

IZA DP No. 4215

Human Capital Accumulation and Labour Market Equilibrium

Ken Burdett
Carlos Carrillo-Tudela
Melvyn G. Coles

June 2009

Human Capital Accumulation and Labour Market Equilibrium

Ken Burdett

University of Pennsylvania

Carlos Carrillo-Tudela

University of Leicester and IZA

Melvyn G. Coles

University of Essex

Discussion Paper No. 4215

June 2009

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0

Fax: +49-228-3894-180

E-mail: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ABSTRACT

Human Capital Accumulation and Labour Market Equilibrium^{*}

We analyse an equilibrium labour market with on-the-job search and experience effects (where workers learn-by-doing). The analysis yields a standard Mincer wage equation with worker fixed effects and endogenously determined firm fixed effects. It shows that learning-by-doing increases equilibrium wage dispersion consistent with the data. Equilibrium sorting - where over time more experienced workers also tend to find and quit to better paid employment - has a significant impact on wage inequality. As the model yields a cross section distribution of wages paid with the 'right' structure (the density of wages paid is single peaked with a 'fat' Pareto right tail) and yields the 'right' time profile of worker wage outcomes (the initial 10 years of a worker's career are characterised by several job changes and rapid wage growth) it yields a new, coherent statistical structure for future applied work.

JEL Classification: J24, J42, J64

Keywords: search, wage dispersion, human capital accumulation

Corresponding author:

Carlos Carrillo-Tudela
Department of Economics
University of Leicester
Astley Clarke Building
University Road
Leicester, Leicestershire, LE1 7RH
United Kingdom
E-mail: cct9@leicester.ac.uk

^{*} We would like to thank Pieter Gautier, John Kennan, Jean-Marc Robin and Yoram Weiss for their useful comments and insights. This paper has also benefited from comments and suggestions of participants at the "Structural Models of the Labour Market and Policy Analysis" conference held at the Institute of Fiscal Studies (UCL), the "Heterogeneous Labor: Search Frictions and Human Capital Investment" conference held at the University of Konstanz and the "Advances in the Analysis of Labor Markets with Frictions" conference held at the University of Malaga. We would also like to thank participants in seminars at the University of Essex, University of Leicester, University of Vienna, RWI-Essen and at the SED (2008) and EEA (2008) meetings. All errors are our responsibility.

1 Introduction

Modern labour markets are characterised by a significant amount of churning. Workers change status frequently, wages vary through time, and jobs are continually being created and destroyed. To illustrate, in the US approximately 30 percent of workers change jobs each year, whereas something like 10 percent of all jobs are destroyed in a year (and 10 percent are usually created). Further the average real wage of high school graduates increases by about 50 percent in the first ten years of their working life. After those first ten years, however, further wage gains are much more moderate (Rubinstein and Weiss, 2007). Explaining such variation has long been a central issue in labour economics. Surprisingly, however, little work has been done in constructing a coherent theoretical framework in which to study this variation. Indeed without such a framework it seems difficult to have a consistent statistical description of market behaviour. This paper provides a useful and tractable framework to study such labour market outcomes. We do this by integrating two central pillars of modern labour economics: (i) the theory of human capital accumulation and (ii) equilibrium turnover in labour markets where workers search for better paid employment. As we shall see, the marriage of these two areas leads to new insights on how wages change over time, how wages of different workers differ at a moment in time and the interrelationship between job turnover and changes in wages.

Since the pioneering work of Becker (1975) and Mincer (1974), human capital theory has been used to explain the wage growth of workers over the life cycle. According to this theory, wages increase as workers accumulate firm specific and general skills. Workers who change job, or those who are laid off, lose their firm specific human capital but keep their general human capital. The assumed competitive environment, however, implies this approach has little to say about either why or when a worker changes job. This seems somewhat unfortunate as a significant percentage of observed wage increases occur precisely when the worker changes job (see Topel and Ward, 1992). Further, this approach has difficulty in explaining why workers with seemingly identical characteristics are paid differently (see Mortensen, 2003).

This paper considers the impact of human capital accumulation on equilibrium market outcomes in a non-competitive labour market with search frictions. The approach is motivated in large part by the following empirical fact. It is well known from the Mincer literature that, on average across all workers, log wages are an increasing concave function of experience. There is, however, enormous variation in individual wage outcomes. Table 1 below is a simple modification of results recently described in Rubinstein and Weiss (2007) using the Current Population Surveys (CPS). Workers were asked whether, over the previous year, they had had a rise, no change, or decline in their nominal wage. Those that reported a wage gain are termed Gainers, those who reported a wage decline are termed Losers, the others reported no change in nominal wage. The average real

average wage change was then calculated for each group.

Table 1: Annual Growth Rates of Real Wages of College Graduates: CPS-ORG 1998-2002			
	Experience	Percentage	Wage Growth
Gainers	0-10	64.3	26.3
	11-15	60.2	25.9
	16-25	56.7	26.8
	26-40	53.6	28.7
No Change	0-10	5.5	-2.5
	11-15	8.1	-2.6
	16-25	9.0	-3.1
	26-40	8.5	-2.6
Losers	0-10	30.1	-35.0
	11-15	31.7	-34.3
	16-25	34.2	-35.9
	26-40	37.8	-40.5

Table 1 implies that of workers with 0-10 years experience and in any one year, 64% enjoyed a wage gain and Gainers on average enjoyed a 26% increase in wage. 30% instead suffered a wage loss and Losers on average suffered a loss of 35%. The rest had no change. Across experience groups, these wage changes are surprisingly large: the average real wage growth of Gainers is over 25%, whereas the average loss among Losers is even larger, above 34%. Conditional on being a Gainer or Loser, however, Table 1 shows that the effect of experience on the expected wage change is small. Instead the main effect of experience is that *as workers become more experienced, they are less likely to be Gainers and more likely to be Losers*. Thus it is the change in these probabilities which causes average wage gains to decline with experience. The challenge is to explain why.

A search approach provides an elegant explanation. When jobs are not for life, unemployed workers are willing to take low paid employment as they can continue to search for better paid work while employed. Indeed when experience is valuable, as is the case in this paper, (young) inexperienced unemployed workers are willing to accept very low starting wages in order to gain work experience. When frictions are not too large, however, the early years of a worker's career are then characterised by rapid wage growth as each poorly paid worker eventually finds and quits to better paid employment. Of course such early wage growth is relatively short-lived as each worker eventually exhausts the possibility of finding ever-better paid employment. This approach thus explains why, on average, wages increase over time but at a decreasing rate, and why those wage gains are

directly related to the probability of finding and quitting to better paid employment.¹

Of course learning-by-doing remains an important factor in the determination of wages. Typists become better typists while working as typists, economists become more productive by doing economics, etc. This seems both an important and intuitive idea. The issue, however, is to identify what portion of wage increases arise through productivity increases (through learning-by-doing) and what portion is due to job search, where over time workers simply find and quit to better paid employment. The difficulty is doing this within an equilibrium framework where all behave optimally and wages are endogenously determined.

This paper extends the Burdett and Mortensen (1998) framework, henceforth B/M, to assess the impact of learning-by-doing on equilibrium market outcomes. A most useful feature of our approach is that individual worker wages remain consistent with a standard Mincer equation. Furthermore the firm fixed effects are endogenously determined and quit turnover is an equilibrium outcome. The model not only explains why quit rates decline with age, it is also consistent with the fact that quit turnover is a major source of early wage gains for young workers. By detailing how wages are dispersed (and thus unequal) across all employed workers, this framework also provides a direct link between studies of wage determination and studies of wage inequality.

Indeed this paper yields a remarkably insightful decomposition of the variance of log wages across employed workers. Consistent with Abowd, Kramarz and Margolis (1999), the theory implies that worker and firm fixed effects are not correlated with each other - there is no sorting by types. Nevertheless equilibrium sorting, where over time workers become more experienced and tend to find better paid employment, has a significant impact on overall wage inequality. In essence young workers face a double whammy: they not only have little experience but most job offers typically pay low wages (the equilibrium distribution of pay offers is skewed to the bottom end of the distribution). Over time (young) workers become better paid not only because they become more experienced, and thus more productive, they also eventually find and quit to better paid employment. The positive correlation between these two processes magnifies measured wage inequality: the young typically earn much less than the old.

Hornstein et. al. (2007) demonstrates that when search frictions are calibrated to turnover data, the B/M framework generates too little wage dispersion. Learning-by-doing, however, reduces unemployed worker reservation wages - unemployed workers are willing to accept low wages as work experience is valuable. Learning-by-doing, however, also increases wage competition across firms for experienced workers. Indeed firms which pay higher wages attract and retain a more experienced, and thus more productive, work-

¹Our model also explains why the probability of being a loser increases with experience. With job destruction shocks, a well paid (experienced) worker is likely to be next employed on a significantly lower wage. A poorly paid (inexperienced) worker however, when laid-off, may well find a better paid job on re-employment and is thus less likely to be a Loser.

force. We show that learning-by-doing increases equilibrium wage dispersion consistent with the observations described in Hornstein et. al. (2007). Furthermore we show the implied cross section distribution of wages paid has the right structure (the density of wages paid is single peaked with a ‘fat’ Pareto right tail). Given this framework yields the ‘right’ time profile of worker wage outcomes - the initial 10 years of a worker’s career are characterised by several job changes and rapid wage growth - and a Mincer wage equation, this structure provides an important, coherent benchmark for future empirical work.

There are a few papers which have investigated learning-by-doing effects within a similar turnover framework as studied here. Bunzel et. al. (2000) analyzed a B/M model with human capital accumulation. Unlike our approach, they assume agents are initially homogeneous and workers lose all their human capital when laid off. This leads to very different results. Rubinstein and Weiss (2007) analyse human capital accumulation and on-the-job search but do not consider equilibrium. In an interesting application of record statistics, Barlevy (2008) estimates the wage process identified here. Using a similar model, Fu (2009) asks when firms will provide general human capital for its employees.

Bagger et. al. (2006) instead extend the offer matching framework developed by Postel-Vinay and Robin (2002a,b) to incorporate learning on-the-job with individual productivity shocks. Their focus is on estimating the resulting wage process over a worker’s life cycle. Our focus here is to study both individual wage dynamics and cross-section wage dispersion in the case when firms do not respond to outside offers (see Mortensen, 2003, for a full discussion). Indeed by identifying closed form solutions, our framework yields new and clear insights on the impact of learning-by-doing and on-the-job search on market outcomes.

The paper is outlined as follows. Section 2 describes the model. Section 3 describes optimal search behaviour by workers. Section 4 formally defines an equilibrium and describes the steady state distributions of worker experience across unemployed and employed workers. Section 5 identifies the unique equilibrium and obtains its closed form solution. Section 6 then describes the equilibrium wage outcomes and focuses on two equilibrium effects - positive sorting across employed workers on experience and pay rates earned, and describes how learning-by-doing changes the equilibrium wage setting incentives of firms. Using numerical examples, Section 7 describes the impact of on-the-job learning on equilibrium wage outcomes. Most proofs are relegated to a technical Appendix.

2 The Model

Time is continuous with an infinite horizon. Keeping things as simple as possible only steady-states are considered. There is a continuum of both firms and workers, each

of measure one. All firms are equally productive and have a constant return to scale technology. There is turnover of workers: each worker permanently exits the labour market at rate $\phi > 0$, while ϕ also describes the inflow of new labour market entrants. This yields an overlapping generations structure where worker ages are distributed according to the exponential distribution. Assume there are I types of workers, where each type is defined by his/her initial productivity. In particular, let y_i denote the initial productivity of a type i labour market entrant and assume $y_1 < y_2 < \dots < y_I$. Let A denote the distribution function of these initial productivities and γ_i denote the proportion of new entrants who are type i . As all have the same exit rate ϕ , steady-state turnover implies γ_i is also the number of type i workers in the market.

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

A worker with productivity y generates flow output y while employed. We normalize the price of the production good to one, so y also describes flow revenue. Each firm pays each of its employees the same piece rate θ . Thus given an employee with productivity y , the worker is paid flow wage $w = \theta y$. Each firm's total profit flow is simply total flow output from its employees multiplied by $(1 - \theta)$. As different firms may offer different piece rates, let $F(\theta)$ denote the proportion of firms offering a piece rate no greater than θ . Further, let $\underline{\theta}, \bar{\theta}$ denote the infimum and supremum of the support of F . There are job destruction shocks in that each employed worker is displaced into unemployment according to a Poisson process with parameter $\delta > 0$. For tractability we follow Postel-Vinay and Robin (2002b) and assume a worker with productivity y enjoys flow income by while unemployed, where $0 < b < 1$. We discuss further the role of this assumption below.

Employed and unemployed workers receive job offers according to a Poisson process with parameter $\lambda > 0$. Of course empirical work finds these arrival rates differ across employment states (e.g. Jolivet et. al., 2006). Although the extension to different arrival rates is straightforward (see for example B/M), it does not change the underlying insights and makes the exposition unnecessarily cumbersome.

Search is random and so any job offer θ can be considered as a random draw from F . If a job offer is rejected, the worker remains in his/her current state and there is no recall. We make the standard tie-breaking assumptions: an unemployed worker accepts a job offer if indifferent to accepting it or remaining unemployed, while an employed worker quits only if the job offer is strictly preferred. Note this structure implicitly makes two critical restrictions. First we assume that if an employee receives an outside offer, there is no Bertrand competition for the worker's services (Postel-Vinay and Robin, 2002a,b). The worker instead simply accepts or rejects the offer. Second we simplify by ruling out wage/tenure contracts as considered in Burdett and Coles (2003) which would severely complicate the analysis.

All agents are risk neutral. For simplicity we assume a zero rate of time preference but then require $\phi > \rho$ to ensure total expected lifetime payoffs are finite. Each worker thus maximizes expected lifetime income. Each firm chooses piece rate θ to maximize steady state flow profit, taking into account the search strategies of workers.

3 Worker Behavior

In this section we take the distribution of piece rate offers F as given and characterize optimal worker behavior. For notational ease we suppress reference to F in the following functions. As workers are heterogeneous, let $W^U(y)$ denote the expected lifetime payoff of an unemployed worker with productivity y using an optimal search strategy. $W^E(y, \theta)$ denotes the expected lifetime payoff of a worker with productivity y , currently employed at piece rate θ , when using an optimal search strategy.

Consider first an unemployed worker with productivity y . As there is no learning-by-doing while unemployed (and no depreciation), standard arguments imply the flow Bellman equation describing $W^U(y)$ is

$$\phi W^U(y) = by + \lambda \int_{\underline{\theta}}^{\bar{\theta}} \max[W^E(y, \theta') - W^U(y), 0] dF(\theta'). \quad (1)$$

While unemployed the worker enjoys flow income by . Job offers are received at rate λ and, conditional on the realized draw θ' , the worker either accepts it and enjoys welfare gain $W^E(y, \theta') - W^U(y)$, or remains unemployed with productivity y .

Consider now an employed worker with productivity y employed at a firm paying piece rate θ . As it is always better to be employed at a firm paying a higher piece rate, it is immediate that $W^E(y, \theta)$ is increasing in θ . Thus an employed worker quits to an outside offer θ' if and only if $\theta' > \theta$. Assuming for the moment that this worker never voluntarily quits into unemployment (which is true in equilibrium), standard arguments imply the flow Bellman equation describing W^E is:

$$(\phi + \delta)W^E(y, \theta) = \theta y + \rho y \frac{\partial W^E}{\partial y} + \lambda \int_{\theta}^{\bar{\theta}} [W^E(y, \theta') - W^E(y, \theta)] dF(\theta') + \delta W^U(y). \quad (2)$$

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

Characterising the solution to these Bellman equations is straightforward. As is standard, all unemployed workers use a reservation piece rate strategy. Proposition 1 below, however, shows that all use the same reservation piece rate θ^R and identifies the conditions

which fully determine θ^R . As a useful shorthand, define

$$q(\theta) = \phi + \delta + \lambda(1 - F(\theta)),$$

which is the rate at which any employee exits a firm offering piece rate θ , and note $q(\theta) > \rho$ for all θ (as $\phi > \rho$).

Proposition 1: *Optimal job search implies:*

(i) *all unemployed workers have the same reservation piece rate θ^R ; i.e. an unemployed worker y accepts job offer θ if and only if $\theta \geq \theta^R$, where*

(ii) *the reservation piece rate θ^R is jointly determined by the following pair of equations for (θ^R, α^U) :*

$$\rho\alpha^U = b - \theta^R \tag{3}$$

$$\phi\alpha^U = b + \int_{\theta^R}^{\bar{\theta}} \frac{\lambda(1 - F(\theta))}{q(\theta) - \rho} d\theta. \tag{4}$$

Further for any F a solution exists, is unique, implies $\theta^R < b$ and θ^R is strictly decreasing in ρ .

The proof of Proposition 1 is in the Appendix. Here we present the main arguments. As all payoffs are linear in y , the value functions are also linear in y ; i.e.

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

where $\alpha^U, \alpha^E(\cdot)$ are determined by the above Bellman equations. The proof of Proposition 1 essentially solves for $\alpha^U, \alpha^E(\cdot)$ and θ^R , noting that the reservation piece rate θ^R solves $\alpha^U = \alpha^E(\theta^R)$. As this latter condition is independent of y , it follows that all unemployed workers utilize the same reservation piece rate.²

With no learning-by-doing, (3) implies $\theta^R = b$: when experience has no value, unemployed workers reject all offers below b . But with learning-by-doing, $\rho > 0$, experience is valuable as it increases future productivity. Unemployed workers then have a reservation piece rate below b . Indeed for a given offer distribution F , Proposition 1 shows a higher rate of learning-by-doing implies a strictly lower reservation piece rate of unemployed workers. θ^R may even become negative: workers may be willing to purchase work experience. As in Rosen (1972), this change in reservation wages of unemployed workers affects the wage posting incentives of firms. The next step is to characterise equilibrium wage competition and so determine equilibrium F .

²This very useful property disappears if instead we assume unemployment benefits are proportional to the previous wage; say unemployed workers receive benefit $b\theta^0 y$ where θ^0 denotes the piece rate paid by previous employer and $b < 1$ is the replacement rate. As such payments are proportional to y , the value functions W^E and W^U remain linear in y . But the optimal reservation piece rate θ^R now varies depending on θ^0 . Although describing optimal worker behavior remains straightforward, the additional recursive element yields a much more complicated equilibrium structure. We leave this problem for future research.

4 Profits and Steady-States

Notice first that offering a piece rate $\theta < \theta^R$ implies the firm makes zero profit (the firm attracts no workers). As offering $\theta = b$ generates strictly positive profit (as $b < 1$), this strictly dominates offering $\theta < \theta^R$. Thus in any market equilibrium, we must have $\underline{\theta} \geq \theta^R$ and each unemployed worker, regardless of type, always accepts the first job offer received. We shall show that that this implies each type i will not only have (a) the same unemployment rate U in a Market Equilibrium, but also (b) the same distribution of experiences across unemployed workers, denoted $N(x)$, and (c) the same joint distribution of experience and piece rates across employed workers, which we denote $H(x, \theta)$. Proposition 2 below fully characterises these distribution functions.

As there is no discounting, the arguments in Burdett and Coles (2003) imply steady state flow profit equals the hiring rate of the firm, multiplied by the expected profit of each hire. Given offer $\theta \geq \theta^R$, steady state flow profit is therefore

$$\pi(\theta) = \sum_i \left[\begin{aligned} &\lambda\gamma_i U \int_{x=0}^{\infty} \left[\int_{\tau=0}^{\infty} e^{-q(\theta)\tau} (1-\theta)y_i e^{\rho(x+\tau)} d\tau \right] dN(x) \\ &+ \lambda\gamma_i (1-U) \int_{\theta'=\underline{\theta}}^{\theta} \int_{x=0}^{\infty} \left[\int_{\tau=0}^{\infty} e^{-q(\theta)\tau} (1-\theta)y_i e^{\rho(x+\tau)} d\tau \right] dH(x, \theta') \end{aligned} \right].$$

Recall γ_i is the number of workers in the economy who are type i , and thus $\gamma_i U$ is the number of type i workers who are unemployed. For each i , the first term in the above equation is the steady state flow profit due to attracting type i unemployed workers whose experience x is drawn from $N(x)$. For each such hire, learning-by-doing implies the new hire has productivity $y = y_i e^{\rho(x+\tau)}$ at each tenure $\tau \geq 0$. As the worker leaves employment at rate $q(\theta)$, the integrand thus describes expected total profit per hire. The second term is the flow profit due to attracting type i employed workers who are on lower piece rates $\theta' < \theta$; $\lambda\gamma_i(1-U)dH(x, \theta')$ describes the hiring inflow of each such worker and the inside bracketed integral is again the expected profit per hire. Integrating over τ and simplifying yields:

$$\pi(\theta) = \frac{\lambda(1-\theta)\bar{y}}{q(\theta) - \rho} \left[U \int_{x=0}^{\infty} e^{\rho x} dN(x) + (1-U) \int_{\theta'=\underline{\theta}}^{\theta} \int_{x=0}^{\infty} e^{\rho x} dH(x, \theta') \right], \quad (5)$$

where $\bar{y} = \sum_i \gamma_i y_i$ is the mean ability of labour market entrants. We now formally define an equilibrium.

A Market Equilibrium is a set $\{\theta^R, U, N(\cdot), H(\cdot), F(\cdot)\}$ such that

- (i) θ^R is the optimal reservation piece rate of any unemployed worker;
- (ii) $U, N(\cdot), H(\cdot)$ are consistent with steady state turnover given piece rate offers $F(\cdot)$ and optimal worker search strategies;

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

$$\begin{aligned}\pi(\theta) &= \bar{\pi} > 0 \text{ for all } \theta \text{ where } dF(\theta) > 0; \\ \pi(\theta) &\leq \bar{\pi} \quad \text{for all } \theta \text{ where } dF(\theta) = 0.\end{aligned}$$

The constant profit condition requires that all equilibrium offers, those with $dF(\theta) > 0$, enjoy the same profit $\bar{\pi}$, while all other offers make no greater profit. Lemma 1 presents a useful preliminary result. As the proof is well known, it is relegated to the Appendix.

Lemma 1: *A Market Equilibrium implies (i) $F(\cdot)$ contains no mass points; (ii) $F(\cdot)$ has a connected support and (iii) $\underline{\theta} = \theta^R$.*

The aim now is to construct equilibrium $\pi(\theta)$. To do this we need to solve for U and the distribution functions N and H . We consider each of these objects in turn.

First consider steady state turnover in the pool of type i unemployed workers, where $\gamma_i U$ is the number of type i unemployed workers. As each unemployed worker accepts the first job offer received (because all firms offer $\theta \geq \theta^R$), the total outflow from this pool is $(\lambda + \phi)\gamma_i U$. As steady state requires this equals the total inflow, which is $\phi\gamma_i + \delta\gamma_i(1 - U)$, then the equilibrium unemployment rate is

$$U = \frac{\phi + \delta}{\phi + \delta + \lambda},$$

which is the same for all types.

Next consider the pool of type i unemployed workers with experience no greater than $x \geq 0$. Steady-state turnover requires

$$\gamma_i \phi + \delta\gamma_i[1 - U]H(x, \bar{\theta}) = [\phi + \lambda]\gamma_i U N(x),$$

where the left hand side describes the inflow (new labour market entrants with zero experience) and laid-off employed workers with experience less than x , and the right hand side describes the outflow. Solving for $N(x)$, using the above solution for U , yields

$$N(x) = \frac{\phi(\phi + \delta + \lambda) + \lambda\delta H(x, \bar{\theta})}{(\phi + \lambda)(\phi + \delta)} \text{ for all } x \geq 0. \quad (6)$$

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

$$\gamma_i(1 - U)H(x, \theta)q(\theta)dt + \gamma_i(1 - U)[H(x, \theta) - H(x - dt, \theta)] + O(dt^2).$$

The first term is the outflow of workers in this pool who either leave employment, or quit to a job with $\theta' > \theta$. The second is the outflow of those who accumulate experience greater

than x , while the last term corrects for the fact that some do both but this term has the property $O(dt^2)/dt \rightarrow 0$ as $dt \rightarrow 0$. The inflow into this pool is simply $\gamma_i UN(x)\lambda F(\theta)dt$: those type i unemployed workers with experience no greater than x who find a job no better than θ . Setting inflow equal to outflow, rearranging appropriately and then letting $dt \rightarrow 0$ yields the following partial differential equation for H

$$q(\theta)H + \frac{\partial H}{\partial x} = (\phi + \delta)F(\theta)N(x). \quad (7)$$

By solving this partial differential equation, Proposition 2 now obtains closed form solutions for N and H . Its proof is relegated to the Appendix.

Proposition 2: *A Market Equilibrium implies distribution functions*

$$N(x) = 1 - \frac{\lambda\delta}{(\phi + \lambda)(\phi + \delta)} e^{-\frac{\phi(\phi + \delta + \lambda)x}{(\phi + \lambda)}} \text{ for all } x \geq 0, \quad (8)$$

$$H(x, \theta) = \frac{(\phi + \delta)F(\theta)}{q(\theta)} [1 - e^{-q(\theta)x}] - \frac{\delta F(\theta)}{q(\theta) - \phi F(\theta)} \left[e^{-\frac{\phi(\phi + \delta + \lambda)x}{(\phi + \lambda)}} - e^{-q(\theta)x} \right] \quad (9)$$

for all $\theta \in [\underline{\theta}, \bar{\theta}]$ and $x \geq 0$.

This explicit characterisation of the joint cdf $H(\cdot)$ is a major contribution of the paper. By detailing how experience and pay is distributed across employed workers, it provides a precise understanding of wage inequality in labour markets. We discuss the resulting insights in much greater detail below. Before doing that, however, we first complete the characterisation of a Market Equilibrium.

5 Market Equilibrium

Although algebraically tedious, solving for the Market Equilibrium is remarkably simple. By Lemma 1 (connected support), all we need do is solve the constant profit condition, that $\pi(\theta) = \bar{\pi}$ for all $\theta \in [\underline{\theta}, \bar{\theta}]$, for equilibrium F . To begin, first note equation (5) implies

$$\pi(\theta) = \frac{\lambda(1 - \theta)\bar{y}}{q(\theta) - \rho} \left[U \int_{x=0}^{\infty} e^{\rho x} dN(x) + (1 - U) \int_{\theta'=\underline{\theta}}^{\theta} \int_{x=0}^{\infty} e^{\rho x} \frac{\partial^2 H(x', \theta')}{\partial x' \partial \theta'} d\theta' dx' \right]. \quad (10)$$

We already know $U = (\phi + \delta)/(\phi + \delta + \lambda)$ while Proposition 2 describes the distribution functions N, H . Calculating the integrals described in (10) finds after some work:

$$\int_{x=0}^{\infty} e^{\rho x} dN(x) = \frac{\phi(\phi + \delta + \lambda)}{(\phi + \delta)} \left[\frac{(\phi + \delta - \rho)}{\phi(\phi + \delta + \lambda) - \rho(\phi + \lambda)} \right],$$

$$\int_{\theta'=\underline{\theta}}^{\theta} \int_{x=0}^{\infty} e^{\rho x} \frac{\partial^2 H(x', \theta')}{\partial x' \partial \theta'} d\theta' dx' = \frac{\phi F(\theta)}{q(\theta) - \rho} \left[\frac{(\phi + \delta + \lambda)(\phi + \delta - \rho)}{\phi(\phi + \delta + \lambda) - \rho(\phi + \lambda)} \right].$$

Now substitute these expressions into (10). Further simplification finds the constant profit condition, $\pi(\theta) = \bar{\pi}$ for all $\theta \in [\underline{\theta}, \bar{\theta}]$, then reduces to

$$\frac{\lambda(1-\theta)\bar{y} \left[\frac{\phi(\phi+\delta-\rho)(\phi+\delta-\rho+\lambda)}{\phi(\phi+\delta+\lambda)-\rho(\phi+\lambda)} \right]}{[\phi-\rho+\delta+\lambda(1-F(\theta))]^2} = \bar{\pi}. \quad (11)$$

Somewhat miraculously, this yields a quadratic equation for equilibrium $F = F(\theta)$. As F must be increasing over θ , (11) implies the following closed form solution for F .

Proposition 3: *A Market Equilibrium implies*

$$F(\theta) = \frac{\phi + \delta + \lambda - \rho}{\lambda} - \left[\frac{\bar{y}\phi(\phi + \delta - \rho) [\phi + \delta - \rho + \lambda]}{\lambda\bar{\pi}(\phi(\phi + \delta + \lambda) - \rho(\phi + \lambda))} \right]^{1/2} (1 - \theta)^{1/2} \quad \text{for all } \theta \in [\underline{\theta}, \bar{\theta}]. \quad (12)$$

All that remains now is to determine equilibrium $\bar{\pi}$. It is easier, however, to instead solve for equilibrium $\underline{\theta}$. By Lemma 1 (no mass points), then F is zero at $\theta = \underline{\theta}$ and so (12) implies $\bar{\pi}$ and $\underline{\theta}$ are related as

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

Use this condition to substitute out $\bar{\pi}$ in (12). The equilibrium offer distribution can then be written as

$$F = \widehat{F}(\theta | \underline{\theta}) = \left(\frac{\phi + \delta - \rho + \lambda}{\lambda} \right) \left[1 - \left(\frac{1 - \theta}{1 - \underline{\theta}} \right)^{1/2} \right] \quad \text{for all } \theta \in [\underline{\theta}, \bar{\theta}]. \quad (13)$$

The final step is to determine equilibrium $\underline{\theta}$. Suppose $\underline{\theta} < b$ is an equilibrium value, and (13) then implies $F = \widehat{F}(\theta | \underline{\theta})$ is the equilibrium offer distribution. Given $F = \widehat{F}(\theta | \underline{\theta})$, the conditions of Proposition 1 determine θ^R , which we denote $\theta^R = \widehat{\theta}^R(\underline{\theta})$. Lemma 1(iii) further requires $\theta^R = \underline{\theta}$ in a Market Equilibrium. Thus a Market Equilibrium requires solving the fixed point condition $\widehat{\theta}^R(\underline{\theta}) = \underline{\theta}$. This condition also has a closed form solution.

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

$$\underline{\theta} = \theta^R = b - \rho \frac{(\phi + \delta - \rho + \lambda)^2 b + \lambda^2(1 - b)}{\phi(\phi + \delta - \rho + \lambda)^2 - \rho\lambda^2}. \quad (14)$$

Equilibrium F is given by (13) with $\underline{\theta}$ given by (14), and the steady state distribution functions N, H are as described in Proposition 2.

Proof: In the Appendix we show that solving $\widehat{\theta}^R(\underline{\theta}) = \underline{\theta}$ yields the unique solution (14) for $\underline{\theta}$. Given the equilibrium distribution F as described in Theorem 1 then, by construction, the unemployed worker's optimal reservation piece rate $\theta^R = \underline{\theta}$. Further U, N , and H are consistent with steady state turnover while F ensures the constant profit condition holds for all $\theta \in [\underline{\theta}, \bar{\theta}]$. All that remains is to show there is no other offer which

is profit increasing. Offering $\theta < \underline{\theta}$ yields zero profit as all workers reject such offers (as $\theta^R = \underline{\theta}$). Conversely offering $\theta > \bar{\theta}$ yields strictly less profit than offering $\theta = \bar{\theta}$ as it attracts no additional workers and yields strictly less profit per hire. Thus the above identifies the Market Equilibrium. This completes the proof of Theorem 1.

6 Equilibrium wage outcomes

The above model yields the following Mincer wage equation: the wage earned by a type i worker employed at firm j at date t with experience x_{it} satisfies:

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

where θ_j is the piece rate offered by firm j . Hence each worker's wage w_{ijt} is composed of a worker fixed effect (initial ability y_i), a firm fixed effect (the firm's piece rate θ_j) and experience. Wages are dispersed because of cohort effects (some workers have been in the labour market for 40 years and have accumulated lots of experience) and within-cohort effects where individual labour market histories evolve stochastically (some may be lucky and quickly find highly paid employment, others might get stuck in low paid employment or be laid-off). The analysis yields two important pieces of information. First it describes $H(x, \theta)$, the joint distribution of experience and piece rates across all employed workers in the labour market. Thus we obtain a detailed description of equilibrium wage dispersion across employed workers. The analysis also determines endogenously the distribution of firm fixed effects, F . We can thus consider how changes in the rate of on-the-job learning ρ affects these firm fixed effects, where $\theta_j \sim F$.

Below we use numerical examples to understand qualitatively the impact of these effects on market wage outcomes. Here we detail each of these effects in turn.

6.1 Equilibrium Sorting

First note consistent with the findings of Abowd, Kramarz and Margolis (1999) and Abowd, Finer and Kramarz (1999), that there is no sorting by underlying types - the worker and firm fixed effects are orthogonal (see Mortensen, 2003, for further discussion). Also as in Altonji and Shakotko, 1987, and Topel, 1991, there is no relationship between experience and worker type. A fundamental insight, however, is there is (strong) positive sorting between experience and piece rates earned.

For example consider a worker with experience x . Using (9) it can be shown that the

conditional distribution of piece rates enjoyed by such workers is³

$$H(\theta | x) = F \left[\frac{\delta + (\phi + \lambda)(1 - F)e^{-\frac{\lambda(q - \phi F)}{\phi + \lambda}x}}{\delta + (\phi + \lambda)(1 - F)} \right].$$

Putting $x = 0$ finds $H(\theta | 0) = F$; the first job of new entrants is a random draw from F . But it is immediate that $H(\theta | x)$ is strictly decreasing in x ; i.e. more experienced workers typically enjoy higher piece rates (first order stochastic dominance). This occurs as workers not only accumulate experience while employed but also continue search for better paid employment. As experience becomes large, this distribution converges to its ergodic distribution

$$H(\theta | \infty) = \frac{\delta F}{\delta + (\phi + \lambda)(1 - F)},$$

which is non-degenerate. Job destruction shocks imply employed workers are occasionally displaced from their jobs and must search anew.

Such sorting generates an important composition effect - firms which offer a higher piece rate also enjoy, in a Market Equilibrium, a more experienced and thus more productive workforce. This has a direct impact on wage competition. For example, consider a firm which offers piece rate θ . Using (9) it can be shown that the conditional distribution of experience across employed workers is:

$$H(x | \theta) = (1 - e^{-qx}) - \frac{\delta q^2}{(\phi + \delta)[q - \phi F]^2} [e^{-\frac{\phi(\phi + \delta + \lambda)x}{(\phi + \lambda)}} - e^{-qx}] - \frac{\lambda F \phi (1 - F) qx}{(\phi + \delta)[q - \phi F]} e^{-qx}.$$

This expression is most easily interpreted by first considering the firm offering the lowest piece rate $\theta = \underline{\theta}$. This firm only attracts unemployed workers. The first term describes its distribution of worker experience through only hiring new market entrants (with no previous experience), taking into account that such employees depart at rate q . Given $\delta > 0$, the second term captures the composition effect by also hiring unemployed workers who have some previous experience. The third term takes into account that a firm offering $\theta > \underline{\theta}$, attracts workers from firms offering $\theta' < \theta$. The added restriction $\delta > \phi$ (which implies each worker expects to be laid off at least once over a working lifetime) is

³The easiest way to obtain this condition is to note

$$\begin{aligned} H(\theta | x) &= \int_{\theta' = \underline{\theta}}^{\theta} \frac{\partial^2 H(x, \theta')}{\partial x \partial \theta'} d\theta' / \int_{\theta' = \underline{\theta}}^{\bar{\theta}} \frac{\partial^2 H(x, \theta')}{\partial x \partial \theta'} d\theta' \\ &= \left[\frac{\partial H(x, \theta)}{\partial x} - \frac{\partial H(x, \underline{\theta})}{\partial x} \right] / \left[\frac{\partial H(x, \bar{\theta})}{\partial x} - \frac{\partial H(x, \underline{\theta})}{\partial x} \right] \end{aligned}$$

and grind away using (9).

sufficient (but not necessary) to establish $\partial[H(x|\theta)]/\partial\theta < 0$;⁴ i.e. firms that pay higher piece rates attract a more experienced, and thus more productive workforce (first order stochastic dominance). By further raising the value of offering a higher piece rate, this composition effect directly increases wage competition between firms for employees.

The market variance of log wages can be decomposed as:

$$var(\log w) = var(\log y_i) + var(\log \theta) + \rho^2 var(x) + 2\rho cov(x, \log \theta).$$

The first term of the right-hand side describes wage variation due to ex-ante worker heterogeneity in initial ability. The second term specifies the variance due to ex-post firm heterogeneity, where on-the-job search implies different firms post different piece rates. The third is largely due to the overlapping generations structure, where older workers typically have greater labour market experience. The last term is due to equilibrium sorting: more experienced workers, having spent more time in the labour market, tend to earn higher piece rates. In the numerical examples below this covariance term is large. It suggests that a significant fraction of observed wage variation arises as young inexperienced workers start their careers on low wages but, over time, accumulate both greater human capital and find better paid employment. Of course with no on-the-job learning, $\rho = 0$, this covariance term has no additional effect on wage dispersion. Similarly if there are no search frictions, the competitive outcome implies $\theta = 1$ and this latter covariance term is zero. Conversely this sorting mechanism with frictions implies learning-by-doing has a potentially large impact on equilibrium wage dispersion.

6.2 Equilibrium Wage Competition

With no value to experience, $\rho = 0$, workers are not willing to work for a piece rate below b (Theorem 1 implies $\theta^R = b$). But using (14) in Theorem 1, some algebra establishes the equilibrium reservation piece rate is strictly decreasing with ρ (even when F is endogenously determined). As first argued in Rosen (1972), firms might extract the rents associated with on-the-job learning by setting relatively low wages. In the Market Equilibrium identified here, (some) firms indeed respond by cutting wages; i.e. those firms offering the least generous piece rate $\theta = \theta^R$.

But market competition here is not just about extracting full rents from unemployed

⁴To obtain this result, differentiation yields

$$\begin{aligned} \frac{\partial H(x|\theta)}{\partial \theta} &= -\frac{2\delta F' \phi (\phi + \lambda + \delta) q}{(\phi + \delta) [q - \phi F]^3} [e^{-\frac{\phi(\phi + \delta + \lambda)x}{(\phi + \lambda)}} - e^{-qx}] - \left[\frac{\phi q \lambda^2 F' F [1 - F] x^2}{(\phi + \delta) [q - \phi F]} \right] e^{-qx} \\ &\quad - e^{-qx} \lambda F' x \left[1 + \frac{\phi^2 (\phi + \lambda + \delta) F [1 - F]}{(\phi + \delta) [q - \phi F]^2} + \frac{\phi q [1 - 2F]}{(\phi + \delta) [q - \phi F]} \right] \end{aligned}$$

The first two terms are guaranteed to be negative. The third term is guaranteed to be negative if $F < \delta/\phi$. Thus $\delta > \phi$ is sufficient for first order stochastic dominance.

workers - by offering a higher wage, a firm also attracts employees from low wage firms. On-the-job learning increases wage competition by directly increasing the value of hiring new workers. It also increases wage competition indirectly through the composition effect identified above: a higher piece rate both attracts and retains a more experienced, and thus more productive, workforce. Consider then the equilibrium piece rate offered by the firm at the n^{th} decile; i.e. the firm which sets piece rate θ so that $F(\theta) = n$. (13) implies this piece rate is given by:

$$\theta(n) = 1 - (1 - \underline{\theta}) \left[\frac{\phi + \delta - \rho + \lambda(1 - n)}{\phi + \delta - \rho + \lambda} \right]^2.$$

An increase in ρ implies equilibrium $\underline{\theta}$ decreases, which lowers $\theta(n)$. But for $n > 0$ and $\rho < \phi$, an increase in ρ decreases the squared bracketed term which increases $\theta(n)$. This reflects the increased competition effect described above. The overall effect of learning-by-doing on $\theta(n)$ is therefore ambiguous (and depends on n). It is not the case that all firms set lower wages with an increase in ρ . There is, however, an increase in the range of offers $\bar{\theta} - \underline{\theta}$, suggesting that greater on-the-job learning increases price dispersion.

Finally we quickly consider the Mm measure of wage dispersion as defined in Hornstein et. al. (2007), which is the ratio between the mean wage paid to the minimum observed wage (or reservation wage). Hornstein et. al. (2007) argue that, for plausible parameter values and for the U.S. economy, the B/M model generates too little wage dispersion; that is, worker reservation wages are too high to be consistent with the data. The introduction of learning-by-doing, which increases equilibrium wage dispersion, corrects this feature of the B/M framework. Indeed as clearly demonstrated in Theorem 1, a $\rho < \phi$ exists where $\theta^R = 0$ and the Mm ratio is then unboundedly large. In the following numerical examples, we find the value of ρ necessary to generate empirically plausible values of the Mm ratio is entirely reasonable.

7 Simulations

We now perform some numerical simulations to illustrate the model's implications for wage dispersion. Using a year as the reference time unit, we set $\phi = 0.025$ so that workers have a 40 year expected working lifetime. Note this also implies workers discount the future at 2.5 percent per annum. Following Jolivet et. al. (2006), who estimate turnover parameters for the US, we set $\delta = 0.055$ and $\lambda = 0.15$.⁵ Following Hall and Milgrom (2008) we set $b = 0.71$. We set $\rho = 0.009$ which ensures that the ratio between the minimum and the average piece rate earned by workers, the Mm ratio, is consistent with the evidence reported in Hornstein et. al. (2007). A one percent per annum rate of learning-by-doing

⁵As our focus is on equilibrium wage dispersion and not equilibrium unemployment rates, we use the offer arrival rate estimate for employed workers.

is clearly reasonable, though it is below the estimate of the marginal impact of experience on log wages as reported by Altonji and Williams (2005) for the US.

Suppose first A is degenerate (identical entrants) and that each labour market entrant has initial productivity $y_i = 1$. Market Equilibrium then implies $\theta^R = 0.35$, $\bar{\theta} = 0.93$ and a Mm ratio $\theta^M/\theta^R = 2.27$.⁶ Figure 1 describes the resulting equilibrium wage density $G'(w)$.⁷

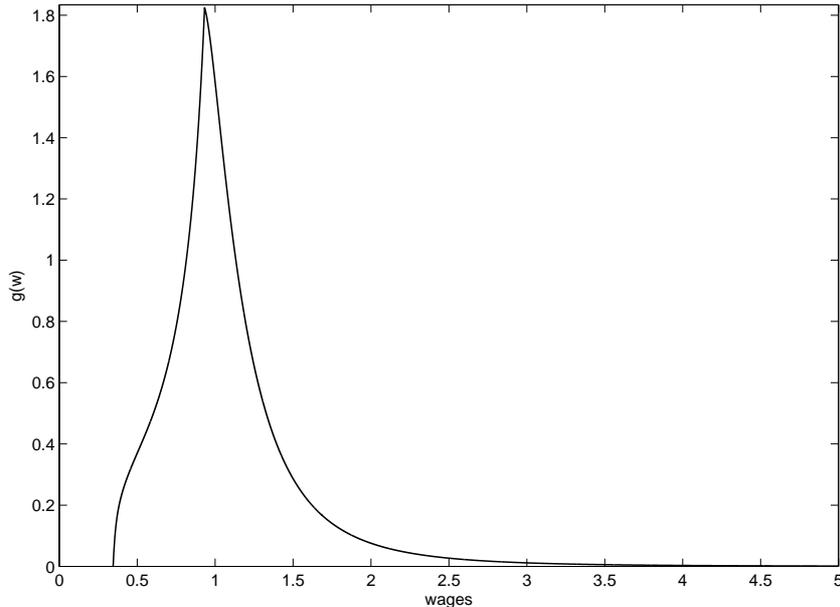


Figure 1: Wage Density when A is a Mass

Despite assuming all entrants are identical, there is significant equilibrium wage dispersion. The insights are twofold. At long experiences (i.e. at high productivities), the conditional distribution of piece rates earned converges to the ergodic distribution $H(\theta | \infty)$. It can be formally shown that at high productivities, the density of wages paid (asymptotically) mirrors the density of worker productivities. But the steady state distribution of worker productivities is Pareto.⁸ Learning-by-doing thus implies the right tail of the wage density function is asymptotically Pareto distributed and so has a ‘fat’ right tail. This of course is a well known property of empirical wage distributions; see for example von Weizsäcker (1993) and Neal and Rosen (2000).

Conversely at short experiences, search frictions imply there is a lot of randomness in

⁶Although Hornstein, et. al. (2007) report an average Mm ratio of 1.7 based on all the data sets used (PSID, 1990 Census and OES), our Mm ratio is overall consistent with their results. In particular, based on the 1990 Census data (Table 3) and using the wage at the first or fifth decile of the wage distribution as the lowest observed wage gives Mm ratios above 2.

⁷In the Appendix we provide a formal derivation of the wage density.

⁸This arises as the distribution of experience, $H(x, \bar{\theta})$, is exponential, while each worker’s productivity is $y = y_i e^{\rho x}$.

initial worker employment outcomes. But note that the job search process (finding better paid employment) is a relatively rapid process (a job offer arrival rate of 15% per annum) compared to on-the-job learning (where productivity increases at only 1% per annum). The increasing left tail of the wage distribution resembles the B/M wage density and arises as young workers typically start their careers in low wage jobs but quickly move to better paid work through on-the-job search. Wage dispersion thus reflects the interaction between (fast) worker on-the-job search and (slow) learning-by-doing.

Smoothing out this wage distribution clearly requires some worker heterogeneity. For ease of exposition the theory section assumed a finite number of types. But the analysis extends straightforwardly if $A(\cdot)$ instead describes a continuum of underlying abilities.⁹ Suppose then that $A(\cdot)$ is distributed according to a Gamma distribution:

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

where $k_0, k_1 > 0$ and $\Gamma(\cdot)$ is the gamma function. The mean and variance of $A(\cdot)$ are $\mu = k_0 k_1$ and $\sigma^2 = (k_0)^2 k_1$.¹⁰ We consider values of k_0 and k_1 such that the density A' roughly resembles the shape of the distribution of ex-ante worker heterogeneity used in Bontemps et. al. (1999) and estimated by Postel-Vinay and Robin (2002b). Figure 2 describes three ability distributions with the same mean but different spreads: specifically (a) $k_0 = 6, k_1 = 3.3$; (b) $k_0 = 4, k_1 = 5$ and (c) $k_0 = 2, k_1 = 10$. Figure 3 describes the resulting equilibrium wage distributions.

It is immediately apparent that the overall shape of the wage density closely resembles that of the density of initial productivities. This is perhaps not very surprising as wages paid are directly linked to a worker's productivity, and a worker's productivity is directly related to his/her initial productivity when first entering the labour market. Nevertheless we know by Figure 1 that there is significant wage dispersion and non-trivial wage dynamics for each type. Table 2 below decomposes the wage variation described in Figure 3 into its underlying components:

$$var(\log w) = var(\log y_i) + var(\log \theta) + \rho^2 var(x) + 2\rho cov(x, \log \theta).$$

To construct this table, we first constructed the Market Equilibrium for each case and then computed each of the above components. Their total sum is reported in the second column, while their relative contribution to this sum is given in the subsequent columns. Noting that $\log y = \log y_i + \rho x$, row (a) implies that 82% in the variation in log productivity is due to disperse initial abilities, the remaining 18% arises as workers accumulate human capital with experience, and worker experience is disperse in a steady state. The lower

⁹The only difference is that we instead integrate over dA rather than sum over i in the profit function.

¹⁰The alternative specifications of A we used were Generalised Pareto, Three parameter Weibull and Uniform. The insights described below are robust to these functional forms of A .

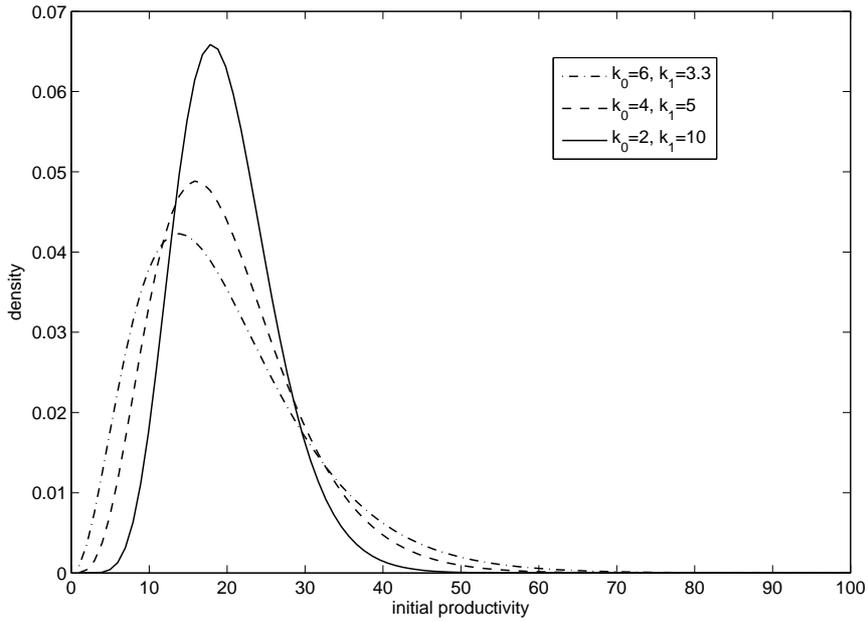


Figure 2: Initial productivity densities

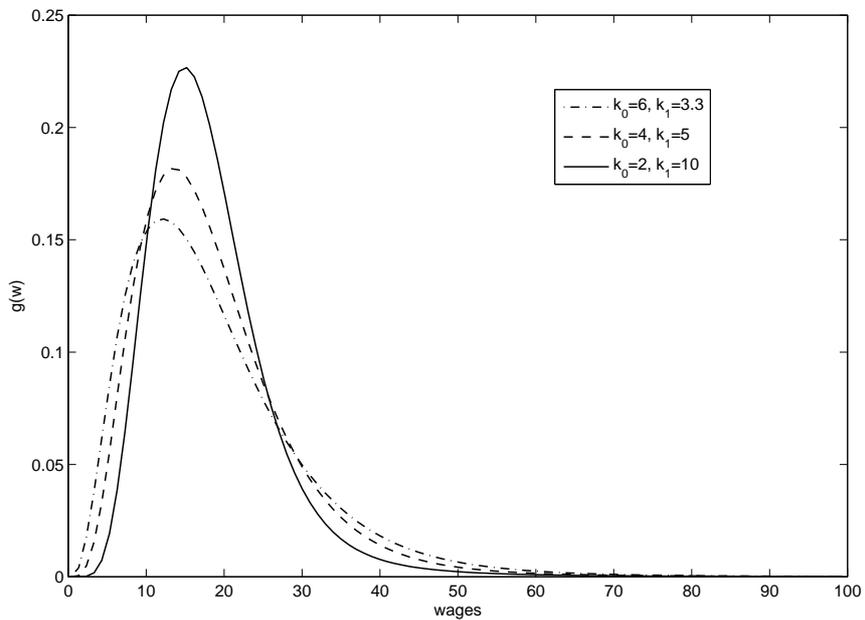


Figure 3: Wage density when A is continuous

rows (b)-(d) impose smaller variation in the distribution of initial abilities.¹¹

¹¹Since (13) and (9) are independent of workers initial abilities, the values of $var(\log \theta)$, $\rho^2 var(x)$ and $2\rho cov(x, \theta)$ do not change across the cases studied. Namely, the above parameter values imply $var(\log \theta) = 0.0744$, $\rho^2 var(x) = 0.075$ and $2\rho cov(x, \log \theta) = 0.0936$ for all rows (a)-(d) in Table 1. The implied correlation between x and $\log \theta$ is 0.63.

Density of A	Total variation	Relative contribution (%)			
	$var(\log w)$	$var(\log y_i)$	$var(\log \theta)$	$\rho^2 var(x)$	$2\rho cov(x, \log \theta)$
(a)	0.596	59.2	12.5	12.6	15.7
(b)	0.464	47.6	16.0	16.2	20.2
(c)	0.348	30.2	21.4	21.6	26.8
Degenerate	0.243	0	30.6	30.9	38.5

The most interesting feature of Table 2 is the final column which describes the impact of equilibrium sorting on wage dispersion. Note that if there were no learning-by-doing, $\rho = 0$, then the final two columns would both be zero. In these examples, equilibrium sorting more than doubles the impact of learning-by-doing on wage dispersion/inequality. Young unemployed workers not only have little experience, most job offers typically pay low wages. Over time young workers become better paid not only because they become more experienced, and thus more productive, they also eventually find and quit to better paid employment. The positive correlation between these two processes significantly magnifies measured wage inequality: the young typically earn much less than the old.

8 Conclusion

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

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

There are several lines for further research. A simple extension is to allow different arrival rate of offers for employed and unemployed workers. This changes the equation for the reservation piece rate of unemployed workers θ^R but, given θ^R , this does not otherwise affect competition between firms for employed workers. We believe the insights of the paper go through automatically.

A more difficult challenge is to note that standard Mincer wage equations typically assume decreasing returns to experience; i.e. the return to experience, $\rho = \rho(x)$, is quadratic rather than linear. For tractability the model here has assumed productivity grows with experience at a constant rate. Without this assumption unemployed workers with different experiences would have different returns to further experience. Intuition suggests more experienced workers, having a lower return to further experience, would then have higher reservation piece rates. Thus when laid-off, more experienced workers would have longer (expected) spells of unemployment (which appears a reasonable prediction). Such an extension would be valuable as it would yield a more flexible Mincer wage equation

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

Unfortunately disperse reservation piece rates severely complicates the aggregation problem and solving for the equilibrium distribution functions F, H becomes problematic. Nevertheless these latter distribution functions might still be usefully estimated on data.

A difficult but more tractable extension is to incorporate wage/tenure contracts into the analysis. Burdett and Coles (2003) supposes all firms and workers are ex-ante identical but, in contrast to B/M, firms post contracts where wages paid depend on tenure. It identifies an equilibrium where firms offer different contracts, but each firm offers a contract where wages paid increase smoothly with tenure. In such a market environment, workers are promoted by seniority: as more senior employees quit or retire, junior employees are promoted to take their place. By rewarding loyalty, this promotion mechanism reduces the quit incentives of junior employees and so allows the firm to extract even greater search rents (and so increase profit). Learning-by-doing and a piece rate tenure contract $\theta = \theta(\tau)$ would then yield a Mincer wage equation of the following form

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

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

Finally one might incorporate firm heterogeneity into the model. With no on-the-job learning and identical workers, Van den Berg and Ridder (1998) and Bontemps et. al.

(1999) show that for a given distribution of firm productivities, the B/M framework yields a unique equilibrium distribution of posted wages. By selecting an appropriate distribution of firm productivities it is possible to generate an equilibrium wage distribution that has an interior mode and is skewed the right way. But as noted by Mortensen (2003) and Postel-Vinay and Robin (2006), it is difficult to construct a distribution of firm productivities that approximates the empirical wage density - one needs to assume a distribution of firm productivities that has an implausible long right tail. Although introducing firm heterogeneity would seem useful for empirical work, it is not clear that it will yield new and useful theoretical insights.

References

- [1] Abowd, John M., Hampton Finer and Francis Kramarz. 1999. "Individual and Firm Heterogeneity in Compensation." in John C. Haltiwanger, Julia I. Lane, James R. Spletzer, Jules Theeuwes and Kenneth R. Troske. eds., *The Creation and Analysis of Employer-Employee Matched Data*, Amsterdam: North-Holland Publishing.
- [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. Shaktoko. 1987. "Do Wages Rise with Job Seniority?" *Review of Economic Studies*, 54(3): 437-459.
- [4] Altonji, Joseph G. and Nicolas Williams. 2005. "Do Wages Rise with Job Seniority? A Reassessment." *Industrial and Labor Relations Review*, 58(3): 370-397.
- [5] Bagger, Jesper, Francois Fontaine, Fabien Postel-Vinay and Jean-Marc Robin. 2006. "A Feasible Equilibrium Search Model of Individual Wage Dynamics with Human Capital Accumulation." <http://www.efm.bris.ac.uk/ecfybpv/wp/humancapital.pdf>
- [6] Barlevy, Gadi. 2008, "Identification of Search Models using Record Statistics." *Review of Economic Studies*, 75(1): 29-64.
- [7] Becker, Gary S. 1964. *Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education*. Chicago, IL: University of Chicago Press.
- [8] 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.
- [9] 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.

- [10] Burdett, Kenneth and Dale T. Mortensen. 1998. "Wage Differentials, Employer Size, and Unemployment." *International Economic Review*, 39(2): 257-273.
- [11] Burdett, Kenneth and Melvyn G. Coles. 2003. "Equilibrium Wage-Tenure Contracts." *Econometrica*, 71(5): 1377-1404.
- [12] Fu, Chao. 2009. "Training, Search and Wage Dynamics." Unpublished.
- [13] Hall, Robert and Paul Milgrom. 2008. "The Limited influence of Unemployment on the Wage Bargain." *American Economic Review*, 98(2): 1653-1674.
- [14] Hornstein, Andreas, Per Krusell and Giovanni L. Violante. 2007. "Frictional Wage Dispersion in Search Models: A Quantitative Assessment." http://www.econ.nyu.edu/user/violante/Workingpapers/frictional_main_september4.pdf.
- [15] 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.
- [16] Mincer, Jacob. 1974. *Schooling, Experience and Earnings*. New York: National Bureau of Economic Research.
- [17] Mortensen, Dale T. 2003. *Wage Dispersion: Why are Similar Workers Paid Differently?*, Cambridge, MA: MIT Press.
- [18] Neal, Derek A. and Sherwin Rosen. 2000. "Theories of the Distribution of Earning." in Anthony B. Atkinson and Francios J. Bourguignon. eds., *Handbook of Income Distribution*, Vol. 1, Amsterdam: Elsevier Science B.V.
- [19] Postel-Vinay, Fabien and Jean-Marc Robin. 2002a. "The Distribution of Earnings in a Equilibrium Search Model with State-Dependent Offers and Counteroffers." *International Economic Review*, 43(4): 989-1016.
- [20] Postel-Vinay, Fabien and Jean-Marc Robin. 2002b. "Equilibrium Wage Dispersion with Worker and Employer Heterogeneity." *Econometrica*, 70(6): 2295-2350.
- [21] Postel-Vinay, Fabien and Jean-Marc Robin. 2006. "Microeconomic Search-Matching Models and Matched Employer-Employee Data." in Richard Blundell, Whitney Newey and Torsten Persson. eds., *Advances in Economics and Econometrics, Theory and Applications, Ninth World Congress*, Vol. 2, Cambridge: Cambridge University Press.
- [22] Rosen, Sherwin. 1972. "Learning and Experience in the Labor Market." *Journal of Human Resources*, 7(3): 326-342.

- [23] 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.
- [24] Topel, Robert. 1991. "Specific Capital, Mobility, and Wages: Wages Rise with Job Seniority." *Journal of Political Economy*, 99(1): 145-176.
- [25] Topel, Robert and Michael P. Ward. 1992. "Job Mobility and the Careers of Young Men." *Quarterly Journal of Economics*, 107(2): 439-479.
- [26] Van den Berg, Gerard and Geert Ridder. 1998. "An Empirical Equilibrium Search Model of the Labor Market." *Econometrica*, 66(5): 1183-1221.
- [27] von Weizsäcker, Robert K. 1993. *A Theory of Earnings Distribution*, Cambridge: Cambridge University Press.

APPENDIX

Proof of Proposition 1:

Note that a worker's income, whether unemployed or employed, is always proportional to y . As on-the-job learning is also proportional to y and workers are risk neutral, (1) and (2) imply there exists a number α^U and a function $\alpha^E(\cdot)$ such that $W^U(y) = \alpha^U y$ and $W^E(y, \theta) = \alpha^E(\theta)y$. Since θ^R satisfies $W^E(y, \theta^R) = W^U(y)$ it is immediate that θ^R is independent of y and common across all unemployed workers. (1) then implies α^U satisfies

$$\phi\alpha^U = b + \lambda \int_{\theta^R}^{\bar{\theta}} [\alpha^E(\theta') - \alpha^U] dF(\theta'), \quad (15)$$

while (2) implies

$$(\phi + \delta)\alpha^E(\theta) = \theta + \rho\alpha^E(\theta) + \lambda \int_{\theta}^{\bar{\theta}} [\alpha^E(\theta') - \alpha^E(\theta)] dF(\theta') + \delta\alpha^U. \quad (16)$$

Differentiating (16) with respect to θ then implies α^E is determined by the following differential equation

$$\frac{d\alpha^E}{d\theta} = \frac{1}{q(\theta) - \rho} \quad (17)$$

and by evaluating (16) at $\theta = \bar{\theta}$ we obtain the boundary condition $\alpha^E(\bar{\theta}) = (\bar{\theta} + \delta\alpha^U)/(\delta + \phi - \rho)$.

Next we solve for θ^R and α^U . First let $\theta = \theta^R$ in (16). As $\alpha^E(\theta^R) = \alpha^U$ we obtain

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

Comparing this equation with (15) establishes (3) as described in Proposition 1. Next integrate (15) by parts. Using (17) and $\alpha^E(\theta^R) = \alpha^U$ then yields (4) in Proposition 1. Thus (3) and (4) describe a pair of equations for (α^U, θ^R) . Note that (3) is linear, has slope $-1/\rho$ and $\alpha^U = 0$ at $\theta^R = b$. Differentiating (4) with respect to θ^R , on the other hand, implies this equation is continuous and strictly decreasing with slope

$$\left[\frac{d\alpha^U}{d\theta^R} \right]_{eqn(4)} = -\frac{1}{\phi} \left(\frac{\lambda(1 - F(\theta^R))}{q(\theta^R) - \rho} \right).$$

The parameter restriction $\phi > \rho$ implies that the locus described in (4) is flatter (strictly) than the locus described in (3) for all θ , even when $\theta^R < \underline{\theta}$. Moreover, as the former locus implies α^U is strictly positive and finite at $\theta^R = b$, continuity now implies these two loci must have a unique intersection at some $\theta^R < b$.

Finally, solving for θ^R using (3) and (4) and then (implicitly) differentiating establishes that θ^R is strictly decreasing in ρ . This completes the proof of Proposition 1.

Proof of Lemma 1:

As the arguments are already well known (see B/M) we only sketch the proof. A contradiction argument establishes there cannot be a mass point in F . If there were, say at $\theta = \theta^m$, steady state would imply a mass of employees on piece rate θ^m . But offering piece rate $\theta = \theta^m + \varepsilon$ where $\varepsilon > 0$ but very small would yield a slightly lower profit per hire but a large increase in the hiring rate, and this deviation would strictly increase profit (and so contradict equilibrium). A contradiction argument also establishes the support of F must be connected. Otherwise if there were a hole, say for $\theta \in [\theta_L, \theta_H]$ then the offer $\theta = \theta_L$ would yield strictly greater profit (as both offers attract the same number of workers) which contradicts θ_H being an optimal offer. Finally a contradiction argument also establishes $\underline{\theta} = \theta^R$. $\underline{\theta} < \theta^R$ would imply a firm offering $\theta = \underline{\theta}$ obtains zero profit, which contradicts strictly positive profit. $\underline{\theta} > \theta^R$ would instead imply offering $\theta = \theta^R$ makes strictly greater profit than offering $\underline{\theta}$ which contradicts optimality of $\underline{\theta}$. This completes the proof of lemma 1.

Proof of Proposition 2:

Evaluating (7) at $\theta = \bar{\theta}$ and using (6) to substitute out $N(\cdot)$ yields

$$q(\bar{\theta})H(x, \bar{\theta}) + \frac{\partial H(x, \bar{\theta})}{\partial x} = \frac{\phi(\phi + \delta + \lambda) + \lambda\delta H(x, \bar{\theta})}{(\phi + \lambda)}.$$

Integrating this linear differential equation, noting $H(0, \bar{\theta}) = 0$, finds

$$H(x, \bar{\theta}) = 1 - e^{-\left(\frac{\phi(\phi + \delta + \lambda)x}{(\phi + \lambda)}\right)}.$$

Using this in (6) and simplifying yields (8).

Next for a given θ , integration of (7) using the integrating factor $e^{q(\theta)x}$ yields:

$$\left[H e^{q(\theta)x'} \right]_{x'=0}^x = \int_{x'=0}^x e^{q(\theta)x'} (\phi + \delta) F(\theta) N(x') dx'.$$

As $H(0, \theta) = 0$, employed workers instantly accumulate positive experience, we obtain

$$H(x, \theta) = (\phi + \delta) F(\theta) \int_{x'=0}^x e^{-q(\theta)(x-x')} N(x') dx'$$

for all $x \geq 0$, $\theta \in [\underline{\theta}, \bar{\theta}]$. Using (8) and integrating yields (9). This completes the proof of Proposition 2.

Proof of Theorem 1 - Derivation of θ^R :

First note that (3) and (4) imply θ^R is determined by

$$\theta^R = \frac{b(\phi - \rho)}{\phi} - \frac{\rho}{\phi} \int_{\theta^R}^{\bar{\theta}} \frac{\lambda(1 - F(\theta))}{q(\theta) - \rho} d\theta,$$

for any given F . Fix a $\underline{\theta}$ and let $F = \widehat{F}(\theta | \underline{\theta})$, where $\widehat{F}(\theta | \underline{\theta})$ is described by (13).

Substituting for F in the expression describing θ^R yields

$$\begin{aligned}\theta^R(\underline{\theta}) &= \frac{b(\phi - \rho)}{\phi} - \frac{\rho(\bar{\theta} - \theta^R(\underline{\theta}))}{\phi} + \frac{\rho(\phi + \delta - \rho)}{\phi(\phi + \delta - \rho + \lambda)} \int_{\theta^R(\underline{\theta})}^{\bar{\theta}} \left(\frac{1 - \underline{\theta}}{1 - \theta}\right)^{1/2} d\theta \\ &= \frac{b(\phi - \rho)}{\phi} - \frac{\rho(\bar{\theta} - \theta^R(\underline{\theta}))}{\phi} + \frac{2\rho(\phi + \delta - \rho)(1 - \underline{\theta})^{1/2}}{\phi(\phi + \delta - \rho + \lambda)} [(1 - \bar{\theta})^{1/2} - (1 - \theta^R(\underline{\theta}))^{1/2}].\end{aligned}$$

Noting that (13) implies

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

solving for $\hat{\theta}^R(\underline{\theta}) = \underline{\theta}$ then gives the expression in (14).

Derivation of the wage density:

Differentiating (9), we obtain the joint density of workers characteristics, y and θ for each i . Namely,

$$\begin{aligned}h(x, \theta) &= \phi F'(\theta) e^{-q(\theta)x} [1 + \lambda x F(\theta)] \left[\frac{(\phi + \delta + \lambda)(1 - F(\theta))}{q(\theta) - \phi F(\theta)} \right] \\ &\quad + \frac{\delta \phi (\phi + \delta + \lambda) F'(\theta)}{(q(\theta) - \phi F(\theta))^2} \left[\left(\frac{\phi + \delta + \lambda}{\phi + \lambda}\right) e^{-\left(\frac{\phi + \delta + \lambda}{\phi + \lambda}\right)x} - F(\theta) e^{-q(\theta)x} \right],\end{aligned}$$

for all $\theta \in [\underline{\theta}, \bar{\theta}]$ and $x \geq 0$, otherwise, $h(x, \theta) = 0$. As employed workers earning a wage below w must have experience x satisfying $\theta y_i e^{\rho x} \leq w$, the wage distribution can be written as

$$G(w) = \sum_i \gamma_i \int_{\theta=\underline{\theta}}^{\bar{\theta}} \int_{x=0}^{\frac{1}{\rho} \ln \frac{w}{\theta y_i}} h(x, \theta) dx d\theta.$$

The wage density is then obtained by differentiating wrt w .