The Cost of Job Loss*

Kenneth Burdett †
University of Pennsylvania

Carlos Carrillo-Tudela ‡
University of Essex,
CEPR, CESifo and IZA

Melvyn Coles §
University of Essex

July 2017

Abstract

This paper identifies an equilibrium theory of wage formation and quit turnover in a labour market where risk averse workers accumulate human capital through learning-by-doing while employed, there is skill loss while unemployed and there is on-the-job search. Firms set optimal company wage policies which, in equilibrium, have the property that the wage paid increases with experience and tenure. Using indirect inference, our quantitative analysis shows the model not only reproduces the large and persistent fall in wages following job loss as found in UK data, it also explains why those wage losses differ so markedly across skill groups.

Keywords: Job search, human capital accumulation, job loss.
JEL: J63, J64, J41, J42.

*We would like to thank Giovanni Mastrobuoni, Guiseppe Moscarini, Jean-Marc Robin, Ludo Visschers and seminar participants at the Essex SaM workshop, Bristol SaM workshop, St Louis FED and the Universities of Vienna and Konstanz for their comments and suggestions. The usual disclaimer applies.

†Corresponding author: Department of Economics, University of Pennsylvania, 3718 Locust Walk, Philadelphia, PA 19104, USA. Email: kennethb@econ.upenn.edu

‡Correspondence: Department of Economics, University of Essex, Wivenhoe Park, Colchester, CO4 3SQ, UK. Email: cocarr@essex.ac.uk. Carlos Carrillo-Tudela acknowledges financial support from the UK Economic and Social Research Council (ESRC), award reference ES/I037628/1.

§Correspondence: Department of Economics, University of Essex, Wivenhoe Park, Colchester, CO4 3SQ, UK. Email: mcoles@essex.ac.uk. Melvyn Coles acknowledges financial support from the UK Economic and Social Research Council (ESRC), award reference ES/I037628/1.

1 Introduction

Following Burdett and Mortensen (1998), there is a large literature on equilibrium wage formation and labour turnover in frictional labour markets. The approach is important since it provides a structural interpretation not only for how wages evolve over individual worker careers, it also explains the surprisingly large variation in wage outcomes across workers; e.g. Mortensen (2003). Much of this literature adopts the sequential auction approach of Postel-Vinay and Robin (2002) for it yields a highly tractable econometric framework (see Robin, 2011, Bagger et al., 2014, Lise and Robin, 2015, Jarosch, 2015, Bagger and Lentz, 2015). Here instead we follow the approach of Moscarini and Postel-Vinay (2013) where firms can contract on future wages and do not match outside offers. We extend that approach by allowing firms to compete in optimal contracts (as in Stevens, 2004, and Burdett and Coles, 2003), there is also learning-by-doing and skill loss while unemployed but we simplify by only considering equilibrium along a stationary growth path. The theoretical results lead to rich predictions about how a worker's wage evolves within and across employment spells.

Following seminal work by Topel (1990) and Ruhm (1991), there has also been extensive, empirical research on the large and persistent wage losses faced by workers who are displaced into unemployment.[1] The second part of our paper uses UK data to quantitatively assess our equilibrium framework. It identifies equilibrium wage outcomes which are consistent with (i) the estimated returns to tenure and experience in the UK, (ii) worker job turnover rates and (iii) measures of wage dispersion as suggested by Hornstein et al. (2011). We find that the model not only reproduces the large and persistent fall in wages following job loss as found in UK household longitudinal data, it also explains why those wage losses differ so markedly across skill groups.

Burdett and Coles (2003) considers equilibrium wage contracting in a steady state environment with on-the-job search and risk averse workers. It shows the optimal contract increases the wage paid with tenure for, by backloading wages paid in the job spell, the contract reduces employee quit incentives. Here we extend that approach in three important ways. First we allow learning-by-doing while employed so that wages exhibit both experience and tenure effects (e.g. Altonji and Shakotko, 1987, Topel, 1991, Dustmann and Meghir, 2005). Second we allow that layoff may generate (or be correlated with) loss in human capital (e.g. Pissarides, 1992, Ljungqvist and Sargent, 1998). This yields a framework which is ideal for identifying the cost of job loss in a frictional labour market. Finally we follow the ideas of Moscarini and Postel-Vinay (2013) and do not restrict attention to steady state. Instead at date zero, each firm j precommits to a company wage policy which pays wage $w_{ijt} = w_{jt}(k_{it}, \tau_{it}, \Omega_t)$ to each employee i at date $t \geq 0$ depending on the employee's (general) human capital k_{it}, the worker's tenure τ_{it} and the aggregate state of the economy Ω_t. A well known difficulty with such precommitment games, however, is that the state of the economy Ω_0 is payoff relevant when the optimal wage policy is written down at $t = 0$. This introduces a non-stationary element into the equilibrium outcome, for the equilibrium path of payments $w_{jt}(.)$ then depends on initial Ω_0. Here we follow Woodford

[1] See for example Jacobson et al. (1993), Stevens (1997), Kletzer and Fairlie (2003), Couch and Placzek (2010) and Davis and von Wachter (2011).

(2003) and only consider timeless equilibria. In such an equilibrium, each firm $j \in [0,1]$ precommits to an optimal company wage policy $w = w_j(k, \tau, \Omega)$ which does not depend directly on time (though individual wage payments vary over time as a worker's tenure and human capital change within the employment spell and as the (stationary) economy evolves). One can consider such equilibria either as the ergodic limit of the precommitment contracting game, or that the startpoint is sufficiently far in the past that initial values no longer distort the market outcome.[2] In fact Burdett and Mortensen (1998) provides an early illustration of this equilibrium concept: there each firm precommits to a fixed wage w_j which is profit maximising given the economy is in its (ergodic) steady state.

Even with the restriction to timeless equilibria the framework remains complex for each company wage policy must be profit maximising given the set of company wage policies posted by all other firms, the worker quit strategies (which must also be optimal given the set of company wage policies) and the endogenous distributions of worker tenures and general human capital. Fortunately equilibrium has a sufficiently simple structure that it can be estimated on the data using indirect inference. To do this, the empirical section uses the British Household Panel Survey (BHPS) which provides wage and turnover data for the U.K. over 1991-2004, a period of stable economic growth. Using this data, we first identify the structural parameters of the model by using various moments of the data directly related to our model (described in the text). Following Stevens (1997) we then estimate the long-term cost of job loss where, as a test of the model, we verify that the reduced form estimates of the cost of job loss using model-generated data are indeed the same as those obtained on the actual data.[3]

As in Ljungqvist and Sargent (1998), the permanent cost of job loss arises due to human capital loss and foregone skills accumulation while unemployed. The temporary component is due to job ladder effects, where displaced workers must once again re-climb the job ladder. As job ladder effects differ markedly across different career-types, here we follow Burdett et al. (2016) and divide the sample into two education/skill categories: low e and high e workers. The distinction is important for low e workers, on average, face high layoff probabilities, rarely receive outside job offers and endure relatively long durations of unemployment. As these workers rarely receive outside offers, there is little purpose to backloading their wages within a job spell. Consistent with reduced form estimates, we find the short run fall in wages due to job loss is relatively small for low e workers. Skill loss effects imply the permanent fall in wages due to job loss is, in expectation, around 5.8%.

Job ladder effects for high e workers are instead very large: high e workers, on average, enjoy long employment spells and frequently receive outside offers. For such workers, the model finds the average, temporary, fall in wages due to job loss is 19.1%. Although this wage loss is "temporary", it is long-lived for it takes time to re-climb the job ladder: estimates suggest the wage recovery process has a half-life of around 4 years. The permanent loss in wages for high

[2]Moscarini and Postel-Vinay (2013) does not define timeless equilibria. Rather it assumes incumbent employees make side-payments at date zero so that continuation payoffs at that date are consistent with those implied by a timeless equilibrium with $\Omega = \Omega_0$.

[3]Stevens (2007) uses the Panel Survey of Income Dynamics (PSID) to estimate the cost of job loss in the U.S.. See also Kletzer and Fairlie (2003) who instead use the National Longitudinal Survey of Youth.

e workers through job loss is estimated at 4.4% which, though quite large, is smaller than that for low e workers. These results echo the empirical findings of Stevens (1997) and Kletzer and Fairlie (1997) in the US.[4]

There are several closely related papers. Davis and von Wachter (2011) evaluate a search model based on Mortensen and Pissarides (1994) augmented to allow on-the-job search. They find that framework fails to match the observed wage losses of displaced workers in the US. Perhaps closest in spirit are Bagger et al. (2014) and Jarosch (2015) which consider stochastic human capital dynamics with instead sequential auctions, but also see Krolikowski (2014) and Jung and Kuhn (2014). Those papers model heterogeneity as a characteristic of the firm or of the firm-worker match; i.e. some firms or matches offer more stable jobs than others. Here instead we emphasise how heterogeneity in unemployment risk *and* job ladder differences across skill groups are important for explaining the cost of job loss.

Section 2 describes the model and fully characterises timeless equilibria. Section 3 describes the data, the estimation procedure and the fit of the model. Section 4 presents the estimates of the cost of job loss for the UK and compares it to the one generated by the model. The Appendix contains all tedious proofs, the full description of the data, simulations and the estimation procedure.

2 The Model

Time is continuous with an infinite horizon. There is a continuum of both firms and workers, each of measure one. All are infinitely lived and discount the future at rate $r > 0$. Firms, indexed by $j \in [0, 1]$, are equally productive with a constant returns to scale technology. Workers are ex-ante heterogenous with general human capital $k \in (0, \infty)$. A worker type k generates revenue flow $Ak > 0$ while employed and home production flow bAk while unemployed where $b \in [0, 1)$ implies a gain to trade exists. $A > 0$ is an aggregate productivity parameter which grows at exogenous rate $\gamma_A \geq 0$.

Learning-by-doing implies a worker's human capital grows at rate $\rho \geq 0$ while employed. While unemployed there is skill loss whereby the worker's human capital falls at rate $\phi > 0$. Unemployed workers receive job offers at exogenous Poisson rate $\lambda_0 > 0$, on-the-job search implies employed workers receive outside offers at rate $\lambda_1 > 0$ and job search is random in that any job offer is considered a random draw from the set of all job offers in the market. There is no recall of rejected job offers.

So what is a job offer? We generalise Burdett and Coles (2003) by allowing firms to compete in optimal contracts. Each firm $j \in [0, 1]$ precommits at date zero to a company wage policy which pays wage $w = w_{jt}(\tau, k, A)$ to each employee at any future date $t \geq 0$ depending on the employee's tenure (or seniority) τ, human capital k and aggregate productivity A at that date.

[4] Our estimates also complement the existing evidence on the cost of job loss for the UK. In particular, Arulampalam (2001) uses the BHPS for the period 1991-1997 and estimates a standard mincer wage equation augmented by variables that capture the effects of a non-employment spell on wages during the re-entry employment spell. She finds that 4 years after re-employment, workers face a reduction of 11% in log wages. See also Gregory and Jukes (2001) for evidence on the impact of the duration of unemployment on re-employment wages.

Thus given contact with a potential hire k_0 at date $t' \geq 0$, the company's wage policy implies a promised sequence of wages $w_{jt}(t-t', k_0 e^{\rho(t-t')}, A(t))$ at future dates $t > t'$ where, should the worker remain employed at the firm by that date, the worker will have accumulated tenure $\tau = t - t'$, human capital $k_0 e^{\rho(t-t')}$ and aggregate productivity is $A(t)$. Should an employee (τ, k, A) at firm j at date t receive a (random) outside job offer from firm $j' \sim U[0,1]$, the worker calculates the continuation value of remaining at current firm j on contract $w_{jt}(.)$ with current tenure τ, and compares it to the value of being employed at the outside firm j' on contract $w_{j't}(.)$ but with zero tenure. No recall implies the worker quits if the latter contract yields a higher payoff. Note this contracting approach rules out offer matching; e.g. Postel-Vinay and Robin (2002). The simplest justification is an equal treatment rule - that anti-discrimination legislation requires the firm's wage policy must pay the same wage to equally productive workers with the same seniority.[5] Thus should an employee receive a preferred outside offer, the worker is simply let go and the firm hires replacement employees on the company contract.

Exogenous job destruction shocks occur at gross rate δ^G but we suppose either that an employee "gets wind" of an impending layoff, or the firm is able to give a period of notice. We assume fraction δ/δ^G of workers who are given notice of layoff fail to obtain an outside offer during the notice period and so enter unemployment upon layoff. The remaining fraction $[1 - \delta/\delta^G]$, however, generate one random outside offer during the notice period and, for simplicity, we assume employment at the new firm starts at the layoff event. Thus δ describes the exogenous rate at which an employed worker is laid-off into unemployment and we let $\lambda_q \equiv \delta^G - \delta$ denote the rate at which employed workers are laid-off but successfully remain employed through an immediate (random) job hire. Some refer to this latter process as exogenous quits.

Although job destruction shocks imply risk averse workers have a precautionary motive to save, for tractability we simplify by assuming consumption equals earnings at all points in time; i.e. there are no savings. We further assume constant relative risk aversion; i.e. $u(w) = w^{1-\sigma}/(1-\sigma)$ with $\sigma > 0$.

The equilibrium framework is complex for each company wage policy $w_{jt}(.)$ must be a best response to those set by competing firms, the quit strategies of workers, the endogenous distribution of employed worker tenures and to the distribution of skills across workers which evolves over time. An added complexity with precommitment games is that initial values generate complex and uninteresting non-stationary dynamics; e.g. Woodford (2003). For tractability we consider "timeless" equilibria where each firm $j \in [0,1]$ precommits to a company wage policy $w = w_j(\tau, k, A)$ which does not change with time (though individual wage payments vary over time as an employee accumulates greater tenure and experience). In effect we characterise the stationary growth path of the economy where each firm j chooses a timeless company wage policy $w_j(\tau, k, A)$ to maximize expected discounted profit.

We further focus the analysis by only characterising timeless equilibria in which the optimal contract $w_j(\tau, k, A)$ has a particularly useful structure, that $w_j(.) = Ak\widetilde{\theta}_j(\tau)$. This equilibrium

[5] Another important contract restriction is we rule out job fees (see for example Stevens, 2004).

outcome is empirically useful for it yields the log-linear wage equation:

$$\log w_{ijt} = \log k_i + \log \widetilde{\theta}_j(0) + \rho x_{it} - \phi \mathrm{Z}_{it} + \log \frac{\widetilde{\theta}_j(\tau)}{\widetilde{\theta}_j(0)} + \log A_t,$$

which has both worker and firm fixed effects (worker i's initial human capital, firm j's starting wage rate $\log \widetilde{\theta}_j(0)$), experience effects ($x_{it}$ is worker $i's$ total work experience and Z_{it} is worker $i's$ time spent unemployed), as well as firm specific tenure effects. As discussed in Section 6, firm fixed effects and firm specific tenure effects determine the temporary losses workers face when transiting into unemployment, while the human capital dynamics determine the permanent losses of job displacement.

Definition of Equilibrium: A timeless equilibrium is a set of wage rate contracts $\{\widetilde{\theta}_j(\tau)\}_{\tau\in[0,\infty)}$ for each firm $j \in [0,1]$ such that along the stationary growth path:
(i) company wage policy $w_j(\tau,k,A) = Ak\widetilde{\theta}_j(\tau)$ is profit maximising for each firm $j \in [0,1]$ given workers use optimal job search strategies, and
(ii) the joint distributions of employment, tenures, wages and human capital are consistent with optimal job search and the set of optimal employment contracts.

We identify such equilibria using the following approach. The next section considers optimal worker behavior given all firms post a company wage policy consistent with a timeless equilibrium; i.e. we describe optimal quit turnover given $w_j(\tau,k,A) = Ak\widetilde{\theta}_j(\tau)$ for each $j \in [0,1]$. Given such worker turnover, the following section identifies the set of contracts $\{\widetilde{\theta}_j(.)\}$ such that there is no deviating (general) contract $w = w(\tau,k,A)$ which is profit increasing.

3 Worker Optimality in a Timeless Equilibrium

Suppose each firm $j \in [0,1]$ posts a company wage policy of the form $w_j(.) = Ak\widetilde{\theta}_j(\tau)$. Along the stationary growth path, let $V = V(\tau,k,A|\theta)$ denote the employment value enjoyed by a worker with tenure τ, human capital k in aggregate state A on representative contract $\theta(\cdot)$. Let $V^U(k,A)$ denote the value of being unemployed.

As there is a gain to trade, it is never optimal for a firm to post a contract $\theta(.)$ which induces workers to quit into unemployment. For any such contract $\theta(.)$, standard arguments imply $V(.)$ is identified by the Bellman equation:

$$\begin{aligned} rV(\tau,k,A|\theta) \quad = \quad & \frac{\theta(\tau)^{1-\sigma}(Ak)^{1-\sigma}}{1-\sigma} + \frac{\partial V}{\partial \tau} + \rho k \frac{\partial V}{\partial k} + \gamma_A A \frac{\partial V}{\partial A} \\ & + \delta[V^U(k,A) - V(\tau,k,A|\theta)] \\ & + \lambda_1 E_j[\max\{V(\tau,k,A|\theta), V(0,k,A|\widetilde{\theta}_j)\} - V(\tau,k,A|\theta)] \\ & + \lambda_q E_j[V(0,k,A|\widetilde{\theta}_j) - V(\tau,k,A|\theta)]. \end{aligned}$$

In words, the flow value of being employed on contract θ equals the flow value of the current wage paid plus the capital gains due to (i) the wage rate paid varying with tenure (picked up by the $\partial V/\partial\tau$ term), (ii) the worker's productivity increases through learning-by-doing (at rate ρ), (iii) aggregate productivity increases (at rate γ_A), (iv) a separation shock occurs at rate δ, (v) a randomly drawn outside offer $\widetilde{\theta}_j$ is received at rate λ_1 and (vi) an exogenous quit occurs at rate λ_q.

The value of being unemployed satisfies

$$\begin{aligned} rV^U(k,A) &= \frac{b^{1-\sigma}(Ak)^{1-\sigma}}{1-\sigma} - \phi k\frac{\partial V^U}{\partial k} + \gamma_A A\frac{\partial V^U}{\partial A} \\ &\quad +\lambda_0 E_j[\max\{V^U(k,A), V(0,k,A|\widetilde{\theta}_j)\} - V^U(k,A)]. \end{aligned}$$

The restriction to a CRRA utility function and a timeless equilibrium imply a critical simplifying property: the value functions are separable in productivity k where

$$\begin{aligned} V(.|\theta) &= (Ak)^{1-\sigma}U(\tau|\theta) \\ V^U(.) &= (Ak)^{1-\sigma}U^U \end{aligned} \tag{1}$$

with $U(\tau|\theta)$, U^U as defined below. $U(.)$ is central to the analysis for it is the measure by which all workers value (or rank) any contract $\theta(.)$ and so determines equilibrium quit turnover. In what follows we refer to $U(\tau|\theta)$ as the value of contract θ (at tenure τ) and U^U as the value of unemployment.

Let $U_0 = U(0|\theta)$ denote the starting value of representative contract $\theta(.)$. As search is random, let $F(\widetilde{U}_0)$ denote the fraction of offered contracts $\{\widetilde{\theta}_j\}$ whose starting value is no greater than $\widetilde{U}_0$. Substituting out $V(.|\theta) = (Ak)^{1-\sigma}U(\tau|\theta)$ and $V^U(.) = (Ak)^{1-\sigma}U^U$ in the Bellman equations above yields the following expressions for $U(.|\theta)$ and U^U :

$$[r+\delta+\lambda_q-[\rho+\gamma_A](1-\sigma)]\,U - \frac{dU}{d\tau} = \frac{\theta(\tau)^{1-\sigma}}{1-\sigma} + \delta U^U + \lambda_1\int_U^{\overline{U}}[1-F(U_0)]dU_0 + \lambda_q\int_{\underline{U}}^{\overline{U}} U_0 dF(U_0) \tag{2}$$

$$(r+\phi(1-\sigma)-\gamma_A(1-\sigma))U^U = \frac{b^{1-\sigma}}{1-\sigma} + \lambda_0\int_{U^U}^{\overline{U}}[U_0-U^U]dF(U_0) \tag{3}$$

which are independent of k, A (as required) To guarantee bounded solutions exist, we assume r satisfies both $r > [\rho+\gamma_A](1-\sigma)$ and $r > (\gamma_A - \phi)(1-\sigma)$. The above expressions now imply Claim 1.

Claim 1: *Optimal job search for any worker k in a timeless equilibrium implies:*
(a) while unemployed, the worker accepts a contract offer $\widetilde{\theta}_j(.)$ if and only if its starting value $U_0 \geq U^U$;
(b) while employed with contract value $U(\tau|\theta)$, the worker accepts a job offer $\widetilde{\theta}_j$ if and only if it offers greater contract value $U(0|\widetilde{\theta}_j) > U(\tau|\theta)$. The worker will quit into unemployment whenever value $U(\tau|\theta) < U^U$.

Claim 1 yields an important corollary. Let $G(U)$ denote the fraction of employed workers who enjoy contract value no greater than U and let $[\underline{U}, \overline{U}]$ denote its support. As Claim 1 implies equilibrium turnover is independent of k in a timeless equilibrium, it further implies for any given worker type k, that the distribution of contract values *across* workers of type k is also $G(.)$; i.e. the distribution of contract values across the entire population is independent of k.

4 Optimal Contracts in a Timeless Equilibrium

Consider now the optimal contract $w_j(\tau, k, A)$ of firm j in a timeless equilibrium. Clearly with no loss of generality any such contract can be rewritten as $w_j(.) = Ak\theta_j(\tau, k, A)$. Consider now a representative hire, where k_0 denotes the worker's human capital when first hired and A_0 the aggregate productivity level at that date. As $k = k_0 e^{\rho\tau}$ and $A = A_0 e^{\gamma_A \tau}$ within the employment spell, there is no loss in generality by further restricting attention to contracts of the form $\theta_j = \theta_j(\tau|k_0, A_0)$. In other words, any contract $w_j(.)$ is equivalent to a wage rate paid $\theta_j(\tau|k_0, A_0)$ which varies with tenure but firm j potentially discriminates contracts across types (k_0, A_0) when hired.

Consider then any such contract $\theta(\tau) = \theta_j(\tau|k_0, A_0)$. The critical insight is that a timeless equilibrium implies any such contract $\theta(.)$ yields corresponding contract value $U(.|\theta)$. If the starting value of this contract $U(0|\theta) < U^U$, the offer is rejected (worker (k_0, A_0) prefers being unemployed) and so this contract makes zero profit. Suppose instead it yields starting value $U(0|\theta) \geq U^U$. If u denotes the steady state unemployment level then, given a random contact with a worker (k_0, A_0), Bayes rule implies

$$\alpha = \frac{\lambda_0 u + \lambda_q(1-u)}{\lambda_0 u + \lambda_1(1-u) + \lambda_q(1-u)}$$

is the probability that the worker is either unemployed or an exogenous quitter. In either case, $U(0|\theta) \geq U^U$ implies the worker accepts the job offer. Instead with complementary probability $1-\alpha$ this worker is employed and Claim 1 implies $G(.)$ describes the distribution of contract values earned by such workers. Hence $\alpha + (1-\alpha)G(U_0)$ with $U_0 = U(0|\theta)$ is the probability this contract offer is accepted.

Suppose the worker accepts the job offer and $U(\tau|\theta)$ is the value of this contract at tenure τ. As $F(.)$ describes the distribution of starting contract values offered by all other firms in a timeless equilibrium, the probability this new hire remains employed by tenure τ is

$$\psi(\tau|\theta) = e^{-\int_0^\tau \{\delta + \lambda_q + \lambda_1[1-F(U(s|\theta))]\}ds}. \tag{4}$$

To determine the contract $\theta(.)$ that maximizes expected profit, we first identify the optimal contract which maximizes expected discounted profit conditional on the worker accepting the job with starting value $U_0 \geq U^U$; i.e. we solve

$$\max_{\theta(.)} \int_0^\infty \psi(\tau|\theta) A_0 k_0 e^{(\rho+\gamma_A - r)\tau}[1-\theta(\tau)]d\tau,$$

subject to $U(0|\theta) = U_0$. As $\psi(.)$ defined by (4) does not depend on (k_0, A_0) then, given choice of starting value U_0, the optimal profit maximising contract is independent of (k_0, A_0) (the optimisation problem is simply multiplicative in $A_0 k_0$). Let $\theta = \theta^*(\tau|U_0)$ denote that optimal contract and let

$$\Pi^*(U_0) = \int_0^\infty \psi(\tau|\theta^*) e^{(\rho+\gamma_A - r)\tau}[1 - \theta^*(\tau|.)]d\tau,$$

which we will refer to as contract profit. As $\alpha + (1-\alpha)G(U_0)$ is the probability this contract offer is accepted then, given contact with (k_0, A_0), an optimal contract yields expected profit,

$$\Omega(U_0|A_0, k_0) = A_0 k_0 \left[\alpha + (1-\alpha)G(U_0)\right] \Pi^*(U_0),$$

and the firm now chooses U_0 to maximise $\Omega(U_0|A_0, k_0)$. As the profit maximisation problem is again simply multiplicative in (k_0, A_0) we have established Claim 2.

Claim 2: *In any timeless equilibrium, it is always optimal to offer contracts $\theta_j(\tau|.)$ which are independent of A_0, k_0.*

Given there is no value to discriminate contract offers by (k_0, A_0), it is consistent with optimality to only consider timeless equilibrium in which each firm offers the same contract $\theta_j(.)$ to all potential hires (k_0, A_0).

Consider now an optimal contract $\theta = \theta(\tau)$ conditional on that contract yielding a starting value $U(0|\theta) = U_0 \geq U^U$. The optimal contract $\theta(.)$ thus solves the program:

$$\Pi^*(U_0) = \max_{\theta(.) \geq 0} \int_0^\infty \psi(\tau|\theta) e^{(\rho+\gamma_A - r)\tau}[1 - \theta(\tau)]d\tau, \tag{5}$$

subject to $U(0|\theta) = U_0$, where $\psi(.)$ is given by (4) and $U(.)$ by (2). For ease of exposition we only consider contracts for which the constraint $\theta(.) \geq 0$ is never binding (we discuss this further below).

Theorem 1: *In a timeless equilibrium, an optimal contract $\theta^*(.|U_0)$ and corresponding worker and firm payoffs U^* and Π^* are solutions to the following dynamical system $\{\theta, U, \Pi\}$ where, at any tenure $\tau \geq 0$,*
(a) $\theta(\tau) > 0$ is given by the implicit function

$$\begin{aligned} & \frac{\theta^{1-\sigma}}{1-\sigma} + \theta^{-\sigma}\left[(1-\theta) - [r + \delta + \lambda_q - \rho - \gamma_A + \lambda_1[1 - F(U)]]\,\Pi\right. \\ = \; & [r + \delta + \lambda_q - [\rho + \gamma_A](1-\sigma)]\,U - \delta U^U - \lambda_1 \int_U^{\overline{U}} [1 - F(U_0)]dU_0 - \lambda_q \int_{\underline{U}}^{\overline{U}} U_0 dF(U_0) \end{aligned} \tag{6}$$

(b) contract profit

$$\Pi(\tau) = \int_\tau^\infty e^{-\int_\tau^s [r+\delta+\lambda_q - \rho - \gamma_A + \lambda_1[1-F(U(t))]dt}[1 - \theta(s)]ds, \tag{7}$$

(c) and contract value U evolves according to the differential equation

$$\frac{dU}{d\tau} = -\theta^{-\sigma}\frac{d\Pi}{d\tau} \tag{8}$$

with initial value $U(0|.) = U_0$.
Proof: See the Appendix.

The structure of the optimal contract is similar to Burdett and Coles (2003). Differentiating (6) and (7) with respect to τ yields the system of differential equations for $\{\theta, \Pi, U\}$:

$$\dot{\theta} = \frac{\lambda_1\left[\theta^{1-\sigma}\right]}{\sigma}F'(U)\Pi - (\rho+\gamma_A)\theta, \tag{9}$$

$$\dot{\Pi} = [r+\delta+\lambda_q-\rho-\gamma_A+\lambda_1[1-F(U)]]\Pi-(1-\theta), \tag{10}$$

with $\dot{U}$ given by (8).

Equation (9) describes how the wage rate paid changes with tenure, where the actual wage paid is $w(\tau) = A_0k_0e^{(\rho+\gamma_A)\tau}\theta(\tau)$. Thus along the optimal contract, (9) implies the wage paid changes as:

$$\left[\frac{-u''}{u'^2}\right]\frac{dw}{d\tau} = \left[A_0k_0e^{(\rho+\gamma_A)\tau}\right]^{\sigma}\lambda_1F'(U)\Pi. \tag{11}$$

A quit implies the firm loses continuation profit Π. While $\Pi > 0$, (11) implies the wage paid increases within the employment spell, where $F'(U)$ measures the number of firms whose outside offer will marginally attract this worker. If $F'(U) = 0$ then marginally raising the wage paid at tenure τ has no impact on the worker's quit rate and optimal consumption smoothing then implies the firm pays a (locally) constant wage. If $F'(U) > 0$, however, a slightly higher wage results in a slightly lower marginal quit rate and it is optimal for the firm to increase the wage paid. The scaling term $\left[A_0k_0e^{(\rho+\gamma_A)\tau}\right]$ arises as the worker's value of employment at tenure τ is $V(\tau,.) = [A_0k_0e^{(\rho+\gamma_A)\tau}]^{1-\sigma}U(\tau|\theta)$ while the firm's continuation profit is $\left[A_0k_0e^{(\rho+\gamma_A)\tau}\right]\Pi(\tau)$. (11) thus describes the optimal trade-off in wage space. As workers compare contracts by contract value $U(.)$, however, Theorem 1 describes the choice-relevant objects. Most importantly conditional on any $U_0 \geq U^U$, Theorem 1 describes the optimal contract for all worker types (A_0, k_0) in any timeless equilibrium.

A constant wage (perfect consumption smoothing) implies $\theta(\tau)$ declines at rate $\rho+\gamma_A$. Thus although an optimal contract implies wages paid always increase within an employment spell, it is no longer the case that tenure effects are necessarily positive. Let $(\theta^{\infty}, \Pi^{\infty}, U^{\infty})$ denote the stationary point of this dynamical system. Figure 1 illustrates the possible set of optimal contracts $\theta^*(.)$.

Consider first the optimal contract for the firm offering the least generous contract, i.e. one which yields starting value $U_0 = \underline{U} \geq U^U$ and suppose $\underline{U} < U^{\infty}$. As wage rate $\theta(.)$ increases with tenure, contract value $U(.)$ also increases with tenure and so $U(\tau|.)$ converges to U^{∞} from below. Let $\theta_1(\tau)$ denote this optimal contract which we define as the lower baseline scale. Consider instead the optimal contract offered by firms offering the most generous contract $U_0 = \overline{U}$ and

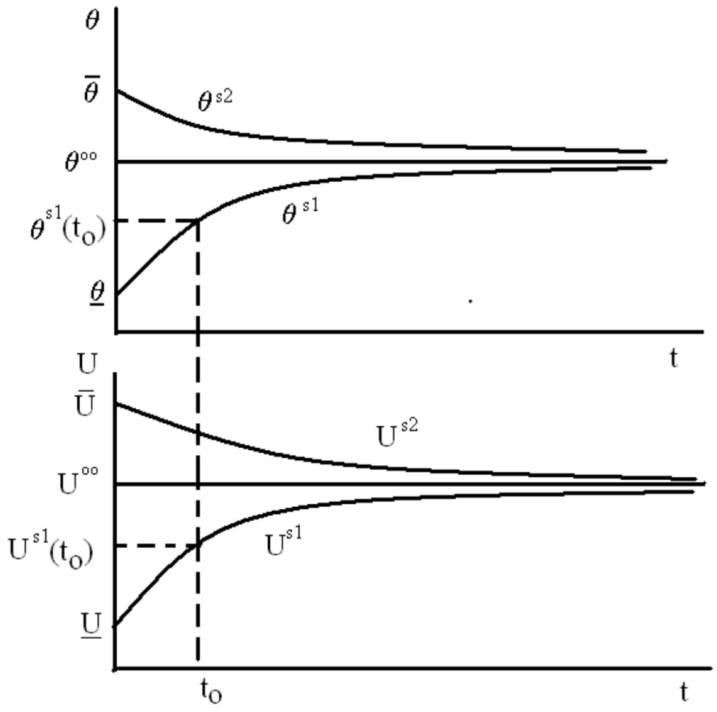


Figure 1: Possible Baseline Salary Scales

suppose $\overline{U} > U^{\infty}$. Although the wage paid increases within the employment spell, $\theta(.)$ decreases with tenure. Contract value thus falls with tenure and so $U(\tau|.)$ converges to U^{∞} from above. Let $\theta_2(\tau)$ denote this optimal contract which we refer to as the upper baseline scale.

Consider now a firm that offers a contract which yields an initial piece rate value U_0 such that $\underline{U} < U_0 < U^{\infty}$. As depicted in Figure 1, define t_0 as the point on the lower baseline scale where $U_1(t_0) = U_0$. Optimality of the lower baseline scale yields an important simplification: the optimal contract yielding U_0 is the continuation contract starting at point t_0 on the lower baseline scale; i.e., the optimal contract $\theta^*(\tau|U_0)$ is $\theta_1(t_0+\tau)$ where the wage rate paid at tenure τ corresponds to point $(t_0+\tau)$ on the lower baseline scale. We denote its corresponding contract profit as $\Pi_1(t_0)$. Suppose instead $U^{\infty} < U_0 < \overline{U}$. This time the optimal contract yielding U_0 is the continuation contract starting at point t_0 on the upper baseline scale where $U_2(t_0) = U_0$ and so yields profit $\Pi_2(t_0)$. It is this baseline property of the optimal contract structure which makes tractable the complete characterisation of timeless equilibria.

4.1 Characterisation and Existence of Timeless Equilibria.

In any timeless equilibrium and given contact with any worker (k_0, A_0), an optimal contract corresponds to a wage rate contract $\theta(.)$ with starting point t_0 on either the upper or lower baseline scale $i = 1, 2$. Corresponding to any such starting point is a starting contract value $U_i(t_0)$ which, if accepted, generates expected profit $A_0 k_0 \Pi_i(t_0)$. Thus any such contract offer generates expected profit

$$\Omega_i(t_0|A_0, k_0) = A_0 k_0 \left[\alpha + (1-\alpha)G(U_i(t_0))\right] \Pi_i(t_0)$$

per worker contact. As expected profit is simply proportional to $k_0 A_0$, identifying a timeless equilibrium reduces to solving the constant profit condition:

$$\left[\frac{\lambda_0 u + \lambda_q(1-u) + \lambda_1(1-u)G(U_i(t_0))}{\lambda_0 u + \lambda_q(1-u) + \lambda_1(1-u)}\right]\Pi_i(t_0) \;=\; \overline{\Omega} \; > 0 \text{ if } dF(U_i(t_0)) > 0$$
$$\left[\frac{\lambda_0 u + \lambda_q(1-u) + \lambda_1(1-u)G(U_i(t_0))}{\lambda_0 u + \lambda_q(1-u) + \lambda_1(1-u)}\right]\Pi_i(t_0) \;\leq\; \overline{\Omega} \text{ otherwise.}$$

Claim 3: *Suppose $dF(U_i(t_0))) > 0$. A timeless equilibrium implies*

$$\Pi = \frac{\sqrt{[r-\rho-\gamma_A]^2\overline{\Omega}^2 + 4[\delta+\lambda_q]\overline{\Omega}(1-\theta)} - [r-\rho-\gamma_A]\overline{\Omega}}{2[\delta+\lambda_q]} \tag{12}$$

where $\Pi = \Pi_i(t_0)$ *and* $\theta = \theta_i(t_0)$.
Proof: See the Appendix.

The standard recursive contracting approach (e.g. Spear and Srinivastan, 1987) supposes a contract "promises" continuation value U to an employer and then identifies $\theta = \theta(U)$ as the optimal wage rate paid with corresponding contract profit $\Pi = \Pi(U)$. Claim 3, however, identifies a much more convenient approach: we instead let $U = \widehat{U}(\theta)$ describe the contract value enjoyed by a worker when the optimal contract pays θ, and $\widehat{\Pi}(\theta)$ describes the firm's corresponding contract profit. (12), of course, identifies $\Pi = \widehat{\Pi}(\theta)$. Furthermore (8) in Theorem 1 implies $\widehat{U}(\theta)$ solves

$$\frac{d\widehat{U}}{d\theta} = -\theta^{-\sigma}\frac{d\widehat{\Pi}}{d\theta}, \tag{13}$$

and so we are very nearly done.

We thus transform the analysis from the time domain [how wage rates vary with tenure] to the domain of wage rates paid $\theta \in [\underline{\theta}, \overline{\theta}]$. Let $F_\theta(\theta)$ denote the distribution of starting wage rates paid by firms. As (12) and (13) imply $\widehat{U}(.)$ is strictly increasing, the definition of $F(.)$ implies :

$$F_\theta(\theta) = F(\widehat{U}(\theta)) \text{ for } \theta \in [\underline{\theta}, \overline{\theta}].$$

Let $G_\theta(\theta)$ denote the distribution of wage rates paid across employed workers and so

$$G_\theta(\theta) = G(\widehat{U}(\theta)).$$

The constant profit condition is thus equivalent to

$$\left[\frac{\lambda_0 u + \lambda_q(1-u) + \lambda_1(1-u)G_\theta(\theta)}{\lambda_0 u + \lambda_q(1-u) + \lambda_1(1-u)}\right]\widehat{\Pi}(\theta) \;=\; \overline{\Omega} \; > 0 \text{ if } dF_\theta(\theta) > 0 \tag{14}$$
$$\left[\frac{\lambda_0 u + \lambda_q(1-u) + \lambda_1(1-u)G_\theta(\theta)}{\lambda_0 u + \lambda_q(1-u) + \lambda_1(1-u)}\right]\widehat{\Pi}(\theta) \;\leq\; \overline{\Omega} \text{ otherwise.}$$

As an optimal contract implies the worker never quits into unemployment and strict positive profit implies all firms offer starting contracts which are strictly preferred to being unemployed [otherwise the firm makes zero profit], steady state unemployment is given by $u = \delta/(\delta + \lambda_0)$.

We first solve for $\overline{\Omega}$. Let $\overline{U} = \widehat{U}(\overline{\theta})$ denote the highest value enjoyed by workers. A simple contradiction argument establishes $G(\overline{\theta}) = 1$ and so putting $\theta = \overline{\theta}$ in (14) finds $\widehat{\Pi}(\overline{\theta}) = \overline{\Omega}$. Putting $\theta = \overline{\theta}$ and substituting out $\widehat{\Pi}(\overline{\theta}) = \overline{\Omega}$ in (12) now determines

$$\overline{\Omega} = \frac{1-\overline{\theta}}{\delta + \lambda_q + r - \rho - \gamma_A}. \tag{15}$$

Thus (12) and (15) imply (16) described in Theorem 2 below. The following rules out the existence of the upper baseline scale.

Claim 4: *A timeless equilibrium implies* $\theta^{\infty} = \overline{\theta}$.
Proof: As $\widehat{\Pi}(\overline{\theta}) = \overline{\Omega} = \frac{1-\overline{\theta}}{\delta+\lambda_q+r-\rho-\gamma_A}$, (10) implies $\dot{\Pi} = 0$ at $\theta = \overline{\theta}$. Hence $\overline{\theta}$ is a stationary point of the differential equation system implied by Theorem 1.

Conditional on an equilibrium value for $\overline{\theta}$ [determined by Theorem 3 below] backward induction on $\theta \leq \overline{\theta}$ now fully characterises a timeless equilibrium.

Theorem 2: *For $\theta \leq \overline{\theta}$ and while $F'_\theta > 0$* a timeless equilibrium implies $\left\{\widehat{\Pi}, \widehat{U}, G_\theta, F_\theta\right\}$ satisfy:

$$\widehat{\Pi}(\theta) = \frac{1-\overline{\theta}}{2[\delta+\lambda_q]}\left[\sqrt{\left[\frac{r-\rho-\gamma_A}{\delta+\lambda_q+r-\rho-\gamma_A}\right]^2 + \frac{4[\delta+\lambda_q]}{\delta+\lambda_q+r-\rho-\gamma_A}\frac{1-\theta}{1-\overline{\theta}}} - \frac{r-\rho-\gamma_A}{\delta+\lambda_q+r-\rho-\gamma_A}\right] \tag{16}$$

$$\widehat{U}(\theta) = \overline{U} - \int_\theta^{\overline{\theta}} \frac{[\theta']^{-\sigma}}{\left[[r-\rho-\gamma_A]^2 + 4[\delta+\lambda_q]\left[\delta+\lambda_q+r-\rho-\gamma_A\right]\frac{1-\theta'}{1-\overline{\theta}}\right]^{1/2}} d\theta' \tag{17}$$

$$\lambda_0 u + \lambda_q(1-u) + \lambda_1(1-u)G_\theta(\theta) = \frac{\left[1-\overline{\theta}\right]\left[\lambda_0 u + \lambda_q(1-u) + \lambda_1(1-u)\right]}{\left[\delta+\lambda_q+r-\rho-\gamma_A\right]\widehat{\Pi}(\theta)} \tag{18}$$

$$1 - F_\theta = \theta^\sigma \int_\theta^{\overline{\theta}} \sigma[\frac{1}{\theta'}]^{\sigma+1}\Psi(\theta')d\theta', \tag{19}$$

where

$$\Psi(\theta) = \frac{\lambda_q+\delta}{\lambda_1}\left[\frac{\widehat{\Pi}(\theta)-\overline{\Omega}}{\overline{\Omega}}\right] - \left[\frac{(\rho+\gamma_A)}{\lambda_1}\frac{\theta d\widehat{\Pi}/d\theta}{\widehat{\Pi}(\theta)}\right] > 0.$$

Proof: See the Appendix.

All that remains is to determine equilibrium $\overline{\theta}$. Claim 5 establishes the relevant boundary condition.

Claim 5: *A timeless equilibrium implies $\underline{U} = \widehat{U}(\underline{\theta}) = U_U$ where:*

$$1 - \underline{\theta} = \frac{(\delta+\lambda_q+\lambda_1)\left[\delta+\lambda_q+\lambda_1+r-\rho-\gamma_A\right]}{(\delta+\lambda_q)[\delta+\lambda_q+r-\rho-\gamma_A]}\left[1-\overline{\theta}\right]. \tag{20}$$

Proof: Standard contradiction arguments establish $\underline{U} = U_U$ and $G_\theta(\underline{\theta}) = 0$. Putting $\theta = \underline{\theta}$ in the constant profit condition (14), with $\widehat{\Pi}(\underline{\theta})$ given by (16), $\overline{\Omega}$ given by (15) and $u = \delta/(\delta + \lambda_0)$ yields the result.

To identify a timeless equilibrium, fix a candidate equilibrium value for $\overline{\theta}$ in the range

$$\overline{\theta} \in (1 - \frac{(\delta+\lambda_q)[\delta+\lambda_q+r-\rho-\gamma_A]}{(\delta+\lambda_q+\lambda_1)\,[\delta+\lambda_q+\lambda_1+r-\rho-\gamma_A]}, 1). \tag{21}$$

Such a candidate value implies strictly positive profit ($\overline{\Omega} > 0$) and $\underline{\theta} > 0$ (strictly positive wage rates). Given this candidate choice of $\overline{\theta}$, let $\widetilde{F}_\theta(.|\overline{\theta})$ denote the unique candidate distribution function F_θ implied by (19) in Theorem 2. Given the implied distribution of contract offers, Claim 6 now identifies the implied values of $\underline{U}$ and U^U at $\underline{\theta}$, which we denote $\widetilde{\underline{U}}(\overline{\theta})$, $\widetilde{U}^U(\overline{\theta})$ respectively.

Claim 6: *Given $\overline{\theta}$ and the implied candidate distribution function $\widetilde{F}_\theta$ then:*

$$[r-[\rho+\gamma_A](1-\sigma)]\,\widetilde{\underline{U}} = \frac{\overline{\theta}^{1-\sigma}}{1-\sigma} - \int_{\underline{\theta}}^{\overline{\theta}} \frac{\left[r+\delta-[\rho+\gamma_A](1-\sigma)+\lambda_q\widetilde{F}_\theta(\theta|.)\right]\theta^{-\sigma}d\theta}{\left[[r-\rho-\gamma_A]^2+4[\delta+\lambda_q]\,[\delta+\lambda_q+r-\rho-\gamma_A]\,\frac{1-\theta}{1-\overline{\theta}}\right]^{1/2}}. \tag{22}$$

$$[r+\phi(1-\sigma)-\gamma_A(1-\sigma)]\widetilde{U}^U = \frac{b^{1-\sigma}}{1-\sigma}+\lambda_0\int_{\underline{\theta}}^{\overline{\theta}} \frac{[1-\widetilde{F}_\theta(\theta|.)]\theta^{-\sigma}d\theta}{\left[[r-\rho-\gamma_A]^2+4[\delta+\lambda_q]\,[\delta+\lambda_q+r-\rho-\gamma_A]\,\frac{1-\theta}{1-\overline{\theta}}\right]^{1/2}}. \tag{23}$$

Proof: See the Appendix.

Identifying a timeless equilibrium reduces to finding a $\overline{\theta}$ such that $\widetilde{\underline{U}}(\overline{\theta}) = \widetilde{U}^U(\overline{\theta})$.

Theorem 3: *For $\sigma > 1$, a $\overline{\theta}$ satisfying (21) exists such that $\widehat{U}(\underline{\theta}) = U^U$.*
Proof: See the Appendix.

If F_θ identified by (19) is a positive increasing function (i.e. has the properties of a distribution function), then Theorems 2 and 3 fully characterise timeless equilibria. By construction, all optimal contracts which offer $\theta \in [\underline{\theta}, \overline{\theta}]$ yield the same expected profit $\overline{\Omega} > 0$. Consider then any deviating contract. Clearly, a suboptimal contract which offers $U_0 \in [\widehat{U}(\underline{\theta}), \widehat{U}(\overline{\theta})]$ yields lower profit. Further as $U^U = \widehat{U}(\underline{\theta})$, any contract which offers value $U_0 < \widehat{U}(\underline{\theta})$ yields zero profit as all workers reject such an offer. Finally any contract which offers $U_0 > \widehat{U}(\overline{\theta})$ attracts no more workers than the optimal contract which offers $U_0 = \widehat{U}(\overline{\theta})$ while the latter contract earns strictly greater profit per hire. As no deviating contracts exist which yield greater profit, Theorems 2 and 3 identify timeless equilibria.

Theorem 2 describes all equilibrium objects apart from the (lower) baseline scale. Equations

(9) and (13) imply that scale is identified by the initial value problem:

$$\dot{\theta} = \frac{\lambda_1 \left[\theta^{1-\sigma}\right]}{\sigma} \frac{F'_\theta \widehat{\Pi}}{d\widehat{U}/d\theta} - (\rho + \gamma_A)\theta$$

with $\theta(0) = \underline{\theta}$, where the proof of Theorem 2 establishes:

$$\frac{\theta F'_\theta}{\sigma} = \Psi - (1 - F_\theta).$$

In Appendix B we describe the algorithm to compute timeless equilibria. Note that the $\theta \geq 0$ constraint may bind if $\sigma < 1$ and b sufficiently small. For example suppose $\lambda_0 = \lambda_1$: as experience is valuable, a worker will accept a lower starting wage rate $\theta < b$ and thus $\theta \geq 0$ binds if $b = 0$. Whenever this occurs, the baseline scale pays a zero wage rate for tenures $\tau \leq \overline{\tau}$ and a positive (increasing) wage rate thereafter. We find this constraint never binds in our quantitative analysis.

5 Quantitative Analysis

We estimate the model using indirect inference (see Gourieroux, Monfort, and Renault, 1993, also Bagger et al., 2014, for recent related work and Appendix B for a full description of our approach). An important feature of the data, however, is that high skilled and low skilled workers exhibit very different turnover patterns. We therefore suppose these types participate in separate markets; i.e. we assume two labour markets, one for low skilled workers and one for high skilled workers, and estimate the model separately for each type. We shall show the distinction plays an important role when trying to identify the costs of job loss.

5.1 Data

We use information contained in the British Household Panel Survey (BHPS) and the Labour Force Survey (LFS). The BHPS is an annual survey of individuals, age 16 years or more, in a nationally representative sample of about 5,500 households. Approximately 10,000 individuals are interviewed each year. It started in 1991 and was subsumed by the new and bigger survey "Understanding Society" in 2010. The BHPS contains socioeconomic information, including information about household organization, the labour market, income and wealth, housing, health and socioeconomic values. Using this information one is able to reconstruct the labour market histories individuals since leaving full-time education. Using Maré (2006) we derive consistent histories that summarize individual's labour market histories; and several socio economic characteristics that are standard in household survey data. This data set allows us to estimate workers' average wage-experience profiles and measures of wage dispersion. For these exercises, we consider real hourly (gross) wages and trim the wage distribution by 5 percent on each side to reduce measurement error and to consider all jobs that pay above the national minimum wage, introduced in the UK in 1999.[6]

The LFS is a quarterly survey of individuals, aged 16 years or more. It has a rotating panel structure, in which individuals that live on the sampled address are followed for a maximum of 5 quarters, also referred to as waves. Each quarter, one-fifth of the sample of addresses is replaced by an incoming rotation group. In each wave, the respondents provide information about, among other things, their labour market status. We use the two-quarter longitudinal sample of the LFS to derive the transitions rates between different employment status and use them as part of our calibration procedure. Relative to the BHPS, the LFS gives more accurate estimates of these empirical transitions rates as it offers higher frequency information and consists of about 60,000 individuals each quarter.

For both data sets information on white male workers for the period 1991-2004 is used. The data is stratified into two educational or skill groups. We consider workers to be low skilled if they reported having no qualification, other qualifications, apprenticeship, CSE, commercial qualifications or no O-levels. We consider workers to be high skilled if they reported having achieved O-levels, A-levels, nursing qualifications, teaching qualifications, university degree or higher and other higher qualifications. As the model, for tractability, assumes that learning-by-doing accrues at a constant rate, we focus on young workers as this assumption may be less reasonable across a worker's entire lifetime.[7] We further restrict attention to paid (dependent) full-time employment spells in the private sector and unemployment spells that lasted at least one month. Keeping the sample as homogeneous as possible only those employment and non-employment spells are considered 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. Individuals that re-entered full-time education or had a spell in government training were dropped.

5.2 Estimation Procedure

The reference period of time is a month and we set $r = 0.005$ (an annual discount rate of 6%). We set $\gamma_A = 0.00183$ to match the estimated slope of a linear trend on output per worker in the UK over the relevant time period.[8] This parametrisation leaves us with a vector $\Lambda = \{\delta, \lambda_0, \lambda_1, \lambda_q, \rho, \phi, \sigma, b\}$, consisting of 8 parameters that we jointly recover by minimizing the sum of squared (percentage) difference between a set of simulated moments from the model and their counterparts in the data (see Appendix B for full details).

We target 14 statistics based on the main characteristics of the labour market to which the model is directly related. Table 1 describes those statistics for each skill group. The average

[6] Following Dustmann and Pereira (2008), we construct real hourly wages by dividing monthly (gross) earnings by 4.33 weeks and then by the average number of hours worked in a week in full-time jobs. We also take into account overtime hours and use the CPI to deflate nominal wages.

[7] To account for difference in age of graduation, we consider those low skilled workers that have 16 and 30 years of age, and those high skilled who have between 22 and 36 years of age.

[8] We estimate the slope of the linear trend through OLS, by regressing the log of quarterly output per worker on quarterly dummies and a linear trend. Output per worker is obtained by dividing total output (GDP) over total employment.

duration of spells provides direct information for $(\lambda_0, \lambda_1, \delta, \lambda_q)$.[9] With optimal contracting, (11) implies wages evolve within an employment spell according to

$$\left[\frac{-u''}{u'^2}\right]\frac{dw}{d\tau} = \left[A_0 k_0 e^{(\rho+\gamma_A)\tau}\right]^{\sigma} \lambda_1 F'(U)\Pi.$$

Wages increase within the spell for the worker is becoming more productive through experience (learning-by-doing implies $k = k_0 e^{\rho\tau}$), outside competition induces the firm to raise wages paid as aggregate productivity increases ($A = A_0 e^{\gamma_A \tau}$) and equilibrium tenure effects are strictly positive. In an optimal contract, the magnitude of these wage effects depend on the degree of risk aversion σ. To identify the returns to experience and to tenure, we estimate a standard Mincer OLS wage regression as an auxiliary equation. Indirect inference requires that the same estimation procedure, but on the simulated data, yields equivalent parameter estimates. As there is equilibrium non-degenerate wage dispersion we also use as identifying information (i) the estimated ratio of re-employment wages to average wage and (ii) the Mm ratio (using the methodology proposed in Hornstein et al., 2007). The former statistic measures the short-term wage loss of displaced workers (see also Farber, 1997, and Manning, 2003), the latter requires the reservation wage rate (identified as the lowest wage rate paid) is consistent with observed wage dispersion and frictions (for full details on the importance of this statistic see Hornstein et al., 2011). The model is thus overidentified with 14 targets and 8 parameter values.

Table 1: Targeted Moments

Moments	Low Skilled		High Skilled	
	Data	Model	Data	Model
Average transitions / duration				
Non-employment spell (months)	12.67	11.91	8.74	8.52
Employment spell (years)	3.78	3.78	12.60	12.63
Job spell (years)	3.10	2.89	5.03	5.43
Invol/vol EE transitions	0.57	0.58	0.50	0.54
OLS returns to experience (%)				
2 years	4.47	3.86	4.58	4.21
4 years	8.62	7.82	8.75	8.40
6 years	12.47	11.87	12.49	12.57
8 years	16.00	16.02	15.81	16.73
10 years	19.23	20.26	18.71	20.86
OLS returns to tenure (%)				
2 years	4.19	4.62	2.27	2.46
4 years	7.29	8.64	3.84	4.65
6 years	9.31	12.03	4.70	6.58
Wage dispersion				
Mm ratio	1.62	1.63	1.62	1.63
re-emp wage / mean wage	0.91	0.92	0.84	0.85

[9]The LFS provides information that allows us to classify whether a employer-to-employer transition was voluntary or not. This information useful as it helps identify λ_q relative to λ_1. In Appendix B we detail how we classified voluntary and involuntary employer-to-employer transitions.

5.3 Model Fit

Table 1 shows the fit of the model is very good and captures the main features of worker early careers in the labour market (though tenure effects would seem too strong at high tenures; i.e. there is not enough curvature in the firm specific tenure effects). The model does an excellent job in fitting both the Mm ratio and the short-term impact of job losses on wages.

Table 2 describes the estimated parameters for each skill group. Reflecting the duration information provided in Table 1, low skilled workers experience unemployment spells much more often and take longer to find work than do high skilled workers. The corresponding low skill unemployment rate, u, is consequently much higher (21.8% rather than 5.5%).

The estimated return to experience, i.e. the rate of learning-by-doing ρ, is the same for both groups and equals 4.2% per annum. Note this estimated return to experience is approximately twice those identified by the Mincer wage equations in Table 1, which instead suggest a return to experience of around 2.2% in both skill groups. One important reason for this gap is that, because of data limitations, the typical Mincer approach (as adopted here) uses potential experience to proxy for actual experience. Unfortunately when there is significant skill loss while unemployed (and unemployment spells are long) potential experience becomes a poor proxy for time spent learning-by-doing while employed. In a stationary environment, the average growth rate of skills by potential experience is instead $\gamma_k \equiv (1-u)\rho - u\phi$. For the low skill group Table 2 parameters imply $\gamma_k = 2.3\%$ $p.a.$ which is remarkably close to the OLS Mincer estimate of 2.2% for this group. As u and ϕ are relatively large for the low skill group, however, this statistic far underestimates the actual return to experience ρ.

For the high skill sector, Table 2 instead implies $\gamma_k = 3.2\% p.a.$ which is appreciably higher than the OLS Mincer estimates. The important insight for what follows is that the high skill sector has high (endogenous) quit rates and large job ladder effects, effects which are not properly controlled for by a simple OLS Mincer wage equation. Of course indirect inference requires the same OLS Mincer wage regression but estimated on model-based data must yield the same low estimated return to experience. Table 1 shows the match is very good.[10]

The estimated relative risk aversion parameter implies $\sigma = 2.58$ for high skilled workers and is close to standardly used values. For low skilled workers, the estimated risk aversion parameter is appreciably higher at $\sigma = 4.02$, though still not unreasonable.[11] Figure 2 describes the resulting baseline scales. Note that low skilled workers are paid lower wage rates and enjoy smaller wage tenure effects. This occurs for two reasons. First low skilled workers are unlikely to receive outside offers and so the incentive for firms to backload wages is relatively weak. Second their inferred degree of relative risk aversion is large and the imputed welfare loss through a very low starting wage is appreciably higher.

It is perhaps surprising, then, that the estimated Mincer wage equation in Table 1, both

[10] The OLS estimates clearly suggest that learning-by-doing begins to decline after around 8 years and so, given the model cannot consider declining rates of learning-by-doing by experience, we do not consider the returns to experience beyond year 10. Nevertheless for skilled workers the framework infers a higher average rate of learning-by-doing than is suggested by simple OLS regressions.

[11] Gandelman and Hernandez-Murillo (2015) survey some of the literature and find that estimates of relative risk aversion vary widely, going from from 0.2 to 10. They argue that the accepted range for this parameter seems to be between 1 and 3.

Table 2: Parameters

	Low Skilled	High Skilled
Parameters		
δ	0.0215	0.0066
λ_0	0.0768	0.1108
λ_1	0.0096	0.0285
λ_q	0.0021	0.0030
ρ	0.0035	0.0036
ϕ	0.0022	0.0016
σ	4.015	2.575
b	0.3252	0.5326
Endog. variables		
$F(\underline{\theta})$	0.071	0.16
u	0.218	0.055
γ_k	0.0023	0.0032

on actual data and the simulated data, suggest that low skilled workers enjoy higher returns to tenure. This result occurs because job ladder effects are very different in the two skill sectors. Table 2 implies high skilled workers enjoy long expected employment spells (12.7 years) and outside job offers arrive relatively frequently (roughly once every three years). Thus job-to-job quits are a relatively frequent occurrence in the high skill sector. Conversely low skilled workers have much shorter expected employment spells (3.8 years) and are twice as likely to be laid-off than receive an outside offer. Job-to-job quits in this sector are comparatively rare. Reflecting the poor opportunities for low skilled workers to climb the job ladder, the model finds that two thirds of low skilled workers are predominantly employed at low points on the baseline scale (described by the square brackets in Figure 2).[12] Furthermore over this region the slope of the baseline scale is reasonably linear and the marginal return to tenure along the baseline scale is approximately 2.1%*p.a.* This implied return to tenure is strikingly close to the OLS Mincer estimate of the average return to tenure reported in Table 1.

Turnover in the high skill sector is instead very different. Coming out of unemployment, a high skilled worker is most likely to start the employment phase on a low starting wage where Table 2 suggests 16% of firms offer starting wage $\theta = \underline{\theta}$. Workers employed on such low starting wages enjoy steep marginal returns to tenure. But in a steady state, most currently employed high skilled workers have long employment durations and, for them, the early low wage phase is a distant memory. In steady state, we find that two thirds of high skilled workers are employed at relatively high points on the baseline scale (described by the square brackets in Figure 2). Furthermore over this region, the marginal return to tenure falls quickly with tenure. For the high skill sector, estimating the (average) return to tenure does not adequately reflect the underlying job ladder structure.

[12] The squared brackets describe a symmetric interval such that one sixth of the employed workers earn wages rates to the left of the interval and one sixth of employed workers earn wages rates to the right of the interval.

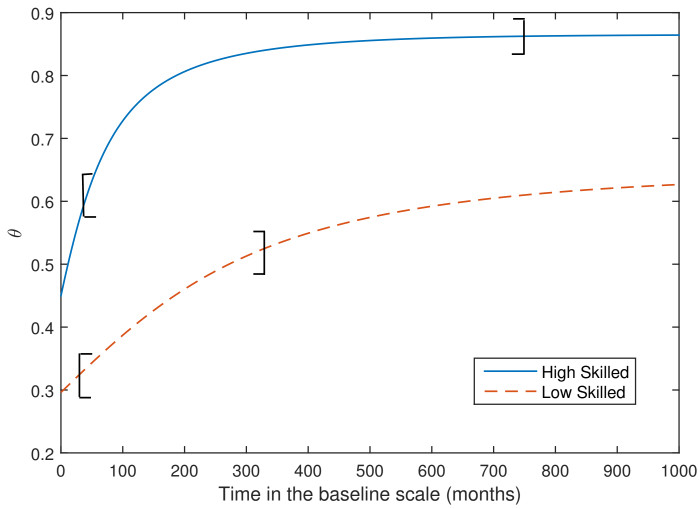


Figure 2: Baseline Salary Scales

5.4 Decomposing the Cost of Job Loss

Job loss affects future wages in two ways:

(i) job ladder effects: by being laid-off, the worker must search anew for employment and re-climb the job ladder, and

(ii) foregone human capital accumulation while unemployed implies a permanent fall in future human capital.

As the model implies $\log w = \log \theta + \log k$, we now identify the permanent and temporary cost of job displacement on future log wages.

Permanent Loss in Human Capital through Displacement Consider two workers at date $t = 0$, one is employed with current human capital k_0, the other has just been displaced and so is unemployed but with the same k_0. Their human capital levels subsequently evolve according to a stationary (geometric) Markov process. Let $G_t^u(k)$ denote the distribution of skills of the initially unemployed worker by date t, and $G_t^e(k)$ those of the initially employed worker. The expected permanent loss in human capital due to job displacement is thus

$$\Phi^P = \lim_{t\to\infty} \left\{ E_t^e \left[\log k\right] - E_t^u[\log k] \right\},$$

where the expectations operators use the respective probability distributions $G_t^e(.)$, $G_t^u(.)$.

Proposition 1: *A timeless equilibrium implies*

$$\Phi^P = \frac{\rho + \phi}{\lambda_0 + \delta}.$$

Proof: See the Appendix.

The permanent loss in human capital through job displacement depends on foregone skill accumulation rates $\rho + \phi$ while unemployed and on the turnover parameters. Not surprisingly if finding work is fast, λ_0 large, then skill loss through becoming unemployed is small. Less obviously, however, if job loss rates δ are very high, then skill loss through displacement *relative to the control group of those workers currently employed* is also small, for currently employed workers are also very likely to become unemployed in the near future.[13] Using the parameter values in Table 2, Table 3 below shows the Φ^P for each skill type.

Temporary Wage Loss through Displacement On the assumption that job separation shocks are random, a displaced employed worker who had previously enjoyed wage rate $\theta \sim G_\theta$, must search anew for a job. Random job search implies starting wage rate $\theta \sim F_\theta$ on re-employment. Subsequent promotion effects and on-the-job search imply the worker enjoys higher wage rates in the future. Of course, in the ergodic limit as $t \to \infty$, the worker again enjoys wage rate paid $\theta \sim G_\theta$. We thus define the temporary fall in (log) wages due to job loss as

$$\Phi^T = [E_G \log[\theta] - E_F [\log[\theta]],$$

where the expectations operators use the respective probability distributions G_θ, F_θ. Using the parameters reported in Table 2, Table 3 shows the Φ^T for each skill type.

Table 3: The Cost of Job Loss - Model

	Low Skill	High Skill
Temporary Wage Loss Φ^T	4.6%	19.1%
Permanent Wage Loss Φ^P	5.8%	4.4%

We discuss these estimates in detail in the next section.

6 The Cost of Job Loss - A Test

Jacobson et al. (1993) describes a (reduced form) statistical approach to measure the cost of job displacement on future wages. That methodology provides a direct test of the model: does that statistical methodology on the simulated data generate the same estimates of the cost of job loss as obtained on the original data? The following not only finds the fit is very good, we identify an important composition bias that arises when using the reduced form methodology.

The Jacobson et al. (1993) approach selects a given year y and considers those who were displaced into unemployment in that year y. Wage losses due to job separation are estimated using the following distributed-lag wage equation:

$$log\ w_{it} = \alpha_i + \gamma_t + \beta X_{it} + \sum_{j=-T_s}^{T_e} \varepsilon_j D_{ij} + \mu_{it}, \tag{24}$$

[13]This result is consistent with the finding of Stevens (1997) for the US, where she finds a significant role for subsequent job loss of the displaced population in explaining average long-term wage losses.

where we use the log of real hourly wages as the dependent variable and the D_{ij} are a set of dummy variables which take value 1 if worker i was displaced in year y and $j = t - y$. Thus the estimated parameter values ε_j describe the displaced worker's average loss of wages j years after displacement relative to those who were not displaced in year y. The wage equation includes a worker fixed effect α_i, year dummies γ_t, X_{it} is a quadratic on potential experience and μ_{it} is assumed a white noise error term. We specify (24) in log wages so that the results are directly comparable with the theory.

We estimate the above equation on the BHPS sample described in Section 5.1 for the period 1991-2004 with $T_s = 2$ and $T_e = 6$. As displacement is a relatively rare occurrence, we increase the number of displacement events by considering three sets of displaced workers where, to avoid double counting: (i) workers who were displaced in year 1995 have $y = 1995$; (ii) workers who were not displaced in year 1995 but were displaced in year 1996 have $y = 1996$; (iii) workers who were not displaced in years 1995, 1996 but were displaced in year 1997 have $y = 1997$. We consider displaced workers as those employed workers who transited into non-employment.[14] The sample pools together both high and low skilled workers and (24) is estimated with worker fixed effects (see also Stevens, 1997, Kletzer and Fairlie, 2003, and Jarosh, 2015, among others). Table 4 reports the corresponding results.[15]

Table 4 describes the parameter estimates $\widehat{\varepsilon_j}$ and standard errors for three cases: (i) pooled sample and (ii) estimated separately for each skill group (high and low education). The results on the pooled sample exhibit two main features: (i) young workers have large and persistent displacement wage losses and (ii) the size of these losses do not seem to recover over time. When estimated separately by skill group, Table 4 shows that high skilled workers suffer a large *wage* loss immediately following displacement ($\widehat{\varepsilon_1} = -14\%$). There is evidence of a large recovery in wage loss for intermediate j while, at long j, the wage loss seems to again increase with j (though the estimated standard errors are also large). Instead for low skilled workers, the post-displacement point estimates for $\widehat{\varepsilon_j}$ are typically negative and much smaller in absolute magnitude to those estimated for high skilled workers. The estimated standard errors, unfortunately, are large and so these estimates are not significantly negative.

We now describe the results obtained when using the same estimation procedure but instead with simulated (model-based) data with the parameters values described in Table 2. For the

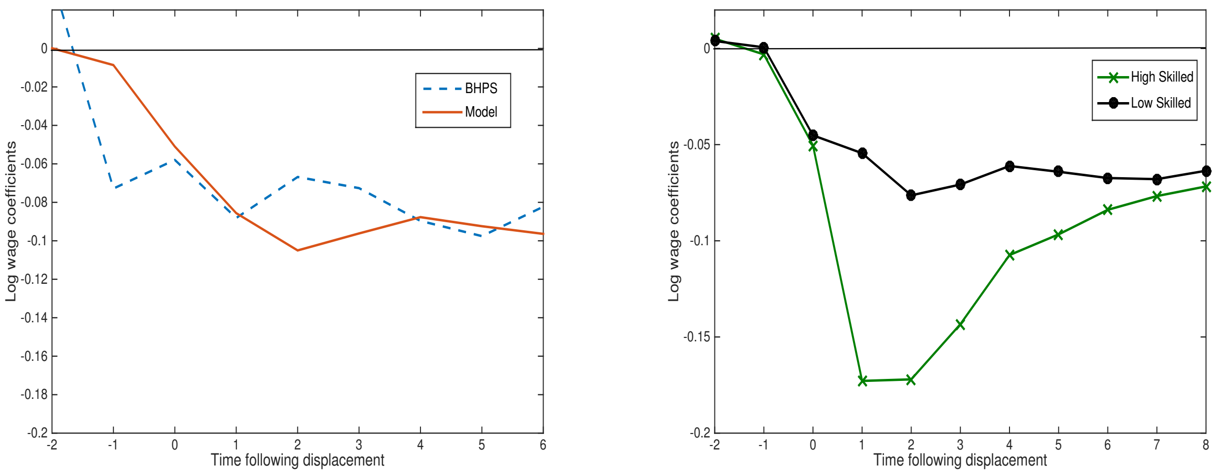
[14]We consider transitions into non-employment as transitions into unemployment and into non-participation, excluding transitions due to maternity leave, retirement, further education or training and self-employment. From those who transited to the non-participation category we are left with those that transited due to health reasons or to care for a family member. Using the LFS we find that on average only 20% of those employed workers who transited into non-employment every quarter declared that they did so because they resigned, the rest did so involuntarily, where we define voluntary and involuntary transitions in Appendix B. Further, note that our definition of who is a displaced worker differs from the one used in Jacobson et al. (1993), Couch and Placzek (2010) and others who use administrative data. In such papers a displaced worker is considered to be someone who lost his/her job due to a mass layoff. Jacobson et al. (1993) showed that these workers are the ones that face large and persistent earnings losses, while the earnings of those workers that went to non-employment due to other reasons recover quite quickly. Topel (1990), Ruhm (1991) and Stevens (1997), using household panel survey data for the US (PSID), and Kletzer and Fairlie (2003), using cohort data for the US (NLSY), apply a similar definition of displaced workers as we do here, finding large and persistent wage losses for these workers. More recently Jarosch (2015), using administrative data for Germany, finds large and persistent wage losses for all those employed workers that transited into non-employment. We find a similar result as these papers.

[15]In Appendix B we provide further details of the sample used, estimation results and robustness exercises.

Table 4: Wage losses using log wages, BHPS: 1991-2004

	All Workers		High Skilled		Low Skilled	
log real wage	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
pot exp	0.0060	2.44e-04	0.0064	2.80e-04	0.0052	5.27e-04
pot exp2	-1.22e-05	7.57e-07	-1.28e-05	8.09e-07	-7.52e-06	1.62e-06
ε_{-2}	0.037	0.023	0.013	0.027	0.122	0.045
ε_{-1}	-0.073	0.022	-0.086	0.027	0.007	0.045
ε_0	-0.058	0.027	-0.050	0.031	-0.022	0.057
ε_1	-0.088	0.021	-0.140	0.024	0.044	0.048
ε_2	-0.067	0.023	-0.069	0.029	-0.029	0.046
ε_3	-0.073	0.025	-0.074	0.030	-0.035	0.052
ε_4	-0.090	0.027	-0.063	0.033	-0.065	0.052
ε_5	-0.098	0.031	-0.115	0.037	-0.032	0.060
ε_6	-0.082	0.033	-0.114	0.040	0.036	0.063
const.	1.279	0.028	1.326	0.031	1.014	0.061

pooled sample case, we weight the sample with high skill and low skill workers to reflect the weights found in the original data. The test of the model is whether the results estimated using the simulated data with worker fixed effects are consistent with those obtained above. For ease of comparison, Figure 3a graphs the point estimates $\widehat{\varepsilon_j}$ described in Column1, Table 4 and the equivalent estimates using the pooled simulated data.



(a) All Workers: Data and Model

(b) Decomposition by Skill Groups: Model

Figure 3: Post displacement wage losses

Results on the BHPS data identify an "Ashenfelter dip" prior to the displacement event; i.e. $\widehat{\varepsilon}_{-1} = -7.3\%$ is significantly negative. This dip suggests the displacement event may (sometimes) be foreseeable, a reasonable interpretation being that employees might previously accept wage cuts in an attempt to forestall displacement. Although the forecasting of possible displacement lies outside the model (the point estimates of $\widehat{\varepsilon}_{-2}$ and $\widehat{\varepsilon}_{-1}$ using simulated data are not statistically different from zero), the model otherwise captures both the extent of the wage losses and their persistence following job displacement.

Figure 3.b instead graphs the estimated $\widehat{\varepsilon}_j$ using the simulated data for each skill group. Reflecting $\Phi^T = 19.1\%$ as identified by the model, high skilled workers face a large, temporary wage loss following job loss. High skilled workers also face a permanent wage loss $\Phi^P = 4.4\%$. Although the temporary wage loss gradually dissipates through time, Figure 3.b suggests the wage recovery rate is slow: a half-life of approximately 4 years. In contrast, job ladder effects for low skilled workers are small: the temporary wage loss through job displacement is only $\Phi^T = 4.6\%$ and the permanent wage loss $\Phi^P = 5.8\%$ is only slightly higher.

6.1 Composition Bias

Using the model-based data and estimated separately by education group, equation (24) identifies a wage loss of around 6-7% by year $j = 6$ for both skill groups. This estimated loss is not very far away from the long term permanent losses Φ^P as described in Table 3. Yet using a pooled sample (of the simulated data), and despite including worker fixed effects, Figure 3.a shows the same estimation procedure suggests the (long term) wage loss is larger, at around 9% by year $j = 6$ and this loss is not decreasing as j becomes large. Estimating (24) on a pooled sample overestimates the long run cost of wage loss.

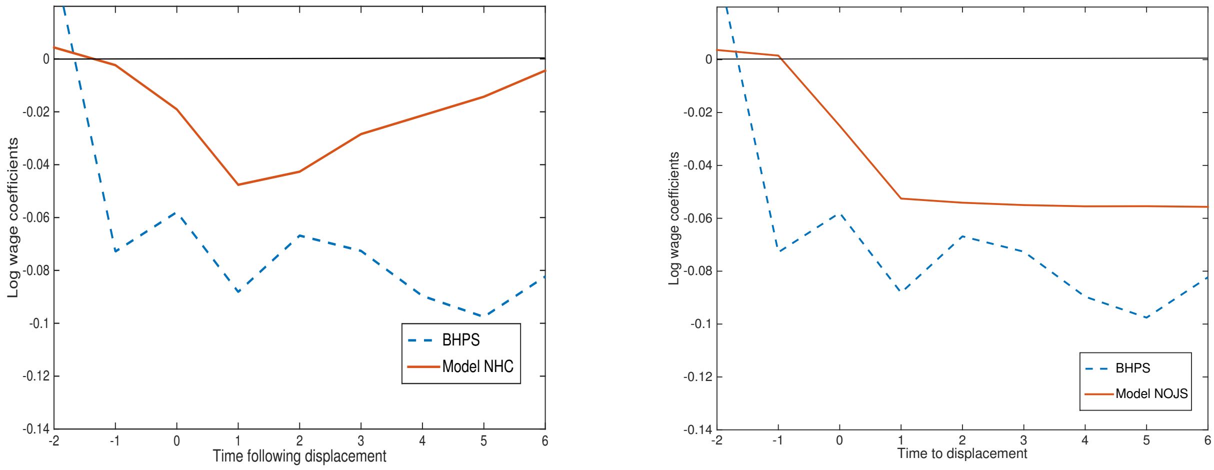
The bias arises here as low educated workers not only have higher layoff rates (and so are over-represented in the job-displacement pool), they also have lower average growth rates of skills, $\gamma_k = 2.3\%$, compared to $\gamma_k = 3.2\%$. As the wage equation (24) attributes the growing skills gap between displaced and non-displaced workers as the result of individual job displacement, this composition effect exaggerates the estimated cost of job loss. It is not clear how a reduced form methodology can control for this composition bias (e.g. Davis and von Wachter, 2011). The issue would appear fundamental. For example, one could argue that high educated workers in the data might also be further decomposed into, say, high and low ability types. If low ability types are both more likely to be laid off and have lower average skills growth, the same composition bias applies and potentially explains why the estimated cost of job loss seems to grow with j large.

6.2 The Importance of Job-Ladder Effects for Explaining Wage Losses

Finally we establish the roles played by both the "job ladder" and human capital dynamics in explaining wage loss dynamics. Consider instead two restricted cases: (a) pure job ladder [i.e no human capital dynamics $\rho = \phi = 0$], (b) pure human capital dynamics [i.e. no job ladder with $\lambda_1 = \lambda_q = 0$]. For each case we re-estimate the model on the relevant subset of targets described in Table 1 and then re-estimate (24) on the (pooled) simulated data.[16] Figure 4 depicts the resulting estimated values of ε_j compared to the results on actual data (as described in Figure

[16]For the no human capital case we target all the original moments except the returns to experience, while for the no on-the-job search case we target only the average unemployment and employment durations, the returns to experience, the ratio of re-employment to average wages for each skill group. For further details see Appendix B.

3.a).



(a) No Human Capital ($\rho = \phi = 0$)

(b) No On-the-job Search ($\lambda_1 = \lambda_q = 0$)

Figure 4: Post displacement wage losses: counterfactual exercise

Without human capital dynamics, the pure job-ladder case (Figure 4.a) fails to explain both the magnitude of the wages losses and the permanent fall in wages following displacement. Without the job ladder, however, the pure human capital dynamics case (Figure 4.b) fails to explain the magnitude of wage losses at short j. Both the job ladder and human capital dynamics play important roles in explaining the magnitude and pattern of the cost of job loss.

7 Conclusion

This paper has identified an equilibrium theory of the labour market in which risk averse workers accumulate human capital through learning-by-doing and engage in on-the-job search. Firms precommit to optimal company wage policies which, in equilibrium, have the property that the wage paid increases with tenure and experience. We show how the restriction to a "timeless equilibrium" generates a highly tractable equilibrium framework of wage formation and turnover. Using indirect inference, we show the model reproduces the returns to tenure and experience as observed in the data, it generates wage dispersion consistent with the Hornstein et al (2011) critique, and reproduces the reduced form estimated losses in wages after a job separation observed for young workers in the UK and for different skill groups. We establish that job ladder effects are large for the highly educated group and so workers in this group face large costs to job loss. In contrast, workers in the low educated sector rarely receive outside job offers and face relatively short employment spells. As outside offers are rare for these workers, job ladder effects are consequently small and their short run fall in wages due to job loss is comparatively small.

We have also highlighted a potentially important composition bias which arises when trying to estimate the cost of job loss using reduced form statistical techniques. The difficulty is that some workers in the sample may not only be more likely to be laid-off into unemployment than others, those types may also have lower average wage growth rates over their working lifetimes. Including worker fixed effects is not sufficient to control for the resulting composition bias: in our example, low skill workers are over-represented in the set of displaced workers and, having low earnings growth rates, a reduced form statistical approach overstates the long run cost of job loss.

References

[1] Altonji, Joseph G. and Robert A. Shakotko. 1987."Do Wages Rise with Job Seniority?" *Review of Economic Studies*, 54(3): 437-459.

[2] Arulampalam, Wiji. 2001. "Is Unemployment Really Scarring? Effects of Unemployment Experiences on Wages," *Economic Journal*, 111(475): F585-606.

[3] Autor, David and Lawrence Katz. 1999. "Changes in the Wage Structure and Earnings Inequality," in Ashenfelter O, Card D *Handbook of Labor Economics*, vol. 3A: 1463-1555.

[4] Bagger, Jesper and Rasmus Lentz. 2015."An Empirical Model of Wage Dispersion with Sorting."Mimeo, University of Wisconsin-Madison, USA.

[5] Bagger, Jesper, Francois Fontaine, Fabien Postel-Vinay and Jean-Marc Robin. 2014. "A Feasible Equilibrium Search Model of Individual Wage Dynamics with Human Capital Accumulation." *American Economic Review*, 104(6): 1551-1596.

[6] Burdett, Kenneth, Carlos Carrillo-Tudela and Melvyn G. Coles. 2016. "Wage Inequality: A Structural Approach." *Review of Economic Dynamics*, forthcoming.

[7] Burdett, Kenneth, Carlos Carrillo-Tudela and Melvyn G. Coles. 2011. "Human Capital Accumulation and Labour Market Equilibrium." *International Economic Review*, 52(3): 657-677.

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

[9] Burdett, Kenneth and Dale T. Mortensen. 1998. "Wage Differentials, Employer Size, and Unemployment."*International Economic Review*, 39(2): 257-273.

[10] Carrillo-Tudela, Carlos, Bart Hobijn, Powen She, and Ludo Visschers. 2015. "The Extent and Cyclicality of Career Changes: Evidence for the U.K,"*European Economic Review*, forthcoming.

[11] Couch, Kenneth A. and Dana W. Placzek. 2010. "Earnings Losses of Displaced Workers Revisited," *American Economic Review*, 100(1): 572-89.

[12] Davis, Steven J. and Till Von Wachter. 2011. "Recessions and the Costs of Job Loss," *Brookings Papers on Economic Activity*, 43(2): 1-72.

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

[14] 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.

[15] Farber, Henry S. 1997. "The Changing Face of Job Loss in the United States, 1981-1995" in *Brookings Papers on Economics Activity*. Microeconomics: 55-128.

[16] Gandelman, Nestor and Ruben Hernandez-Murillo. 2015. "Risk Aversion at the Country Level". *Federal Reserve Bank of St. Louis Review*, 97(1): 53-66.

[17] Gourieroux, Christian, Alain Monfort, and Eric Renault. 1993. "Indirect Inference." *Journal of Applied Econometrics*, 8(S): S85-118.

[18] Gregory, Mary and Robert Jukes. 2001. "Unemployment and Subsequent Earnings: Estimating Scarring among British Men 1984-94," *Economic Journal*, 111(475): F607-25.

[19] 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.

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

[21] Jacobson, Louis, Robert LaLonde and Daniel Sullivan. 1993. "Earnings Losses of Displaced Workers. "*American Economic Review*, 83(4): 685-709.

[22] Jarosh, Gregor. 2015. "Searching for Job Security and the Consequences of Job Loss". Mimeo, University of Chicago, USA.

[23] Jung, Philip and Moritz Kuhn. 2014. "Earnings Losses and Labor Mobility Over The Lifecycle". Mimeo, University of Bonn, Germany.

[24] Kletzer, Lori G. and Robert W. Fairlie. 2003. "The Long-term Costs of Job Displacement for Young Adult Workers," *Industrial and Labor Relations Review, ILR Review*, 56(4): 682-698.

[25] Kletzer, Lori G. and Robert W. Fairlie. 1997. "The Long-term Costs of Job Displacement for Young Adult Workers," Mimeo, University of California, Santa Cruz.

[26] Krolikowski, Pawel. 2014. "Job Ladders and Earnings of Displaced Workers", Mimeo, University of Michigan, USA.

[27] Lise, Jeremy and Jean-Marc Robin. 2015. "The Macro-dynamics of Sorting between Workers and Firms". Mimeo, University College London, UK.

[28] Ljungqvist, Lars and Thomas Sargent. 1998. "The European Unemployment Dilemma," *Journal of Political Economy*, 106(3): 514-550.

[29] Manning, Alan. 2003. *Monopsony in Motion: Imperfect Competition in Labor Markets.* Princeton: Princeton University Press, USA.

[30] 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.

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

[32] Mortensen, Dale and Pissarides, Christopher A. 1994. "Job Creation and Job Destruction in the Theory of Unemployment," *Review of Economic Studies*, 61(3): 397-415.

[33] Moscarini, Giuseppe and Fabien Postel-Vinay. 2013. "Stochastic Search Equilibrium," *Review of Economic Studies*, 80(4): 1545-1581.

[34] Neal, Derek. 1995. "Industry-Specific Human Capital: Evidence from Displaced Workers." *Journal of Labor Economics*, 13(4): 653-677.

[35] Pissarides, Christopher A. 1992. "Loss of Skill During Unemployment and the Persistence of Employment Shocks." *Quarterly Journal of Economics*, 107(4): 1371-1391.

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

[37] Robin, Jean-Marc. 2011. "On the Dynamics of Unemployment and Wage Distributions". *Econometrica*, 79(5): 1327-1355.

[38] Ruhm, Christopher J. 1991. "Are Workers Permanently Scarred by Job Displacements?" *American Economic Review*, 81(1): 319-324.

[39] Spear, Stephen E. and Sanjay Srivastava. 1987. "On Repeated Moral Hazard with Discounting," *The Review of Economic Studies*, 54(4): 599-617.

[40] Stevens, Ann Huff. 1997. "Persistent Effects of Job Displacement: The Importance of Multiple Job Losses." *Journal of Labor Economics*, 15(1): 165-188.

[41] Stevens, Margaret. 2004. "Wage-Tenure Contracts in a Frictional Labour Market: Firms' Strategies for Recruitment and Retention," *Review of Economic Studies*, 71(2): 535-551.

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

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

[44] Topel, Robert. 1990. "Specific Capital and Unemployment: Measuring the Costs and Consequences of Job Loss." *Carnegie-Rochester Conference Series on Public Policy*, 33: 181-214.

[45] Williams, Nicolas. 2009. "Seniority, experience, and wages in the UK." *Labour Economics*, 16(3): 272-283.

[46] Woodford, Michael. 2003. *Interest and Prices: Foundations of a Theory of Monetary Policy.* Princeton: Princeton University Press, USA.

Appendix

A Proofs

Proof of Theorem 1: Substituting out ψ in the objective functions gives the dynamic optimisation problem:

$$\max_{\theta(.)} \int_0^\infty e^{-\int_0^\tau \{r+\delta+\lambda_q-\rho-\gamma_A+\lambda_1[1-F(U(s|\theta))]\}ds}[1-\theta(\tau)]d\tau,$$

subject to starting value $U(0|\theta) = U_0$ where (2) describes how $U(.)$ evolves with tenure. Define transformed variable

$$\psi_0 = e^{-\int_0^\tau \{r+\delta+\lambda_q-\rho-\gamma_A+\lambda_1[1-F(U(s|\theta))]\}ds}$$

and note it satisfies the differential equation

$$\dot{\psi}_0 = -\left[r+\delta+\lambda_q-\rho-\gamma_A+\lambda_1[1-F(U)]\right]\psi_0. \tag{25}$$

The dynamic optimisation problem is equivalently rewritten as

$$\max_{\theta(.)} \int_0^\infty \psi_0[1-\theta]d\tau, \tag{26}$$

where ψ_0, U are state variables which evolve according to the autonomous, first order differential equations (25) and (2) respectively with initial values $\psi_0 = 1$, $U = U_0$ at $\tau = 0$. We can solve this dynamic optimisation problem using the Hamiltonian approach. Define

$$\begin{aligned} H &= \psi_0[1-\theta] - \xi_{\psi_0}\left[r+\delta+\lambda_q-\rho-\gamma_A+\lambda_1[1-F(U)]\right]\psi_0 \\ &\quad +\xi_U\left[\begin{array}{c} [r+\delta+\lambda_q-[\rho+\gamma_A](1-\sigma)]\,U \\ -\left[\frac{\theta^{1-\sigma}}{1-\sigma}+\delta\phi^{1-\sigma}U^U+\lambda_1\int_U^{\overline{U}}[1-F(U_0)]dU_0+\lambda_q\int_{\underline{U}}^{\overline{U}}U_0dF(U_0)\right] \end{array}\right] \end{aligned}$$

where ξ_{ψ_0}, ξ_U are the respective costate variables. The Maximum principle yields the following necessary conditions for optimality:

$$\theta^{-\sigma} = -\frac{\psi_0}{\xi_U} \tag{27}$$

$$\dot{\xi}_{\psi_0} = -[1-\theta] + \xi_{\psi_0}\left[r+\delta+\lambda_q-\rho-\gamma_A+\lambda_1[1-F(U)]\right] \tag{28}$$

$$\dot{\xi}_U = -\left[\xi_{\psi_0}\lambda_1F'(U)\psi_0 + \xi_U\left[[r+\delta+\lambda_q-[\rho+\gamma_A](1-\sigma)]+\lambda_1[1-F(U)]\right]\right] \tag{29}$$

along with autonomous differential equations (25), (2) for $\dot{\psi}_0$ and $\dot{U}$. As we do not wish to assume F is differentiable, however, we drop condition (29) and instead note that as the objective function in (26) does not depend explicitly on tenure, optimality also implies

$$H = 0 \tag{30}$$

(e.g. p298, Leonard and Long, 1992). Now integrating (28) forward yields:

$$\begin{aligned}\xi_{\psi_0}(t) &= \int_t^{\infty} e^{-\int_t^s [r+\delta+\lambda_q-\rho-\gamma_A+\lambda_1[1-F(U(\tau))]d\tau}(1-\theta(s))ds + B_0 e^{\int_0^t [r+\delta+\lambda_q-\rho-\gamma_A+\lambda_1(1-F(U(x)))]dx} \\ &= \Pi(t) + B_0 e^{\int_0^t [r+\delta+\lambda_q-\rho-\gamma_A+\lambda_1(1-F(U(x)))]dx}\end{aligned}$$

where B_0 is the constant of integration and $\Pi(.)$ is the firm's continuation profit as defined in Theorem 1. (27) implies $\xi_U = -\psi_0\theta^{\sigma}$. Substituting out ξ_U and ξ_{ψ_0} in the definition of H, the restriction $H = 0$ yields the optimality condition:

$$\begin{aligned}0 &= [1-\theta] - [r+\delta+\lambda_q-\rho-\gamma_A+\lambda_1[1-F(U)]]\left[\Pi(t) + B_0 e^{\int_0^t [r+\delta+\lambda_q-\rho-\gamma_A+\lambda_1(1-F(U(x)))]dx}\right] \\ &\quad -\theta^{\sigma}\left[[r+\delta+\lambda_q-[\rho+\gamma_A](1-\sigma)]\,U - \frac{\theta^{1-\sigma}}{1-\sigma} - \delta\phi^{1-\sigma}U^U - \lambda_1\int_U^{\overline{U}}[1-F(U_0)]dU_0 - \lambda_q\int_{\underline{U}}^{\overline{U}} U_0 F(U_0\right]\end{aligned} \tag{31}$$

Now the restriction $r+\delta-\rho-\gamma_A > 0$ ensures the exponential term becomes arbitrarily large as $\tau \to \infty$. As Π and U must be bounded, then (31) implies $B_0 = 0$. (31) now yields (6) given in the Theorem. Using this to substitute out $\frac{\theta^{1-\sigma}}{1-\sigma}$ in (2) then yields (8). This completes the proof of Theorem 1.

Proof of Claim 3: Equation (12) follows by solving the constant profit condition. To do so, note that standard turnover arguments imply $G(U)$ satisfies

$$\begin{aligned}& u\lambda_0[1-F(U)] + (1-u)G(U)[\lambda_q+\lambda_1][1-F(U)] + (1-u)G'(U)\dot{U} \\ =\;& (1-u)(1-G(U))[\delta+\lambda_q F(U)],\end{aligned}$$

where the left hand side describes the flow of workers into employment with piece rate value more than U while the right hand side describes the flow out through job separation. As (8) and (10) together imply

$$\dot{U} = \widehat{\theta}^{-\sigma}\left\{1-\theta-[r+\delta+\lambda_q-\rho-\gamma_A+\lambda_1[1-F]]\Pi\right\},$$

rearranging the previous expression yields

$$\frac{dG}{dU} = \frac{(1-u)\delta(1-G(U)) - u\lambda_0[1-F(U)] - (1-u)G(U)\lambda_1[1-F(U)]}{(1-u)\theta^{-\sigma}\left[1-\theta-[r+\delta-\rho-\gamma_A+\lambda_1(1-F(U))]\Pi\right]}, \tag{32}$$

where $\Pi = \Pi_i(t_0)$ and $\theta = \theta_i(t_0)$.

While $dF(U) > 0$, differentiating the constant profit condition implies:

$$[\lambda_0 u + \lambda_q(1-u) + \lambda_1(1-u)G(U)]\,\widehat{\Pi}'(U) + \lambda_1(1-u)G'(U)\widehat{\Pi}(U) = 0.$$

As (8) implies

$$\frac{d\widehat{U}}{d\theta} = -\theta^{-\sigma}\frac{d\widehat{\Pi}}{d\theta}, \tag{33}$$

and using (32) to substitute out $\widehat{\Pi}'(U)$ and $G'(U)$ gives

$$
\begin{aligned}
& [\lambda_0 u + \lambda_q(1-u) + \lambda_1(1-u)G]\left\{1-\widehat{\theta} - [r+\delta+\lambda_q-\rho-\gamma_A+\lambda_1[1-F]]\widehat{\Pi}\right\} \\
= \; & \lambda_1(1-u)\widehat{\Pi}\left[\delta(1-G) - \frac{u}{1-u}\lambda_0[1-F] - G\lambda_1[1-F] + \lambda_q[F-G]\right].
\end{aligned}
$$

Inspection finds the F-terms all cancel out and so:

$$
\begin{aligned}
& [\lambda_0 u + \lambda_q(1-u) + \lambda_1(1-u)G]\left\{1-\widehat{\theta} - [r+\delta+\lambda_q-\rho-\gamma_A+\lambda_1]\widehat{\Pi}\right\} \\
= \; & \lambda_1(1-u)\widehat{\Pi}\left[\delta(1-G) - \frac{u}{1-u}\lambda_0 - G\lambda_1 - G\lambda_q\right].
\end{aligned} \tag{34}
$$

But the constant profit condition also implies

$$
G(U) = \frac{[\lambda_0 u + \lambda_q(1-u) + \lambda_1(1-u)]\frac{\overline{\Omega}}{\widehat{\Pi}} - [\lambda_0 u + \lambda_q(1-u)]}{\lambda_1(1-u)}.
$$

Using this to substitute out G in (34) and substituting out $u = \delta/(\delta+\lambda_0)$ yields the quadratic equation

$$
\widehat{\Pi}^2[\delta+\lambda_q] + [r-\rho-\gamma_A]\overline{\Omega}\widehat{\Pi} - (1-\widehat{\theta})\overline{\Omega} = 0. \tag{35}
$$

As $dF(U) > 0$ implies the firm must make positive profit $\widehat{\Pi} > 0$, the positive root to this quadratic equation yields the result. This completes the proof of Claim 3.

Proof of Theorem 2: (13) and (16) imply

$$
\frac{d\widehat{U}}{d\theta} = \frac{\theta^{-\sigma}}{\left[[r-\rho-\gamma_A]^2 + 4[\delta+\lambda_q][[\delta+\lambda_q+r-\rho-\gamma_A]]\frac{1-\theta}{1-\overline{\theta}}\right]^{1/2}}, \tag{36}
$$

whose solution is given by (17). Given $\widehat{\Pi}(\theta)$, the constant profit condition (**??**) implies (18).

To determine the equilibrium distribution of offers F_θ, standard turnover arguments imply G_θ must satisfy

$$
u\lambda_0[1-F_\theta(\theta)] + (1-u)G_\theta(\theta)[\lambda_1+\lambda_q][1-F_\theta(\theta)] + (1-u)G'_\theta(\theta)\dot{\theta}(\theta) = (1-u)(1-G_\theta(\theta))\left[\delta+\lambda_q F_\theta(\theta)\right],
$$

where the left hand side describes the flow of workers into employment with piece rate more than θ while the right hand side describes the flow out through job separation. Now (9), (13) and $F'_\theta(\theta) = F'(\widehat{U})d\widehat{U}/d\theta$ together imply

$$
\dot{\theta} = \frac{\lambda_1 F'_\theta}{\sigma}\left[-\frac{\theta\widehat{\Pi}}{\frac{d\widehat{\Pi}}{d\theta}}\right] - (\rho+\gamma_A)\theta.
$$

Using this solution for $\dot{\theta}$ and $G_\theta, \widehat{\Pi}$ described in the Theorem, a lot of algebra finds the turnover equation for G implies the following first order differential equation for F:

$$\frac{\theta F_\theta'}{\sigma} + (1 - F_\theta) = \Psi(\theta),$$

where

$$\Psi(\theta) = \frac{\lambda_q + \delta}{\lambda_1}\left[\frac{\widehat{\Pi}(\theta) - \overline{\Omega}}{\overline{\Omega}}\right] - \left[\frac{(\rho + \gamma_A)}{\lambda_1}\frac{\theta d\widehat{\Pi}/d\theta}{\widehat{\Pi}(\theta)}\right] > 0.$$

Integration now yields the stated solution for F_θ while using (12) it is easy to show that $\Psi(\theta) > 0$ for all $\theta \in [\underline{\theta}, \overline{\theta}]$. This completes the proof of Theorem 2.

Proof of Claim 6: Integration by parts finds

$$\begin{aligned}\int_{\underline{\theta}}^{\overline{\theta}} \widehat{U}(\theta) dF_\theta(\theta) &= \overline{U} - \int_{\underline{\theta}}^{\overline{\theta}} \frac{d\widehat{U}}{d\theta} F_\theta(\theta) d\theta \\ &= \overline{U} - \int_{\underline{\theta}}^{\overline{\theta}} \frac{[\theta]^{-\sigma} F_\theta(\theta) d\theta}{\left[[r - \rho - \gamma_A]^2 + 4[\delta + \lambda_q]\left[\delta + \lambda_q + r - \rho - \gamma_A\right] \frac{1-\theta}{1-\overline{\theta}}\right]^{1/2}}. \quad (37)\end{aligned}$$

Putting $\theta = \underline{\theta}$ in (17) implies

$$\overline{U} = \underline{U} + \int_{\underline{\theta}}^{\overline{\theta}} \frac{\theta^{-\sigma} d\theta}{\left[[r - \rho - \gamma_A]^2 + 4[\delta + \lambda_q]\left[\delta + \lambda_q + r - \rho - \gamma_A\right] \frac{1-\theta}{1-\overline{\theta}}\right]^{1/2}}. \quad (38)$$

Putting $\theta = \overline{\theta}$ in (6), noting $U^U = \underline{U}$ in a timeless equilibrium (Claim 5) implies:

$$\frac{\overline{\theta}^{1-\sigma}}{1-\sigma} = [r + \delta + \lambda_q - [\rho + \gamma_A](1 - \sigma)]\,\overline{U} - \delta\underline{U} - \lambda_q \int_{\underline{\theta}}^{\overline{\theta}} \widehat{U}(\theta) dF_\theta(\theta). \quad (39)$$

Using (37) and (38) to substitute out $\overline{U}$ implies:

$$[r - [\rho + \gamma_A](1 - \sigma)]\,\underline{U} = \frac{\overline{\theta}^{1-\sigma}}{1-\sigma} - \int_{\underline{\theta}}^{\overline{\theta}} \frac{[r + \delta - [\rho + \gamma_A](1 - \sigma) + \lambda_q F_\theta(\theta)]\,\theta^{-\sigma} d\theta}{\left[[r - \rho - \gamma_A]^2 + 4[\delta + \lambda_q]\left[\delta + \lambda_q + r - \rho - \gamma_A\right] \frac{1-\theta}{1-\overline{\theta}}\right]^{1/2}}. \quad (40)$$

Equation (3) with $U^U = \underline{U}$ [Claim 5], (37) and substituting out $\overline{U}$ using (38) implies

$$(r + \phi(1-\sigma) - \gamma_A(1-\sigma))U^U = \frac{b^{1-\sigma}}{1-\sigma} + \int_{\underline{\theta}}^{\overline{\theta}} \frac{\theta^{-\sigma}\lambda_0[1 - F_\theta(\theta)]}{\left[[r - \rho - \gamma_A]^2 + 4[\delta + \lambda_q]\left[\delta + \lambda_q + r - \rho - \gamma_A\right] \frac{1-\theta}{1-\overline{\theta}}\right]^{1/2}} d\theta. \quad (41)$$

As a timeless equilibrium requires $\underline{U} = U^U$, we obtain the equilibrium condition stated with $F = \widetilde{F}$. This completes the proof of Claim 6.

Proof of Theorem 3: Note that as $\overline{\theta} \to 1$, (20) implies $\underline{\theta} \to 1$ and so all piece rates paid lie in a neighbourhood of 1. Frictions $\lambda_0 < \infty$, $b < 1$ and $\phi \geq 0$ ensure the value of being unemployed $\widetilde{U}^U(\overline{\theta}) < \underline{\widetilde{U}}(\overline{\theta})$ in this limit.

Suppose instead $\overline{\theta} \to 1 - \frac{\delta[\delta+r-\rho-\gamma_A]}{(\delta+\lambda_1)[\delta+\lambda_1+r-\rho-\gamma_A]}$ and so $\underline{\theta} \to 0^+$. As

$$\begin{aligned}
&\int_{\underline{\theta}}^{\overline{\theta}} \frac{[r+\delta-(\rho+\gamma_A)(1-\sigma)][\theta']^{-\sigma}}{\left[[r-\rho-\gamma_A]^2+4\delta\left[\delta+r-\rho-\gamma_A\right]\frac{1-\theta'}{1-\overline{\theta}}\right]^{1/2}}d\theta' \\
> &\int_{\underline{\theta}}^{\overline{\theta}} \frac{[r+\delta-(\rho+\gamma_A)(1-\sigma)][\theta']^{-\sigma}}{\left[[r-\rho-\gamma_A]^2+4\delta\left[\delta+r-\rho-\gamma_A\right]\frac{1-\underline{\theta}}{1-\overline{\theta}}\right]^{1/2}}d\theta' \\
= &\frac{[r+\delta-(\rho+\gamma_A)(1-\sigma)]}{\left[[r-\rho-\gamma_A]^2+4(\delta+\lambda_1)\left[\delta+\lambda_1+r-\rho-\gamma_A\right]\right]^{1/2}}\left[\frac{\overline{\theta}^{1-\sigma}}{1-\sigma}-\frac{\underline{\theta}^{1-\sigma}}{1-\sigma}\right]
\end{aligned}$$

then $\underline{\theta} \to 0^+$ and (40) implies $\underline{\widetilde{U}}(\overline{\theta}) \to -\infty$ in this limit. As (41) implies $\widetilde{U}^U \geq \frac{b^{1-\sigma}}{1-\sigma}/(r + \phi(1-\sigma) - \gamma_A(1-\sigma))$ and so is finite, then $\widetilde{U}^U(\overline{\theta}) > \underline{\widetilde{U}}(\overline{\theta})$ in this limit. As these are continuous functions for $\overline{\theta}$ satisfying (21), a $\overline{\theta}$ satisfying (21) exists such that $\underline{\widetilde{U}}(\overline{\theta}) = \widetilde{U}^U(\overline{\theta})$. This completes the proof of Theorem 3.

Proof of Proposition 1.

Consider $t = \Delta$ where $\Delta > 0$ but very small. At time Δ, the initially employed worker with probability $1 - \delta\Delta$ remains employed and has human capital $k^e_\Delta = k_0 \exp[\rho\Delta]$. At time Δ, the initially unemployed worker with probability $1-\lambda_0\Delta$ remains unemployed and has human capital $k^u_\Delta = \exp[-\phi\Delta]k_0$. Note that if the employed worker becomes unemployed there is no further difference in their expected human capital dynamics, similarly if the unemployed worker finds a job, while the probability they both switch states is $0(\Delta^2)$. Furthermore as $\log k^e_\Delta = \rho\Delta + \log k_0$ while $\log k^u_\Delta = -\phi\Delta + \log k_0$, then the law of iterated expectations implies Φ^P must satisfy the recursive relationship

$$\Phi^P = \rho\Delta + \phi\Delta + [1-\delta\Delta][1-\lambda_0\Delta]\Phi^P + 0(\Delta^2).$$

Rearranging and taking the limit $\Delta \to 0$ yields the result. This completes the proof of Proposition 1.

B Quantitative Analysis

B.1 Simulation

To recover the parameters of the model we solve $min_\Omega \sum_{i=1}^{14} \left[\frac{M_i^S - M_i^D}{M_i^D}\right]^2$, where M_i^S denotes the i^{th} moment obtained from the simulations and M_i^D the corresponding data moment. To obtain the empirical moments, M_i^D, we use data drawn from the LFS and the BHPS as mentioned in the main text. To obtain the simulated moments at each iteration, we first compute the equilibrium of our model, then simulate workers' employment histories, and then compute each M_i^S from this data.

For a set of parameter values, computing the equilibrium implies picking a $\overline{\theta}$ satisfying (21), then using Theorem 2 to compute F_θ over $[\underline{\theta}, \overline{\theta}]$ with $\underline{\theta}$ given by (20). Then computing $\underline{\widetilde{U}}(\overline{\theta}), \widetilde{U}^U(\overline{\theta})$ as defined in Claim 6. The equilibrium value of $\overline{\theta}$ is then determined by $\underline{\widetilde{U}}(\overline{\theta}) =$

$\widetilde{U}^U(\overline{\theta})$. Using the corresponding value of $\underline{\theta}$ we then solve the differential describing the evolution of θ to obtain the baseline salary scale.

Given these equilibrium outcomes, we simulate the employment histories of 10,000 workers. We assume that all workers start unemployed and experience different types of shocks during their lifetime depending on the worker's employment status. In particular, every time a worker is unemployed, he receives a job offer at rate λ_0. We obtain his unemployment duration by drawing a random number, $r1 \in [0, 1]$ and then exploiting 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 $tu = -log(1 - r1)/\lambda_0$. After deriving tu, we sample a position in the baseline salary scale from the offer distribution F, by choosing a random number between 0 and 1 and interpolating between the sample value of F and the corresponding value of θ.

When the worker is employed, he faces three shocks: a reallocation shock, a job offer shock and a displacement shock. All these shocks follow Poisson process with rates, λ_q, λ_1 and δ, respectively. What is important here is to track 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. To obtain these durations we need to compute the durations until the worker receives a job offer tj, receives a displacement shock, tu, or receives a reallocation shock, tr. We do this by drawing three random number between 0 and 1 and using the inverse of the corresponding exponential distribution. The job duration until the worker experiences one of these three events is equal to the $min\{tj, tu, tr\}$. If the worker becomes unemployed, $tu = min\{tj, tu, tr\}$, we repeat the corresponding procedure described in the above paragraph. If the worker receives an outside offer, $tj = min\{tj, tu, tr\}$, we draw a new position in the baseline salary scale. If the current position is greater than the one drawn, the worker stays employed in his current job. Otherwise, we move the worker to the new position and compute a new set $\{tj, tu, tr\}$. If the worker receives a reallocation shock, $tr = min\{tj, tu, tr\}$, we obtain a new position in the baseline salary scale, move the worker to the new position and compute a new set $\{tj, tu, tr\}$. When we compute these events, we also compute workers' labour market experience defined as the sum of employment spells. This information, together with the length of the worker's unemployment spells, can then be used to compute wages at each point in which an event has occurred taking into account that human capital accumulation occurs at rate ρ and human capital depreciation occurs at rate ϕ.

We follow workers for 40 years as this guarantees that we converge to the ergodic distributions for each M_i^S. To compute the transition moments, we use average durations, except for the average Invol/vol EE transitions, for which we compute the average number of involuntary and voluntary transitions. To compute the returns to experience and tenure, we construct a yearly panel resembling the BHPS structure and regress log wages on a constant, a quadratic on experience and a quadratic on tenure. We also use this panel to compute the Mm ratio and the ratio between the average re-employment wage to average wage. The latter two follow the same procedure as we use to compute these moments from the data.

After the the simulated moments are computed, we use the loss function described earlier to compute the sum of square percentage distance for each data moment and its simulated counterpart. If the value of the loss function is high enough, a new set of parameter values are chosen and the above procedure is repeated, iteratively until the value of the loss function is sufficiently closed to zero. To choose the new set of parameter values at each iteration we use the Nelder-Meade (simplex) algorithm as implemented in Matlab.

B.2 Data Moments

In this section we describe the procedure we followed to compute the data moments. To compute the returns to experience and tenure we estimate workers' wage-experience profile using the BHPS for the period 1991-2004. In particular, we obtain these returns from estimating the following equation

$$\log w_{ijt} = \beta X_{it} + \eta_{ijt}, \tag{42}$$

where X is a vector of covariates consisting of a quadratic on potential experience, a quadratic on tenure, a dummy for marital status, 8 regional dummies, 8 (one-digit) occupational dummies, 8 (one-digit) industry dummies, dummies for cohort effects and a time trend, and η denotes white noise and is assumed to be normally distributed. One potential worry with the above specification is that workers' unobservable characteristics might be biasing the estimated returns to general experience because, for example, more able workers could be more likely to receive outside offers than less able workers in the data. The results of Dustmann and Pereira (2008), however, suggest that any potential bias of this sort is very small. These authors estimate returns to experience for the UK using the BHPS by skill/education categories. When controlling for worker and job match (unobservable) fixed effects, their estimated experience effects hardly change across specification and estimation methods. See also Williams (2009). Furthermore, as suggested by Dustmann and Pereira (2008) and Williams (2009) we incorporate yearly dummies into the wage regressions to control for the presence of a macro trend. We find that the latter is important for high skilled workers as without it the estimated returns to experience nearly double in size. In addition we also incorporate cohort dummies into our wage regressions. In this case, we find that these dummies have a very small impact on the estimated returns to experience across skill groups.

To compute the Mm ratio we follow Hornstein, et al. (2007) and first estimate the wage equation (42) for each year of the sample period and skill group using OLS and the same covariates in X. We then 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. Finally, we 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. For each skill group, we estimate a set of three Mm ratios using the minimum observed wage, the wage at the first percentile, and fifth percentile. 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 for the high skilled category, we use as a target the Mm ratio obtained from averaging the ones obtained for the first and fifth percentile.

As pointed out by Hornstein, et al. (2007) the danger with their approach is that one may underestimate the amount of frictional wage dispersion when controlling for those worker characteristics that also provide information on generate wage dispersion due to productivity differentials among workers. Further, by introducing a polynomial on experience and tenure in (42) one is reducing the effects of on-the-job search and human capital accumulation on wage dispersion. However, in the data this reduction is not very strong and hence the downward bias does not have a mayor impact on the estimated parameters. Indeed, when estimating (42) without controlling for experience or tenure effects, the resulting average Mm ratios are 1.57 and 1.54, for low and high skilled workers, respectively. These Mm ratios are only slightly different than the ones used as targets in the simulations, reported in the main text.

We also use the BHPS to compute the ratio between the average re-employment wage to average wage. This is used as a measure of the short-term wage losses due to job displacement.

To compute the labour market transition moments we use the 2-quarter sample of the U.K. LFS. The non-employment to employment transition rate is computed as the ratio between the flow of those non-employed workers in a given quarter q gaining employment the following quarter and the stock of non-employed workers in quarter q. The employment to non-employment transition rate is computed as the ratio between the flow of those employed workers in a given quarter q that lost their jobs the following quarter and the stock of employed workers in quarter q. The inverse of these rate gives the average duration of a non-employment and of an employment spell. We also use the LFS to obtain the average duration of a job. In this case, we use directly the question asked to employed workers about the length of their current job spell. To construct the ratio of involuntary to voluntary employer to employer transitions, we use the reported reasons of why workers left their current job and construct three groups. The *voluntary* group consists of those employed workers who changed jobs because they "resigned". The *involuntary* group consists of those employed workers who left their last job because they were "dismissed", "made redundant/took voluntary redundancy", "temporary job finished", "gave up work for health reasons" or "gave up for family or personal reasons". The *other* group consists of those employed worker who left their last job because they "took early retirement", "retire" or due to "other reasons". We then use the first two for our statistic.

B.3 Empirical Wage Losses

Our analysis of wage losses focus on a sample of workers that in 1991 were between 16 and 30 years old and that we follow until 2004. From this sample we dropped all non-white workers, workers with employment spells in the government, with spells in training or full-time education and with spells in self-employment. We also drop reported spells of employment and non-employment that were shorter than a month. This selection left us with 1,911 individuals and 21,875 individual/spell observations. Using this sample we estimate (24) using fixed effects, which controls for differences in gender, skills and other fixed observable and unobservable characteristics. See Stevens (1997) for a similar procedure using the PSID.

Table 4, in the main text, shows the full regression estimates, based on the set of displacement years $y = \{1995, 1996, 1997\}$. The set of displacement years considered in y allows us to control for the 2 years previous the displacement event and go forward for 6 years after displacement. As an alternative we also considered the set of displacement years $y = \{1994, 1995, 1996, 1997\}$ without much variation in our results. We also considered a cutoff of 8 years after displacement instead of 6 years. In this case, the estimates for ε_7 and ε_8 are statistically significant when we consider wage levels, but lose significance when considering log wages. In both cases the pattern of the ε_j documented in the baseline case is kept unchanged when considering $T_e = 8$.

In addition we follow the insights of Neal (1995) and also estimate equation (24), controlling for 2-digit industry dummies and once again the pattern of the ε_j documented in the baseline case is kept unchanged. Following Davis and von Wachter (2011), we also run separate regressions for different value of y. In particular, we considered $y = \{94/95, 95/96, 96/97\}$ and then average out the resulting estimated ε_j. This approach gives similar estimates when estimating (24) using wage levels but reduces the significance of the estimates. However, it does give statistically significant coefficients when using log wages. Finally, the estimates presented in Table 4 are obtained by imputing to the displacement year the wage the worker earned if employed at the end of that year. If the worker was still unemployed, then a missing observation is assigned. This approach is similar to the one used in the empirical literature that measures the long-term earnings losses of displaced workers (see the reference in the main text). As an alternative way of treating the data, we re-estimated (24) without imputing a wage to the displacement year and treating wages in the displacement year as missing values. The pattern of wage losses that emerges in this case is very similar to the one documented for the baseline case.

Finally we estimate (24) based on wage levels instead of log wages. In this case we also observe the two features highlighted in the main text: (i) Young workers in the UK have large and persistent displacement wage losses. (ii) When measuring wage losses in relation to pre-displacement wages, the size of these losses increases over time and do not seem to recover. Once again we find that the fit of the model is very good when pooling the data sets for high and low skilled workers together, capturing the two features described above.

B.4 Estimation of the Restricted Models

To understand the importance of modelling human capital dynamics together with workers' on-the-job search in matching the observe wage losses obtained from the BHPS, we re-estimate our model for two restricted cases: (a) no human capital dynamics $\rho = \phi = 0$, but on-the-job search as previously assumed; (b) no job ladder; i.e. $\lambda_1 = \lambda_q = 0$, but human capital dynamics as previously assumed. The targets in the first case are the same as in the baseline case with the exception of the moments characterising returns to experience. For the second case we target only the average unemployment and employment durations, the returns to experience, the ratio of re-employment to average wages. The parameters to estimate for the no human capital case are all of the previous parameters except for ρ and ϕ. The parameters for the case without on-the-job search are δ, λ_0, ρ, ϕ and b. For this case we consider a logarithmic utility function (i.e. $\sigma = 1$) as this implies we can let λ_1 go to zero while generating a flat baseline scale and guaranteeing existence of equilibrium. As in the baseline case, we perform the estimation separately for each skill group.

Table 5: Targeted Moments for the Restricted Models

Moments	Low Skilled			High Skilled		
	Data	NHC	NOJS	Data	NHC	NOJS
Average transitions / duration						
Non-employment spell (months)	12.67	12.76	13.04	8.74	8.64	8.69
Employment spell (years)	3.78	3.94	3.76	12.60	12.52	12.73
Job spell (years)	3.10	2.97	n.a.	5.03	4.93	n.a.
Invol/vol EE transitions	0.57	0.55	n.a.	0.50	0.51	n.a.
OLS returns to experience (%)						
2 years	4.47	n.a.	3.93	4.58	n.a.	4.03
4 years	8.62	n.a.	7.97	8.75	n.a.	8.08
6 years	12.47	n.a.	12.11	12.49	n.a.	12.14
8 years	16.00	n.a.	16.35	15.81	n.a.	16.21
10 years	19.23	n.a.	20.70	18.71	n.a.	20.29
OLS returns to tenure (%)						
2 years	4.19	3.21	n.a.	2.27	2.99	n.a.
4 years	7.29	6.10	n.a.	3.84	5.61	n.a.
6 years	9.31	8.68	n.a.	4.70	7.87	n.a.
Wage dispersion						
Mm ratio	1.62	1.60	n.a.	1.62	1.61	n.a.
re-emp wage / mean wage	0.91	0.95	0.99	0.84	0.89	0.97

Note n.a.= not applicable.

Table 5 show the fit of the model for each skill group for each of the cases discussed above, where NHC refers to the model without human capital dynamics and NOJS refers to the model without a job ladder. Note that the model without human capital, generates even higher returns to tenure for high skilled workers, but improves its fit on this dimension for low skilled workers. Further, the model without on-the-job search has a very hard time matching the re-employment to average wage ratio for both skill groups. That is, the model without on-the-job search misses completely the short-term losses of job displacement. The model without human capital is able to capture this feature somewhat better, albeit still generating too little short-term losses. Table 6 shows the parameter values resulting from the estimations. Interestingly, this table shows that the models without human capital accumulation need very higher values of risk aversion relative to the baseline case to be able to match the data.

Table 6: Parameters for the Restricted Models

	Low Skilled		High Skilled	
Parameters	NHC	NOJS	NHC	NOJS
δ	0.0212	0.0222	0.0067	0.0065
λ_0	0.0752	0.0744	0.1116	0.1138
λ_1	0.0115	0	0.0426	0
λ_q	0.0023	0	0.0040	0
ρ	0	0.0038	0	0.0037
ϕ	0	0.0022	0	0.0017
σ	8.039	1.001	7.030	1.001
b	0.2958	0.3357	0.4389	0.3349