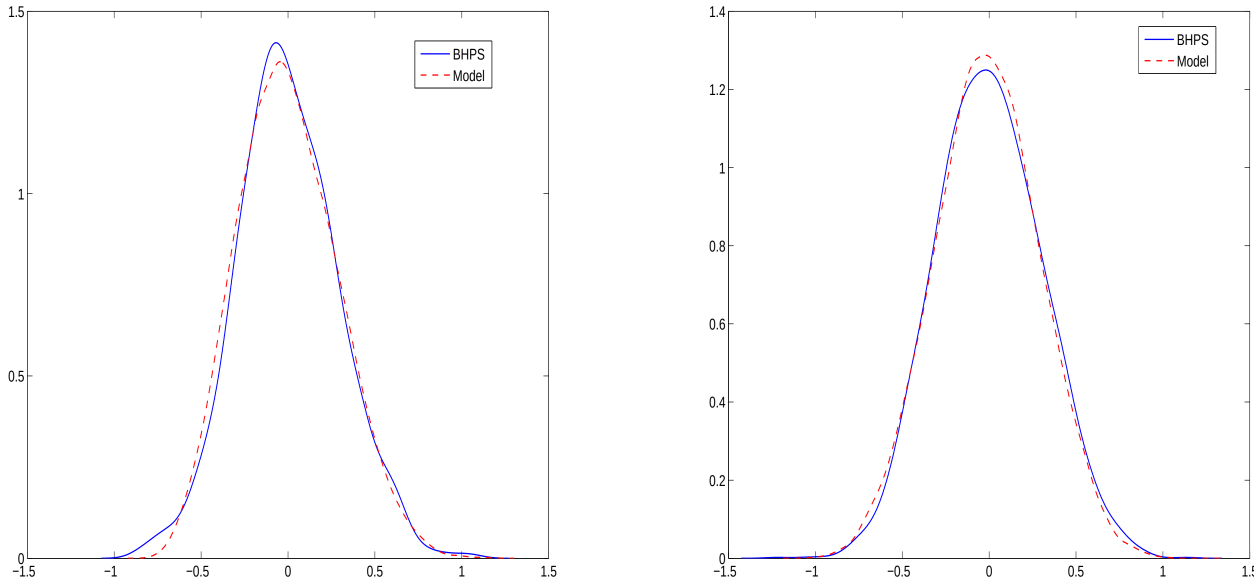
Workers Ability Distribution

Given the above parameters, the last step is to recover A, the worker ability distribution. I do so by matching the mean and variance of cross sectional (log) wage distribution implied by the data. As before, I parametrize the ability distribution to follow a Weibull distribution. In particular, let

$$A(\varepsilon) = 1 - e^{-\left(\frac{\varepsilon - \underline{\varepsilon}}{\alpha_2}\right)^{\alpha_1}},$$

where α_1 describes the shape parameter, α_2 describes the scale parameter and $\underline{\varepsilon}$ describes the

[23]The latter can be observed in Figure 2.b by noting that the piece rate function, $\theta(p)$, is higher for medium skilled and by noting that in the B/M framework a lower λ_e implies there is less competition among firms and hence the offer distribution F is closer to the one predicted by Diamond (1971).



(a) Low ability workers

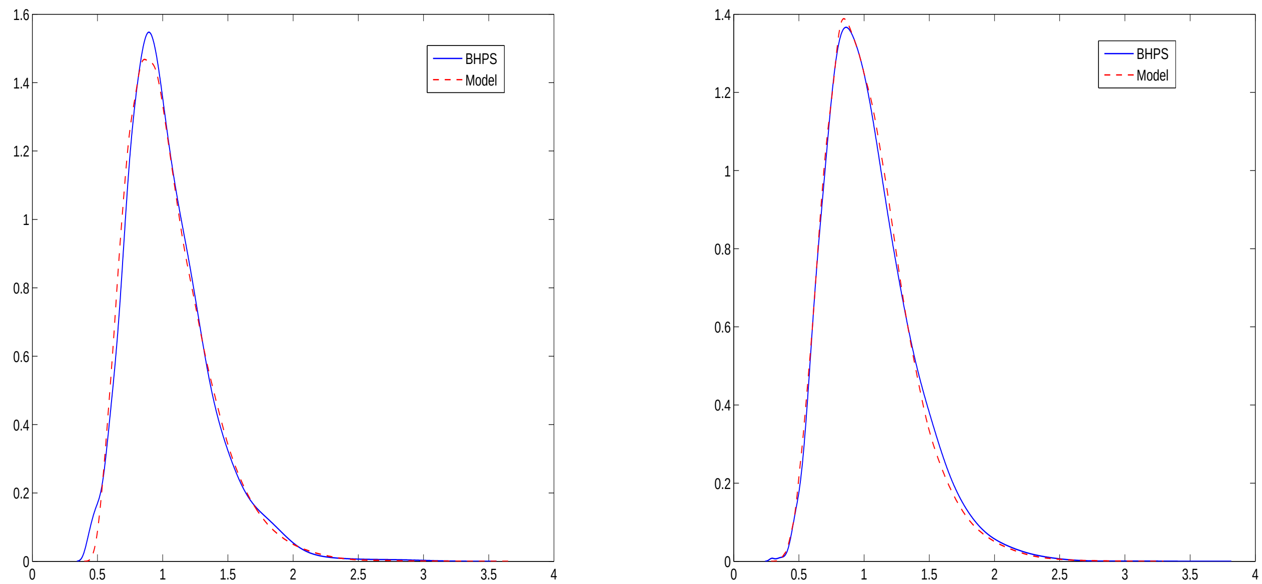
(b) Medium ability workers

Figure 3: Log Wage Densities

location parameter. I choose these parameters such that the log wage density implied by the model approximates the kernel density of the log wage residuals obtained from estimating the same log wage equation used to calibrate the firm productivity distribution parameters, but without the experience effects (see Appendix C for details of this wage equation). In particular, I choose $\underline{\varepsilon}$ to match the lowest observable wage residual in the data for each skill group. Further, the shape and scale parameters are calibrated to match the mean and variance of the kernel density of log wage residuals, such that one cannot reject the hypothesis that data generated by the model and the one obtained from the BHPS are drawn from different cdf based on the two-sample Kolmogorov-Smirnov goodness-of-fit hypothesis test. Table 4 shows the parameter values of the ability distributions and the corresponding moments.

6 Wage Dispersion

Figure 3 show the kernel densities of log wages obtained from the model and the BHPS, where the latter is based on the log wage residuals as explained in the previous section. Figure 4 show the kernel densities in levels. The model is able to fit the overall shape of the empirical wage density distribution quite well without relying on firm productivity distributions that exhibit unrealistic long tails. This result is important as the B/M model has been criticised by its inability to match the empirical wage distribution without firm productivity distributions that exhibit very long tails.



(a) Low ability workers　　(b) Medium ability workers

Figure 4: Wage Densities

Table 5 considers the wage variance decomposition implied by the model. The first four columns present the variance decomposition described in (12) and the last three the one described in (13). For each skill level the first two rows show both the variance of each of its components in levels and as a proportion of the variance of log wages reported in Table 4. The third row shows the variance decomposition (13) as a proportion of the variance of $\log(z)$. These figures show that frictional wage dispersion and experience effects are able to explain between 60 and 70 percent of overall wage dispersion across skill groups. Unobserved heterogeneity (ability), unrelated to firm's wage posting strategies, workers' job search and human capital accumulation process, explain the rest of the variation: 30 for medium skilled workers and 40 percent low skilled workers.

Frictional wage dispersion, as measured by the variance of $\log(z)$, accounts for around 40 percent of the variation in log wages across medium and low skilled workers. The contribution of frictional wage dispersion for medium skilled workers, however, is higher than that for low skilled workers. The last three columns of Table 5 show that this difference arises due to contribution of earned piece rates differentials, $\log(\theta)$. Piece rate differentials are 3 times more important in accounting for frictional wage dispersion among medium skilled workers than among low skilled workers, while they are also 3 times more important in accounting for overall wage dispersion among the medium skilled than among low skilled workers. On the other hand, firm productivity differentials contribute a substantial but similar proportion to frictional wage dispersion and to overall wage dispersion

Table 5: Variance Decomposition of Log Wages

var(log w)	var(logε)	ρ^2 var(x)	var(log$z$)	2ρcov(x,log$z$)	var(log$p$)	var(logθ)	cov(log$p$,logθ)
Low skilled							
0.0830	0.0326	0.0165	0.0316	0.0023	0.0344	0.0039	-0.0067
100.00 (%)	39.28	19.88	38.07	2.77	41.45	4.70	-8.0730
			100.00 (%)		108.86	12.34	-21.20
Medium skilled							
0.0918	0.0281	0.0169	0.0390	0.0078	0.0418	0.0125	-0.0153
100.00 (%)	30.61	18.14	42.48	8.50	45.53	13.62	-16.67
			100.00 (%)		107.18	32.05	-36.60

across these skill groups. To obtain their overall effect I consider the contributions of both firm productivities (on their own) and their covariance with earned piece rates as this last term would not arise when all firms have the same productivity (Burdett et al., 2011). The two terms together imply that firm productivity differentials account for 71 and 88 percent of frictional wage dispersion, and 29 percent and 33 percent of total wage variation for medium skilled and low skilled workers, respectively.

Note that if one controls for experience effects on wage differentials and only considers the importance of workers' unobserved heterogeneity (ability), firms unobserved heterogeneity (productivity) and search frictions (piece rate), as is the objective in Abowd, et. al (1999) and Postel-Vinay and Robin (2002), the above variance decomposition shows workers' unobserved heterogeneity explains 42 percent of the wage variation among medium skilled workers, while it explains 51 percent of the wage variation among low skilled workers. Firm productivity differentials explain 39 percent of the wage variation among medium skilled workers and 43 percent of the variation among low skilled workers. Search frictions explain 19 percent of the wage variation among medium skilled workers and only 6 percent of the wage variation of low skilled workers. Interestingly, these contributions seem closer to the ones obtained by Abowd, et. al (1999) than to those obtained by Postel-Vinay and Robin (2002).

Now consider the importance of labour market experience in accounting for overall wage dispersion. Taking together the contribution of human capital accumulation and sorting dynamics, Table 5 shows that labour market experience is more important for medium skilled workers, accounting for 27 percent of overall wage dispersion; while for low skilled workers it accounts for 22 percent. Human capital accumulation on its own explains very similar proportions of the variation of log wages for both skill groups. The difference arises due to the importance of sorting dynamics between these groups. The contribution of sorting dynamics to medium skilled workers' wage dispersion is about 4 times that of low skilled workers, reflecting the relative importance job-to-job transitions have on the job reallocation and wage dynamics of medium skilled workers vis á vis low skilled workers. Overall, however, human capital accumulation is more important than sorting dynamics in accounting for the contribution of labour market experience for both skill groups. Human capital accounts for 68 percent of the experience effects, while sorting dynamics explain 32 percent of the experience effects among medium skilled workers. Among low skilled workers this difference is much more pronounced. Human capital accumulation explains 88 percent of the experience effects, while sorting dynamics only explains 12 percent. This suggest that among young workers human capital accumulation plays a more important role than job-to-job transitions in explaining wage differentials, consistent with the findings of Bagger et al. (2011) and Menzio et al. (2011), and this role is more pronounced among low skilled workers.

In summary, the results in Table 5 show that experience effects account for an important proportion of overall wage dispersion for both skill groups and these are mostly driven by human capital accumulation rather than on-the-job search. Workers' unobserved heterogeneity plays a more important role in accounting for wage dispersion among the low skilled, while frictional wage dispersion is more important in accounting for wage differences among medium skilled workers. These differences imply that wage dispersion among low skilled workers is driven mostly by productivity differentials among workers (ability and human capital accumulation), while wage dispersion among medium skilled workers is equally explained by workers' productivity differentials and search frictions (frictional wage dispersion and sorting dynamics).

7 Further Discussion

This paper constructs and quantitatively assesses a tractable equilibrium model with two sided heterogeneity in which workers accumulate general human capital and engage in on-the-job search. The analysis shows that a search model with these characteristics is able to replicate the shape of cross-sectional wage distribution, the average wage-experience profile, employment and unemployment durations and the Mean-min ratio for a sample young workers in the UK. Decomposing the variation of log wages I find important differences among low and medium skilled workers that reflect the the way these workers reallocate across jobs over time.

The results show that low skilled workers reallocate across jobs mainly through spells of unemployment. When employed these workers find jobs in low productive firms that offer low wages and exert a higher monopsony power (vis *á* vis medium skilled workers). The employment dynamics of these worker are similar to the ones Steward (2007) has described in an earlier paper using the BHPS, showing that low paid workers, who also tend to be low skilled, are caught in "low-pay, no-pay cycles". These employment dynamics imply that the on-the-job search model with general human capital accumulation proposed in this paper captures around 60 percent of the variation in log wages, after controlling for several observable characteristics of workers, while the residual variation is attributed to workers' unobserved heterogeneity. This result stands in contrast with the finding of Postel-Vinay and Robin (2002) using French data. They find that for workers performing low skilled jobs, unobserved heterogeneity performs a minor role in explaining the variations in log wages. However, in their case, the employment dynamics of these workers do not exhibit the pattern described above and hence their search model (without human capital accumulation) is able to explain a larger proportion of the variation in wages.

As noted above medium skilled workers present different employment dynamics. These workers change jobs both through unemployment and directly through job-to-job transitions. When employed they find better paid jobs in firms with higher productivity than their less skilled counterparts. Further, longer employment spells enables these workers to increase their wages over time through both human capital accumulation and on-the-job search. This implies that sorting dynamics become more important in explaining both the average wage-experience profile and the cross-sectional wage distribution. Overall, the model is able to capture 70 percent of their observed log wage variation.

In this paper I have analysed wage differentials among low and medium skilled workers assuming the economy is in steady state. An important question, however, is to what extend the employment patterns and the relative contributions of workers' productivity differentials vis *á* vis firm productivity differentials and frictional wage dispersion change over the business cycle. Recent work by Coles and Mortensen (2012) suggests that such an extension is possible and can remain tractable. I leave this important extension for future research.

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. 2011. "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] Bowlus, Audra and Huju Liu. 2012. "The Contribution of Search and Human Capital to Earnings Growth over the Life Cycle." CIBC Working Paper Series 2012-2, University of Western Ontario.

[12] 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.

[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. 2010. "Tenure and Experience Effects on Wages: A Theory." CESifo Working Paper Series 3203, CESifo Group Munich.

[15] Burdett, Kenneth and Melvyn G. Coles. 2003. "Equilibrium Wage-Tenure Contracts." *Econometrica*, 71(5): 1377-1404.

[16] Burdett, Kenneth, Carlos Carrillo-Tudela and Melvyn G. Coles. 2009. "Human Capital Accumulation and Labour Market Equilibrium." *International Economic Review*, 52(3): 657-677.

[17] Carrillo-Tudela, Carlos. 2009. "An Equilibrium Search Model with Optimal Wage-experience Contracts. " *Review of Economic Dynamics*, 12(1): 108-128.

[18] Coles, Melvyn G. and Dale Mortensen. 2012. "Equilibrium Labour Turnover, Firm Growth and Unemployment " ISER Working Paper No. 2012-07, University of Essex, UK.

[19] Diamond, Peter. 1971. "A Model of Price Adjustment."*Journal of Economic Theory*, 3(2): 156-168.

[20] Dustmann, Christian and Costas Meghir. 2005. "Wages, Experience and Seniority."*Review of Economic Studies*, 72 (1): 77-108.

[21] 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.

[22] Fu, Chao. 2011. "Training, Search and Wage Dynamics." *Review of Economic Dynamics*, 14(4): 650-666.

[23] Hagedorn, Marcus and Iourri Manovskii. 2010. "Search Frictions and Wage Dispersion. " Mimeo, Department of Economics, University of Pennsylvania.

[24] Hall, Robert. 1982. "The Importance of Lifetime Jobs in the U.S. Economy. " *American Economic Review*, 72(4): 716-24.

[25] Hornstein, Andreas, Per Krusell and Giovanni L. Violante. 2011. "Frictional Wage Dispersion in Search Models: A Quantitative Assessment." *American Economic Review.* 101(7): 2873-2898.

[26] Hornstein, Andreas, Per Krusell and Giovanni L. Violante. 2007. "Frictional Wage Dispersion in Search Models: A Quantitative Assessment." Mimeo. Department of Economics, NYU.

[27] 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.

[28] Kircher, Philipp and Jan Eeckhout. 2011. "Identifying Sorting in Theory. " *Review of Economic Studies*, 78(3): 872-906.

[29] Lopez de Melo, Rafael. 2009. "Sorting in the Labor Market: Theory and Measurement. " Mimeo, Department of Economics, University of Chicago.

[30] Lentz, Rasmuz. 2010. "Sorting by Search Intensity. "*Journal of Economic Theory*, 145(4): 1436-1452.

[31] Lentz, Rasmuz and Dale T. Mortensen. 2005. "Productivity Growth and Worker Reallocation."*International Economic Review*, 46(3): 731-749.

[32] Manning, Alan. 2003. *Monopsony in Motion: Imperfect Competition in the Labour Market*, Princeton, NJ, USA: Princeton University Press.

[33] 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, UK.

[34] Menzio, Guido and Shouyong Shi. 2011. "Efficient Search on the Job and the Business Cycle. " *Journal of Political Economy*, 119(3): 468-510.

[35] Menzio, Guido, Irena Telyukova and Ludo Visschers. 2011. "Directed Search over the Life cycle." NBER Working papers 17746, National Bureau of Economic Research, Inc.

[36] Mincer, Jacob. 1974. *Schooling, Experience and Earnings.* New York: National Bureau of Economic Research.

[37] Mortensen, Dale T. 2003. *Wage Dispersion: Why are Similar Workers Paid Differently?*, Cambridge, MA, USA: MIT Press.

[38] Moscarini, Guiseppe and Francis Vella. 2008. "Occupational Mobility and the Business Cycle. " Mimeo, Department of Economics, Yale University.

[39] Ortego-Marti, Victor. 2012. "Unemployment History and Frictional Wage Dispersion. " Mimeo, Department of Economics, University of California Riverside.

[40] Pissarides, Christopher. 2000. *Equilibrium Unemployment Theory*, 2nd Edition, Cambridge, MA, USA: MIT Press.

[41] Postel-Vinay, Fabien and Jean-Marc Robin. 2002. "Equilibrium Wage Dispersion with Worker and Employer Heterogeneity." *Econometrica*, 70(6): 2295-2350.

[42] Postel-Vinay, Fabien and Helene Turon. 2008. "On-the-job Search, Productivity Shocks, and the Individual Earnings Process." *International Economic Review*, forthcoming.

[43] Rosen, Sherwin. 1972. "Learning and Experience in the Labor Market." *Journal of Human Resources*, 7(3): 326-342.

[44] 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.

[45] Steward, Mark. 2007. "The Interrelated Dynamics of Unemployment and Low-wage Employment." *Journal of Applied Econometrics*, 22: 511-531.

[46] Topel, Robert. 1991. "Specific Capital, Mobility, and Wages: Wages Rise with Job Seniority." *Journal of Political Economy*, 99(1): 145-176.

[47] Topel, Robert and Michael P. Ward. 1992. "Job Mobility and the Careers of Young Men." *Quarterly Journal of Economics*, 107(2): 439-479.

[48] Tjaden, Volker and Felix Wellschmied. 2012. "Exploring the Causes of Frictional Wage Dispersion. " IZA Discussion Paper Series DP 6299, Institute for the Study of Labor, Bonn.

[49] Yamaguchi, Shitaro. 2010. "Job Search, Bargaining, and Wage Dynamics." *Journal of Labor Economics*, 28(3): 595-631.

Appendix

A Proofs

Proof of Proposition 1:

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

$$(r+\phi)\alpha^U = z_b + \lambda_u \int_{z_R}^{\overline{z}} [\alpha^E(z') - \alpha^U] dF(z') \tag{17}$$

and the Bellman equation for W^E is equivalent to

$$(r+\phi+\delta)\alpha^E(z) = z + \rho\alpha^E(z) + \lambda_e \int_{z}^{\overline{z}} [\alpha^E(z') - \alpha^E(z)] dF(z') + \delta\alpha^U, \tag{18}$$

which is a functional equation for $\alpha^E(.)$. Differentiating (18) with respect to z yields the differential equation describing α^E and evaluating (18) at $z = \overline{z}$ yields its boundary value $\alpha^E(\overline{z})$.

I now solve the conditions for z_R and α^U. First evaluate (18) at $z = z_R$. As $\alpha^E(z_R) = \alpha^U$ one obtains

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

Comparing this equation with (17), integrating by parts and using the differential equation for α^E establishes (1) described in Proposition 1. Similarly, $\alpha^E(z_R) = \alpha^U$, integration by parts and using the differential equation for α^E then yields (2) in Proposition 1. Thus (1) and (2) 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 (1) has slope

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

and implies $\alpha^U = (z_b - \overline{z})/\rho$ at $z_R = \overline{z}$. On the other hand, the equation described by (2) has slope

$$\left[\frac{d\alpha^U}{dz_R}\right]_{eqn(2)} = -\frac{1}{r+\phi}\left[\frac{\lambda_u(1-F(z_R))}{q(z_R)+r-\rho}\right] < 0,$$

for $z_R < \overline{z}$ and zero otherwise and implies that $\alpha^U = z_b/(r+\phi)$ at $z_R = \overline{z}$. Note that

$$\left[\frac{d\alpha^U}{dz_R}\right]_{eqn(2)} > \left[\frac{d\alpha^U}{dz_R}\right]_{eqn(1)}$$

for all z_R and hence (2) is always flatter than (1). Continuity of (1) and (2) and the restriction $\overline{z} > z_b(r+\phi-\rho)/(r+\phi)$ then guarantee there exists a single crossing between these two functions such that $\alpha^U > 0$ and $z_R < \overline{z}$.‖

Proof of Lemma 1:

Consider the pool of type ε unemployed workers with productivity no greater that y. It is straightforward to verify that steady-state turnover implies

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

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,

$$\frac{\partial H_\varepsilon(y,z)}{\partial y} + \frac{q(z)}{\rho y} H_\varepsilon(y,z) = \frac{(\phi+\delta)F(z)N_\varepsilon(y)}{\rho y}, \tag{19}$$

for $z \in [\underline{z}, \overline{z}]$ 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

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

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

Using these formulae we now solve for steady state $N_\varepsilon(.)$ and $H_\varepsilon(.,.)$. In particular, using the above expression for $N_\varepsilon(.)$ and simplifying yields

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

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

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

Using this and simplifying yields the expression for $N_\varepsilon(.)$. Using the latter to substitute out $N_\varepsilon(.)$ in the above expression for $H_\varepsilon(y,z)$ and some algebra then establishes (5).‖

Proof of Proposition 2:

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

$$\Omega(z;p) = \widetilde{\varepsilon} l(z)(p-z), \tag{20}$$

where $l(z)$ is given in the text. Let $z^* = \zeta(p)$ denote the solution to the above maximisation problem (if one exists). Assume the second order condition for a maximum holds. The envelope theorem then implies that $\Omega'(\zeta(p)) = l(\zeta(p))$, which describes a first order differential equation for $\Omega(.)$ in terms of p subject to the boundary condition $\Omega(\zeta(p_0)) = \widetilde{\varepsilon} l(\zeta(p_0))(p_0 - \underline{z})$. Noting that

$F(\zeta(p)) = \Gamma_0(p)$, its solution is given by

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

where $q(x) = \phi + \delta + \lambda_e(1 - \Gamma_0(x))$. Since $\zeta(p) = p - \Omega(\zeta(p))/\widetilde{\varepsilon} l(\zeta(p))$, substituting out for $\Omega(\zeta(p))$ and $l(\zeta(p))$ and some algebra yields (7), the expression for $\zeta(p)$ in the text.

Next I show that (7) indeed satisfies the first order condition of the firm's maximisation problem and then that the second order condition for a maximum is indeed met at $z = z^*$. First note that differentiation of (7) wrt p implies ζ satisfies the differential equation

$$\zeta'(p) = (p - \zeta(p))\Gamma_0'(p) \left[\frac{2\lambda_e}{q(p) - \rho} - \frac{k_2(\lambda_u - \lambda_e)[(q(p) - \phi\Gamma_0(p))(1 - \Gamma_0(p)) - \delta\Gamma_0(p)]}{[q(p) - \phi\Gamma_0(p)]\,[k_1 q(p) - \Gamma_0(p)[k_1\phi + k_2(\lambda_u - \lambda_e)(1 - \Gamma_0(p))]]} \right], \quad (21)$$

given the boundary condition $\zeta(p_0) = z_R$. Noting that the first order condition for a maximum implies that for a given p

$$l'(z)(p - z) - l(z) = 0$$

at $z^* = z$, using the expression for $l(z)$ and $F'(\zeta(p))\zeta'(p) = \Gamma_0'(p)$, some algebra establishes that (21) is indeed obtained from the above first order condition. Hence the function ζ implied by (21) satisfies the first order condition for a maximum given the boundary condition $\zeta(p_0) = z_R$. Substitution of $p - \zeta(p)$ from (7) 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 $l'(.)$ is increasing in z.

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

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

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

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

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

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

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 $p = \overline{p}$, on the other hand, will not offer a $z > \zeta(\overline{p})$ as doing so does not attract or retain any additional worker, but strictly decreases flow profit $\overline{p} - z$ and hence steady state profits.‖

Proof of Theorem 1:

Step 1: The first step to proof existence is to solve for p_0. Note that for any p_0, $T(z_R; p_0)$ gives the solution to $z_R = z_R(p_0)$ when both (3) and (7) are satisfied. Given $p_0 = \max\{\underline{p}, z_R(p_0)\}$, we have that $p_0 = \underline{p}$ if and only if $\underline{p} > z_R(\underline{p})$. Using (8) the latter condition can be expressed as

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

On the other hand, if the above condition does not hold (i.e. $\underline{p} \leq z_R(\underline{p})$), then $p_0 \geq \underline{p}$. In this case (8) implies p_0 satisfies $p_0 = \widetilde{T}(p_0)$, where

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

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

Step 2: Given a $p_0 \leq \overline{p}$ always exists, z_R is described by the solution to (8). $\zeta(.)$ is then characterised by (7) in Proposition 2 and $F(\zeta(p)) = \Gamma_0(p)$ for all $p \in [p_0, \overline{p}]$, where Γ_0 is given by (4). Furthermore, since Proposition 2 implies no firm with productivity $p \in [p_0, \overline{p}]$ will offer a different z, as doing so yields lower steady state profits, this establishes existence of a Market Equilibrium.‖

B Omitted Derivations

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

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

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 the experssion for N_ε 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 $F(\underline{z})=0$, (5) implies $H_\varepsilon(y',\underline{z})=0$. (19) 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 (5) implies

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

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 (6) in the text.‖

Sorting Dynamics:

Following similar arguments as in Burdett et. al. (2009), I first show that $H_\varepsilon(p \mid y)$ is first order stochastically increasing in y. Using (5) to construct

$$H_\varepsilon(z\mid y) = \frac{\int_{z'=\underline{z}}^{z}\frac{\partial^2 H_\varepsilon(y,z')}{\partial y\partial z'}dz'}{\int_{z'=\underline{z}}^{\overline{z}}\frac{\partial^2 H_\varepsilon(y,z')}{\partial y\partial z'}dz'} = \frac{\frac{\partial H_\varepsilon(y,z)}{\partial y}-\frac{\partial H_\varepsilon(y,\underline{z})}{\partial y}}{\frac{\partial H_\varepsilon(y,\overline{z})}{\partial y}-\frac{\partial H_\varepsilon(y,\underline{z})}{\partial y}},$$

and noting that $F(\zeta(p)) = \Gamma_0(p)$ implies that

$$H_\varepsilon(p \mid y) = \Gamma_0(p)\left[\frac{\delta(\phi+\delta+\lambda_u) + (\phi+\delta+\lambda_e)(\phi+\lambda_u)(1-\Gamma_0(p))\left(\frac{y}{\varepsilon}\right)^{-\left(\frac{\delta\lambda_u+\lambda_e(\phi+\lambda_u)(1-\Gamma_0(p))}{\rho(\phi+\lambda_u)}\right)}}{(\phi+\delta+\lambda_u)[(\phi+\lambda_e)(1-\Gamma_0(p))+\delta]}\right].$$

Inspection establishes that $H_\varepsilon(. \mid y)$ is first order stochastically increasing in y; i.e. $\partial H_\varepsilon(. \mid y)/\partial y < 0$.

Next consider $H_\varepsilon(y \mid p)$. Using (5) to construct

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

and noting that $F(\zeta(p)) = \Gamma_0(p)$ then implies that

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

Differentiation of $H_\varepsilon(. \mid p)$, 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. $\partial H_\varepsilon(. \mid p)/\partial p < 0$. ‖

Derivation of the wage density:

Differentiating (5) one obtains the joint density of workers characteristics, y and z, for each ability type ε,

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

where $h_\varepsilon(y,z) \geq 0$ for all $z \in [\underline{z}, \overline{z}]$ 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 $\widehat{h}_\varepsilon(w,z) = h_\varepsilon\left(\frac{w}{z}, z\right)\frac{1}{z}$, where $\widehat{h}_\varepsilon(w,z)$ describes the joint density of w, z. Using the expression for $h_\varepsilon(.,.)$ then gives

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

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

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

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

$$G'_\varepsilon(w) = \int_{\underline{\varepsilon}}^{\overline{\varepsilon}} \int_{\underline{z}}^{w/\varepsilon} \widehat{h}_\varepsilon(w,z)dzd\varepsilon.$$

Integrating over w gives (9).

Derivation of the mean-min ratio:

Since frictional wage dispersion concerns wage dispersion that is not driven by difference is abilities without loss of generality consider the case in which all workers enter with initial productivity $\varepsilon = 1$. 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

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

Putting $y = \infty$ in (5) implies

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

Since r and ρ are of the same order of magnitude (see numerical examples and the quantitative section of the paper), I follow Hornstein et al. (2009) and approximate $H(\infty, z)$ by

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

Solving for $1 - F(z)$ and using (3) yields

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

Dividing both sides by z^M we have that

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

where $\chi = z_b / z^M$ as described in the text.

C Simulation Procedure

Here I detail the simulation procedure used to generate data from the model and obtain the average wage-experience profile and cross-sectional wage distribution. I start by describing the simulation that computes the parameters of the firm distribution, $\underline{p}$, κ_1 and κ_2. Given the values of the interest rate, the transition parameters, the rate of human capital accumulation, I compute the employment histories of 20,000 workers for each combination of κ_1 and κ_2 between $(0, 2]$ and $\underline{p}$ between $(0, 10]$.[24] Given $b = 1$, each tuple ($\underline{p}$,κ_1,κ_2) implies a value of p_h and z_r and functions $\zeta(.)$, $\theta(.)$, $H(.,.)$ and $N(.)$. In simulating the employment histories I assume that all workers start unemployed and experience different types of shocks during their lifetime depending on the worker's employment status.

When unemployed, workers face a retirement and a job offer shock, where both process are Poisson with rates ϕ and λ_u and the shocks are mutually exclusive. What is important here is to capture the time spend in unemployment for each individual. To obtain the unemployment duration, I draw two random numbers, $r1 \in [0, 1]$ and $r2 \in [0, 1]$, using a uniform distribution and then exploit the fact that the inter arrival time between events in a Poisson process follows an exponential distribution with parameter equal to the rate of the process. That is, the duration until the worker receives a job offer is determined by $tu = -log(1 - r1)/\lambda_u$ and the duration until the worker experience a retirement shock is $td = -log(1 - r2)/\phi$. Since this a competing risk model the unemployment duration of an individual is given $\min\{td, tu\}$. If the individual retires, $td < tu$, then he leaves the sample. If the individual becomes employed, $td \geq tu$, then a firm productivity is sampled from Γ by, once again, choosing a random number between 0 and 1 and using the inverse of Γ to recover the corresponding productivity p. The latter then allows us to compute the corresponding z and θ.

Given the individual becomes employed in a firm with productivity p, the value of $z(p)$ is computed using (7) and the value of $\theta = z(p)/p$. This individual now faces three shocks: a retirement shock, a job offer shock and a displacement shock. All these shocks follow Poisson process with rates, ϕ, λ_e and δ, respectively. As for the case of unemployed workers, what is important is the duration of the job and the employment spells, where the latter is defined as the sum of job spells that start with the worker transiting from unemployment to employment and end with the worker becoming unemployed or leaving the labour market. I use the same procedure as before to obtain the durations until the worker receives a job offer tj, receives a displacement shock, tu, and receives a retirement shock, td. The job duration until the worker experiences one of these three events in then $\min\{tj, tu, td\}$. If the worker becomes unemployed, $tu = min\{tj, tu, td\}$, then the procedure described above for unemployed workers is repeated. If the worker leaves the labour market, $td = min\{tj, tu, td\}$, then he drops from the sample. If the worker receives an outside offer, $tj = min\{tj, tu, td\}$, a new p' is drawn using the same procedure described above and the values

[24] I bound the possible values of firm productivity to $\overline{p} = 25$. This upper bound does not affect the results as the CDF of both productivities distributions reach unity at lower values of p.

of $z(p)$ and $z(p')$ are compared. If $z(p) \geq z(p')$ the worker stays employed in his current job, while if $z(p) < z(p')$ the worker moves to the new firm and I repeat the process given a the new firm productivity p'. During this procedure I calculate the labour market experience of workers as the sum of employment spells. This information can then be used to compute workers' wages at each point in which an event has occurred taking into account that workers accumulate human capital at rate ρ.

The above procedure generates the full labour market histories of workers for an average life of $1/\phi$ months. However, the BHPS sample is restricted to workers that in 1991 were between 16 and 30 years of age and by 2004 were between 30 and 44 years of age. Hence one needs to create a sample of the simulated data that resembles that of the BHPS in terms of the age structure and has the same variance of actual experience (this is crucial for the variance decomposition exercise). It is only after creating such a sample that I compute the average wage-experience profiles by using an OLS regression on log wages on a constant and a quadratic on experience. The parameters of the productivity distribution κ_1 and κ_2 are then chosen to minimize the sum of the square differences between the simulated wage-experience profile and the wage-experience profile generated using the experience coefficients obtained from the BHPS by regressing log wages on a constant, a quadratic on experience, a quadratic on tenure, a dummy for marital status, 8 regional dummies, 8 (one-digit) occupational dummies and 8 (one-digit) industry dummies. The parameter $\underline{p}$ is chosen such that variance of $\log(z)$, obtained from the simulations, matches the variance of the log wage residuals used to compute the Mean-min ratio.

The final step in the simulation procedure is to recover the parameters of the ability distribution. Given the value of $\underline{\varepsilon}$, α_1 and α_2 are chosen to minimize the sum of squared differences $(\mu_m - \mu_d)^2 + (\sigma_m^2 - \sigma_d^2)$ under the restriction that one cannot reject the hypothesis that data generated by the model and the one obtained from the BHPS are drawn from different cdf based on the two-sample Kolmogorov-Smirnov goodness-of-fit hypothesis test. Here μ_m and σ_m^2 denote the mean and variance of the kernel density of log wages obtained from the same simulated data sets described above; and μ_d and σ_d^2 denote the mean and variance of the kernel density of residual log wages obtained from the BHPS by estimating an OLS regression on log wages on a constant, a quadratic on tenure, a dummy for marital status, 8 regional dummies, 8 (one-digit) occupational dummies and 8 (one-digit) industry dummies. Note that the latter is the same specification used to recover the firm productivity parameters with the exclusion of the experience terms. The data used and the STATA and MATLAB codes that implement this procedure are available upon request.